USC-SIPI REPORT #461

NEW THEORY AND METHODS FOR ACCELERATED MRI RECONSTRUCTION

by

Rodrigo A. Lobos

A dissertation submitted to
the USC Graduate School
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in

Electrical Engineering

University of Southern California
Los Angeles, California

December 2022

Doctoral Committee:

Associate Professor	Justin P. Haldar, Chair
Professor	Richard M. Leahy
Professor	John Wood

Abstract

Magnetic Resonance Imaging (MRI) is a safe, versatile, and noninvasive imaging modality that has revolutionized medicine by providing high-quality images of living tissue. Even though MRI technology has been constantly evolving since its origins in the 1970's, the time needed to acquire MRI data can still be restrictively long in real clinical applications. This has limited MRI from reaching its full potential, reason why constant efforts have been made in order to accelerate the data-acquisition process. Most of the adopted approaches to reduce the data-acquisition time involve the individual or combined use of advanced hardware and partial data-acquisition, however, conventional approaches based on these principles can be subject to negative effects in the signal-to-noise ratio, resolution, and/or the presence of undesired artifacts. In this study we propose new computational imaging methods which are able to successfully reconstruct partially acquired data obtained using advanced MRI hardware. Additionally, in this study we propose powerful computational techniques which allow dramatic improvements in computational efficiency when performing the reconstruction task.

In the first part of this study we propose reconstruction methods for partially acquired data which are able to automatically identify multiple linear predictability relationships present in the MRI data. For this purpose we rely on modern structured low-rank modeling theory and advanced optimization techniques. Based on these tools, we make the novel observation that reference data can be used to learn additional linear predictability relationships which considerably improve reconstruction performance in challenging scenarios. For instance, we show that linear relationships learned from reference data allow the reconstruction of highly accelerated data acquired using an undersampling scheme with a uniform structure. Notably, we show that these linear predictability relationships can be learned even in cases where the reference data are not pristine.

In the second part of this study we propose computational techniques in order to improve the efficiency of MRI reconstruction methods. We show that reconstruction methods which are based on the linearly predictable structure of the MRI data can be significantly enhanced in terms of computational complexity by a simple modification in the model used for prediction. Specifically, we show that a missing sample can be predicted using an ellipsoidal neighborhood of samples instead of a rectangular neighborhood of samples, which allows important improvements in computational time and memory usage with negligible effects on performance.

Finally, we study how to improve the efficiency of parallel imaging sensitivity map estimation methods. An accurate estimation of sensitivity maps is a fundamental piece in many modern parallel imaging MRI reconstruction methods, which should also be performed efficiently in order to reduce the overall reconstruction time. In this study we provide a novel theoretical description of the sensitivity map estimation problem by leveraging on the linearly predictable structure of the MRI data. Then, relying on these theoretical results, we propose a powerful set of computational techniques which allow massive improvements in computational complexity when integrated to sensitivity map estimation methods based on subspaces. We show that widely used estimation methods can achieve approximately 100-fold acceleration in computational time and dramatic savings in memory usage without sacrificing estimation accuracy. Remarkably, each of the proposed computational techniques can also be used individually to improve the efficiency of methods in other signal processing applications beyond MRI.

To my sister.

Acknowledgments

This dissertation corresponds to the end of one of the most challenging, enriching, and fascinating experiences that I have had the chance to live, and it has been only possible due to the immense support and love that I have received from outstanding people over the last seven years.

I would like to thank the members of my doctoral committee – Professors Richard Leahy, John Wood, Krishna Nayak, and Antonio Ortega – for their guidance, time, and effort. I am truly grateful to my advisor Professor Justin Haldar for his mentoring and constant support. Thank you Justin for inspiring me every day to keep growing as a scientist.

I am also thankful for my labmates Tae Hyung Kim, Daeun Kim, Divya Varayanan, Yunsong Liu, Chin-Cheng Chan, and Jiayang Wang. Thank you for reminding me every day how great and rewarding is working on science. I would like to specially thank Eduardo Pavez, Karel Mundnich, and Omid Sani for being by my side during all these years (we made it!).

I will be forever grateful to Richard Kittlaus, Hayden Buckley, Alexandra Guardado, Leslie Ramirez, Yecenia De La Rosa, Abdallah AlMutairi, and all the incredible people at Crossfit Muse for reinforcing my consistency, discipline, persistence, and sense of community during the last years of my PhD. You were all a fundamental pillar in the last part of this process.

Finally, I am immensely grateful to my family for their unconditional and infinite love over these years.

Table of Contents

List Of Figures

List Of Tables

Chapter 1

Introduction

1.1 Motivation

Magnetic Resonance Imaging (MRI) is one of the most relevant and flexible medical imaging modalities in clinical medicine, which allows high-quality imaging of living tissues in a safe and noninvasive manner. Even though MRI has been constantly pushing the frontiers of diagnostic clinical medicine over the last 50 years, obtaining high-quality images in MRI is still subject to restrictive long data-acquisition processes. This limits scanning throughput, affects accessibility, and contributes to patient discomfort. Over the years many efforts have been made in order to decrease the acquisition time, and one popular approach has been to develop fast acquisition protocols given the constant improvements on MRI hardware. A second widely used approach, which does not involve hardware modifications, corresponds to acquiring data partially. In this approach, also known as accelerated MRI acquisition, only part of the MRI data is acquired, and effects due to the missing data are compensated in post-acquisition stages by relying on advanced optimization algorithms.

Over the last decades researchers have been constantly trying to extend the limits of accelerated MRI by increasing the degree of missing data in the acquisition [3–12]. However, reconstruction methods are not always able to generate a useful image when the number of acquired samples is too small, and one way to surpass this limitation has been to include prior knowledge in the reconstruction process. A popular approach corresponds to leveraging autoregressive linear predictability relationships which typically exists in the MRI data [3, 13]. Nevertheless, current reconstruction

methods based on this approach have not been able to take full advantage of this type of relationships, and many possibilities to obtain further improvements in image-quality remain still open.

Another essential aspect in the reconstruction of accelerated MRI data is computational complexity. In many clinical applications MRI data can achieve massive dimensions, which challenges the efficiency of reconstruction methods in terms of computation time and memory usage. For instance, the potential high-dimensional characteristics of MRI accelerated data might result in long reconstruction processes, which could negatively affect clinical applications where data need to be analyzed soon afterwards the acquisition. Then, even though many reconstruction methods exhibit powerful characteristics in terms of image-quality performance, their high computational complexity tends to hinder their translation to clinical applications. Therefore, the development of computational techniques to improve reconstruction efficiency has been an active area of research in MRI.

In this work we make novel contributions in order to solve the aforementioned problems in accelerated MRI reconstruction. These contributions are summarized in the following section.

1.2   Main Contributions

- We develop novel reconstruction methods for accelerated MRI data which leverage on multiple linear predictability relationships existent within the data, and between the data and correlated reference datasets. Notably, these linear relationships are learned automatically by relying on structured low-rank modeling theory. We theoretically and empirically show that the integration of linear relationships learned from correlated reference datasets allow the successful reconstruction of data acquired using challenging undersampling schemes, which typically cannot be addressed by conventional methods. Remarkably, we show that the proposed methods are robust to imperfections in the reference data, and that they can be used synergistically with fast acquisition protocols based on advanced MRI hardware. Thus, the proposed methods open new opportunities when aiming at decreasing the total scanning time.

- We provide novel computational techniques to improve the efficiency of reconstruction methods. Specifically, we rely on advanced signal processing and linear predictability theory to develop computational techniques which allow important improvements regarding computational time and memory usage of conventional and state-of-the-art reconstruction methods. Remarkably, the improvements in computational complexity are achieved without compromising reconstruction quality. We show that the efficiency of reconstruction methods based on linear predictability principles can be significantly improved by a simple modification of the model used for prediction. We make a systematic comparison between the previous and the proposed predicting model by considering several conventional and state-of-the-art methods. We show that significant improvements in efficiency are obtained in all the cases when adopting the proposed model.

- We propose novel theory and fast computational methods in the context of parallel imaging sensitivity map estimation. The accurate and efficient estimation of sensitivity maps is a fundamental piece in many conventional and modern MRI reconstruction methods which leverage on parallel imaging constraints. In this work we provide a novel and intuitive theoretical description of the sensitivity map estimation problem which relies on linear predictability and structured low-rank modeling theory. Then, equipped with these theoretical results, we propose a set of computational techniques which allow massive improvements in the computational complexity of sensitivity map estimation methods based on subspaces. Notably, we empirically show that conventional subspace-based sensitivity map estimation methods can achieve approximately a 100-fold acceleration in time when combined with all the proposed computational techniques. Remarkably, this efficiency improvement affects the estimation accuracy negligibly. Moreover, the proposed computational techniques can be used combined or individually in several reconstruction methods and other important signal processing applications beyond MRI.

1.3 Organization of the Document

- In Chapter 2 we describe[1] how we proposed new methods for accelerated MRI reconstruction based on leveraging autoregressive shift-invariant linear predictability relationships existent within the data, and between the data and correlated datasets. Using structured low-rank modeling theory we show how correlated reference datasets allows the reconstruction of accelerated MRI data acquired using undersampling schemes with uniform structure. Based on our theoretical results, we propose two reconstruction methods that leverage on constraints learned from reference data in the spatial and in the Fourier domain, respectively. We empirically show that using constraints in the Fourier domain of the reference data offers better performance than constraints related to the spatial domain. By using these Fourier-based and additional nonconvex low-rank constraints, we show that the proposed method successfully reconstruct MRI data in challenging data-acquisition scenarios. Specifically, we show that echo planar imaging (EPI) data can be reconstructed without the presence of undesired artifacts even when the data has been 5-fold accelerated.

- In Chapter 3 we develop[2] an enhanced accelerated MRI reconstruction method based on structured low-rank modeling principles, which is able to leverage on linear predictability relationships learned from correlated reference datasets even in cases where they present imperfections. We account for these reference data imperfections by proposing a novel optimization approach for the reconstruction problem. By relying on our proposed optimization techniques, we additionally show that the enhanced method is able to leverage on further constraints learned from the reference data, which allows the reconstruction of data under challenging acceleration regimes. Remarkably, we show that the enhanced method is able to reconstruct EPI data even when the data has been 6-fold accelerated.

---

[1]The work presented in this Chapter 2 can also be found in [14]. Copyright belongs to IEEE Transactions on Medical Imaging.

[2]The work presented in this Chapter 3 can also be found in [15]. Copyright belongs to Magnetic Resonance in Medicine.

- In Chapter 4 we study[3] how to improve the efficiency of MRI reconstruction methods based on linear predictability principles. We theoretically and empirically show the effects of adopting a predicting model which assumes an ellipsoidal shape for the set of samples involved in the prediction of missing samples, instead of a rectangular shape as implemented in conventional approaches. We conclude that the ellipsoidal approach offers significant improvements in computational time and memory usage for several MRI reconstruction methods while preserving reconstruction quality.

- In Chapter 5 we provide a novel theoretical framework to describe the parallel imaging MRI sensitivity map estimation problem by relying on linear predictability and structured low-rank modeling theory. We then provide a set of powerful computational techniques which allow massive improvements in efficiency while maintaining performance when integrated to subspace-based sensitivity map estimation methods. We show that approximately a 100-fold reduction in computational time is achieved when the proposed techniques are used.

Finally, we provide our conclusions in Chapter 6.

---

[3]The work presented in this Chapter 4 can also be found in [16]. Copyright belongs to Magnetic Resonance in Medicine.

This page intentionally left blank

Chapter 2

Navigator-free EPI Ghost Correction with Structured Low-Rank Matrix Models: New Theory and Methods

2.1 Introduction

Echo-planar imaging (EPI) [17] is currently one of the fastest MRI pulse sequences and one of the most popular sequences for functional, diffusion, and perfusion imaging. EPI uses a train of gradient echoes to measure multiple lines of k-space from a single excitation, but is prone to artifacts because it employs a long readout, uses rapidly-switching high-amplitude gradients, and measures alternating lines of k-space with different gradient polarities [18].

In conventional single-shot EPI, even and odd lines of k-space are acquired with alternating gradient polarities. In practice, hardware imperfections, eddy currents, field inhomogeneity, concomitant fields, system delays, and similar phenomena can introduce signal phase errors between k-space lines acquired with different readout gradient polarities. If these phase errors are not correctly compensated, a Nyquist (or $N/2$) ghost artifact is observed corresponding to an aliased image that is positioned a half field-of-view (FOV) away from the true spatial position along the phase-encoding direction. In multi-shot EPI, full k-space coverage is achieved by using multiple excitations, where a different segment of k-space is acquired using EPI for each shot. Multi-shot EPI is used to reduce the EPI echo train, which subsequently reduces distortion and spin-dephasing

effects from local field inhomogeneity. Images can then be reconstructed by interleaving the multi-shot data together. As in the single-shot case, the mismatch between different gradient polarities also leads to Nyquist ghost artifacts for multi-shot data. However, multi-shot data may also exhibit additional ghost artifacts if there happen to be inconsistencies between each shot – including system drift, subject motion, and, particularly in the case of gradient-recalled EPI, respiration.

Many approaches to ghost correction have been proposed over the years, which we group into two main categories. The first category contains simple model-based approaches such as Refs. [19–26], which assume a low-dimensional model to describe a systematic phase mismatch between even and odd lines. The parameters of the mismatch model are often estimated using separate navigator data, and can then be used to correct the mismatch in the measured EPI data. While these methods are widely used and can work well when the mismatch model is accurate, phenomena such as eddy currents can lead to more complicated data mismatches that are not fully captured by simple models.

This paper focuses on the second category of methods, which includes those described in Refs. [1, 27–30]. These methods rely on a more flexible model in which the data samples of each gradient polarity/shot are assumed to be coming from different but highly-correlated images. For example, it is often assumed that the images corresponding to different gradient polarities or shots have the same image magnitudes but different image phases.[1] This is similar to how parallel imaging methods like SENSE [6] and GRAPPA [5] assume that the different channels of an array receiver coil acquire images that are different (i.e., modulated by different coil sensitivity profiles) but highly correlated. As a result, it is not surprising that many recent ghost correction approaches can be viewed as adaptations of previous parallel imaging methods to the ghost correction context. An example is the dual-polarity GRAPPA (DPG) method [28], which treats different polarities as if they were different virtual coils, and uses a dual GRAPPA kernel (with the GRAPPA weights divided into two halves corresponding to the two different gradient polarities) to synthesize a ghost-free fully-sampled image. Even though methods from the second category have been shown to have state-of-the-art

---

[1]This assumption is a simplification of the true imaging physics, and may not fully account for any time-dependent effects that evolve dynamically during data acquisition (e.g., due to eddy currents). More detailed modeling has been considered in some previous work, e.g., [19], though the simplified image-domain model described above underlies much of the recent literature.

performance in many challenging scenarios, they can still suffer from artifacts in certain cases. For example, DPG can fail to successfully correct ghost artifacts if there are mismatches between the measured EPI data and the autocalibration signal (ACS) used to train the dual GRAPPA kernel. This type of mismatch can occur because of changes in the measured data as a function of time, e.g., due to respiration [2]. In addition, most of the methods in the second category rely on the use of multichannel data, and are not easily applicable to single-channel ghost-correction.

Recently, novel image reconstruction methods have been proposed that enable calibrationless single-channel and multi-channel image reconstruction from undersampled k-space data using structured low-rank matrix (SLM) completion approaches [12,31–37]. SLM approaches are based on the assumption that there exist linear dependencies in k-space due to limited image support, smooth image phase variations, parallel imaging constraints, and/or transform-domain image sparsity. While such constraints have been used before in EPI ghost correction, e.g., Refs. [19,38], the SLM formulation of these constraints is distinct from classical approaches. SLM approaches were not originally developed for EPI data, but have very recently been adapted to such contexts [35,39–42]. These approaches have demonstrated to yield state-of-the-art performance in highly-accelerated EPI image reconstruction [35] and the ability to perform navigator-free EPI ghost correction [39–42].

In this paper, we analyze theoretical aspects of navigator-free EPI ghost correction using SLM approaches and obtain new insights that have major implications for ghost correction performance. Specifically, we prove that the SLM completion problem associated with ghost correction either has a non-unique solution or a unique solution that is undesirable. Based on this result, we observe that constraints are needed to ensure the performance of SLM-based ghost correction, and investigate two approaches that achieve substantially improved results. A preliminary account of portions of this work was previously given in Ref. [43].

In our first approach, we combine ideas from the LORAKS [12, 32, 33, 35] SLM framework with coil sensitivity maps within the SENSE framework, as has previously been done for EPI reconstruction [35] and EPI ghost correction [40, 41]. Compared to MUSSELS [40, 41] (a similar SENSE-based ghost correction method), our new approach makes use of a nonconvex regularization function from earlier LORAKS work [12, 32–35] which yields improved results, both in theory and

in practice, than the convex approach used in MUSSELS. Additionally, MUSSELS uses one of
the simpler forms of SLM construction (named the **C**-matrix in the terminology of LORAKS
[32, 33]) that incorporates support and parallel imaging constraints, but does not leverage image
phase constraints. Our new SENSE-based approach takes advantage of a more advanced SLM
construction (named the **S**-matrix in the terminology of LORAKS [32, 33]).

Our second approach combines LORAKS with k-space domain parallel imaging linear pre-
dictability constraints, like those used in GRAPPA [5], SPIRiT [7], and PRUNO [9]. In our imple-
mentation, these constraints are imposed within the broader framework of autocalibrated LORAKS
(AC-LORAKS) [11]. To the best of our knowledge, this is the first time that this type of informa-
tion has been combined with structured low-rank matrix completion methods in the context of EPI
ghost correction. This second new approach not only works for multi-channel data as expected,
but remarkably, we observe it also works for ghost correction of single-channel data in some cases.

This paper is organized as follows. Section 2.2 reviews SLM approaches and defines the notation
used in the rest of the paper. Section 2.3 presents our novel theoretical analysis of unconstrained
SLM methods for EPI ghost correction. Section 2.4 describes new constrained SLM formulations
that we propose to overcome the theoretical limitations of unconstrained approaches. Section 2.5
presents a systematic evaluation of these approaches with respect to current state-of-the-art ap-
proaches. Finally, discussion and conclusions are presented in Sec. 2.6.

2.2   Background and Notation

While many SLM descriptions have appeared in the literature, our description of SLM will focus
on the perspectives and terminology from the LORAKS framework. For simplicity, we only present
a high-level review of LORAKS for the 2D case, and refer interested readers to Refs. [12, 32–34] for
more general descriptions and additional details.

LORAKS is a flexible constrained reconstruction framework that uses SLM modeling to unify
and jointly impose several different classical and widely used MRI reconstruction constraints: lim-
ited image support constraints, smooth phase constraints, parallel imaging constraints, and sparsity

constraints [12, 32–34]. LORAKS is based on the observation that if any of these constraints is appropriate for a given image, then the Nyquist-sampled k-space data for that image will possess shift-invariant linear prediction relationships. These relationships mean that missing/corrupted data samples can be extrapolated or imputed as a weighted linear combination of neighboring points in k-space. Such linear prediction relationships imply that the k-space data will lie in a low-dimensional subspace, and SLM approaches can implicitly learn and impose this subspace structure directly from undersampled/low-quality data.

Importantly, while the data will possess linear prediction relationships and lie in a low-dimensional subspace under assumptions about the image support, phase, etc., the LORAKS approach is agnostic to the original source of these relationships. Instead, the approach attempts to identify and utilize all of the linear prediction relationships that are present in the k-space data, regardless of their source. This means that, while LORAKS reconstruction may be easier when support, phase, and parallel imaging constraints are applicable simultaneously, the LORAKS approach can still function when one or more of these constraints is inapplicable, as long as there are sufficient sources of linear predictability in the data.

The basic premise of the LORAKS support constraint [32] is that, if there are large regions of the FOV in which the true image is identically zero and if $s(k_x, k_y)$ represents the Fourier transform of the true image, then there exist infinitely many k-space functions $f(k_x, k_y)$ such that $s(k_x, k_y) * f(k_x, k_y) \approx 0$, where $*$ denotes the standard convolution operation. If we let $\mathbf{k}$ denote the vector of samples of $s(k_x, k_y)$ on the Cartesian Nyquist grid for the FOV and let $\mathbf{f}$ represent the samples of $f(k_x, k_y)$ on the same Cartesian grid, then the convolution relationship can be expressed in matrix-vector form as $\mathcal{P}_C(\mathbf{k})\mathbf{f} \approx 0$, where the operator $\mathcal{P}_C(\mathbf{k})$ forms a Toeplitz-structured convolution matrix (called the LORAKS $\mathbf{C}$-matrix [32]) out of the entries of $\mathbf{k}$. Since there are many such vectors $\mathbf{f}$ that satisfy this relationship, we observe that the LORAKS $\mathbf{C}$-matrix will be approximately low-rank.

The basic premise of the LORAKS phase constraint [32, 33] is that, if the image has smoothly-varying phase and the image has limited support, then there are infinitely many functions $h(k_x, k_y)$

such that $s(k_x, k_y) * h(k_x, k_y) - \bar{s}(-k_x, -k_y) * \bar{h}(k_x, k_y) \approx 0$, where $\bar{s}(k_x, k_y)$ and $\bar{h}(k_x, k_y)$ are respectively the complex conjugates of $s(k_x, k_y)$ and $h(k_x, k_y)$. Similar to the previous case, this convolution relationship can be expressed in a matrix-vector form as $\mathcal{P}_S(\mathbf{k})\mathbf{h} \approx 0$, where the operator $\mathcal{P}_S(\mathbf{k})$ combines a Toeplitz-structured convolution matrix with a Hankel-structured convolution matrix (resulting in what we call the LORAKS **S**-matrix [32, 33]) out of the entries of $\mathbf{k}$, and $\mathbf{h}$ is the vector of Nyquist samples of $h(k_x, k_y)$ and $\bar{h}(k_x, k_y)$. Since there are many such vectors $\mathbf{h}$ that satisfy this relationship, we observe that the LORAKS **S**-matrix will also be approximately low-rank.

These low-rank matrix constructions are easily generalized to the context of parallel imaging. Specifically, assume that data is acquired from N_c channels, and let $\mathbf{k}_n$ denote the vector of k-space samples from the nth channel, and let $\mathbf{k}_{\text{tot}}$ denote the vector containing the k-space samples from all channels. It has been shown that the concatenated matrix

$$\mathbf{C}_P(\mathbf{k}_{\text{tot}}) = \left[\begin{array}{cccc} \mathcal{P}_C(\mathbf{k}_1) & \mathcal{P}_C(\mathbf{k}_2) & \cdots & \mathcal{P}_C(\mathbf{k}_{N_c}) \end{array} \right] \tag{2.1}$$

will generally have low rank [9, 12, 31], and that the concatenated matrix

$$\mathbf{S}_P(\mathbf{k}_{\text{tot}}) = \left[\begin{array}{cccc} \mathcal{P}_S(\mathbf{k}_1) & \mathcal{P}_S(\mathbf{k}_2) & \cdots & \mathcal{P}_S(\mathbf{k}_{N_c}) \end{array} \right] \tag{2.2}$$

will also generally have low rank [12]. Note that Eqs. (2.1) and (2.2) reduce to the standard single-channel case when $N_c = 1$, so we will use these expressions for both the single-channel and the multi-channel cases.

The observation that these matrices have approximately low rank is valuable, because these low-rank characteristics can be exploited to improve image reconstruction quality. Specifically, by enforcing one or more of these low-rank constraints during image reconstruction, it becomes possible to reconstruct high-quality images from highly accelerated and/or unconventionally sampled k-space data.

The preceding paragraphs described SLM approaches for single-channel and multi-channel image reconstruction for general contexts, but without specialization to ghost correction for EPI.

However, as described in the introduction, there is a straightforward analogy between parallel imaging and Nyquist ghost correction. For the sake of simplicity and without loss of generality, we will describe the SLM matrix construction for this case in the context of single-shot imaging with positive and negative readout polarities (denoted RO^+ and RO^-, respectively), noting that the extension to multi-shot imaging is trivial (obtained by concatenating together the SLM matrices for each shot as if the different shots were coming from different receiver coils in a parallel imaging experiment). Let $\mathbf{k}_{\text{tot}}^+$ and $\mathbf{k}_{\text{tot}}^-$ represent hypothetical vectors of Nyquist-sampled Cartesian k-space data for the two different readout gradient polarities from either a single-channel or multi-channel experiment. Under the assumption that we can treat different readout polarities in the same way we treat different receiver coils in parallel imaging, we expect the matrices

$$\left[\begin{array}{cc} \mathbf{C}_P(\mathbf{k}_{\text{tot}}^+) & \mathbf{C}_P(\mathbf{k}_{\text{tot}}^-) \end{array} \right] \tag{2.3}$$

and

$$\left[\begin{array}{cc} \mathbf{S}_P(\mathbf{k}_{\text{tot}}^+) & \mathbf{S}_P(\mathbf{k}_{\text{tot}}^-) \end{array} \right] \tag{2.4}$$

to be approximately low-rank.[2] However, due to the form of single-shot EPI imaging, we only measure a subset of the phase encoding lines of $\mathbf{k}_{\text{tot}}^+$ and $\mathbf{k}_{\text{tot}}^-$. Specifically, let the measured data for the RO^+ and RO^- be respectively denoted as $\mathbf{d}_{\text{tot}}^+$ and $\mathbf{d}_{\text{tot}}^-$, respectively, with $\mathbf{d}_{\text{tot}}^+ = \mathbf{A}_+\mathbf{k}_{\text{tot}}^+$ and $\mathbf{d}_{\text{tot}}^- = \mathbf{A}_-\mathbf{k}_{\text{tot}}^-$, where $\mathbf{A}_+$ and $\mathbf{A}_-$ are simple subsampling matrices that extract the measured entries of $\mathbf{k}_{\text{tot}}^+$ and $\mathbf{k}_{\text{tot}}^-$ (i.e., $\mathbf{A}_+$ and $\mathbf{A}_-$ are formed by concatenating the rows of the identity matrix corresponding to the k-space sampling masks for each polarity).

---

[2]As noted in the introduction, the widely-used assumption that different readout polarities are different modulations of some original image (which forms the basis for the analogy between parallel imaging and ghost correction) is a simplification of the true imaging physics. However, it should be noted that Eqs. (2.3) and (2.4) may still possess low-rank structure even if this assumption is violated. In particular, these LORAKS matrices will have low-rank structure as long as there exist shift-invariant linear prediction relationships in k-space, and the LORAKS approach is agnostic to the original source of such relationships. We believe that deriving the existence of linear prediction relationships in the presence of more realistic models (e.g., accounting for eddy currents) may be feasible, although it is beyond the scope of the present work. Nevertheless, the results we show later with real data seem to imply that linear predictability assumptions are reasonably applicable in the real scenarios we have examined.

With these definitions, we can now describe previous SLM-based EPI ghost correction methods with a consistent language. For example, the earliest and simplest such approach, ALOHA [39], can be viewed as a special case of the optimization problem

$$\left\{ \hat{\mathbf{k}}_{\text{tot}}^+, \hat{\mathbf{k}}_{\text{tot}}^- \right\} = \underset{\left\{ \mathbf{k}_{\text{tot}}^+, \mathbf{k}_{\text{tot}}^- \right\}}{\operatorname{argmin}} J \left(\begin{bmatrix} \mathbf{C}_P(\mathbf{k}_{\text{tot}}^+) & \mathbf{C}_P(\mathbf{k}_{\text{tot}}^-) \end{bmatrix} \right), \tag{2.5}$$

subject to the additional data-consistency constraints that $\mathbf{A}_+ \hat{\mathbf{k}}_{\text{tot}}^+ = \mathbf{d}_{\text{tot}}^+$ and $\mathbf{A}_- \hat{\mathbf{k}}_{\text{tot}}^- = \mathbf{d}_{\text{tot}}^-$. Here, $J(\cdot)$ is a cost function that depends only on the singular values of its matrix argument, and promotes low-rank solutions. In the sequel, we will use the notation

$$L_{\mathbf{C}}(\mathbf{k}_{\text{tot}}^\pm) = J \left(\begin{bmatrix} \mathbf{C}_P(\mathbf{k}_{\text{tot}}^+) & \mathbf{C}_P(\mathbf{k}_{\text{tot}}^-) \end{bmatrix} \right), \tag{2.6}$$

where $\mathbf{k}_{\text{tot}}^\pm$ concatenates $\mathbf{k}_{\text{tot}}^+$ and $\mathbf{k}_{\text{tot}}^-$. Similarly, we will also use $L_{\mathbf{S}}(\cdot)$ to denote the function with the same form as Eq. (2.6), but switching from the LORAKS $\mathbf{C}$-matrix to the LORAKS $\mathbf{S}$-matrix by replacing all instances of $\mathbf{C}_P$ with $\mathbf{S}_P$.

A popular choice for $J(\cdot)$ in the general low-rank matrix completion literature (and the choice made by Ref. [39]) is the nuclear norm, which is a convex function that is known to encourage minimum-rank solutions [44]. The nuclear norm of a matrix $\mathbf{G}$ is defined as

$$\|\mathbf{G}\|_* = \sum_{i=1}^{\operatorname{rank}(\mathbf{G})} \sigma_i(\mathbf{G}), \tag{2.7}$$

where $\sigma_i(\mathbf{G})$ is the ith singular value of $\mathbf{G}$. Another potential choice of $J(\cdot)$, which was proposed in the original LORAKS work [32] but which has not been used in the EPI ghost correction work by other groups, is defined by

$$J_r(\mathbf{G}) = \sum_{i=r+1}^{\operatorname{rank}(\mathbf{G})} (\sigma_i(\mathbf{G}))^2, \tag{2.8}$$

where r is a user-selected parameter. This cost function is nonconvex, and $J_r(\mathbf{G})$ will equal zero whenever $\operatorname{rank}(\mathbf{G}) \leq r$. However, if $\operatorname{rank}(\mathbf{G}) > r$, then $J_r(\mathbf{G})$ will be nonzero, and equal to the squared Frobenius norm error that is incurred when $\mathbf{G}$ is optimally approximated by a rank-r

matrix. As a result, this cost function will encourage the reconstructed image to have a LORAKS

matrix that is approximately rank-r or lower.

The following subsection provides a novel theoretical analysis of the optimization problem from

Eq. (2.5), which reveals that it has several undesirable characteristics.

2.3   Analysis of SLM EPI Ghost Correction

For our analysis, we assume a typical setup in which the nominal fully-sampled k-space dataset has

equally-spaced consecutive phase encoding positions. For the sake of brevity, we will assume fully-

sampled single-channel EPI imaging[3] in which $\mathbf{d}_{\text{tot}}^+$ corresponds to the full set of measured even

phase encoding positions, while $\mathbf{d}_{\text{tot}}^-$ corresponds to the full set of measured odd phase encoding

positions.

Notice that the form of $\mathbf{A}_+$ implies that $\mathbf{A}_+^H \mathbf{A}_+$ is a diagonal projection matrix, and that

multiplying any vector of k-space samples by $\mathbf{A}_+^H \mathbf{A}_+$ is equivalent to preserving the values of the

even phase encoding lines while setting the values of the odd phase encoding lines to zero. Similarly,

$\mathbf{A}_-^H \mathbf{A}_-$ is a diagonal projection matrix, and multiplying any vector of k-space samples by $\mathbf{A}_-^H \mathbf{A}_-$

is equivalent to preserving the values of the odd phase encoding lines while setting the values of

the even phase encoding lines to zero. Additionally, we have that $\mathbf{A}_+^H \mathbf{A}_+ = \mathbf{I} - \mathbf{A}_-^H \mathbf{A}_-$, where $\mathbf{I}$ is

the identity matrix.

Using these facts together with the vector space concepts of orthogonal complements and direct

sums [45], we know that if $\{\hat{\mathbf{k}}_{\text{tot}}^+, \hat{\mathbf{k}}_{\text{tot}}^-\}$ obeys the data fidelity constraint from Eq. (2.5), then there

exist corresponding vectors $\mathbf{y}$ and $\mathbf{z}$ such that we can write

$$\hat{\mathbf{k}}_{\text{tot}}^+ = \mathbf{A}_+^H \mathbf{d}_{\text{tot}}^+ + \mathbf{A}_-^H \mathbf{A}_- \mathbf{y}$$
$$\hat{\mathbf{k}}_{\text{tot}}^- = \mathbf{A}_-^H \mathbf{d}_{\text{tot}}^- + \mathbf{A}_+^H \mathbf{A}_+ \mathbf{z}. \tag{2.9}$$

We have the following theoretical results:

---

[3]Generalized theoretical results for the case of parallel imaging with uniformly undersampled phase encoding can
also be derived using the same principles we used for the single-channel fully-sampled case. We have elected not to
show these derivations because they are intellectually straightforward extensions of the single-channel fully-sampled
case, but require a lot of additional notation to describe.

Theorem 1. *Given the context described above and arbitrary vectors* $\mathbf{y}$ *and* $\mathbf{z}$*, the singular values of the matrix*

$$\left[\ \mathbf{C}_P(\hat{\mathbf{k}}_{\text{tot}}^+) \quad \mathbf{C}_P(\hat{\mathbf{k}}_{\text{tot}}^-) \ \right] \tag{2.10}$$

are identical to the singular values of the matrix

$$\left[\ \mathbf{C}_P(\tilde{\mathbf{k}}_{\text{tot}}^+) \quad \mathbf{C}_P(\tilde{\mathbf{k}}_{\text{tot}}^-) \ \right], \tag{2.11}$$

with

$$\tilde{\mathbf{k}}_{\text{tot}}^+ = \mathbf{A}_+^H \mathbf{d}_{\text{tot}}^+ - \mathbf{A}_-^H \mathbf{A}_- \mathbf{y}$$

$$\tilde{\mathbf{k}}_{\text{tot}}^- = \mathbf{A}_-^H \mathbf{d}_{\text{tot}}^- - \mathbf{A}_+^H \mathbf{A}_+ \mathbf{z}. \tag{2.12}$$

Note that $\tilde{\mathbf{k}}_{\text{tot}}^+$ *and* $\tilde{\mathbf{k}}_{\text{tot}}^-$ *from Eq. (2.12) are identical to the vectors* $\hat{\mathbf{k}}_{\text{tot}}^+$ *and* $\hat{\mathbf{k}}_{\text{tot}}^-$ *appearing in Eq. (2.9), except that the estimates of the unmeasured data samples have been multiplied by -1.*

The proof of this theorem is sketched in the Supplementary Information section. Some basic intuition for this result is that we can multiply our estimates for the unmeasured k-space lines by -1 without impacting fidelity with the measured data. Due to uniform subsampling of each gradient polarity by a factor of 2, this multiplication procedure is equivalent to applying linear phase in k-space, which corresponds to a spatial shift of the image by half the FOV along the phase encoding dimension (and a 180° constant phase offset for the RO$^-$ polarity). This shifting procedure has no effect on the image support or on the correlations that exist between the different coils, and thus has no impact on the singular values or the rank of the LORAKS matrix. This means that if we have one solution to Eq. (2.5), then it is easy for us to construct another solution to Eq. (2.5), and this optimization problem will generally not have a unique useful solution.

The following corollaries formalize some of these statements and provide additional useful insight.

Corollary 1. *Equation (2.5) either has the unique solution* $\{\hat{\mathbf{k}}_{\text{tot}}^+, \hat{\mathbf{k}}_{\text{tot}}^-\} = \{\mathbf{A}_+^H \mathbf{d}_{\text{tot}}^+, \mathbf{A}_-^H \mathbf{d}_{\text{tot}}^-\}$ *which corresponds to zero-filling of the measured data, or it has at least two distinct optimal solutions that share exactly the same cost function value.*

(a) Original Data   (b) Modulated Data
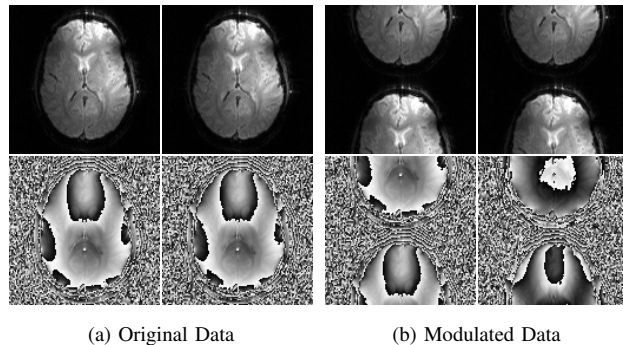
Figure 2.1: (a) EPI magnitude (top) and phase (bottom) images corresponding to (left) RO^+ data and (right) RO^- data. These alias-free images correspond to real data obtained with the PLACE method [1]. (b) Images corresponding to the same data from (a), except that the odd k-space lines for RO^+ and the even k-space lines for RO^- have been multiplied by -1.

Corollary 2. *If the cost function $J(\cdot)$ is chosen to be convex (e.g., the nuclear norm), then the zero-filled solution is always an optimal solution of Eq. (2.5). If Eq. (2.5) has more than one optimal solution, then it has infinitely many optimal solutions.*

Corollary 3. *Theorem 1 and Corollaries 1 and 2 are still true if we replace $\mathbf{C}_P(\cdot)$ in Eqs. (2.5), (2.10), and (2.11) with $\mathbf{S}_P(\cdot)$.*

Corollary 1 is proven in the Supplementary Information section, and implies that the optimization problem of Eq. (2.5) either has a trivial undesirable solution corresponding to zero-filling of the measured data, or it is an ill-posed optimization problem that does not possess a unique solution. While some of the solutions to Eq. (2.5) may be desirable, there are no guarantees that the algorithm we use to minimize Eq. (2.5) will yield one of these desirable solutions. Corollary 2 is also proven in the Supplementary Information section, and suggests that the use of convex cost functions to impose LORAKS constraints is likely to be suboptimal relative to the use of nonconvex cost functions. It should be noted that, unlike recent Nyquist ghost correction methods [39–41] which have made use of the convex nuclear norm, the early structured low-rank matrix completion methods for MRI all made use of nonconvex cost functions [12,31–33]. These nonconvex options are likely to be better for this problem setting. Corollary 3 is stated without proof (but can be proved using an approach that is similar to our proof of Theorem 1), and indicates that the deficiencies of Eq. (2.5) are not alleviated by switching from the LORAKS **C**-matrix to the LORAKS **S**-matrix.
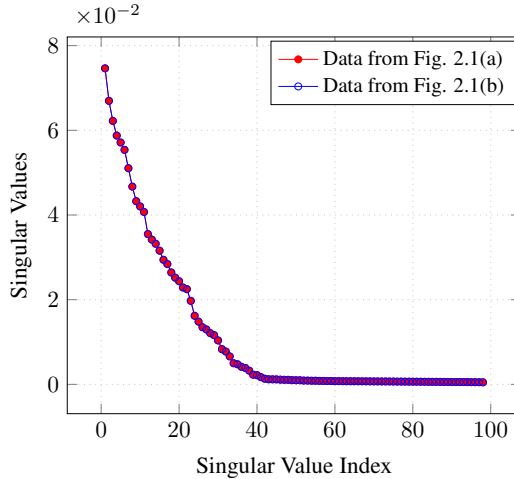
Figure 2.2: Plots of the singular values for the LORAKS matrices from Eq. (2.3) for the k-space datasets from Figs. 2.1(a) and (b).

(a) Convex (Eq. (5) with (7))   (b) Nonconvex (Eq. (5) with (8))   (c) Noncvx. w/SENSE (Eq. (13))   (d) Noncvx. w/AC-LORAKS (Eq. (14))
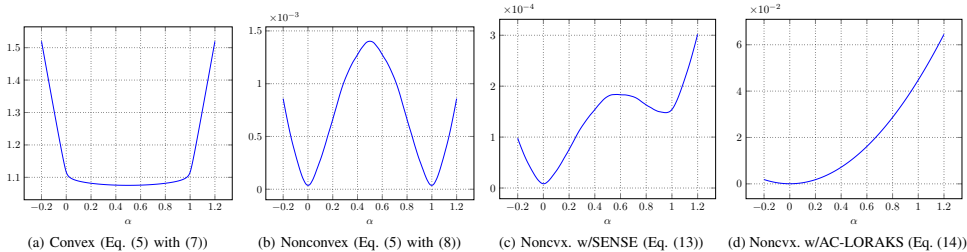
Figure 2.3: Letting $\mathbf{k}_1^{\pm}$ and $\mathbf{k}_2^{\pm}$ denote the k-space data for the images in Figs. 1(a) and 1(b), respectively, we plot the cost function value $L_{\mathbf{C}}\left(\alpha \mathbf{k}_1^{\pm} + (1-\alpha)\mathbf{k}_2^{\pm}\right)$ as a function of α. Setting $\alpha = 0$ yields the cost function value for the images from Fig. 2.1(a), setting $\alpha = 1$ yields the cost function value for the images from Fig. 2.1(b), while setting $\alpha = 0.5$ yields the cost function value for the zero-filled solution. Results are shown for different (a) convex and (b-d) nonconvex choices of $J(\cdot)$. The rank parameter $r = 40$ has been used in the nonconvex cases shown in (b-d), both without (a,b) and with (c,d) additional constraints.

Practical illustrations of these theoretical results are shown in Figs. 2.1–2.3. Figure 2.1 shows two different sets of EPI images that are both perfectly consistent with standard fully-sampled EPI data. The difference between the two datasets is the same as the difference between Eqs. (2.9) and (2.12). As expected, this k-space phase difference leads to shifting of the images for both RO^+ and RO^- by half the FOV, as well as adding a constant phase offset for the RO^- image. Figure 2.2 shows a plot of the singular values of the LORAKS $\mathbf{C}$-matrices corresponding to these two datasets. As expected from Theorem 1, the singular values are identical in both cases. Figure 2.3(a,b) illustrates the difference in behavior between convex and nonconvex cost functions $J(\cdot)$. Figure 2.3(a) shows

that the zero-filled solution is a minimum of Eq. (2.5) in the convex case, as expected from Corollary 2, and that there are many different images with very similar cost function values. Notably, the images from Fig. 2.1(a) and (b) are not optimal solutions in this convex case, even though they both have high quality and appear to be devoid of ghost artifacts. Figure 2.3(b) shows that the cost function has more desirable behavior in the nonconvex case (e.g., the zero-filled solution is no longer a minimum of Eq. (2.5), and there are sharp local minima in the vicinity of the images from Fig. 2.1), although the solution to Eq. (2.5) is still not unique in this case as we should expect based on Corollary 1.

While it may be possible to get a useful result from solving Eq. (2.5), it should be noted that in the presence of multiple global minimizers, it is difficult to ensure that an optimization algorithm will always converge to a desirable minimum. Incorporating additional constraints on the solution is a straightforward way to reduce the ambiguity associated with Eq. (2.5), and we describe two practical approaches for this in the next section.

2.4 Constrained SLM Formulations

2.4.1 Formulation using SENSE Constraints

A natural approach to imposing additional constraints on SLM reconstruction is to use coil sensitivity map information within the SENSE framework [6], assuming that coil sensitivity profiles are available and that data is acquired using a multi-channel receiver array. This style of approach has been used previously for both EPI reconstruction (assuming ghosts have been precorrected using navigator data) [35] and for navigator-free EPI ghost correction [40, 41]. Our proposed approach can be viewed as a combination of these two previous formulations.

In this work, we propose to use the following formulation for navigator-free EPI ghost correction using SENSE:

$$\{\hat{\boldsymbol{\rho}}^+, \hat{\boldsymbol{\rho}}^-\} \quad = \quad \arg \min_{\{\boldsymbol{\rho}^+, \boldsymbol{\rho}^-\}} \|\mathbf{E}_+ \boldsymbol{\rho}^+ - \mathbf{d}_{\text{tot}}^+\|_2^2 \tag{2.13}$$

$$+ \|\mathbf{E}_- \boldsymbol{\rho}^- - \mathbf{d}_{\text{tot}}^-\|_2^2 + \lambda L_{\mathbf{S}}(\mathbf{k}_{\text{tot}}^{\pm}), \tag{2.14}$$

subject to the constraints that $\mathbf{E}\boldsymbol{\rho}^+ = \mathbf{k}_{\text{tot}}^+$, that $\mathbf{E}\boldsymbol{\rho}^- = \mathbf{k}_{\text{tot}}^-$, and that $\mathbf{k}_{\text{tot}}^{\pm}$ is the concatenation of $\mathbf{k}_{\text{tot}}^+$ and $\mathbf{k}_{\text{tot}}^-$. In this formulation, we are using SENSE to reconstruct one image for RO$^+$ ($\boldsymbol{\rho}^+$) and another image for RO$^-$ ($\boldsymbol{\rho}^-$), and the only coupling that occurs between the two comes from the SLM regularization term. We have also used $\mathbf{E}_+$, $\mathbf{E}_-$, and $\mathbf{E}$ to denote the standard SENSE matrices (mapping from the image domain to k-space) corresponding to RO$^+$ subsampling, RO$^-$ subsampling, and full Nyquist sampling, respectively. These three matrices all use exactly the same sensitivity profiles, and differ only in the associated k-space sampling patterns. In addition, λ is a regularization parameter, and we suggest the use of the nonconvex regularizer from Eq. (2.8) for $L_{\mathbf{S}}(\cdot)$.

A major difference between this proposed approach and our previous SENSE-LORAKS work [35] is the separation of the RO$^+$ and RO$^-$ datasets, which enables navigator-free ghost correction. The main difference between this proposed approach and MUSSELS [40, 41] is that MUSSELS used the LORAKS $\mathbf{C}$-matrix and nuclear norm regularization, while we advocate use of the LORAKS $\mathbf{S}$-matrix with nonconvex regularization. Another major difference from MUSSELS is that, in the multi-channel case, we use the multi-channel LORAKS matrices (concatenating $2N_c$ SLMs) of Eqs. (2.3) and (2.4) [35], instead of concatenating only 2 SLMs formed from the SENSE reconstructions of each gradient polarity.

Equation (2.14) has been written assuming single-shot data. In the multi-shot case with phase inconsistencies between different shots, we generalize Eq. (2.14) by reconstructing a separate image for each polarity and each shot, with a separate SENSE encoding matrix and data fidelity term for each. Note that separating the data from different shots increases the effective acceleration factor for each data consistency term and is also associated with additional computational complexity.

There are many ways to solve the optimization problem in Eq. (2.14). In this paper, we use the simple majorize-minimize approach from Ref. [35], as described in the Supplementary Information section.

As shown in Fig. 2.3(c), incorporating SENSE constraints changes the shape of the cost function and can help to resolve the uniqueness issues associated with the inverse problem. In particular,

we observe that in this case, the extra information provided by SENSE constraints reduces the ambiguity between solutions, and causes the local minimum associated with $\alpha = 0$ to be substantially preferred over the local minimum associated with $\alpha = 1$. However, it appears that in this case, the nonconvex cost function may still have multiple local minima, in which case careful initialization may be necessary to ensure that the iterative approach converges to a good local minimum.

2.4.2   Formulation using AC-LORAKS Constraints

Another natural approach is to use k-space constraints like those of GRAPPA [5], SPIRiT [7], and PRUNO [9]. In this work, we use a formulation based on AC-LORAKS [11] (with strong similarities to PRUNO [9]). Specifically, in the single-shot case, we solve

$$\left\{ \hat{\mathbf{k}}_{\text{tot}}^+, \hat{\mathbf{k}}_{\text{tot}}^- \right\} \quad = \quad \arg \min_{\left\{ \mathbf{k}_{\text{tot}}^+, \mathbf{k}_{\text{tot}}^- \right\}} \left\| \mathbf{C}_P(\mathbf{k}_{\text{tot}}^+)\mathbf{N} \right\|_F^2 \tag{2.15}$$

$$+ \left\| \mathbf{C}_P(\mathbf{k}_{\text{tot}}^-)\mathbf{N} \right\|_F^2 + \lambda L_{\mathbf{S}}(\mathbf{k}_{\text{tot}}^\pm), \tag{2.16}$$

subject to the constraints that $\mathbf{d}_{\text{tot}}^+ = \mathbf{A}_+ \mathbf{k}_{\text{tot}}^+$, $\mathbf{d}_{\text{tot}}^- = \mathbf{A}_- \mathbf{k}_{\text{tot}}^-$, and that $\mathbf{k}_{\text{tot}}^\pm$ is the concatenation of $\mathbf{k}_{\text{tot}}^+$ and $\mathbf{k}_{\text{tot}}^-$. In this expression, $\|\cdot\|_F$ denotes the Frobenius norm, and the matrix $\mathbf{N}$ is an estimate of the approximate right nullspace of a LORAKS $\mathbf{C}$-matrix formed from ACS data acquired in a standard parallel imaging calibration pre-scan. The first two terms of Eq. (2.16) are similar to the first two terms of Eq. (2.14), in the sense that they impose support and parallel imaging constraints derived from some form of prescan, but do not make any assumptions about the relationship between $\mathbf{k}_{\text{tot}}^+$ and $\mathbf{k}_{\text{tot}}^-$ or the relationship between the image-domain phase characteristics of the calibration data and the image-domain phase characteristics of the EPI data being reconstructed. Similarly, the third terms in Eqs. (2.14) and (2.16) are the only terms that introduce coupling between $\mathbf{k}_{\text{tot}}^+$ and $\mathbf{k}_{\text{tot}}^-$, and the only terms that use the LORAKS $\mathbf{S}$-matrix to introduce constraints on the image phase. The use of phase constraints is useful both for partial Fourier EPI acquisition and for stabilizing the reconstruction of symmetrically-acquired EPI data [35]. As before, we suggest using the nonconvex regularizer from Eq. (2.8) for $L_{\mathbf{S}}(\cdot)$.

As shown in Fig. 2.3(d), incorporating AC-LORAKS constraints also changes the shape of the cost function. Just like with SENSE constraints, we observe that the additional information provided by the AC-LORAKS constraints leads to less ambiguity, and a clear preference towards the desired solution at $\alpha = 0$. However, the AC-LORAKS approach appears to be even more beneficial than the SENSE approach in this case, since we no longer observe a local minimum associated with $\alpha = 1$ in this case. As a result, we may expect that the AC-LORAKS approach is less sensitive to local minima and yields better solutions than the SENSE approach.

Similar to before, Eq. (2.16) is written for the single-shot case, but the generalization to multi-shot EPI is straightforward by separating and jointly reconstructing images for each polarity and shot. And similar to Eq. (2.14), there are also many ways to solve the optimization problem in Eq. (2.16). These two optimization problems have very similar structure (i.e., the first two terms are least-squares penalties, while the third term encourages low-rank matrix structure), and as a consequence, we have used a minor modification of the algorithm we used for solving Eq. (2.14) to also solve Eq. (2.16).

An interesting feature of our proposed AC-LORAKS formulation is that it can potentially work with single-channel data [11]. This is a major advantage over our proposed SENSE formulation, which is not expected to produce good results unless a multiple-channel receiver array is used. It should be noted that while many human MRI experiments use parallel imaging, single-channel reconstruction is still highly relevant in a variety of situations, including animal studies and studies of human anatomy that make use of specialized receiver coil technology (e.g., single-channel prostate coils).

2.5   Results

This section describes evaluations of our new LORAKS-based EPI ghost correction methods using navigator-free gradient-echo EPI data acquired from phantoms and *in vivo* human brains. The phantom data shown in Fig. 2.6 were acquired using a gradient echo EPI sequence on a Siemens 3T Tim Trio using a standard product 12-channel receiver array (compressed in hardware down to 4 channels). The imaging parameters were: FOV = 24 mm $\times$ 24 mm; matrix size 64$\times$64;
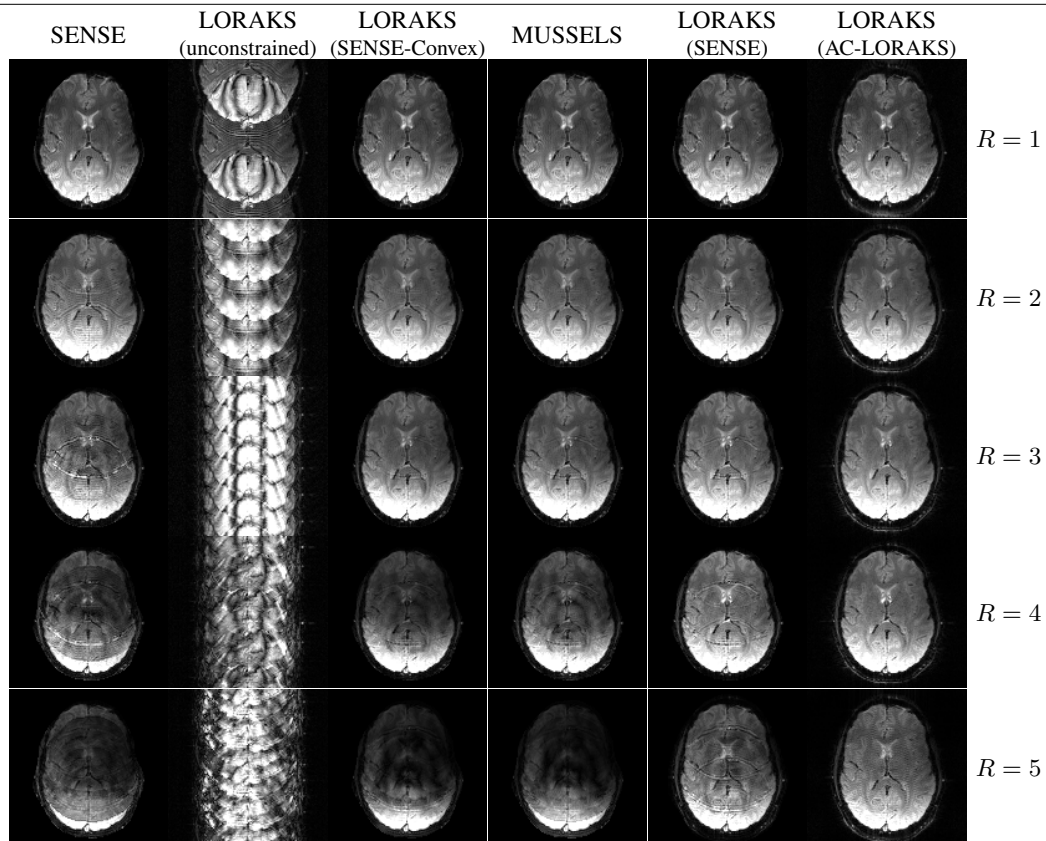
Figure 2.4: Comparison of different reconstruction techniques using prospectively undersampled (except the $R = 5$ case, which is retrospectively undersampled) *in vivo* single-shot EPI data at different parallel imaging acceleration factors.

slice thickness = 5 mm; TR = 1.46 sec; TE = 38 msec. The phantom data used in Figs. 2.15 and 2.16(a) were acquired using a gradient echo EPI sequence on a Siemens 3T Tim Trio using a standard product 12-channel receiver array. The imaging parameters were: FOV = 24 mm × 24 mm; matrix size 96×96; slice thickness = 5 mm; TR = 60 msec; TE = 28 msec. The *in vivo* data were acquired using a gradient echo EPI sequence on a Siemens (Erlangen, Germany) 3T Prisma Fit using a standard product 32-channel receiver array. The imaging parameters were: FOV = 220 mm × 220 mm; matrix size 128×128; slice thickness = 3 mm; TR = 2.08 sec; TE = 47 msec; acceleration factor $R \in \{1, 2, 3, 4, 5\}$. For each dataset, ACS data (used both for estimating nullspaces and for estimating sensitivity maps) was acquired using the same approach as previously used in DPG [28]. Most of the reconstructions we show in this section estimate separate images for each gradient polarity and each shot, and some also estimate separate images

Figure 2.5: Comparison between (a) DPG and (b) LORAKS with AC-LORAKS constraints for the real single-shot EPI *in vivo* brain data from Fig. 2.4. Instead of showing coil-combined images, a single representative channel is shown to avoid contamination of the phase characteristics induced by coil combination.

for each coil. While various approaches exist for combining together multiple images from different coils/polarities/shots for visualization, for simplicity and consistency we have combined the multiple images into a single image using principal component analysis, which is a standard method for parallel imaging coil compression/combination [46, 47]. Unless otherwise specified, our LORAKS-based results also always use the nonconvex regularization penalty from Eq. (2.8), with the rank threshold r chosen based on the singular values of the LORAKS matrix formed from ACS data. Specifically, r was chosen as the point at which the plot of the singular values appears to flatten out, which is a standard approach to matrix rank estimation in the presence of noise. For methods that use regularization parameters, λ was initially set to a small value ($\lambda = 10^{-3}$), and if necessary based on visual assessment of image ghost artifacts, was gradually increased until good reconstructions were observed. LORAKS-based reconstructions were performed based on adaptations of publicly-available code [10].

Figure 2.4 shows a comparison of different parallel imaging reconstruction and EPI ghost correction methods for *in vivo* single-shot EPI data. A gold standard image with fully-sampled RO^+ data and fully-sampled RO^- images was obtained using PLACE [1] with a 32-channel receiver coil and an 128×128 acquisition matrix. We also acquired standard fully-sampled EPI (acceleration factor $R = 1$, with each gradient polarity undersampled by a factor of two) and prospectively accelerated EPI acquisitions for a range of acceleration factors ($R = 2, 3, 4$). Additionally, the acceleration factor of $R = 5$ was simulated by retrospectively undersampling the PLACE data.

Uncorrected	ALOHA	LORAKS (unconstrained)	LORAKS (SENSE)	DPG	LORAKS (AC-LORAKS)
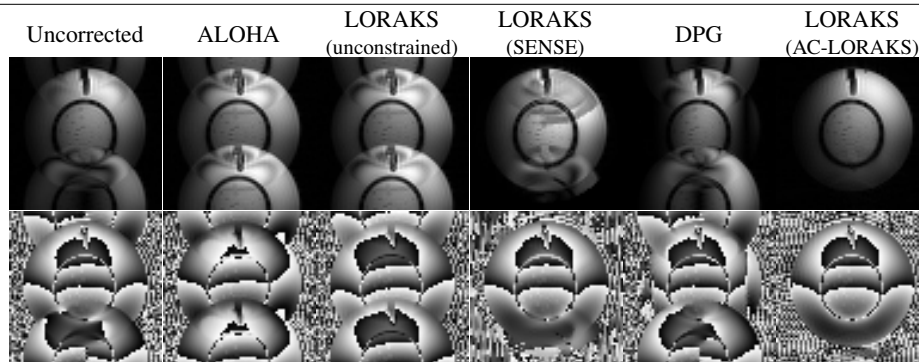
Figure 2.6: (top) Magnitude and (bottom) phase images corresponding to reconstruction of unaccelerated ($R = 1$) single-channel single-shot EPI phantom data.

Reconstructions were performed using unconstrained LORAKS as in Eq. (2.5) but using the LORAKS **S**-matrix, LORAKS with SENSE constraints as in Eq. (2.14) in both convex (Eq. (2.7)) and nonconvex (Eq. (2.8)) variations, and LORAKS with AC-LORAKS constraints as in Eq. (2.16). For SENSE-based reconstruction, sensitivity profiles were estimated using ESPIRiT [8]. For comparison, we also performed MUSSELS reconstruction [40, 41] and independent SENSE reconstruction of each gradient polarity [27] without LORAKS-based regularization (equivalent to setting $\lambda = 0$ in Eq. (2.14)).

The figure shows that SENSE without SLM regularization works well for low-acceleration factors, though faces challenges at high acceleration factors. This behavior is expected as an EPI acceleration factor of $R = 5$ is an effective acceleration factor of $R = 10$ for each readout polarity, which is a very challenging case for SENSE reconstruction. We also observe that unconstrained LORAKS reconstruction has severe problems, as should be expected based on our theoretical analysis of Eq. (2.5). The results also show that the convex SLM approaches (MUSSELS and LORAKS with SENSE and convex regularization) are effective at low acceleration factors, but start demonstrating artifacts as the acceleration factor increases. On the other hand, both of our proposed new formulations are substantially more successful, achieving high quality reconstruction results even at very high acceleration factors. This result is consistent with our theoretical expectations from Corollaries 1-3 that nonconvex cost functions can lead to a better-posed reconstruction problem. At the highest acceleration factors, LORAKS with AC-LORAKS constraints was more effective than LORAKS with SENSE constraints, which displayed unresolved aliasing artifacts.

Table 2.1: NRMSE VALUES FOR THE IMAGES SHOWN IN FIG. 2.8.

R	SENSE	LORAKS (unconstr)	LORAKS (SENSE-Conv)	MUSSELS	LORAKS (SENSE)	LORAKS (AC)
1	0.17	0.70	0.17	0.17	0.09	**0.03**
2	0.23	0.86	0.20	0.20	0.10	**0.05**
3	0.38	1.04	0.27	0.26	0.11	**0.06**
4	0.53	0.94	0.42	0.49	0.15	**0.07**
5	0.65	1.00	0.62	0.67	0.25	**0.12**

The data shown in Fig. 2.4 was also reconstructed using the state-of-the-art DPG method [28] using the same ACS data, and a comparison against LORAKS with AC-LORAKS constraints is shown in Fig. 2.5. While DPG generally works well, a close examination of the reconstructed magnitude and phase images demonstrates that DPG still has small residual ghost artifacts that are not present in the LORAKS-based reconstruction. These artifacts are particularly visible in the phase images, since the phase is highly sensitive to ghosting in regions of the image where the magnitude is small. A deeper examination of the data leads us to believe that the ghost artifacts we see for DPG are the result of systematic changes (between the relative phases of the different gradient polarities) that have occurred in part due to the length of time that passed between the collection of the ACS data and the acquisition of the accelerated EPI data that is being reconstructed.

The data used in Figs. 2.4 and 2.5 was prospectively sampled, but it is hard to quantify accuracy for prospectively sampled data because the image phase characteristics (due to eddy currents, etc.) can vary as a function of sequence parameters like the acceleration factor, and because our subject is living and breathing (which leads to variations over time). An illustration of the phase differences between RO^+ and RO^- for different acceleration factors is shown in Fig. 2.7. To enable a quantitative comparison of different approaches, we also performed reconstructions of retrospectively undersampled versions of the fully sampled PLACE data. Results are shown in supplementary Fig. 2.8, with normalized root-mean-squared error (NRMSE) shown in Table 2.1. The retrospective results are consistent with our prospective results, and the quantitative NRMSE results are consistent with our qualitative evaluations.
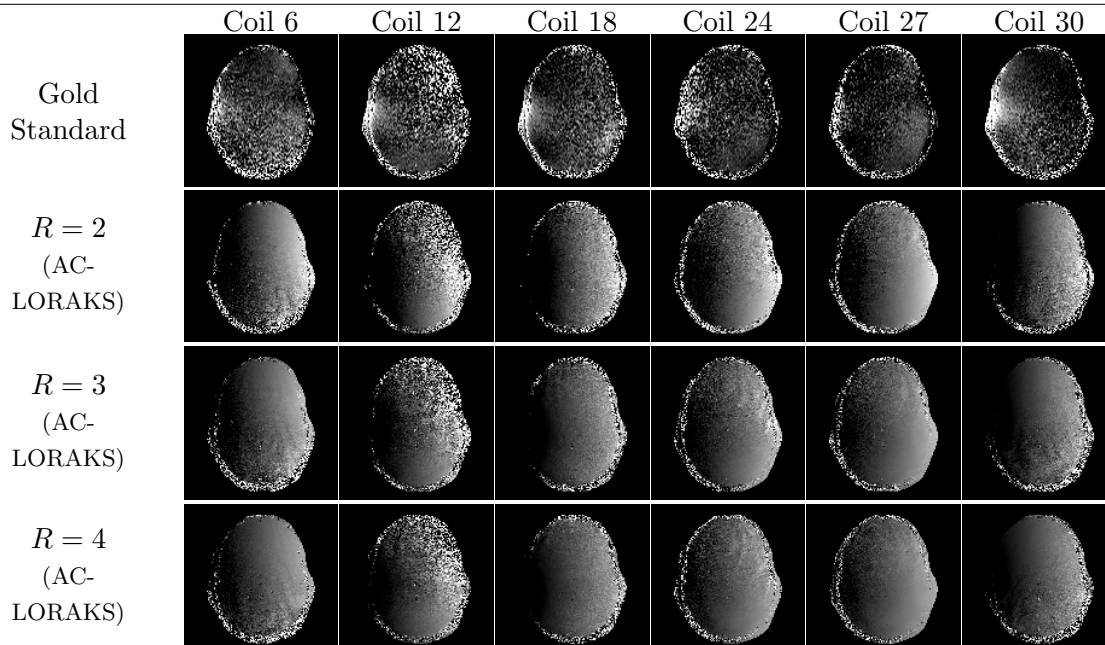
Figure 2.7: Phase differences between RO$^+$ and RO$^-$ for a representative subset of coils from the prospectively undersampled data shown in Fig. 4. In cases with accelerated data (where no gold standard is available), we show the phase difference obtained after AC-LORAKS reconstruction. Unlike most other phase images shown in this paper (which do not include background masking and show the entire phase range from $-\pi$ to π), we have taken steps to make the phase images in this figure easier to visualize. In particular, we have masked the background noise to make it easier to focus on the signal structure. In addition, we have shown a restricted phase range for both the gold standard (black = 0.95π radians, white = 1.02π radians) and the reconstructions using AC-LORAKS (black = 0.89π radians, white = 1.21π radians).

The data shown in Figs. 2.4 and 2.5 was acquired with an axial slice orientation, and therefore had less phase mismatch between RO$^+$ and RO$^-$ than it could have if we had used a less-conventional slice orientation. In addition, the slice we used was far from sources of field inhomogeneity, and therefore had a relatively smooth phase profile. To demonstrate a more complicated case, we have also performed ghost correction of EPI data acquired with a double-oblique slice orientation. Double-oblique orientation is non-traditional for EPI imaging of the brain, but is known to give rise to 2D nonlinear phase-differences between gradient polarities due, e.g., to concomitant field effects [24,28,48–50]. In addition, the double-oblique slice we selected passes close to air-tissue interfaces and has a substantially less-smooth phase profile. We have performed reconstructions of prospectively undersampled double-oblique data (similar to Figs. 2.4 and 2.5), and the results are shown in Figs. 2.10 - 2.12. We have also performed reconstructions of retrospectively undersampled double-oblique data (similar to Fig. 2.8 and Table 2.1), and the results are shown in Fig. 2.13 and
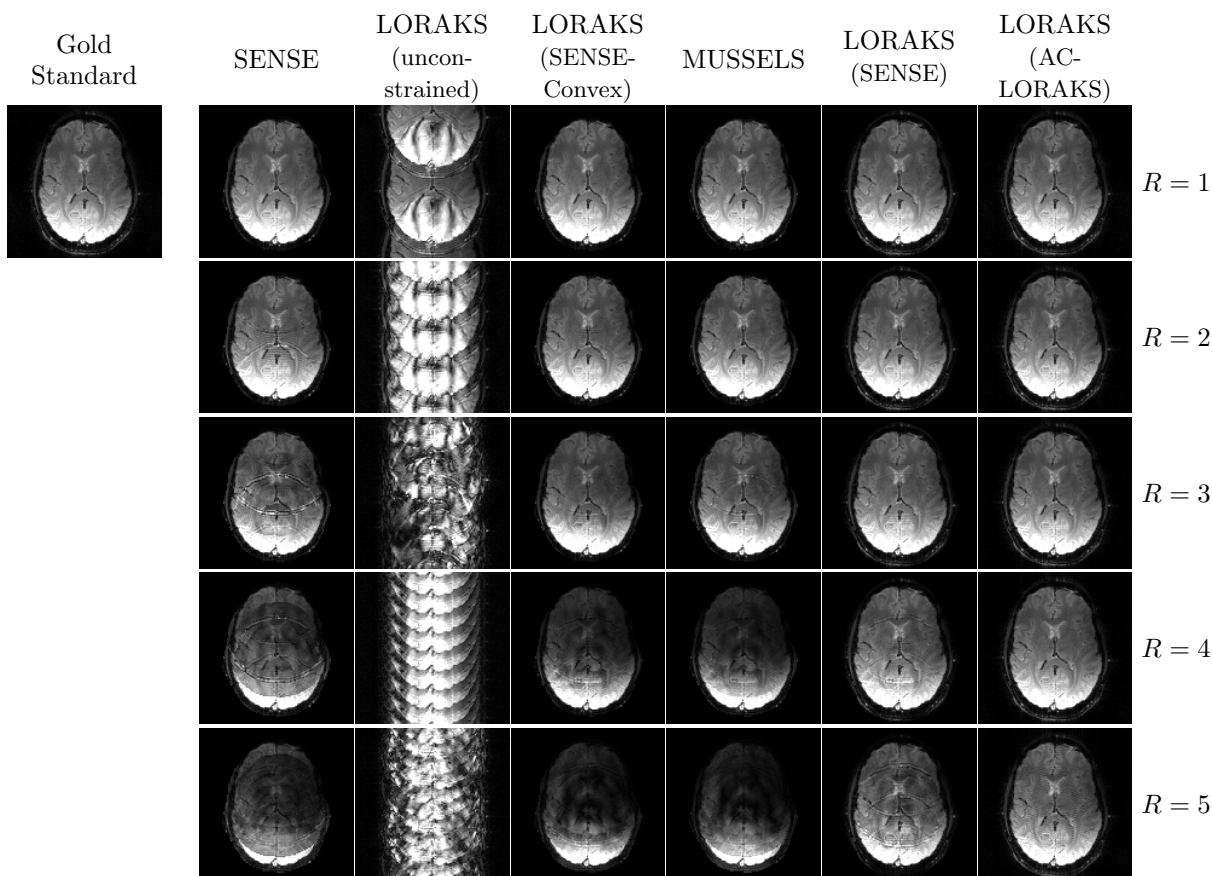
Figure 2.8: Comparison of different reconstruction techniques using retrospectively undersampled *in vivo* single-shot EPI data to simulate different parallel imaging acceleration factors.

Table 2.2. As can be seen from the phase difference maps in Fig. 2.11, we have more significant nonlinear 2D phase mismatches between RO^+ and RO^- in this case. Our results with double-oblique data are consistent with our results from axial data, suggesting that our proposed approaches can still work well in more challenging scenarios.

The results described above were all based on data acquired with a somewhat loose FOV, which may be beneficial for the use of support constraints. However, there are also imaging scenarios of interest in which the FOV is much tighter. To test performance in the presence of a tight FOV, we acquired similar data to that shown in Figs. 2.4 and 2.5, but with half the FOV along the phase encoding dimension. This causes our gold standard image to demonstrate aliasing. To make the reconstruction problem even more challenging, we also simulated an additional 2D nonlinear phase difference between RO^+ and RO^-, and this phase difference was chosen differently for the
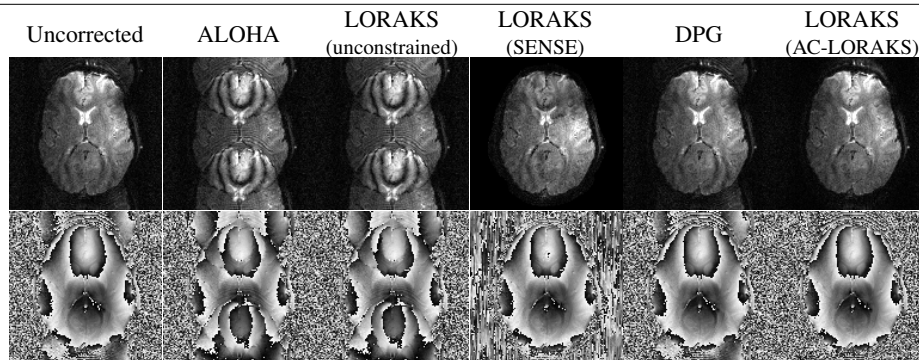
Figure 2.9: (top) Magnitude and (bottom) phase images corresponding to reconstruction of unaccelerated ($R = 1$) single-channel single-shot EPI *in vivo* human brain data.

Table 2.2: NRMSE VALUES FOR THE IMAGES SHOWN IN FIG. 2.13.

R	SENSE	LORAKS (unconstr)	LORAKS (SENSE-Conv)	MUSSELS	LORAKS (SENSE)	LORAKS (AC)
1	0.06	0.70	0.09	0.09	0.07	**0.03**
2	0.07	0.86	0.13	0.15	0.08	**0.07**
3	0.19	0.91	0.20	0.25	0.11	**0.10**
4	0.46	0.99	0.35	0.40	0.18	**0.12**
5	0.61	0.95	0.65	0.74	0.29	**0.11**

ACS data and the gold standard used for simulation. Since it is difficult to apply conventional SENSE-based methods in the presence of a tight FOV (because of the presence of aliasing in the gold standard), we only performed reconstructions using methods that do not use sensitivity maps, i.e., DPG and LORAKS with AC-LORAKS constraints. The results are reported in Fig. 2.14. As can be seen, DPG does not perform very well in this very complicated scenario, particularly because of the phase mismatch between the ACS data and the data being reconstructed. On the other hand, AC-LORAKS is substantially more successful.

In addition to navigator-free multi-channel settings, the proposed methods were also evaluated in navigator-free single-channel settings, which are expected to be substantially more challenging. Single-channel datasets were obtained by isolating the information from a single coil in multi-channel acquisitions. Reconstructions were performed using ALOHA [39] (Eq. (2.5) with Eq. (2.7)), unconstrained LORAKS (Eq. (2.5) with Eq. (2.8) but using the LORAKS **S**-matrix), LORAKS with SENSE constraints (Eq. (2.14)), DPG [28], and LORAKS with AC-LORAKS constraints (Eq. (2.16)). Note that, since sensitivity-map estimation is not feasible in the single-channel setting,
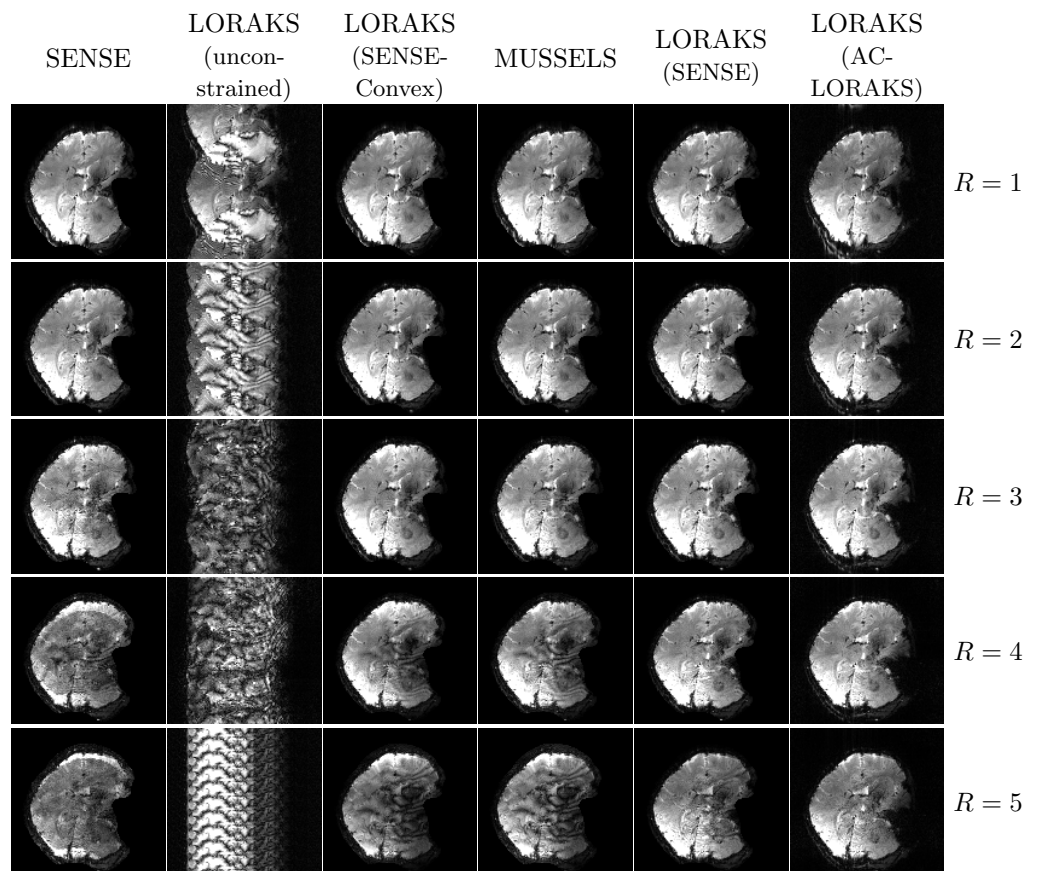
Figure 2.10: Comparison of different reconstruction techniques using prospectively undersampled *in vivo* double-oblique single-shot EPI data at different parallel imaging acceleration factors.
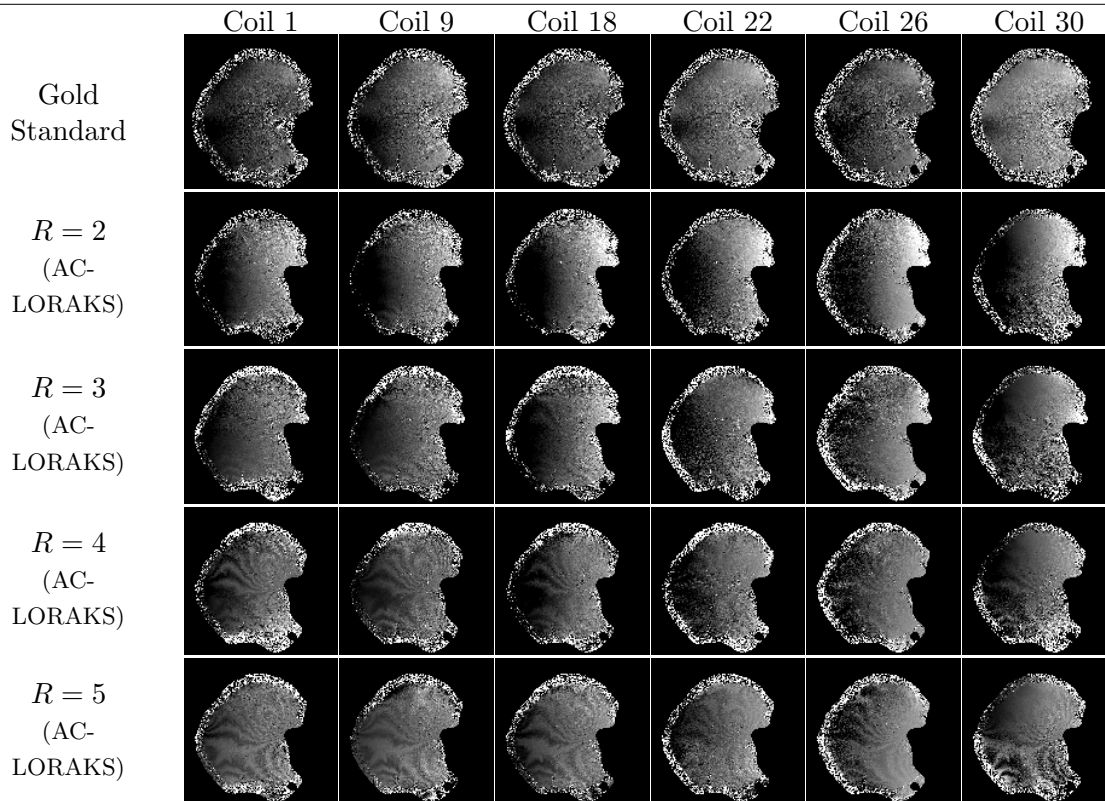
Figure 2.11: Phase differences between RO^+ and RO^- for a representative subset of coils from the prospectively undersampled double-oblique data shown in Fig. 2.10. In cases with accelerated data (where no gold standard is available), we show the phase difference obtained after reconstruction using LORAKS with AC-LORAKS constraints. Unlike most other phase images shown in this paper (which do not include background masking and show the entire phase range from $-\pi$ to π), we have taken steps to make the phase images in this figure easier to visualize. In particular, we have masked the background noise to make it easier to focus on the signal structure. In addition, we have shown a restricted phase range for both the gold standard (black $= 0.95\pi$ radians, white $= 1.02\pi$ radians) and the reconstructions using AC-LORAKS (black $= 0.86\pi$ radians, white $= 1.18\pi$ radians).

our SENSE-based results used a binary support mask (that has value 1 inside the support of the image and value 0 everywhere else) in place of a coil sensitivity map. This has the effect of imposing prior knowledge of the image support on the reconstructed image. Note also that DPG was not originally designed to be used with single-channel data, although the formulation can still be applied to the single-channel case. Unaccelerated ($R = 1$) single-channel single-shot EPI results are shown for a phantom dataset (64×64 acquisition matrix) in Fig. 2.6 and for one channel of the previous *in vivo* human brain dataset in Fig. 2.9. The results are consistent in both cases. Images obtained without compensating the mismatch between RO^+ and RO^- have obvious ghost artifacts, and
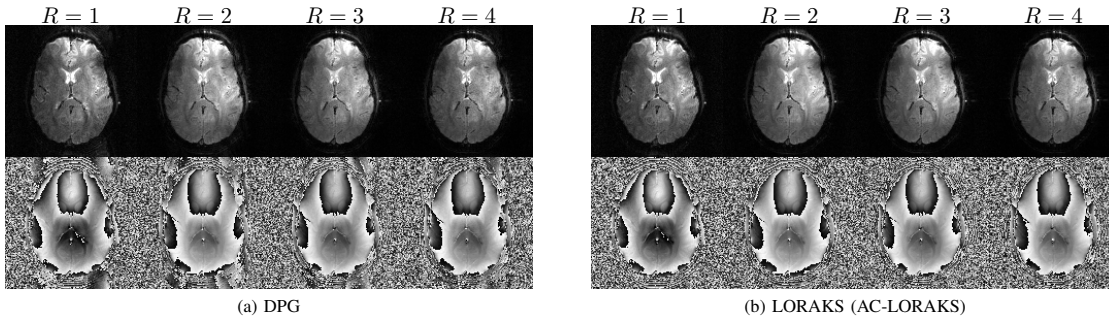
(a) DPG
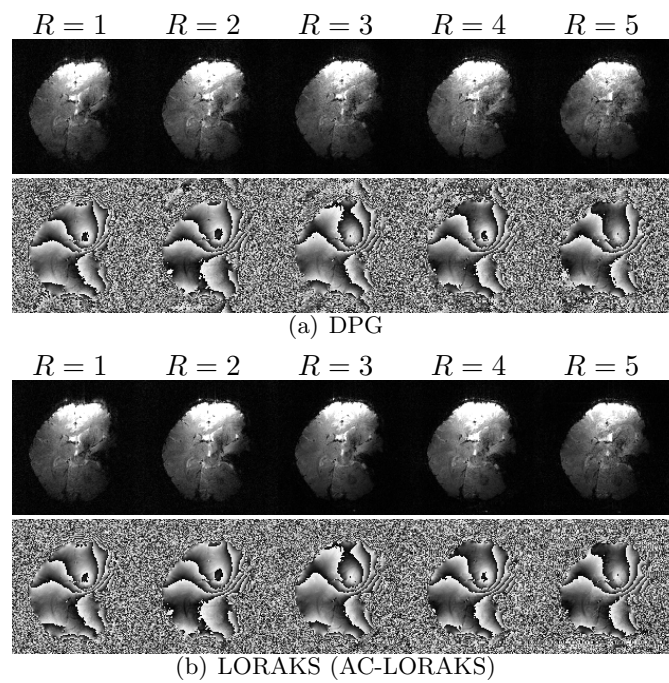
(b) LORAKS (AC-LORAKS)

Figure 2.12: Comparison between (a) DPG and (b) LORAKS with AC-LORAKS constraints for the real single-shot EPI *in vivo* brain double-oblique data from Fig. 2.10. Instead of showing coil-combined images, a single representative channel is shown to avoid contamination of the phase characteristics induced by coil combination.
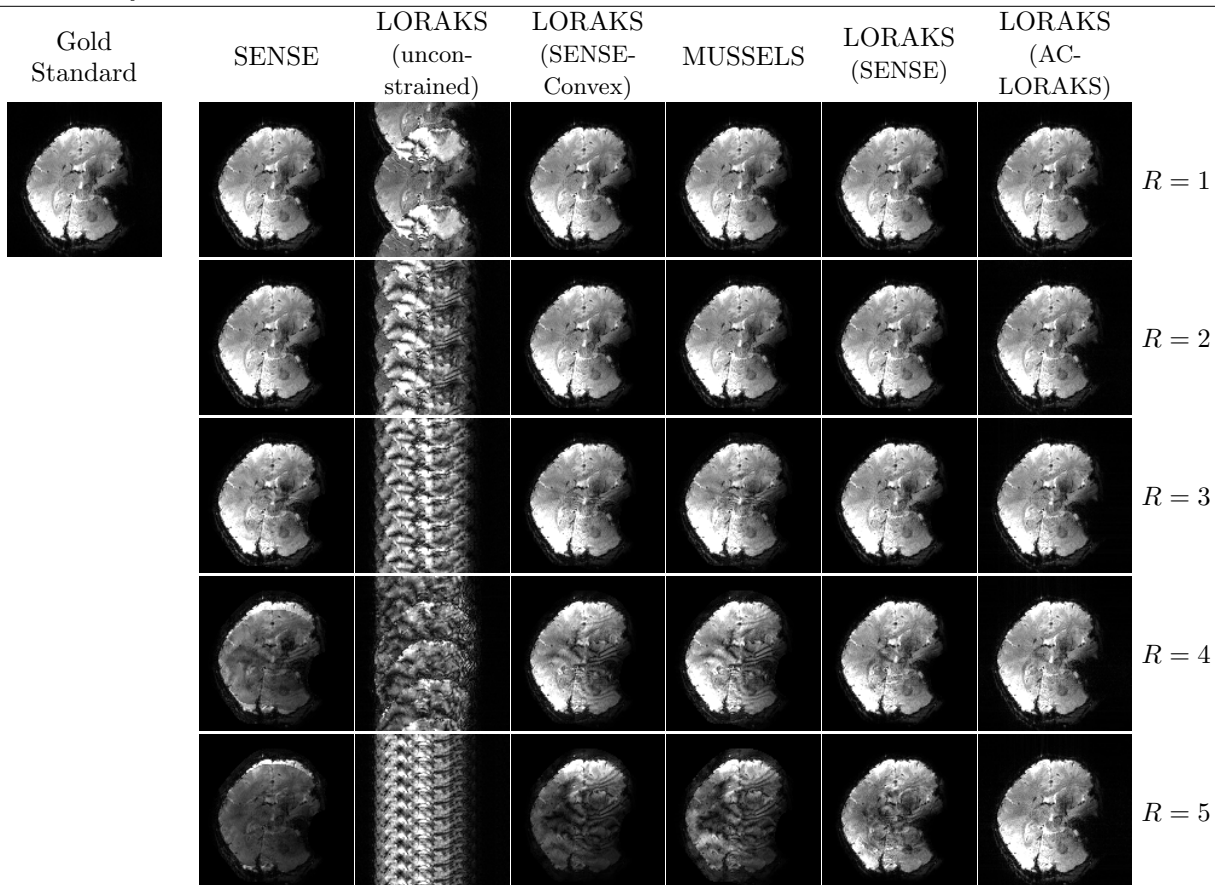
Figure 2.13: Comparison of different reconstruction techniques using retrospectively undersampled *in vivo* double-oblique data to simulate single-shot EPI experiments at different parallel imaging acceleration factors.

these artifacts are not solved (and are potentially even amplified) when using unconstrained SLM approaches (i.e., ALOHA or LORAKS without constraints). The LORAKS reconstruction with "SENSE" constraints (i.e., support constraints) helps to eliminate some of the ghost artifacts that appeared outside the support of the original object, although residual aliasing artifacts are still observed within the support of the object. These artifacts are most visible in the phantom image, though close inspection also reveals the appearance of aliasing artifacts in the brain image. We also observe that DPG is unsuccessful in this single-channel case, which we believe is due both to the difficulties of the single-channel problem as well as systematic differences between the ACS data and the data being reconstructed. On the other hand, the LORAKS results with AC-LORAKS constraints are substantially more successful than any of the previous methods.
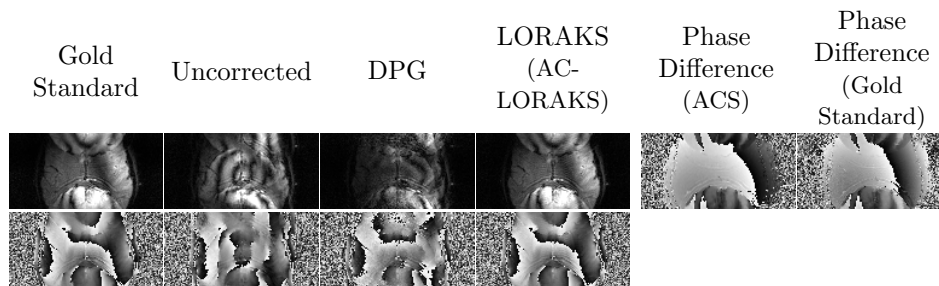
Figure 2.14: (top-left) Magnitude and (bottom-left) phase images corresponding to reconstruction of unaccelerated ($R = 1$) *in vivo* axial multi-channel single-shot EPI data acquired with a small FOV, and with extra simulated nonlinear phase differences between RO^+ and RO^-. The images on the right show the phase differences that were used between RO^+ and RO^-, which were set differently for the ACS data (used for calibration) than they were for the gold standard (used for reconstruction).

We also applied the different reconstruction approaches to single-channel double-oblique data, for the same reasons as in the multi-channel case. Results are shown in Fig. 2.17, and are consistent with the results observed with axial data. One potential criticism of the single-channel results we have shown is that, at least for these datasets, the phase mismatch between RO^+ and RO^- is not very severe, which causes the uncorrected reconstructions to have relatively low levels of ghost artifact. To demonstrate the performance in a more severe case, we performed reconstruction of this dataset again with an exaggerated phase mismatch between RO^+ and RO^-. The results of this simulation are shown in Fig. 2.18. As can be seen, this case demonstrates much stronger ghost artifacts without correction. However, similar to the previous cases, LORAKS with AC-LORAKS constraints is still the most successful at reducing ghost artifacts.

Reconstructions were also performed using multi-shot data. Figure 2.15 shows results using fully sampled $R = 1$ two-shot data (128×128 acquisition matrix, 12-channel receiver coil) at different time points for a phantom with physically-simulated respiration effects [2]. Note that with $R = 1$ and two-shots, each gradient polarity for each shot has an effective undersampling factor of 4. LORAKS with AC-LORAKS constraints is compared against DPG for segmented EPI [51], and to make reconstruction even more challenging for LORAKS, LORAKS reconstruction was performed from single-channel data while DPG was provided with the full set of multi-channel data. Due to simulated respiration, there are mismatches between the measured ACS data and the EPI data being reconstructed. DPG is not robust to these mismatches, and displays residual ghost artifacts

(a) Uncorrected

(b) DPG

(c) LORAKS (AC-LORAKS)

(d) Simulated Respiration
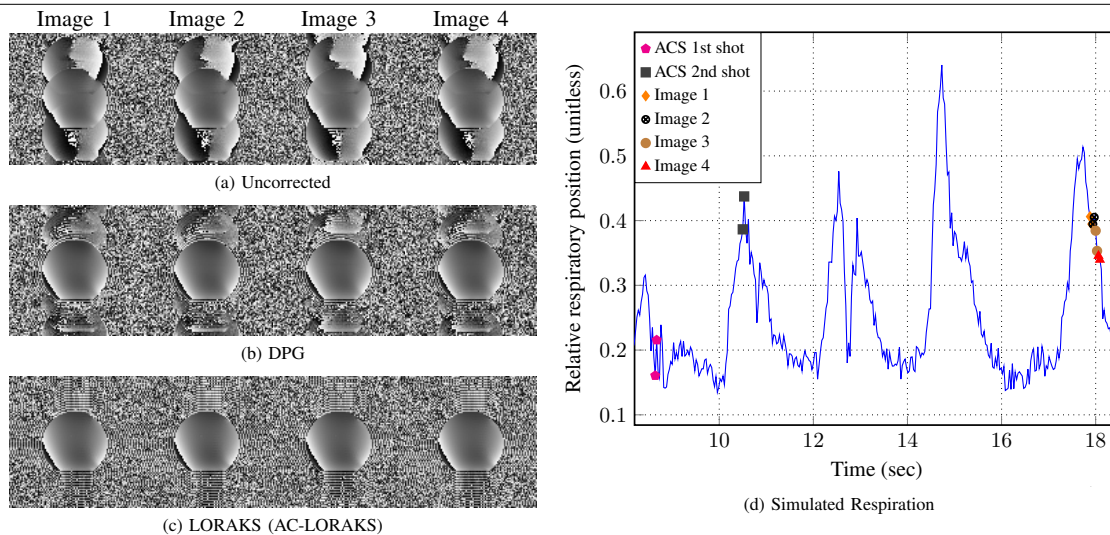
Figure 2.15: Evaluation with multi-shot EPI data for a phantom with simulated respiratory effects. A representative set of four segmented images is shown, extracted from a longer acquisition spanning several minutes. (a) Phase images corresponding to one channel of the data, with images reconstructed without compensating for the mismatches between different shots and different gradient polarities. (b) Phase images for one channel of the DPG reconstruction, with reconstruction performed using multi-channel data. (c) Phase images for single-channel LORAKS reconstruction with AC-LORAKS constraints. (d) Plot showing the relative respiratory position across EPI shots (TR=60msec), as measured with an ultrasound transducer coupled to a respiratory phantom [2]. The line-plot peaks show points where the phantom air bag is maximally inflated. The sampling times for the ACS data and for the shots used to generate each of the two-shot images from (a)-(c) are marked as labeled in the legend.

with time-varying characteristics. On the other hand, LORAKS with AC-LORAKS demonstrates robustness against the time-varying changes in this dataset, even despite the challenging single-channel multi-shot nature of this reconstruction problem.

For further insight, Fig. 2.16(a) shows an even more challenging case where LORAKS with AC-LORAKS constraints is used to reconstruct a single-shot of the previous multi-shot dataset. Note that this corresponds to EPI with an acceleration factor of $R = 2$ (i.e., an effective acceleration factor of $R = 4$ for each gradient polarity). Remarkably, we observe that single-channel LORAKS with AC-LORAKS constraints is still successful in this very difficult scenario, with similar quality to that obtained using multi-channel LORAKS with AC-LORAKS constraints. However, it is also important for us to point out that single-channel reconstruction with $R = 2$ is not always successful, as illustrated in Fig. 2.16(b) using previously described *in vivo* human brain data. In this brain case, we observe that single-channel LORAKS reconstruction is unsuccessful at correctly reconstructing the image, while the multi-channel case (also shown in Fig. 2.5) yields accurate results as described

Zero-filled   LORAKS   LORAKS
Uncorrected   (single-channel)   (multi-channel)

(a) Phantom dataset with simulated respiration

(b) *In vivo* human brain dataset
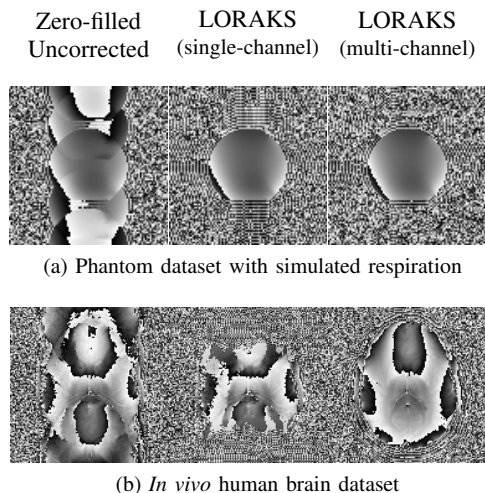
Figure 2.16: Images reconstructed using LORAKS with AC-LORAKS constraints for accelerated ($R = 2$) data in both single-channel and multi-channel contexts. (a) Phantom dataset with simulated respiration. (b) *In vivo* human brain dataset. Phase images from a single channel are shown for zero-filled data reconstructed without compensating for mismatches between the gradient polarities, LORAKS reconstruction with AC-LORAKS constraints from single-channel data, and LORAKS reconstruction with AC-LORAKS constraints from multi-channel data.

previously. We suspect that the difference between the phantom result and the *in vivo* result is explained by differences in the size of the FOV relative to the size of the object. Specifically, the size of the FOV is more than twice the size of the phantom, while the same is not true for the *in vivo* data. We expect that LORAKS-based single-channel image reconstruction will be harder for tight FOVs than it is for larger FOVs that contain a larger amount of empty space.

2.6   Discussion and Conclusions

This paper derived novel theoretical results for EPI ghost correction based on structured low-rank matrix completion approaches. Key theoretical results include the observation that the corresponding matrix completion problem is ill-posed in the absence of additional constraints, and that convex formulations have undesirable characteristics that are somewhat mitigated by the use of nonconvex formulations. These theoretical results led to two novel problem formulations that use additional constraints and nonconvex regularization to avoid the problems associated with ill-posedness. Our results showed that these new approaches are both effective relative to state-of-the-art ghost correction techniques like DPG, and are capable of handling nonlinear 2D phase mismatches between
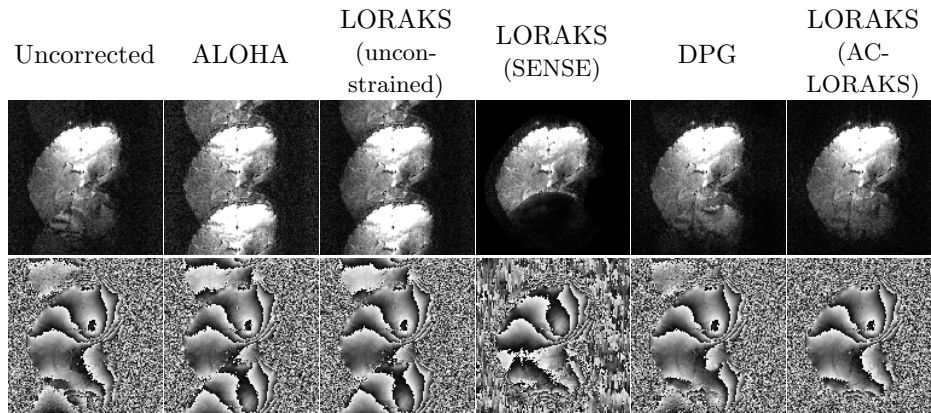
Figure 2.17: (top) Magnitude and (bottom) phase images corresponding to reconstruction of unaccelerated $(R = 1)$ *in vivo* double-oblique single-channel single-shot EPI data.
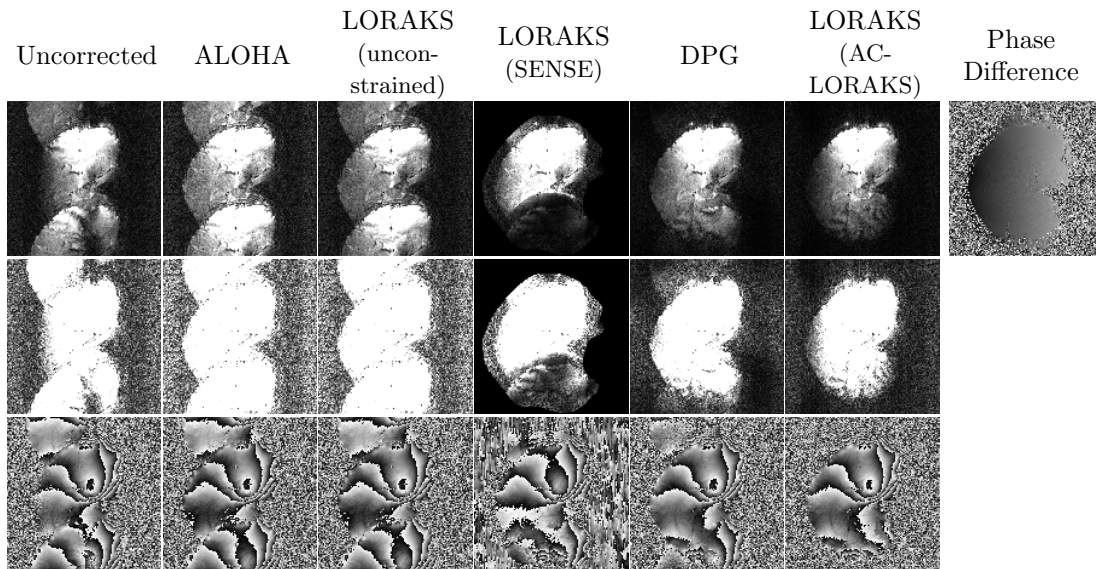
Figure 2.18: (top, middle) Magnitude and (bottom) phase images corresponding to reconstruction of unaccelerated $(R = 1)$ *in vivo* double-oblique single-channel single-shot EPI with strong simulated phase differences between RO^+ and RO^-. The middle row shows a different windowing of the magnitude images to more clearly visualize ghost-artifacts. The top-right image shows the simulated phase difference between RO^+ and RO^-.

RO^+ and RO^-. It was also observed that the proposed variation that uses AC-LORAKS constraints appears to be more effective than the proposed variation that uses SENSE constraints in challenging scenarios with single-channel data or highly-accelerated multi-channel data. In addition, we were surprised to observe that the variation using AC-LORAKS constraints can even be successful when applied to undersampled single-channel data. We believe that these approaches will prove valuable across a range of different applications, but especially those in which navigator-based ghost correction methods are ineffective, in which the need for navigation places undesirable constraints on the minimum achievable echo time or repetition time, in cases where single-channel imaging is unavoidable, or in cases where imaging conditions are likely to change over time over the duration of a long experiment (e.g., functional imaging or diffusion imaging).

This paper focused on simple proof-of-principle demonstrations of the proposed approaches, and there are opportunities for a variety of improvements. For example, the choices of r and λ were made manually using heuristic approaches in this paper, though it is likely that these same decision processes could be made automatically using standard techniques for rank estimation (e.g., [52]) or ghost correction parameter tuning (e.g., [23]). Additionally, extensions to the case of simultaneous multi-slice imaging are possible within the LORAKS framework [53], and are likely to be practically useful in a range of experiments given the advantages and the modern popularity of combining this approach with EPI [54].

While we derived our new theory and methods in the context of EPI ghost correction, we believe that this paper also has more general consequences. For example, we believe that it is straightforward to generalize our theoretical results to show that unconstrained LORAKS-based reconstruction will be ill-posed for any application that uses uniformly-undersampled Cartesian k-space trajectories (also see similar comments in Ref. [35]). We also believe that the combination of LORAKS with additional constraints will always be beneficial when the constraints are accurate, and so encourage the use of additional constraints whenever the LORAKS reconstruction problem is ill-posed. In addition, we believe that our novel LORAKS formulation with AC-LORAKS constraints is an important innovation that is likely to be useful in other applications, similar to how LORAKS with SENSE constraints has already proven to be useful in other settings [35, 55]. For

example, it may be beneficial to use AC-LORAKS constraints to augment existing SLM approaches that have recently been proposed for gradient delay estimation in non-Cartesian MRI [56,57]. Finally, while recent work has proposed the use of convex LORAKS-based formulations, our empirical experience since we first started exploring LORAKS several years ago [32,58] has consistently been that nonconvex formulations are substantially more powerful than convex ones.

2.7 Supplementary Information

2.7.1 Proof of Theorem 1

Assuming the notation of Eq. (2.5), let $\mathbf{B}$ denote the zero-filled LORAKS matrix

$$\mathbf{B} = \left[\begin{array}{cc} \mathbf{C}_P(\mathbf{A}_+^H \mathbf{d}_{\text{tot}}^+) & \mathbf{C}_P(\mathbf{A}_-^H \mathbf{d}_{\text{tot}}^-) \end{array} \right] \qquad (2.17)$$

and let $\mathbf{D}$ denote the matrix corresponding to the unmeasured data samples

$$\mathbf{D} = \left[\begin{array}{cc} \mathbf{C}_P(\mathbf{A}_-^H \mathbf{A}_- \mathbf{y}) & \mathbf{C}_P(\mathbf{A}_+^H \mathbf{A}_+ \mathbf{z}) \end{array} \right]. \qquad (2.18)$$

Due to the way the LORAKS $\mathbf{C}$-matrix is constructed, if the entry in the mth column and nth row of the matrix $\mathbf{B}$ is nonzero, then the corresponding entry of the matrix $\mathbf{D}$ is required to be zero and vice versa. Note also that the matrix from Eq. (2.10) can be written as $\mathbf{B} + \mathbf{D}$, while the matrix from Eq. (2.11) can be written as $\mathbf{B} - \mathbf{D}$.

To avoid additional tedious notation, we will assume in our proof sketch that the rows and columns of the matrix $\mathbf{B}$ have been permuted in such a way that samples from even and odd lines in k-space are never adjacent to one another in the matrix, which is always possible based on the convolutional structure of the LORAKS $\mathbf{C}$-matrix. This allows the $\mathbf{B}$ matrix to be written in a "checkerboard" form

$$\mathbf{B} = \left[\begin{array}{ccccccc|ccccccc} b_{11}^+ & 0 & b_{13}^+ & 0 & b_{15}^+ & \cdots & 0 & b_{12}^- & 0 & b_{14}^- & 0 & \cdots \\ 0 & b_{22}^+ & 0 & b_{24}^+ & 0 & \cdots & b_{21}^- & 0 & b_{23}^- & 0 & b_{25}^- & \cdots \\ b_{31}^+ & 0 & b_{33}^+ & 0 & b_{35}^+ & \cdots & 0 & b_{32}^- & 0 & b_{34}^- & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \right], \qquad (2.19)$$

where b_{ij}^+ and b_{ij}^- are the nonzero entries of the $\mathbf{B}$ matrix corresponding to positive and negative readout polarities, respectively. Using the same permutation scheme, the matrix $\mathbf{D}$ can similarly be written in the corresponding complementary "checkerboard" form:

$$
\mathbf{D} = \left[\begin{array}{ccccccc|ccccccc}
0 & d_{12}^+ & 0 & d_{14}^+ & 0 & \cdots & d_{11}^- & 0 & d_{13}^- & 0 & d_{15}^- & \cdots \\
d_{21}^+ & 0 & d_{23}^+ & 0 & d_{25}^+ & \cdots & 0 & d_{22}^- & 0 & d_{24}^- & 0 & \cdots \\
0 & d_{32}^+ & 0 & d_{34}^+ & 0 & \cdots & d_{31}^- & 0 & d_{33}^- & 0 & d_{35}^- & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array} \right],
\tag{2.20}
$$

Consider the diagonal matrix $\mathbf{Q}_1$ which has the same number of columns as $\mathbf{B}$, and whose diagonal entries alternate in sign in a way that follows the non-zero pattern of the first row of $\mathbf{B}$:

$$
\mathbf{Q}_1 = \left[\begin{array}{cccccc|cccccc}
1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & -1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & -1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
\hline
0 & 0 & 0 & 0 & 0 & \cdots & -1 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & -1 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & -1 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array} \right].
\tag{2.21}
$$

Similarly, consider the diagonal matrix $\mathbf{Q}_2$ which has the same number of rows as $\mathbf{B}$, and whose diagonal entries alternate in sign in a way that follows the non-zero pattern of the first column of $\mathbf{B}$:

$$\mathbf{Q}_2 = \begin{bmatrix} 1 & 0 & 0 & \cdots \\ 0 & -1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{2.22}$$

We make the following observations:

- The matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are unitary and satisfy $\mathbf{Q}_1^{-1} = \mathbf{Q}_1$ and $\mathbf{Q}_2^{-1} = \mathbf{Q}_2$.

- The matrix $\mathbf{B}$ is structured in such a way that $\mathbf{B}\mathbf{Q}_1 = \mathbf{Q}_2\mathbf{B}$.

- The matrix $\mathbf{D}$ is structured in such a way that $\mathbf{D}\mathbf{Q}_1 = -\mathbf{Q}_2\mathbf{D}$.

- The matrix $\mathbf{Q}_2(\mathbf{B} + \mathbf{D})\mathbf{Q}_1$ simplifies according to

$$\begin{aligned} \mathbf{Q}_2(\mathbf{B} + \mathbf{D})\mathbf{Q}_1 &= \mathbf{Q}_2\mathbf{Q}_2(\mathbf{B} - \mathbf{D}) \\ &= \mathbf{B} - \mathbf{D}. \end{aligned} \tag{2.23}$$

From Eq. (2.23), we can infer that if we write the singular value decomposition of $\mathbf{B} + \mathbf{D}$ as $\mathbf{B} + \mathbf{D} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$, then the matrix $\mathbf{B} - \mathbf{D}$ can be written as $\tilde{\mathbf{U}}\boldsymbol{\Sigma}\tilde{\mathbf{V}}^H$, where $\tilde{\mathbf{U}} = \mathbf{Q}_2\mathbf{U}$ and $\tilde{\mathbf{V}} = \mathbf{Q}_1\mathbf{V}$. Since $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$ are matrices with orthonormal columns, we must have that $\tilde{\mathbf{U}}\boldsymbol{\Sigma}\tilde{\mathbf{V}}^H$ is a valid singular value decomposition of $\mathbf{B} - \mathbf{D}$. Thus, we can conclude that $\mathbf{B} - \mathbf{D}$ and $\mathbf{B} + \mathbf{D}$ have identical singular values. This completes the proof of the theorem. $\square$

2.7.2   Proof of Corollary 1

Assume that the vectors $\mathbf{y}$ and $\mathbf{z}$ are chosen such that the expressions in Eq. (2.9) represent an optimal solution to Eq. (2.5). Theorem 1 then tells us that the vectors from Eq. (2.12) represent another optimal solution to Eq. (2.5). These two solutions are identical to one another if and only if $\mathbf{A}_-^H\mathbf{A}_-\mathbf{y} = \mathbf{0}$ and $\mathbf{A}_+^H\mathbf{A}_+\mathbf{z} = \mathbf{0}$, in which case both solutions are equal to the zero-filled solution. In

this case, the zero-filled solution is clearly an optimal solution to Eq. (2.5), and must be the unique

optimal solution if Eq. (2.5) only has a single solution. If either $\mathbf{A}_-^H \mathbf{A}_- \mathbf{y} \neq \mathbf{0}$ or $\mathbf{A}_+^H \mathbf{A}_+ \mathbf{z} \neq \mathbf{0}$, then

Eqs. (2.9) and (2.12) represent two distinct solutions to Eq. (2.5), indicating that Eq. (2.5) does

not have a unique solution. $\square$

2.7.3   Proof of Corollary 2

Based on the proof of Corollary 1, we know that if Eq. (2.5) has a solution that is not equal to the

zero-filled measured data, then we can use Eqs. (2.9) and (2.12) to obtain a pair of two distinct

solutions to Eq. (2.5). Let $\mathbf{k}_1^{\pm}$ and $\mathbf{k}_2^{\pm}$ denote these two solutions, and notice that by the definition

of an optimal solution, we must have that $L_{\mathbf{C}}(\mathbf{k}_1^{\pm}) = L_{\mathbf{C}}(\mathbf{k}_2^{\pm})$ and that $L_{\mathbf{C}}(\mathbf{k}_1^{\pm}) \leq L_{\mathbf{C}}(\mathbf{k})$ for all

possible candidate solutions $\mathbf{k}$.

Corollary 2 is easily proven based on the definition of a convex function. Specifically, if $L_{\mathbf{C}}(\mathbf{y})$

is convex, then it must satisfy [45]

$$L_{\mathbf{C}}\left(\alpha\mathbf{y}_1 + (1-\alpha)\mathbf{y}_2\right) \leq \alpha L_{\mathbf{C}}\left(\mathbf{y}_1\right) + (1-\alpha)L_{\mathbf{C}}\left(\mathbf{y}_2\right), \tag{2.24}$$

for every possible pair of vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ and for every real-valued scalar α between 0 and 1.

Setting $\mathbf{y}_1 = \mathbf{k}_1^{\pm}$ and $\mathbf{y}_2 = \mathbf{k}_2^{\pm}$ in Eq. (2.24) leads to

$$\begin{aligned} L_{\mathbf{C}}\left(\alpha\mathbf{k}_1^{\pm} + (1-\alpha)\mathbf{k}_2^{\pm}\right) &\leq \alpha L_{\mathbf{C}}\left(\mathbf{k}_1^{\pm}\right) + (1-\alpha)L_{\mathbf{C}}\left(\mathbf{k}_2^{\pm}\right) \\ &= L_{\mathbf{C}}\left(\mathbf{k}_1^{\pm}\right), \end{aligned} \tag{2.25}$$

Combining Eq. (2.25) with the previous observation that $L_{\mathbf{C}}(\mathbf{k}_1^{\pm}) \leq L_{\mathbf{C}}(\mathbf{k})$ for all possible candidate

solutions $\mathbf{k}$ implies that $L_{\mathbf{C}}\left(\alpha\mathbf{k}_1^{\pm} + (1-\alpha)\mathbf{k}_2^{\pm}\right) = L_{\mathbf{C}}\left(\mathbf{k}_1^{\pm}\right)$. As a result, $\alpha\mathbf{k}_1^{\pm} + (1-\alpha)\mathbf{k}_2^{\pm}$ must also

be an optimal solution of Eq. (2.5) for every possible choice of $0 \leq \alpha \leq 1$, and we have successfully

proven that there exist an infinite number of solutions.

Specifically, an infinite set of solutions is given by $\{\mathbf{k}_\alpha^{\pm} \equiv \alpha\mathbf{k}_1^{\pm} + (1-\alpha)\mathbf{k}_2^{\pm} \mid \alpha \in [0,1]\}$.

Additionally, the zero-filled solution is obtained as one of these solutions, corresponding to the

specific choice of $\alpha = 0.5$. $\square$

2.7.4   Majorize-Minimize Algorithm

For the sake of completeness, this section provides an operational description of the majorize-minimize algorithm we use to solve Eq. (2.14). The key observation [10, 32] that enables this algorithm is that the function $J_r(\mathbf{G})$ from Eq. (2.8) is majorized at a point $\hat{\mathbf{G}}^{(j)}$ by the function

$$g\left(\mathbf{G}, \hat{\mathbf{G}}\right) = \|\mathbf{G} - \hat{\mathbf{G}}^{(j)}\|_F^2. \tag{2.26}$$

Using this majorant, Eq. (13) can be optimized by iteratively solving easy-to-optimize surrogate problems. Specifically, using $\hat{\boldsymbol{\rho}}^{(j)}$ to denote the estimate of $\{\boldsymbol{\rho}^+, \boldsymbol{\rho}^-\}$ at the jth iteration, and given some initial guess $\hat{\boldsymbol{\rho}}^{(0)}$, we minimize Eq. (2.14) by iteratively solving the surrogate problems

$$\begin{aligned}
\hat{\boldsymbol{\rho}}^{(j+1)} = \arg \min_{\{\boldsymbol{\rho}^+, \boldsymbol{\rho}^-\}} &\|\mathbf{E}_+\boldsymbol{\rho}^+ - \mathbf{d}_{\text{tot}}^+\|_2^2 \\
&+ \|\mathbf{E}_-\boldsymbol{\rho}^- - \mathbf{d}_{\text{tot}}^-\|_2^2 \\
+ \lambda\| \left[\begin{array}{cc} \mathbf{S}_P(\mathbf{E}_+\boldsymbol{\rho}^+) & \mathbf{S}_P(\mathbf{E}_-\boldsymbol{\rho}^-) \end{array} \right] &- \mathbf{G}_r^{(j)}\|_F^2
\end{aligned} \tag{2.27}$$

for $j=0, 1, 2, \ldots$ until convergence, where $\mathbf{G}_r^{(j)}$ is the optimal rank-r approximation (obtained by truncating the singular value decomposition) of the matrix formed by numerically evaluating $\left[\begin{array}{cc} \mathbf{S}_P(\mathbf{E}_+\boldsymbol{\rho}^+) & \mathbf{S}_P(\mathbf{E}_-\boldsymbol{\rho}^-) \end{array} \right]$ at the point $\{\boldsymbol{\rho}^+, \boldsymbol{\rho}^-\} = \hat{\boldsymbol{\rho}}^{(j)}$. These surrogate problems have the form of simple linear least squares problems, and can be solved using standard iterative algorithms like the conjugate gradient algorithm. See Refs. [10, 32] for a more detailed description and justification of this algorithmic approach.

Chapter 3

Robust Autocalibrated Structured Low-Rank EPI Ghost Correction

3.1 Introduction

Echo-planar imaging (EPI) is a widely-used high-speed MRI acquisition strategy [59], but is subject to several undesirable artifacts [18]. Nyquist ghosts are one of the most common EPI artifacts, and occur because of systematic differences between the interleaved lines of k-space that are acquired with different readout gradient polarities, and/or because of systematic differences between interleaved lines of k-space data that are acquired with different shots in a multi-shot acquisition. Despite substantial efforts over several decades to solve this problem [1,14,18–30,39,40,60,61], the widely-deployed modern ghost correction schemes are still prone to incomplete ghost suppression, as illustrated in Fig. 3.1.

Recently, structured low-rank matrix methods for ghost correction [14,39,40,60,61] have received increasing attention for their ability to provide excellent ghost-suppression performance without the need for additional "navigator" information (i.e., reference scans collected alongside each EPI readout that allow estimation of the systematic inconsistencies between different gradient polarities or different shots). These methods can suppress ghosts better than navigator-based methods, and eliminate the need to acquire navigators for each EPI readout. Among different structured low-rank matrix approaches, a ghost correction method based on Autocalibrated LORAKS (AC-LORAKS) [11] was previously demonstrated to yield high-quality results across a range of different scenarios
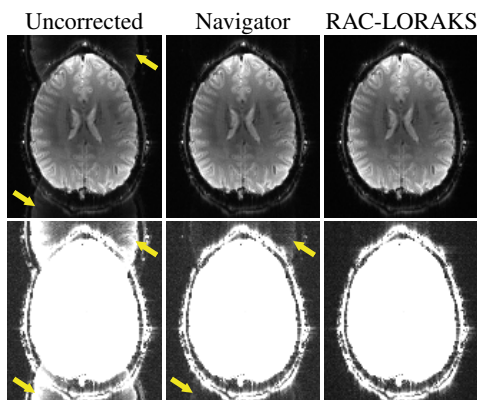
Figure 3.1: Illustration of EPI ghost correction. The top row of this figure shows EPI images obtained from different methods, while the bottom row shows the same images with 10× intensity amplification to highlight the ghost characteristics. If EPI data is naively reconstructed without accounting for the systematic differences between data acquired with positive and negative readout gradient polarities ("Uncorrected"), then strong Nyquist ghosts appear in the image as indicated with arrows. Modern EPI techniques frequently try to eliminate these artifacts using navigator information to estimate the systematic differences between the data collected with different readout polarities. In the navigator-based example we show ("Navigator"), the navigator information was collected using a 3-line EPI acquisition with the phase encoding gradients turned off, and the difference between positive and negative gradient polarities was modeled using constant and 1D linear phase terms. Although this approach substantially reduces Nyquist ghosts, it is common for some amount of residual ghosting to still be present in the images, particularly in cases where simple 1D phase modeling is inadequate to capture the differences between the two gradient polarities. We also show an example of our proposed approach ("RAC-LORAKS"), which can account for more complicated variations between the different gradient polarities, and which is substantially more successful at suppressing Nyquist ghosts in this example.

[14]. To eliminate a fundamental ambiguity in structured low-rank matrix recovery from uniformly undersampled EPI data [14], AC-LORAKS makes use of parallel imaging subspace information estimated from autocalibration (ACS) data acquired in a pre-scan. This ACS-based approach is similar to standard autocalibrated parallel imaging methods like GRAPPA [5], SPIRiT [7], and PRUNO [9].

While the AC-LORAKS approach to ghost correction generally works well when the ACS data is pristine and well-matched to the EPI data to be reconstructed, there are many situations where experimental conditions (e.g., subject motion, eddy currents, etc.) can lead to artifacts within the ACS data or mismatches between the ACS and EPI data. The performance of the AC-LORAKS ghost correction procedure degrades in the presence of these ACS artifacts and mismatches. Note that this kind of issue is not unique to AC-LORAKS or to ghost correction, and imperfect ACS/calibration

data is a longstanding and commonly-reported problem for all calibration-based image reconstruction methods [28, 62–66]. For AC-LORAKS ghost correction, imperfect ACS data can be especially troublesome in contexts where the prescan would be done once before acquiring a long sequence of multiple EPI images (e.g., in BOLD fMRI or diffusion MRI applications), and then used to reconstruct each image in the sequence.

In this paper, we propose an extension of AC-LORAKS for EPI ghost correction that is more robust to imperfections in the ACS data. The new method, called Robust Autocalibrated LORAKS (RAC-LORAKS), has two major differences from the previous AC-LORAKS approach. First, RAC-LORAKS does not completely trust the subspace information learned from the ACS data, but rather uses a novel structured low-rank matrix formulation that learns subspace information jointly from both the (imperfect) ACS data and the EPI data being reconstructed. To the best of our knowledge, no previous methods have used this kind of approach to address the longstanding issue of imperfect ACS data. And second, RAC-LORAKS uses the ACS data to provide additional complementary information for the reconstruction of the EPI data within a multi-contrast joint reconstruction framework [67]. Preliminary accounts of the first strategy were previously reported in recent conferences [68, 69], although we have not previously reported the combination with the second strategy.

3.2 Theory

Due to space constraints, our descriptions in this paper will assume that the reader is already familiar with the basic physics of EPI. Readers interested in a more detailed explanation are referred to classic references [18, 59]. For simplicity, our description of EPI ghost correction will generally assume that we are correcting ghosts associated with the differences between data acquired with different readout gradient polarities in a single-shot EPI experiment. However, since the ghost model for bipolar gradients is nearly identical to the ghost model for multi-shot acquisition, the same approach is easily adapted *mutatis mutandis* to multi-shot acquisition with an arbitrary number of shots [14].

3.2.1   Background: Structured Low-Rank EPI ghost correction

Structured low-rank matrix methods for EPI ghost correction [14, 39, 40, 60, 61] can be viewed as an extension of structured low-rank matrix methods for conventional MR image reconstruction [8, 9, 12, 31, 32, 36, 37, 70], and are based on the same underlying theoretical principles. In particular, it has been shown that when MRI images have limited support, smooth phase variations, multi-channel correlations, or transform-domain sparsity, then the MRI k-space data will be linearly predictable [13], which means that convolutional Hankel- or Toeplitz-structured matrices formed from the k-space data will possess low-rank characteristics. This observation means that MRI reconstruction can be reformulated as structured low-rank matrix recovery. Importantly, these structured low-rank matrix recovery methods can even be successful in calibrationless scenarios where ACS data or other prior information about the spatial support, phase, or multi-channel sensitivity profiles is not available [12, 31, 32].

Structured low-rank EPI ghost correction methods combine these principles with the earlier observation that EPI data acquired from different gradient polarities or different shots can be treated as coming from different effective "channels" in a parallel imaging experiment, where the systematic differences between different polarities or shots lead to different phase or magnitude modulations of the underlying EPI image [27–30]. Since structured low-rank methods for conventional image reconstruction automatically account for the unknown sensitivity maps that modulate the underlying image in a parallel imaging experiment, it is reasonable to apply these same types of methods to handle the unknown polarity- or shot-dependent modulations that manifest in EPI ghost correction.

For the sake of brevity, we will focus the remainder of our review on the AC-LORAKS method for EPI ghost correction [14], since our proposed RAC-LORAKS method is a generalization of AC-LORAKS. The AC-LORAKS method for EPI ghost correction is based on solving the following regularized optimization problem subject to exact data consistency constraints:

$$\left\{\hat{\mathbf{k}}^+, \hat{\mathbf{k}}^-\right\} = \underset{\{\mathbf{k}^+, \mathbf{k}^-\}}{\arg\min} \left\|\mathcal{P}_C(\mathbf{k}^+)\mathbf{N}\right\|_F^2 + \left\|\mathcal{P}_C(\mathbf{k}^-)\mathbf{N}\right\|_F^2 + \lambda J_r([\mathcal{P}_S(\mathbf{k}^+)\ \mathcal{P}_S(\mathbf{k}^-)]). \qquad (3.1)$$

In this expression, $\mathbf{k}^+$ and $\mathbf{k}^-$ respectively represent the ideal fully-sampled multi-channel Cartesian k-space data for the positive and negative readout gradient polarities; $\mathcal{P}_C(\cdot)$ is the LORAKS operator that maps the k-space data into a structured matrix that should possess low-rank if the multi-channel image possess limited support and/or interchannel parallel imaging correlations; $\mathcal{P}_S(\cdot)$ is the LORAKS operator that maps the k-space data into a structured matrix that should possess low-rank if the multi-channel image possess limited support, smooth phase, and/or interchannel parallel imaging correlations; the matrix $\mathbf{N}$ comprises an orthonormal (i.e., $\mathbf{N}^H\mathbf{N} = \mathbf{I}$) basis for the nullspace of the matrix $\begin{bmatrix} \mathcal{P}_C(\mathbf{k}_{\text{acs}+}) \\ \mathcal{P}_C(\mathbf{k}_{\text{acs}-}) \end{bmatrix}$, where $\mathbf{k}_{\text{acs}}^+$ and $\mathbf{k}_{\text{acs}}^-$ respectively represent the ACS data for the positive and negative readout gradient polarities; λ is a regularization parameter; $J_r(\cdot)$ is a regularization penalty that promotes low-rank characteristics; and $\|\cdot\|_F$ denotes the Frobenius norm. Due to space constraints, this paper will not provide a detailed recipe for implementing the LORAKS operators $\mathcal{P}_C(\cdot)$ and $\mathcal{P}_S(\cdot)$, and simply note that our implementations for this paper are identical to those that are described in detail in earlier LORAKS papers [12,32]. There are theoretical benefits to choosing a nonconvex low-rank regularization penalty [14], and the previous AC-LORAKS approach for ghost correction [14] used the nonconvex function proposed in the original LORAKS paper [32] defined by

$$J_r(\mathbf{X}) = \min_{\mathbf{Y}} \|\mathbf{X} - \mathbf{Y}\|_F^2 \ \text{s.t.} \ \text{rank}(\mathbf{Y}) \leq r, \tag{3.2}$$

where r is a user-selected rank parameter, $\mathbf{X}$ is a matrix representing the point at which we are evaluating the function $J_r(\mathbf{X})$, and $\mathbf{Y}$ is an optimization variable of the same size as $\mathbf{X}$. This penalty encourages matrices that have accurate rank-r approximations.

The first two terms appearing on the right hand side of Eq. (3.1) respectively impose limited support and parallel imaging constraints on the reconstructions of the positive and negative readout gradient polarities. The constraints that are used in these terms are implicit in the low-rank characteristics of the structured LORAKS matrices, as captured by the nullspace matrix $\mathbf{N}$. The nullspace matrix is learned in advance from the ACS data, and as a result, there is an implicit assumption that the support and parallel imaging constraints that were valid for the ACS data are

also valid for the EPI data to be reconstructed. Note that if the third term were removed from Eq. (3.1), then these first two terms would reduce to performing separate PRUNO [9] or conventional AC-LORAKS [11] reconstructions of the data from each polarity. Acquiring ACS/calibration data is relatively fast and easy, and is already a standard part of most modern parallel imaging protocols, so is not very burdensome on the acquisition. Using ACS data can also be important in this context, since it has been mathematically proven that structured low-rank matrix methods for ghost correction suffer from fundamental ambiguities unless some form of side information is available [14]. While other options exist for removing ambiguity (e.g., using SENSE-like [6] image-domain constraints [14, 40]), it was previously observed that the AC-LORAKS approach (i.e., using GRAPPA-like [5] Fourier-domain constraints) offered better performance [14].

The third term of Eq. (3.1) couples the reconstruction of the two polarities together, allowing the reconstruction of one polarity to benefit from information from the other polarity, while also introducing phase constraints to allow the reconstruction to benefit from k-space conjugate symmetry characteristics. In particular, the third term implicitly and automatically imposes the following constraints whenever they are compatible with the measured data: limited image support, smooth phase, interchannel parallel imaging correlations, and interpolarity correlations. Notably, these constraints are all imposed implicitly through the nullspace of a structured matrix, and if a given constraint is not compatible with the measured data, then that constraint will automatically not be imposed by the reconstruction procedure [13].

The ACS data for AC-LORAKS ghost correction has typically been acquired using the same process used by Dual Polarity GRAPPA (DPG) [28, 71–73]. In particular, assuming a parallel imaging acceleration factor of R, DPG employs a $2R$-shot EPI prescan. The data from different shots and different gradient polarities is then rearranged and interleaved to form one fully-sampled ACS dataset comprised only of data acquired with a positive readout gradient polarity ($\mathbf{k}_{\mathrm{acs}}^{+}$) and another fully-sampled ACS dataset comprised only of data acquired with a negative readout gradient polarity ($\mathbf{k}_{\mathrm{acs}}^{-}$). Since this ACS acquisition strategy is based on a multi-shot approach, it therefore may be prone to ghosting artifacts due to shot-to-shot variations. In addition, since the ACS data is often acquired only once at the beginning of a long multi-image EPI scan (e.g., in BOLD

fMRI or diffusion MRI experiments), the ACS data acquired at the beginning of the experiment may gradually become mismatched with data acquired at later time points due to scanner drift, subject motion, etc. As noted previously, the ghost correction performance of AC-LORAKS can be substantially degraded when there are mismatches between the ACS data and the EPI data to be reconstructed. Although a pre-processing procedure has been previously developed to correct for shot-to-shot variations in the ACS data for DPG [28], this approach is not sufficient for the present context. In particular, this approach undesirably modifies the magnitude and phase characteristics of the ACS data in ways that are not well-suited for AC-LORAKS, and only addresses ACS artifacts without accounting for mismatches that may exist between the ACS data and the EPI data.

3.2.2   RAC-LORAKS

Our proposed RAC-LORAKS method is based on solving the following optimization problem

$$\left\{\hat{\mathbf{k}}^+, \hat{\mathbf{k}}^-, \hat{\mathbf{N}}\right\} = \underset{\{\mathbf{k}^+, \mathbf{k}^-, \mathbf{N}\}}{\arg\min} \left\|\mathcal{P}_C(\mathbf{k}^+)\mathbf{N}\right\|_F^2 + \left\|\mathcal{P}_C(\mathbf{k}^-)\mathbf{N}\right\|_F^2 \tag{3.3}$$
$$+\eta \left\|\mathcal{P}_C(\mathbf{k}_{\mathrm{acs}}^+)\mathbf{N}\right\|_F^2 + \eta \left\|\mathcal{P}_C(\mathbf{k}_{\mathrm{acs}}^-)\mathbf{N}\right\|_F^2$$
$$+\lambda J_r \left(\begin{bmatrix}\mathcal{P}_S(\mathbf{k}^+) & \mathcal{P}_S(\mathbf{k}^-) & \mathcal{P}_S(\mathbf{k}_{\mathrm{acs}}^+) & \mathcal{P}_S(\mathbf{k}_{\mathrm{acs}}^-)\end{bmatrix}\right)$$

subject to exact data consistency constraints on $\mathbf{k}^+$ and $\mathbf{k}^-$ and subject to orthonormality constraints on $\mathbf{N}$ such that $\mathbf{N}^H\mathbf{N} = \mathbf{I}$. This optimization problem involves four user-selected parameters: the regularization parameters η and λ, the rank parameter r, and the number of columns p of the matrix $\mathbf{N}$ (which determines the dimension of the approximate nullspace).

Equation (3.3) has two main differences from Eq. (3.1). First, instead of choosing a predetermined value of the approximate nullspace matrix $\mathbf{N}$ that depends only on the ACS data, $\mathbf{N}$ is now an optimization variable that depends on both the ACS data and the EPI data to be reconstructed. This allows the reconstruction to be more robust against possible imperfections in the ACS data. The extent to which the ACS data is trusted is controlled by the user-selected parameter η. In the limit as $\eta \to \infty$, the approximate nullspace matrix $\mathbf{N}$ will converge to the fixed matrix from Eq. (3.1).

The second difference is that the final term of Eq. (3.3) now includes structured matrices formed from the ACS data, in addition to the previous structured matrices formed from the EPI data to be reconstructed. By concatenating the ACS data in this way, we are essentially treating the ACS data in the same way that we would treat additional channels in a parallel imaging experiment. Although the ACS data may not have the same contrast as the EPI data to be reconstructed, it has previously been shown that treating multi-contrast information like additional channels in a parallel imaging experiment often leads to improved reconstruction performance [67]. While this improvement has been justified empirically, some level of theoretical justification for this approach can be obtained by modeling different image contrasts as different modulations of some latent image [13].

Algorithmically, Eq. (3.3) can be minimized using existing algorithms for LORAKS optimization [12,14,32,74]. In particular, it is not hard to show that the solution to Eq. (3.3) can be equivalently obtained by solving:

$$
\begin{aligned}
\left\{ \hat{\mathbf{k}}^+, \hat{\mathbf{k}}^- \right\} \;\; = \;\; & \operatorname*{arg\,min}_{\{\mathbf{k}^+, \mathbf{k}^-\}} J_{(C-p)} \left(\begin{bmatrix} \mathcal{P}_C(\mathbf{k}^+) \\ \mathcal{P}_C(\mathbf{k}^-) \\ \sqrt{\eta}\mathcal{P}_C(\mathbf{k}_{\mathrm{acs}}^+) \\ \sqrt{\eta}\mathcal{P}_C(\mathbf{k}_{\mathrm{acs}}^-) \end{bmatrix} \right) \\
& + \lambda J_r \left(\begin{bmatrix} \mathcal{P}_S(\mathbf{k}^+) & \mathcal{P}_S(\mathbf{k}^-) & \mathcal{P}_S(\mathbf{k}_{\mathrm{acs}}^+) & \mathcal{P}_S(\mathbf{k}_{\mathrm{acs}}^-) \end{bmatrix} \right),
\end{aligned}
\tag{3.4}
$$

where $J_{(C-p)}(\cdot)$ is the same as $J_r(\cdot)$ but replacing the rank parameter r with the rank parameter $(C-p)$, where C is the number of columns of the LORAKS matrix formed by $\mathcal{P}_C(\cdot)$. Equation (3.4) is convenient because it takes the same form as previous LORAKS optimization problems involving multiple $J_r(\cdot)$ terms [32]. For this paper, we use a multiplicative half-quadratic majorize-minimize algorithm to minimize this objective function [74], which takes advantage of FFT-based matrix multiplications to improve computational complexity [75].

The RAC-LORAKS solution is obtained through the optimization of a nonconvex cost function. As such, the algorithm has the potential to converge to an undesirable local minimum. For the results shown in this paper, we initialize RAC-LORAKS using a naive initialization with minimal

processing cost as explained in the next section. Other choices could potentially result in even higher performance, but are not considered here.

3.3   Methods

3.3.1   Datasets used for Evaluation

As described below, we evaluated the characteristics of RAC-LORAKS using data from several different contexts. All in vivo data were acquired under IRB-approved written informed consent.

3.3.1.1   Gradient-Echo EPI Brain data

In one set of experiments, we acquired in vivo human brain data using a gradient-echo EPI sequence with parameters that are somewhat similar to a BOLD fMRI experiment. Data was acquired on a Siemens 3T Prisma Fit scanner using a standard 32-channel receiver array. The data was acquired using FOV = 220 mm × 220 mm; matrix size = 128 × 128; slice thickness = 3 mm; and TR = 2 sec. In one subject, data was acquired without acceleration ($R = 1$) with TE = 47 msec. From this same subject, data was also acquired for parallel imaging acceleration factors of $R = 2, 3, 4$ with TE = 35 msec. In a second subject, data was acquired for parallel imaging acceleration factors of $R = 5, 6$ with TE = 35 msec. In all cases, fully-sampled ACS data was acquired using the same interleaved $2R$-shot EPI prescan as used for DPG [28].

The previous datasets were acquired with a conventional axial slice orientation. However, because Nyquist ghost problems tend to be more extreme with oblique acquisitions due to concomitant fields that can produce substantial nonlinear 2D phase differences between positive and negative readout polarities [24, 28, 48–50], we also acquired an additional dataset with a double-oblique slice orientation from a third subject to test performance in a more challenging scenario. The slice orientation in this case is nonstandard and likely difficult to interpret for many readers, so we have depicted its position in Fig. 3.2. For this case the data was acquired with TR = 2.08 sec and TE = 35 msec for parallel imaging acceleration factors of $R = 1, 2, 3, 4, 5, 6$.
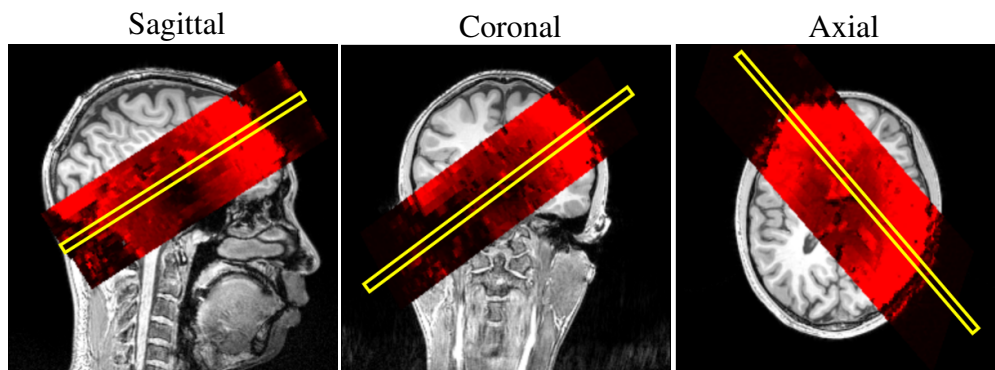
Figure 3.2: Illustration of the orientation of the double-oblique gradient-echo EPI dataset. The double-oblique slices are shown in red, overlaid on a structural T1-weighted image of the same subject. The double-oblique slice used for the results in Fig. 3.7 is shown with a yellow rectangle.

3.3.1.2 Diffusion-encoded EPI Brain Data

In another set of experiments, we acquired in vivo human brain data using a diffusion-encoded spin-echo EPI sequence. Diffusion EPI data might be considered more challenging than the previous gradient-echo EPI data, due to the fact that diffusion MRI data usually suffers from random image-to-image phase variations, and can also have lower SNR than gradient-echo EPI. In addition, the rapid switching of strong diffusion gradients can introduce substantial additional eddy current effects that can cause systematic differences between the ACS data and the diffusion EPI data if they are acquired with different diffusion gradient settings [76].

A first diffusion dataset was acquired on a Siemens 3T Prisma Fit scanner using a standard 32-channel receiver array. For the sake of computational complexity, this data was subsequently reduced to 16 channels using standard coil-compression techniques. The data was acquired using FOV = 220 mm × 220 mm; matrix size = 220 × 220; slice thickness = 5 mm; TR = 2.8 sec; TE = 63 msec; b-values of 0 sec/mm^2 and 1000 sec/mm^2; 6 diffusion encoding directions; parallel imaging acceleration factor $R = 3$; and 6/8ths partial Fourier sampling. ACS data was acquired using the same interleaved $2R$-shot EPI prescan as used for DPG [28], except that the data was acquired with lower resolution along the phase encoding dimension (i.e., we only acquired 45 phase-encoding lines for the ACS data). Due to the random phase variations associated with diffusion encoding gradients, the ACS data was acquired without diffusion weighting, which means that the ACS data has very different contrast characteristics from the EPI data. To show results across a broader

range of acceleration factors, a second set of acquisitions was performed with $R = 2, 3, 4, 5$. Other parameters were identical to the previous case, except for matrix size $= 110 \times 110$; slice thickness $= 2$ mm; TR $= 11.4$ sec; TE $= 73$ msec; and fully-sampled ACS data.

3.3.1.3   Cardiac EPI Data

In a third set of experiments, we acquired in vivo human cardiac data during diastole using a spin-echo EPI sequence with parameters that are typical for a myocardial arterial spin labeling experiment [77]. Data was acquired on a GE 3T Signa HDx scanner with an 8-channel cardiac coil. The acquisition used FOV $= 280$ mm $\times$ 140 mm; matrix size $= 128 \times 64$; slice thickness $= 10$ mm; TR $= 55$ msec; TE $= 32.9$ msec; velocity cutoff $= 5$ cm/s; no parallel imaging acceleration $(R = 1)$; and 5/8ths partial Fourier sampling. ACS data was acquired using the same interleaved $2R$-shot EPI prescan as used for DPG [28], but with 5/8ths partial Fourier sampling. Data was acquired with a double-oblique slice orientation to achieve a mid-short axis view.

3.3.2   Simulations

In addition to in vivo data, the different methods were also evaluated using simulations where a gold standard was present. To form a gold standard with realistic EPI characteristics, we took two in vivo gradient-echo EPI brain datasets (as described in Section 3.3.1.1) with axial slice orientation and $R = 1$ from the same scan session, and reconstructed them both using SENSE. Each gradient polarity was reconstructed separately, providing a realistic representation of typical interpolarity image differences. This procedure provides two sets of fully-sampled multi-channel dual-polarity gold standard images. One of these sets was used for ACS data, while the other was undersampled (including parallel imaging acceleration, along with interleaving the data from positive and negative gradient polarities) to simulate EPI data. These datasets were acquired roughly 5 minutes apart, allowing time for mismatches to evolve. Since ghost correction is frequently more difficult for EPI datasets with 2D nonlinear phase differences between the two polarities, we applied an additional 2D nonlinear phase pattern to make the problem more challenging. This additional phase difference was designed to be roughly $3\times$-larger than we observed in the real data from Fig. 3.7 .

In a first set of simulations, to mimic the situation where a localized image feature is different between the ACS data and EPI data (e.g., as may happen in a dynamic experiment), we added a Gaussian-shaped additive image hyperintensity to the EPI data that we did not add to the ACS data. The hyperintensity was designed to follow both the coil sensitivity maps (obtained by applying ESPIRiT [8]) and the phase characteristics of the original data. We also performed simulations with these two datasets interchanged, i.e., with the hyperintensity in the ACS data but not in the EPI data.

In another set of simulations, to mimic the situation where the ACS data and EPI data have very different contrasts, we inverted the magnitude image for the ACS data to create an image with different contrast [78], while still following the coil sensitivity maps and the phase characteristics of the original data.

In addition to performing multi-channel simulations, we also performed a simulation in a very challenging single-channel setting. For this, single-channel data was obtained by a linear combination of the multi-channel data [79]. Single-channel ghost correction is a difficult setting where only a few previous methods have had any success [14, 39]. This case is hard because even with unaccelerated data ($R = 1$), each polarity has an effective acceleration factor of $R = 2$ when the data for each readout gradient polarity is separated, and it can be difficult to reconstruct $R = 2$ data without multi-channel information.

The fully-sampled ACS and EPI datasets used for all three simulations are illustrated in Fig. 3.3 .

3.3.3 Data Processing

RAC-LORAKS was applied to perform reconstruction and ghost correction on these datasets. For comparison against existing methods, the datasets were also reconstructed using the previous AC-LORAKS ghost-correction method [14], DPG [28], and MUSSELS [40].

For some of the datasets we consider, the ACS data may be incomplete due to low-resolution ACS acquisition (i.e., the first brain EPI diffusion data) or partial Fourier ACS acquisition (i.e., the cardiac EPI data). In such cases, we modify RAC-LORAKS to consider the fully sampled ACS
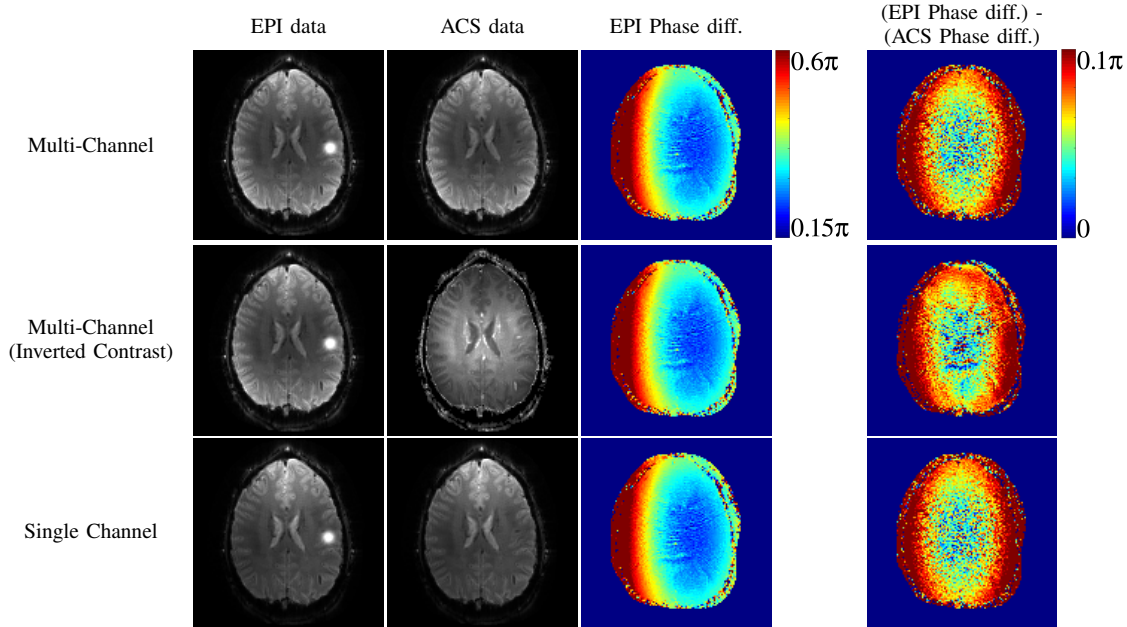
Figure 3.3: Illustration of the EPI and ACS datasets used in simulation. The first and second top rows show coil-combined multi-channel data for the case when the EPI and ACS data have similar and inverted contrast, respectively, while the bottom row shows representative single-channel images. We also show the interpolarity phase difference for the coil-combined EPI data, as well as the difference in the interpolarity phase difference between the coil-combined EPI and ACS data.

data vectors $\mathbf{k}_{\mathrm{acs}}^{+}$ and $\mathbf{k}_{\mathrm{acs}}^{-}$ as additional variables to be optimized in Eq. (3.4), subject to ACS data consistency constraints.

For RAC-LORAKS and AC-LORAKS, the regularization parameters λ and η were selected manually based on subjective visual inspection reconstruction quality and ghost-reduction performance for in vivo data, and to minimize quantitative error measures for simulated data. The rank-related parameters p and r were selected based on the singular value characteristics of LO-RAKS matrices formed from the ACS data. The rank parameters were set based on the points at which the singular value curves begin to flatten out, which is a common rank estimation technique for noisy matrices. This decision was made manually (based on visual inspection) for the results shown later in the paper, although fully automatic approaches would also be viable.

DPG is a ghost correction method that treats different gradient polarities like different coils in a parallel imaging experiment, and uses a dual GRAPPA kernel estimated from ACS data for image reconstruction [28]. In order to use DPG for the initialization of RAC-LORAKS, we have adapted DPG to output two sets of images (with calibration based on the raw uncorrected multi-channel

ACS data), one for the original $\mathbf{k}^+$ data and one for the original $\mathbf{k}^-$ data. This is different than the original DPG implementation, which applies ACS pre-processing to try and correct for errors in the ACS data, and then directly fuses information from the two polarities together into a single virtual "hybrid" output [28]. This hybrid output can have different magnitude and phase characteristics than the original $\mathbf{k}^+$ and $\mathbf{k}^-$ data, so is not useful as an initialization for RAC-LORAKS. We refer to our adapted version as modified DPG (mDPG) from now on. In some cases, we also compare against the original version of DPG (including the original ACS pre-processing procedure to correct for shot-to-shot variations in the ACS data [28]), although note that such comparisons are necessarily qualitative, since the magnitude and phase characteristics of the hybrid output DPG images do not match the images generated using other methods.

MUSSELS is a structured low-rank matrix recovery method that uses SENSE-type parallel imaging constraints together with nuclear norm regularization to impose low-rank constraints [40]. While MUSSELS was originally developed for multi-shot EPI ghost correction, it can apply equally well to the ghost correction problem associated with different gradient polarities. Sensitivity maps for MUSSELS were estimated by applying ESPIRiT [8] to the same ACS data used for the other methods. The regularization parameter for MUSSELS was selected manually based on subjective visual inspection of reconstruction quality and ghost-reduction performance in the case of in vivo data, or to minimize quantitative error measures for simulated data.

Note that DPG and MUSSELS were both developed for the multi-channel setting. We can adapt DPG to the single-channel setting in straightforward ways [14], and we apply this adaptation to the single-channel simulated data. We did not adapt MUSSELS to the single-channel case. Note that the SENSE-based constraints used by MUSSELS would reduce to a simple spatial-domain support constraint in the single-channel case, which is not strong enough to yield good performance results.

For all methods, results were visualized by using a standard square-root sum-of-squares technique to combine the images from different coils and different gradient polarities into a single image. Results from in vivo experiments were evaluated qualitatively, since a gold standard reference was not available in these cases.

Simulation results were evaluated quantitatively using the normalized root mean-squared error (NRMSE):

$$\text{NRMSE} \triangleq \frac{\sqrt{\left\| \hat{\mathbf{k}}^+ - \mathbf{k}_{\text{gold}}^+ \right\|_2^2 + \left\| \hat{\mathbf{k}}^- - \mathbf{k}_{\text{gold}}^- \right\|_2^2}}{\sqrt{\left\| \mathbf{k}_{\text{gold}}^+ \right\|_2^2 + \left\| \mathbf{k}_{\text{gold}}^- \right\|_2^2}}, \tag{3.5}$$

where $\mathbf{k}_{\text{gold}}^+$ and $\mathbf{k}_{\text{gold}}^-$ are respectively the gold standard values for the positive and negative gradient polarities. We also plotted Fourier Error Spectrum Plots (ESPs) to gain further insight into how the errors were distributed across different spatial resolutions scales [80]. An ESP is designed to reveal the spectral characteristics of the error, and for example, can discriminate between methods that make more errors in the low-resolution features of an image versus methods that make more errors in high-resolution features.

3.4   Results

Figure 3.4 shows ACS data and reconstruction results from the in vivo gradient-echo EPI brain data with an axial slice orientation. The ACS data in this case does not have strong artifacts, although close inspection does reveal that ACS ghost artifacts are present. This can be further appreciated in Fig. 3.5 where the same images are shown with amplified image intensity to highlight ghost characteristics in the image background. As can be seen, all ghost correction methods work well at smaller acceleration factors, although performance begins to degrade at larger acceleration factors. We observe that, compared to other methods, the visual quality of the MUSSELS reconstruction seems to degrade most rapidly as a function of acceleration factor, which is consistent with previous observations [14]. The mDPG method had qualitatively better performance than MUSSELS in this case. However, a close inspection of the images reveals that the mDPG results are not entirely ghost-free even for the unaccelerated ($R = 1$) case. This may be expected due to the artifacts and mismatches that are present in the ACS data. Although mDPG does not attempt to correct the ACS artifacts, it should be noted that the original DPG method does try to correct them through pre-processing. Results showing the qualitative performance of the original DPG method are shown in Fig. 3.6, where we observe that the ghost artifacts still exist, though as expected, are less prominent

than were observed for mDPG. In spite of the ACS artifacts, the AC-LORAKS reconstruction still has good performance at low acceleration factors and does a good job of suppressing ghosts in the background regions of the image at all acceleration factors, although exhibits substantial degradation in image quality at the highest acceleration factors (with artifacts similar to those observed for highly-accelerated parallel imaging reconstructions). However, the RAC-LORAKS reconstruction appears to have much higher quality than the other methods, even at very high acceleration factors like $R = 6$. (Note that when $R = 6$, the effective acceleration factor is $R = 12$ when each readout gradient polarity is considered separately. This leads to a highly ill-posed inverse problem).

Figure 3.7 shows results from the in vivo gradient-echo EPI brain data with a double-oblique slice orientation. This case is more challenging than the previous one due to the complicated nonlinear 2D spatial phase differences we observed between data acquired with positive and negative polarities (as visualized in the last column of Fig. 3.7), the proximity to air-tissue interfaces that result in substantial magnetic field inhomogeneity effects, as well as more substantial ghosting artifacts present in the ACS data. Note that the ACS data corresponding to the $R = 5$ case is particularly corrupted, which can be attributed to the unpredictable shot-to-shot variations that frequently occur in these kinds of multi-shot acquisitions. Despite the more extreme scenario, the different ghost reconstruction methods have similar characteristics to those observed in the previous case, with RAC-LORAKS appearing to demonstrate the cleanest overall results.

Figure 3.8 shows reconstruction results from the first set of multi-channel simulations (with similar contrast between ACS and EPI data, but with a hyperintensity added to the EPI data). Quantitative NRMSE values are reported in Table 3.1 with corresponding ESPs shown in Fig. 3.10. Qualitatively, the results from Fig. 3.8 have similar characteristics to the results observed with in vivo data. Notably, RAC-LORAKS is able to consistently reconstruct a high-quality image that bears close resemblance to the gold standard image, while methods like mDPG and AC-LORAKS have artifacts due to the small mismatches between the ACS and EPI data. The visual assessment of reconstruction quality matches well with the quantitative NRMSE assessment shown in Table 3.1.
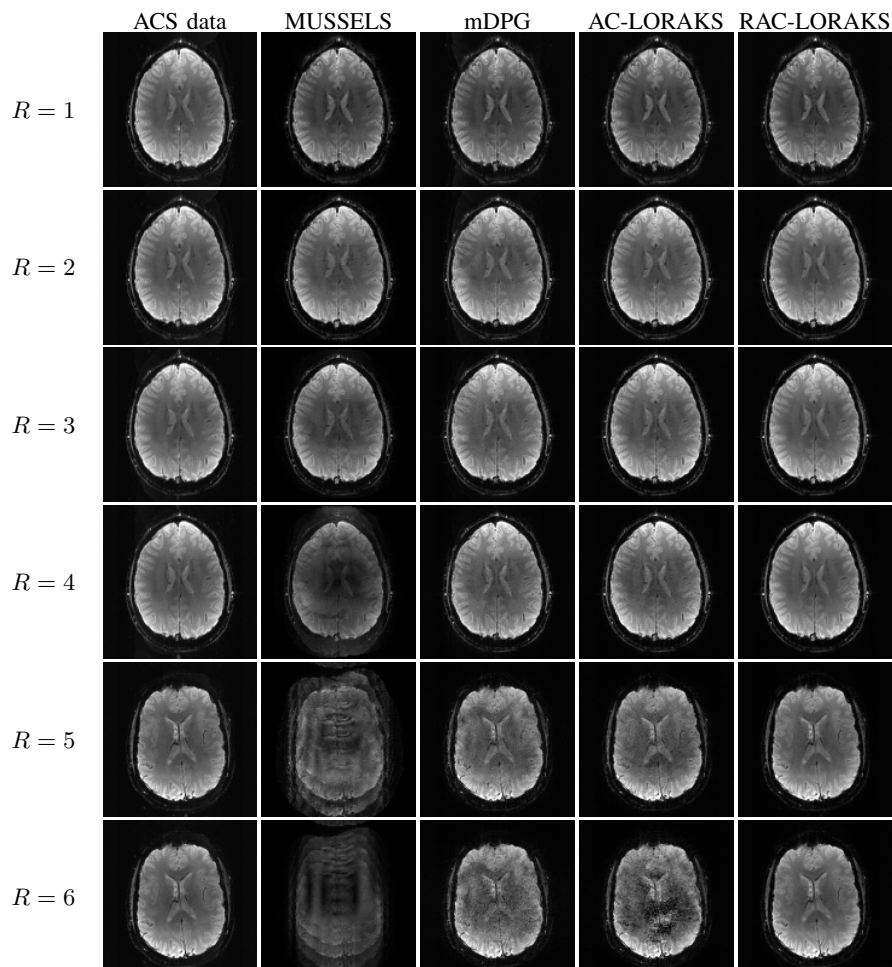
Figure 3.4: ACS data and reconstruction results for in vivo gradient-echo EPI brain data with an axial slice orientation for different parallel imaging acceleration factors. Note that the first four acceleration factors ($R = 1$-4) were acquired from one subject during a single scan session while the last two acceleration factors ($R = 5, 6$) were acquired from a different subject on a different day, which explains the visual discontinuity between these cases.

AC-LORAKS and RAC-LORAKS have a similar performance at $R = 1$ and 2, with RAC-LORAKS having the best performance at high acceleration factors.

Reconstructions were also performed using the original DPG formulation as shown in Fig. 3.9. In this case, DPG has similar ghost artifacts to mDPG, which is expected because there are no artifacts in the ACS data, while there is a problematic mismatch between the ACS data and the EPI data that neither DPG nor mDPG address. Notably, for both DPG and mDPG, we observe aliasing artifacts that seem to be associated with the hyperintensity that was present in the EPI data but was not in the ACS data.
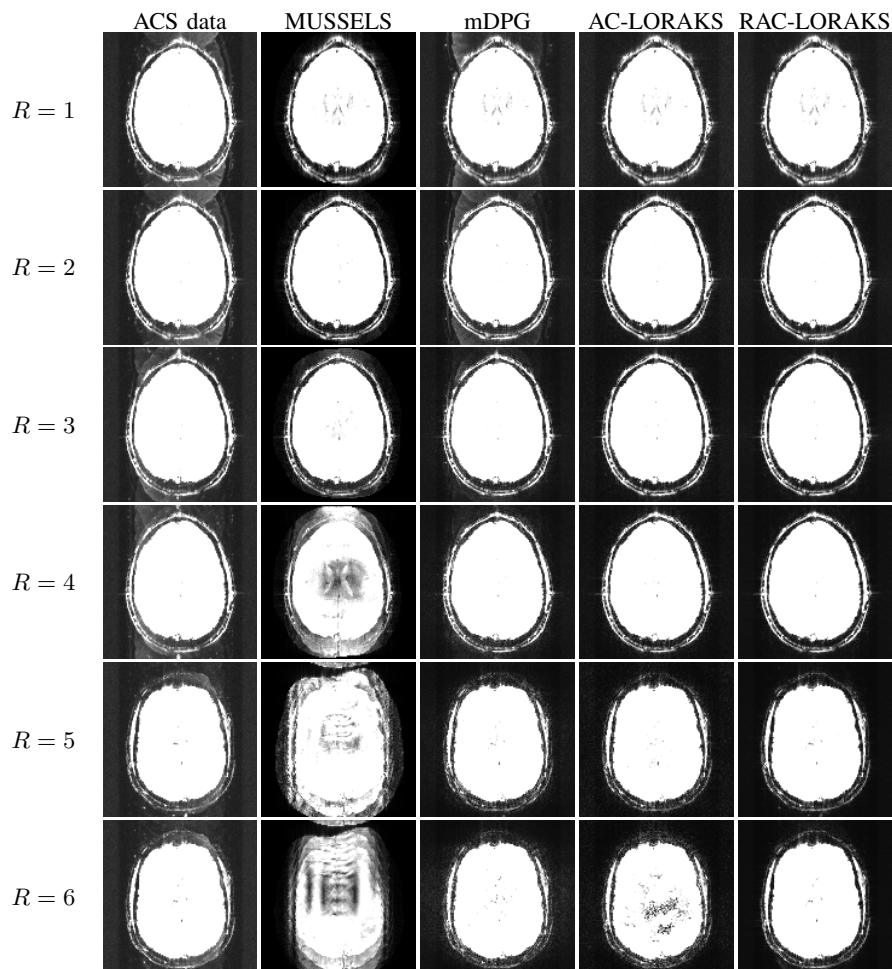
Figure 3.5: The same results shown in Fig. 3.4 , but with a $5\times$ intensity amplification to highlight the ghost characteristics.

The ESP plots in Fig. 3.10 enable a more nuanced analysis. These results suggest that RAC-LORAKS has good (i.e., among the best, even if it is not always the best) performance at all spatial frequencies, meaning that it is good at reconstructing image features across the whole range of resolution scales.

Figure 3.11 shows a similar simulation result to that shown in Fig. 3.8, with the main difference being that the previous EPI images (with the hyperintensity) were used as ACS data and the previous ACS images (without the hyperintensity) were used to generate EPI data. Consistent with the previous case, we observe good performance for RAC-LORAKS, and do not observe the
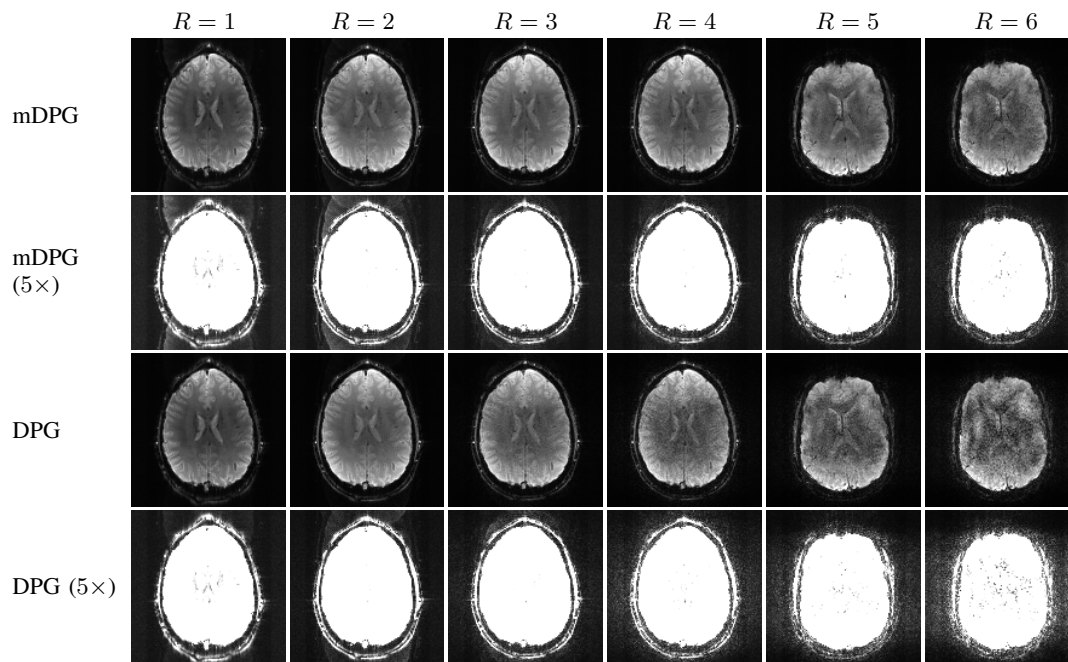
Figure 3.6: DPG results corresponding to the same data shown in Fig. 3.4 and Fig. 3.5. The same mDPG results shown in Fig. 3.4 and Fig. 3.5 are also reproduced in this figure for reference. Note that the processing steps of DPG cause the image intensities to be mismatched from the intensities of mDPG and the other reconstruction methods, which precludes a quantitative comparison.

features of the hyperintensity being erroneously transferred into the RAC-LORAKS reconstruction results.

Figure. 3.12 and Table 3.2 show simulation results for the case where the ACS data has an even more substantial contrast difference (i.e. inverted contrast) with respect to the EPI data. For this case we observe a degradation in performance for all methods compared to the previous cases, although RAC-LORAKS still showed the best overall qualitative and quantitative performance. This result suggests that RAC-LORAKS may have better performance when the contrast is similar between the ACS and EPI data, although can still provide benefits when the contrast difference is substantial.

Figure 3.13 shows reconstruction results from the single-channel simulation, with quantitative NRMSE values reported in Table 3.3. While previous work [14] reported that mDPG and AC-LORAKS can be reasonably successful for single-channel data with $R = 1$ when the ACS data is pristine, our new results demonstrate that this performance can be sensitive to the quality of the ACS data. In particular, we observe strong ghost artifacts for both of these methods, even though
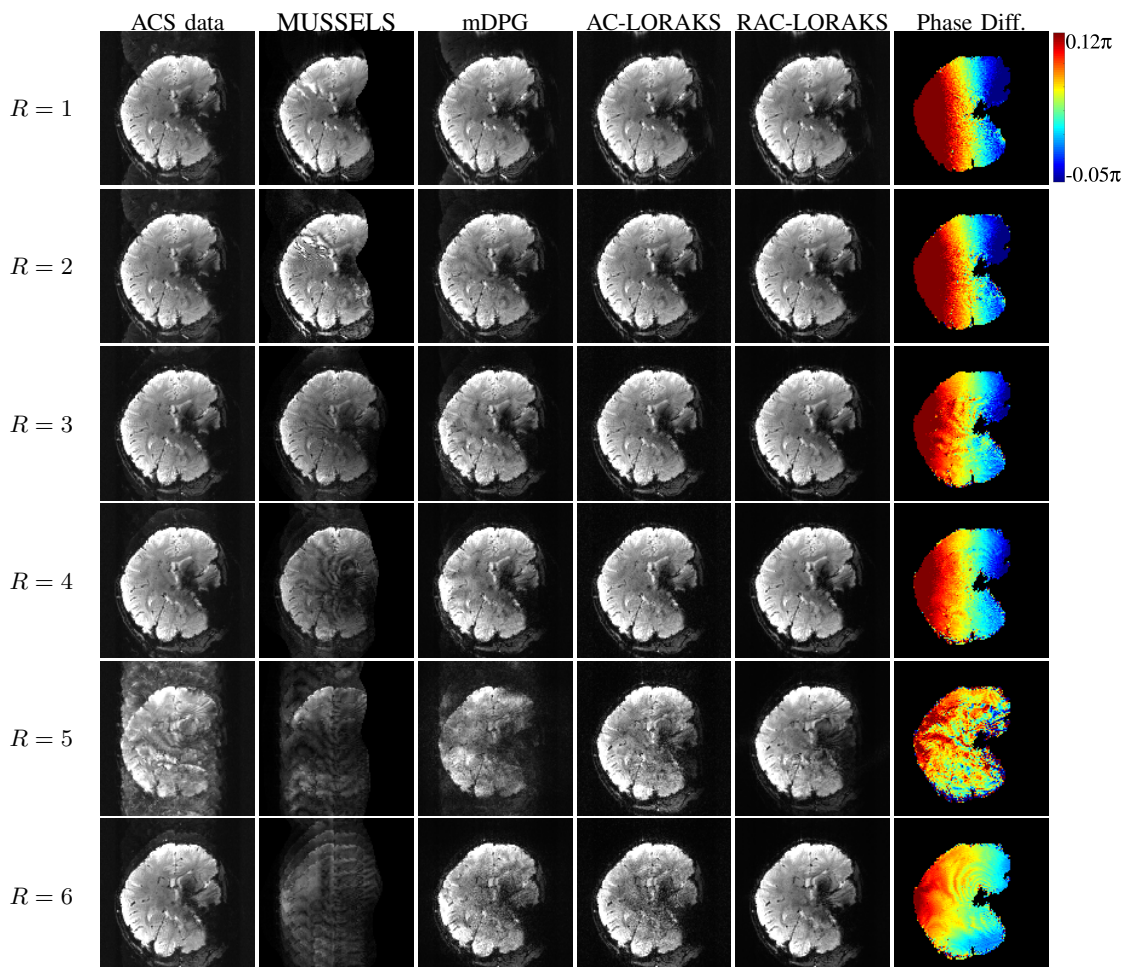
Figure 3.7: ACS data and reconstruction results for in vivo gradient-echo EPI brain data with a double-oblique slice orientation for different parallel imaging acceleration factors. For reference, we also show the interpolarity phase difference as estimated from a coil-combined RAC-LORAKS result. The degree of phase nonlinearity is an indicator of how difficult ghost correction is expected to be. As can be seen, complicated 2D nonlinear phase differences are present in many of these cases.

we do observe that the AC-LORAKS reconstruction has successfully suppressed ghost artifacts in the image background (outside of the support of the true image). In contrast, RAC-LORAKS is substantially more successful for $R = 1$. Notably, RAC-LORAKS also performed well for the even more challenging $R = 2$ case, unlike the other methods. For reference, note that even with high-quality ACS data, the previous AC-LORAKS method did not yield good results with similar single-channel $R = 2$ data [14].

Figure 3.14 shows reconstruction results from the first set of in vivo diffusion EPI brain data, including a $10\times$ intensity amplification to highlight the ghost characteristics. As can be seen,

Table 3.1: NRMSEs for the multi-channel simulation results shown in Fig. 3.8. For each acceleration factor, the smallest values are highlighted in bold.

	MUSSELS	mDPG	AC-LORAKS	RAC-LORAKS
$R = 1$	0.059	0.024	**0.016**	0.020
$R = 2$	0.104	0.045	**0.035**	0.042
$R = 3$	0.271	0.083	0.056	**0.055**
$R = 4$	0.572	0.127	0.132	**0.064**
$R = 5$	0.741	0.161	0.269	**0.085**

Table 3.2: NRMSEs for the multi-channel inverted contrast simulation results shown in Fig. 3.12 . For each acceleration factor, the smallest values are highlighted in bold.

	MUSSELS	mDPG	AC-LORAKS	RAC-LORAKS
$R = 1$	0.098	0.177	0.027	**0.025**
$R = 2$	0.131	0.182	0.065	**0.045**
$R = 3$	0.280	0.190	0.141	**0.077**
$R = 4$	0.532	0.213	0.236	**0.117**
$R = 5$	0.754	0.309	0.373	**0.160**

the ACS data has ghost artifacts in all cases, and both MUSSELS and mDPG reconstructions also exhibit unsuppressed ghosting artifacts. On the other hand, both AC-LORAKS and RAC-LORAKS are relatively ghost-free in this example and have only minor differences from one another (it might be argued that the RAC-LORAKS result has a slightly less-noisy appearance than the AC-LORAKS result, but if so, this difference is very subtle). While this result does not demonstrate an obvious advantage for RAC-LORAKS over AC-LORAKS, it should be observed that this diffusion result is at least consistent with the previous gradient-echo EPI data results, in which we also did not observe a substantial difference between RAC-LORAKS and AC-LORAKS when $R = 3$. In addition, this case involves a very substantial contrast difference between the ACS data and the EPI data. This difference does not appear to have adversely affected the performance characteristics of these methods in substantial ways.

Figure 3.15 shows reconstruction results from the second set of in vivo diffusion EPI brain acquisitions (with different acceleration factors), with zoom-ins shown in Fig. 3.16 for improved visibility. Consistent with the results shown for the gradient-echo EPI data in Fig. 3.4, we observe that all methods perform well for low acceleration factors. As the acceleration factor increases, the performance of each method degrades, with RAC-LORAKS showing a lower qualitative degradation

MUSSELS   mDPG   AC-LORAKS RAC-LORAKS
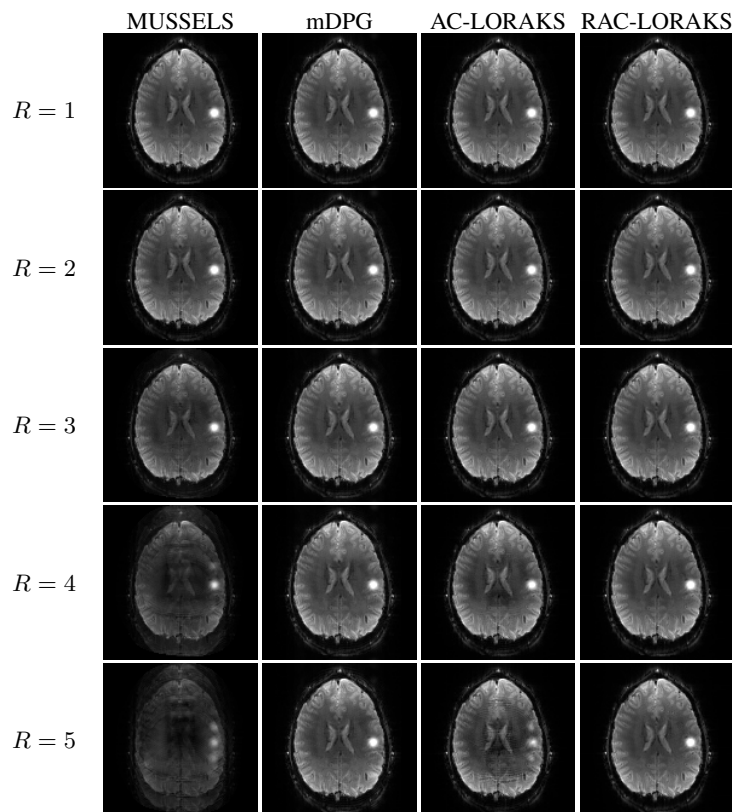
Figure 3.8: Reconstruction results for the first set of multi-channel simulations (with similar contrast between ACS and EPI data, but with a hyperintensity added to the EPI data) with different parallel imaging acceleration factors.

in comparison to the other methods at the very high acceleration factors $R = 4, 5$. Note that at high acceleration factors (e.g., $R = 4, 5$) the reconstruction quality for RAC-LORAKS is not quite as good as for the gradient-echo EPI dataset shown in Fig. 3.4. We believe that this should be expected, since as mentioned before, diffusion EPI data can be considered more challenging than the gradient-echo EPI data due to SNR issues, eddy current effects, motion-induced phase effects, and contrast mismatches between the ACS and EPI datasets.

Finally, Fig. 3.17 shows results from the in vivo cardiac EPI data. While this data was not accelerated ($R = 1$), this case is challenging because of the double-oblique slice orientation as well as the substantial artifacts present in the ACS data resulting from cardiac motion-induced shot-to-shot variations. In addition, this case can also be challenging for SENSE-based methods (like MUSSELS), due to the use of a small FOV with aliasing. When aliasing is present within the FOV, it violates the standard SENSE modeling assumption of one sensitivity map value per spatial
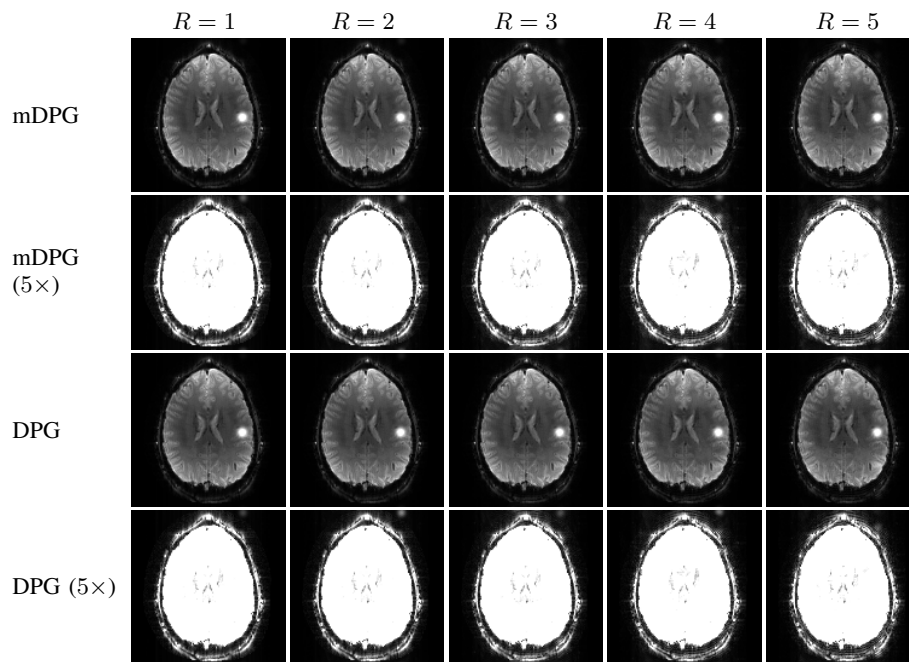
Figure 3.9: mDPG and DPG results corresponding to the same multi-channel simulated data from Fig. 3.8.

Table 3.3: NRMSEs for the single-channel simulation results shown in Fig. 3.13. For each acceleration factor, the smallest values are highlighted in bold.

	mDPG	AC-LORAKS	RAC-LORAKS
$R = 1$	0.170	0.584	**0.046**
$R = 2$	0.494	0.809	**0.073**

location, which generally leads to artifacts if not properly accounted for. The results demonstrate that both MUSSELS and mDPG have substantial residual ghosting artifacts, which might not be surprising given the high degree of corruption that is present in the ACS data. On the other hand, both AC-LORAKS and RAC-LORAKS are more successful at suppressing the ghosts. Without a gold standard reference, it is hard to establish definitively whether AC-LORAKS or RAC-LORAKS is better in this example, although we believe that the RAC-LORAKS result demonstrates slightly less ghosting than AC-LORAKS, particularly on the left side of the image where the ACS data and mDPG both have particularly strong ghost artifacts.

3.5   Discussion

The results in the previous section demonstrated that, in the presence of imperfect ACS data, RAC-LORAKS frequently offers similar or better performance to the previous AC-LORAKS ghost correction method that it generalizes, while both of these methods perform substantially better than methods like MUSSELS or DPG. We also observed that RAC-LORAKS appears to have the biggest advantage over AC-LORAKS in scenarios where the parallel imaging acceleration factor was high. For these cases, we observed that RAC-LORAKS was able to mitigate ghost artifacts both inside and outside the support of the original image, while AC-LORAKS was only able to mitigate ghost artifacts outside the support but not inside. This advantage for RAC-LORAKS is likely the result of its improved robustness to ACS errors combined with the multi-contrast linear predictability constraints which help to make the reconstruction problem less ill-posed. However, it should be noted that RAC-LORAKS has one more regularization parameter than AC-LORAKS (i.e., η, which controls the level of trust placed in the information from the ACS data). In our experience, manual tuning of this parameter is not hard (i.e., we always started from the small value $\eta = 10^{-3}$, and frequently did not have to modify this value to achieve satisfying results). The method would be easier to use if the selection of η were automated.

Both RAC-LORAKS and AC-LORAKS also depend on the choice of rank parameters, and as described previously, the results shown in this work made a heuristic choice based on the empirical rank characteristics of the ACS data. Even though the low-rank characteristics of the structured matrices might vary between the ACS data and the acquired EPI data due to systematic phenomena (e.g., thermal noise, subject motion, respiration, artifacts in the ACS data, etc.), we have not observed major problems associated with inappropriate rank selection in our empirical results. This might be expected, based on the observation that LORAKS reconstruction results are frequently not very sensitive to small variations in the rank parameter [12, 32]. Nevertheless, the development of improved automatic RAC-LORAKS parameter selection methods would be an interesting topic for future work.

Although RAC-LORAKS offers good performance, it should be noted that our current implementation of RAC-LORAKS can be more computationally expensive than existing methods. For

example, for the results shown with $R = 1$ in Fig. 1, RAC-LORAKS used ≈ 45 min of reconstruction time, while MUSSELS, mDPG, and AC-LORAKS respectively used ≈ 15 min, ≈ 2 min, and ≈ 100 min. All methods were implemented in MATLAB on a standard desktop computer with an Intel Xeon E5-1603 2.8 GHz quad core CPU processor and 32GB of RAM. While this relatively long computation time may be a concern, it should be noted that we are reporting the results of a simple proof-of-principle implementation, and we did not spend much time to optimize the computational efficiency of this approach. We believe that major improvements may be possible by leveraging better computational hardware, smarter algorithms, and more efficient implementations. Given the reconstruction performance offered by RAC-LORAKS, we believe that improving its computational performance is a promising topic for future research. However, RAC-LORAKS is notably faster than AC-LORAKS, and it appears that this speed difference results from the fact that RAC-LORAKS has consistently faster convergence than AC-LORAKS in this setting. The reason for this faster convergence is unclear at this stage, although we believe that a detailed analysis of convergence characteristics is beyond the scope of the present paper.

While this paper focused on EPI ghost correction for standard single-slice excitation, we believe that the extension of these ideas to simultaneous multi-slice EPI acquisitions (similar to Refs. [60, 61, 81, 82]) is a very promising research direction.

Finally, although the techniques we developed in this work were described and evaluated in the context of EPI ghost correction, we believe that the overall approach is likely to be useful across a wide range of parallel imaging applications, particularly those for which the measured ACS data is not adequate to resolve all of the reconstruction artifacts. Specifically, we believe that the key principles employed by RAC-LORAKS (i.e., using structured low-rank matrix methods to avoid placing complete trust in the accuracy of ACS data, and leveraging ACS data to provide additional information in a multi-contrast framework) are both novel ideas that are applicable to arbitrary image reconstructions involving ACS data, and are not exclusive to ghost correction settings. In addition, we are encouraged by the high-quality reconstruction results that RAC-LORAKS produces even in very highly-accelerated scans. These results suggest to us that there

may be value in exploring the usefulness of RAC-LORAKS to other parallel imaging experiments in future work.

3.6   Conclusions

This work proposed and evaluated RAC-LORAKS, a new structured low-rank matrix method for EPI ghost correction that integrates multiple constraints (including parallel imaging constraints, support constraints, phase constraints, and inter-image linear predictability constraints) to not only mitigate artifacts resulting from imperfect ACS data and Nyquist ghosts, but also accounting for partial Fourier acquisition and reducing parallel imaging artifacts and noise in an integrated fashion. RAC-LORAKS uses ACS data and k-space domain linear predictive modeling to stabilize the solution of the ill-posed inverse problem, and was observed to offer advantages relative to state-of-the-art ghost correction methods like AC-LORAKS, DPG, and MUSSELS.
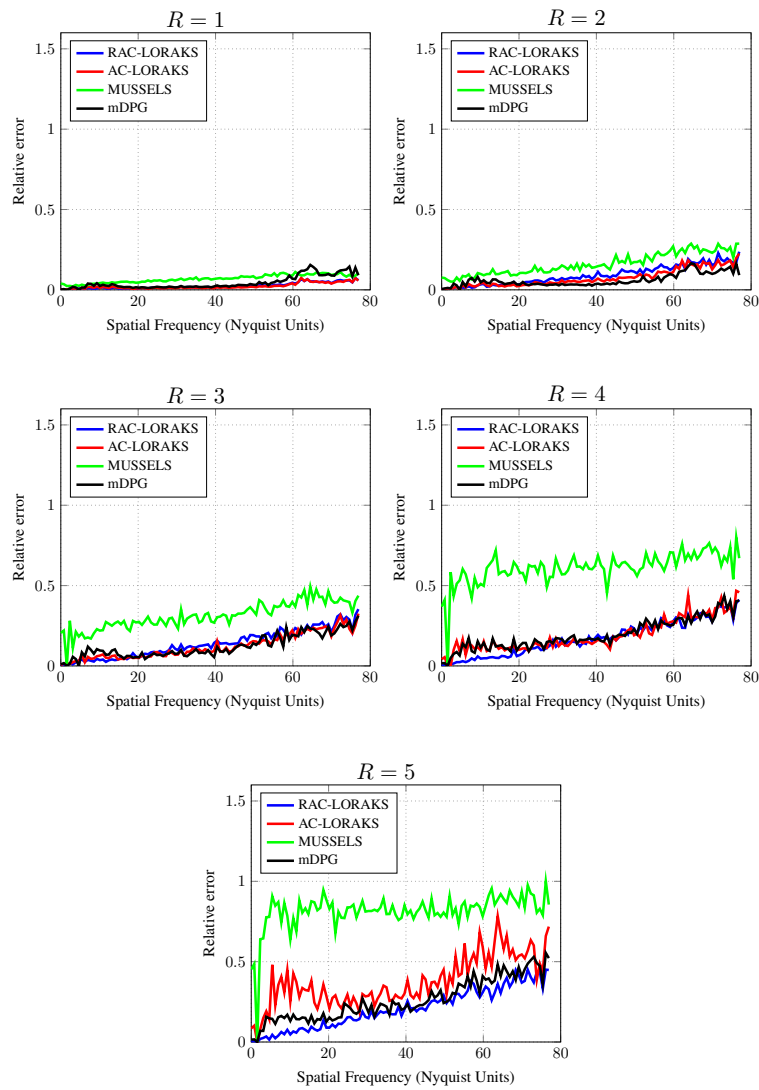
Figure 3.10: ESPs for the multi-channel simulation results shown in Fig. 3.8. The vertical axis of each ESP uses a consistent range to enable comparisons between different acceleration factors.
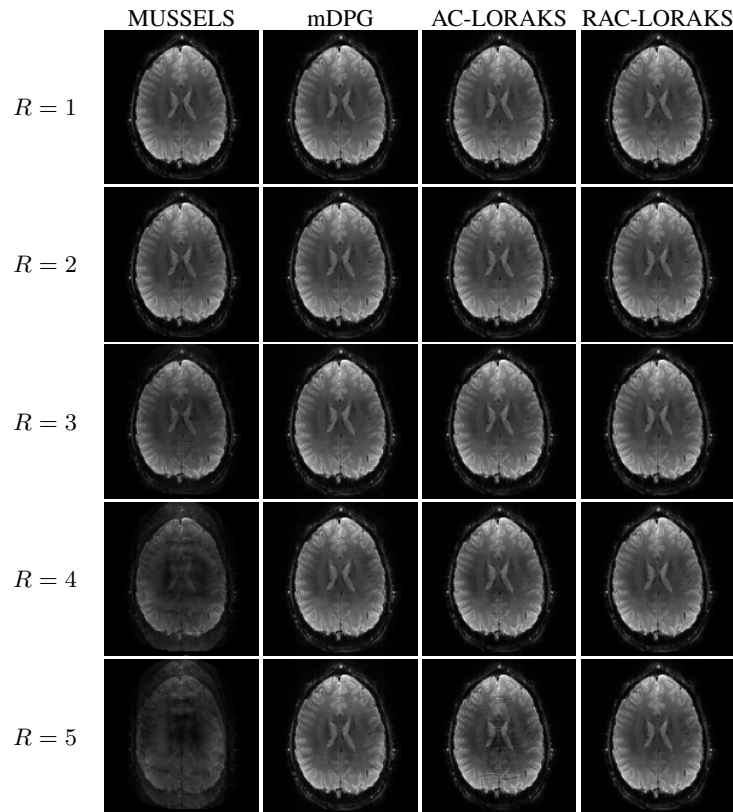
Figure 3.11: Reconstruction results for multi-channel simulated data with different parallel imaging acceleration factors. These simulations are identical to those reported in Fig. 3.8, except that the images used to generate EPI data and the images used to generate ACS data were interchanged.
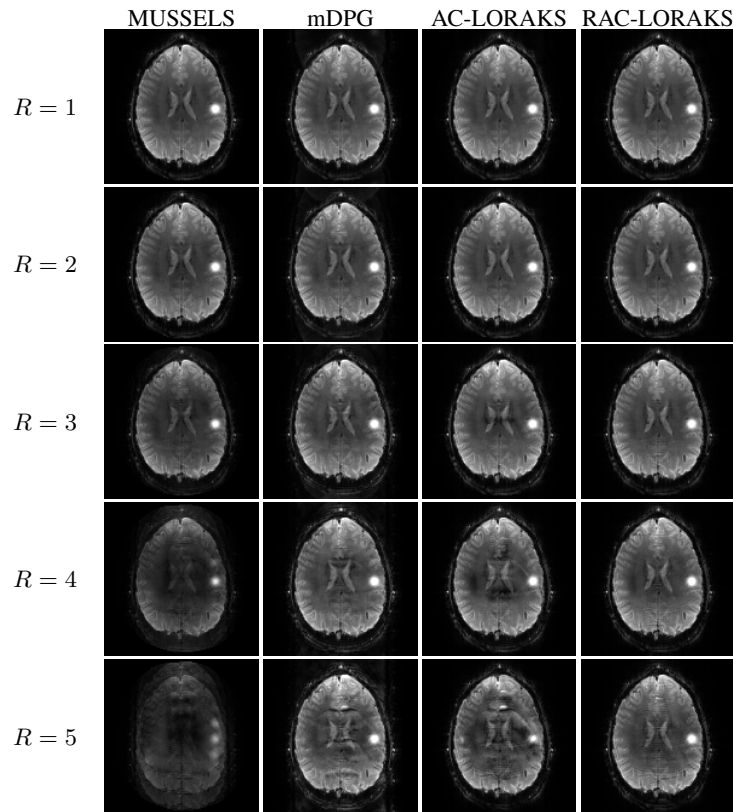
Figure 3.12: Reconstruction results for the second set of multi-channel simulations (with inverted contrast between ACS and EPI data) with different parallel imaging acceleration factors.
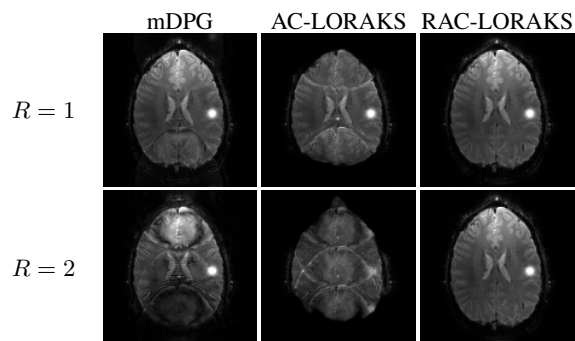
Figure 3.13: Reconstruction results for single-channel simulated data with different acceleration factors.
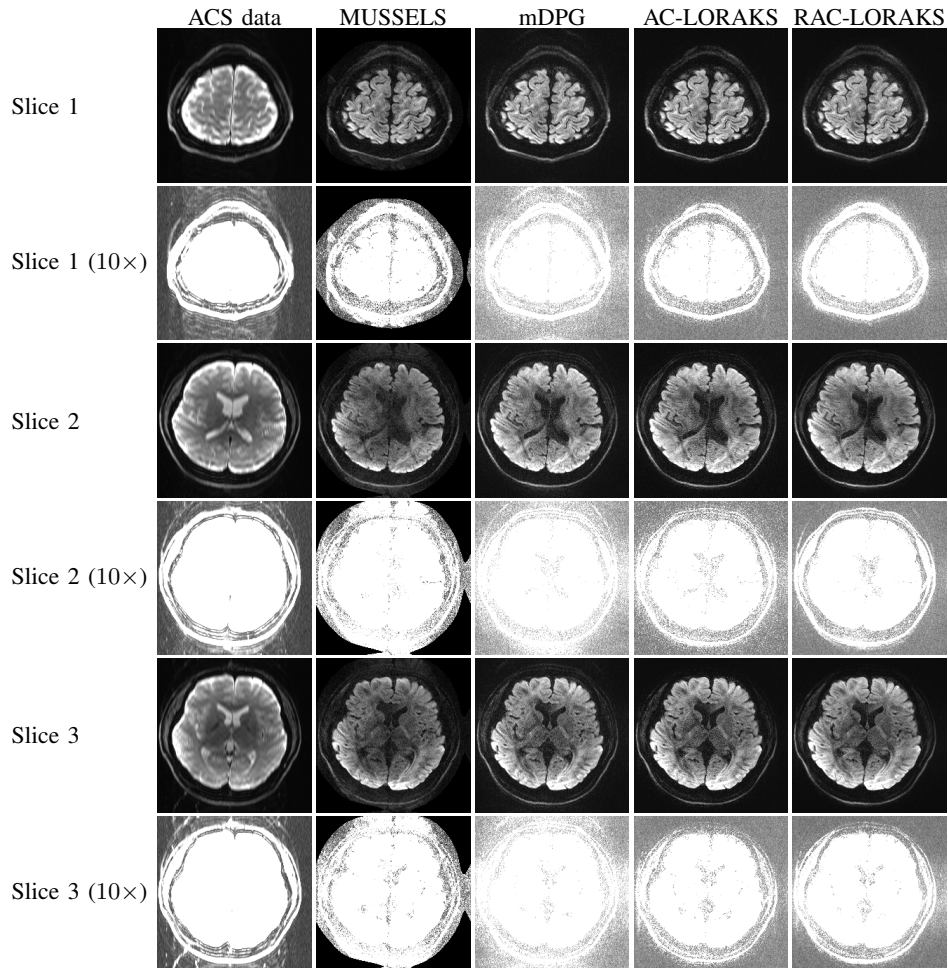
Figure 3.14: ACS data and reconstruction results for three representative slices from in vivo diffusion brain data ($R = 3$). A $10\times$ intensity amplification is also shown for each slice to better highlight the ghosting characteristics.
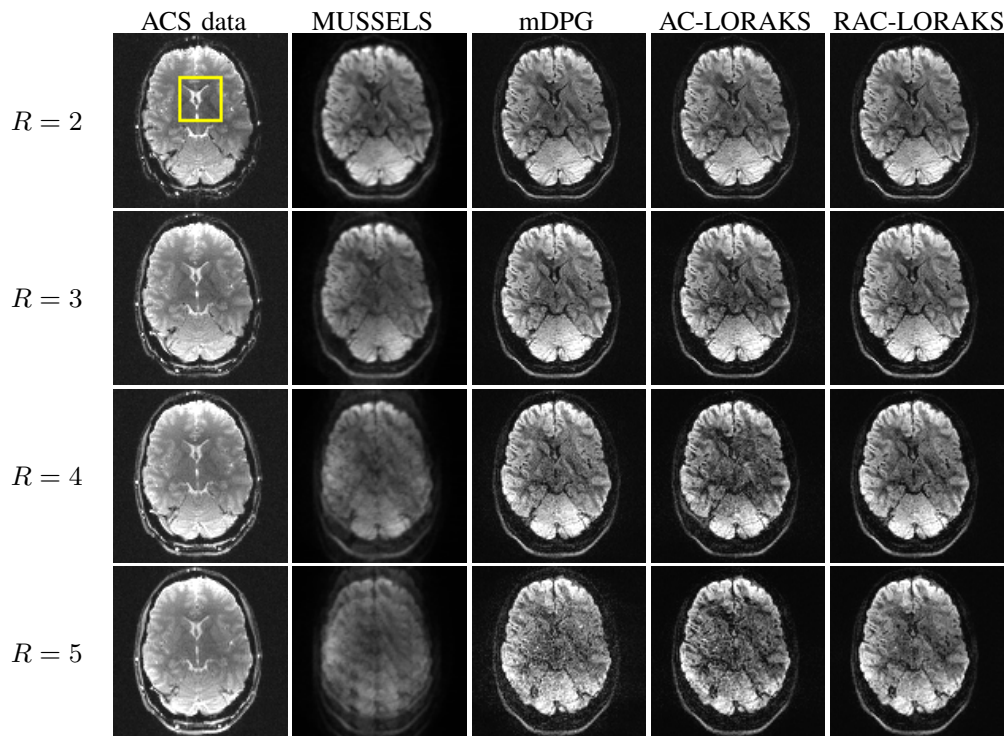
Figure 3.15: ACS data and reconstruction results for in vivo diffusion EPI brain data for different parallel imaging acceleration factors. For improved visualization, zoomed-in versions of these results (corresponding to the spatial region marked with a yellow rectangle in the first column and first row) are shown in Fig. 3.16. It should be noted that the subject appears to have slightly moved between scans, so that there is not perfect correspondence between anatomical image features across different acceleration factors.
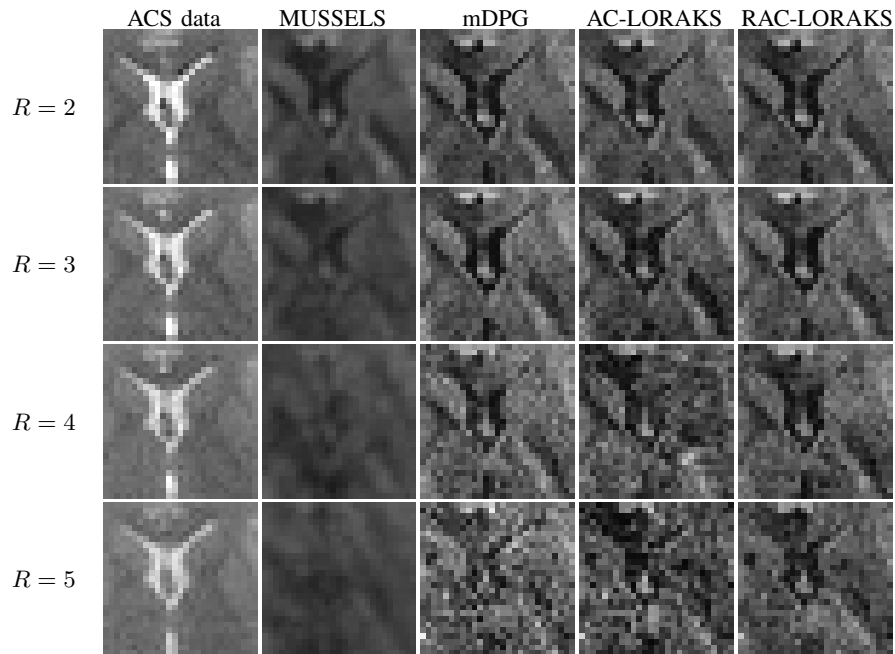
Figure 3.16: The same results shown in Fig. 3.15, but zoomed-in to a region of interest for improved visualization.
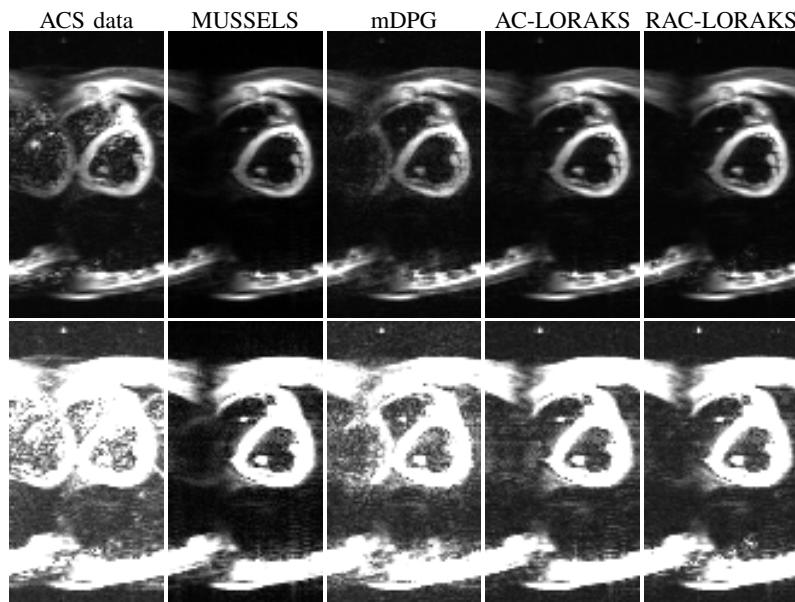
Figure 3.17: ACS data and reconstruction results for unaccelerated in vivo cardiac EPI data. The two rows show the same results, but the second row has $5\times$ intensity amplification to better highlight the ghosting characteristics.

Chapter 4

On the Shape of Convolution Kernels in MRI Reconstruction: Rectangles versus Ellipsoids

4.1 Introduction

Constrained MRI reconstruction methods that rely on shift-invariant convolution models have existed for decades. One of the earliest approaches (which still remains quite popular) is to assume that a missing sample of k-space data can be predicted as a linear shift-invariant combination of neighboring samples [3, 13]. In particular, it is common to assume that

$$\hat{\rho}[\mathbf{k}] = \sum_{\mathbf{z} \in \Omega} w[\mathbf{z}] \rho[\mathbf{k} - \mathbf{z}] \tag{4.1}$$

for an appropriately-chosen shift-invariant interpolation kernel $w[\mathbf{k}]$, where $\hat{\rho}[\mathbf{k}]$ is the missing sample to be interpolated, the interpolation kernel $w[\mathbf{k}]$ has a shape dictated by the support set Ω, and the neighboring samples correspond to $\rho[\mathbf{k} - \mathbf{z}]$ for $\mathbf{z} \in \Omega$. In many cases, the interpolation kernel $w[\mathbf{z}]$ is a quantity that must be learned from some kind of training data. Methods that rely on these kinds of concepts include popular approaches like SMASH [4], GRAPPA [5], SPIRiT [7], and structured low-rank matrix reconstruction methods [31, 32, 36, 37], as well as related subspace methods like ESPIRiT [8]. Another group of methods adopts the same basic structure, but moves from shift-invariant linear prediction to shift-invariant nonlinear prediction relationships to improve

performance [83–86]. And in recent years, due to the growing excitement about machine learning and deep learning in MRI, a huge number of different convolutional neural network (CNN) methods have been proposed [87, 88].

Each of these methods can be implemented using shift-invariant filtering with appropriate convolution kernels. Most of the existing methods use a rectangular kernel shape, in which the support set Ω of the kernel has a rectangular geometry. For example, the support set of an 11×11 rectangular kernel in 2D could be defined as $\Omega = \{\mathbf{k} = (n_x \Delta k_x, n_y \Delta k_y) : \max(|n_x|, |n_y|) \leq 5\}$, where Δk_x and Δk_y are the Nyquist sampling intervals along each dimension. In contrast, an ellipsoidal kernel shape would have a support set Ω with an ellipsoidal geometry, and could be defined for an 11×11 kernel in 2D as $\Omega = \{\mathbf{k} = (n_x \Delta k_x, n_y \Delta k_y) : \sqrt{|n_x|^2 + |n_y|^2} \leq 5\}$. Although we [12, 32, 86] and other researchers [89] have sometimes used ellipsoidally-shaped Ω, we are not aware of any previous systematic comparisons of rectangular and ellipsoidal kernel shapes.

In this work, we perform a detailed evaluation of rectangular versus ellipsoidal kernel shapes, and the results we obtained (to be described later) suggest that ellipsoidal kernels can indeed offer several advantages over more-common rectangular kernels. A preliminary account of portions of this work was previously presented at a recent conference [90].

4.2   Theory

Many of the earliest convolutional reconstruction methods were designed for 1D reconstruction problems, in which case there is no distinction between a rectangular kernel and an ellipsoidal kernel. For more recent higher-dimensional methods, it is easy to access rectangle-shaped subcomponents of a higher-dimensional array in most programming languages (e.g., in Python, for a 2D array denoted by `B`, the code `B[1:3,1:3]` will yield access to a 2D rectangular subcomponent), while accessing an ellipsoid-shaped subcomponent of an array is generally more involved to code. This is likely a major contributing factor to the widespread modern use of rectangular kernels.

However, there are some theoretical reasons why an ellipsoidal kernel might be preferred over a rectangular kernel. One reason is that an ellipsoidal kernel can be viewed as a rectangular kernel with the corners removed, which implies that for ellipsoidal and rectangular kernels with "matched

size" (i.e., the principal axes of the ellipse have the same lengths as the sides of the corresponding

rectangle), an ellipsoidal kernel has smaller area/volume and fewer coefficients. In particular, in

2D, the area of an ellipse is more than 20% smaller than the area of the rectangle that inscribes

it, while in 3D, the volume of an ellipsoid is nearly 50% smaller than the volume of the inscribing

hyperrectangle [91]. This difference in the number of coefficients is illustrated for the 2D case in

Fig. 4.1. Practically, this means that ellipsoidal kernels are associated with fewer degrees-of-freedom

(i.e., fewer coefficients $w[\mathbf{z}]$ that need to be learned in the linear case) than rectangular kernels of

the same basic size.
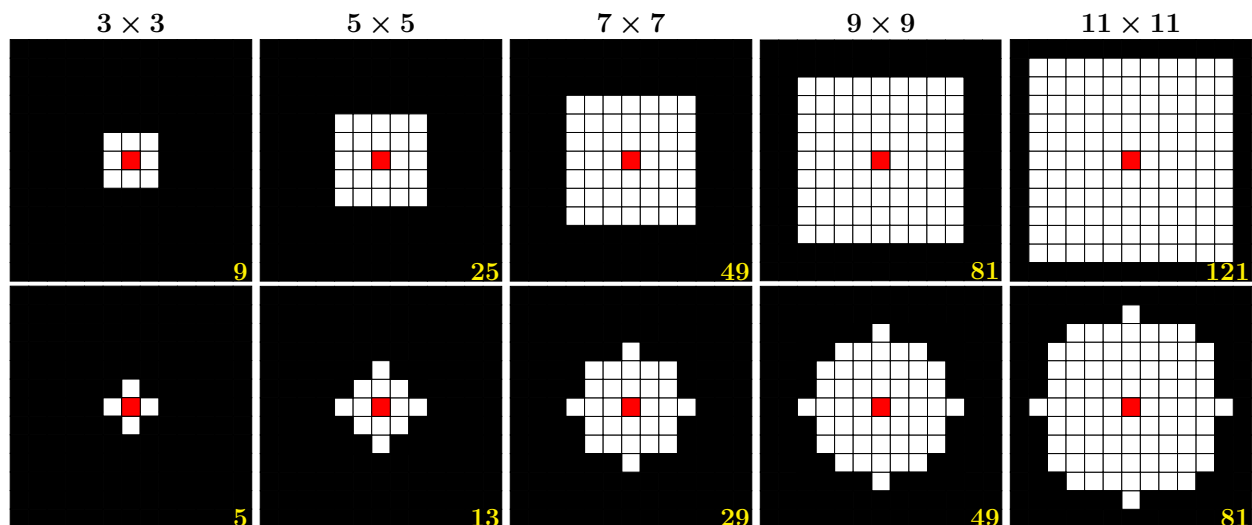
Figure 4.1: Rectangular kernels (first row) versus ellipsoidal kernels (second row) for different kernel sizes.
The center of the kernel is marked in red, while other locations within the support are shown in white.
For each configuration, the total number of kernel coefficients is indicated in yellow in the right-bottom
corner. As the kernel size increases, the ellipsoidal kernels have substantially fewer coefficients than the
corresponding rectangular kernels.

Choosing the number of parameters of a reconstruction model represents a classical trade-off.

A reconstruction model with too few parameters may not have enough representational capabilities

to accurately capture the important features of the desired image to be reconstructed. On the other

hand, a model with too many parameters can be prone to overfitting, can be sensitive to noise, and

can require more training data than models with fewer parameters. As a result, it is important to

choose a kernel shape that balances these factors appropriately.

Another factor to consider is that reconstruction methods that use convolution in k-space can always be interpreted as using multiplication in the image domain [7,8,13,32,92,93]. In that sense, the shape of the support set Ω implicitly determines the characteristics of the spatial-domain function corresponding to the convolution kernel $w[\mathbf{k}]$. Interestingly, it is well-established that functions with rectangular k-space support are expected to have highly-anisotropic spatial resolution (with much higher spatial resolution along diagonal lines), while functions with ellipsoidal k-space support are expected to have more isotropic spatial resolution characteristics [91]. This has led some authors to suggest that, in the context of data sampling to achieve a certain target resolution, it suffices to acquire an ellipsoidal region of k-space and acquiring the corners of k-space is often inefficient [91]. By the same logic, it is reasonable to hypothesize that the corners of rectangular convolution kernels may not be very important in practice, and that ellipsoidal convolution kernels may be able to achieve similar capabilities with better efficiency.

4.3   Methods

We carried out a systematic comparison between the rectangular and the ellipsoidal kernel shapes, by assessing the performance and efficiency of seven representative k-space convolutional reconstruction methods: GRAPPA, SPIRiT, ESPIRiT, SAKE, LORAKS , AC-LORAKS, and a CNN-based approach.

In what follows we give a very brief explanation of the selected methods. GRAPPA [5] is a noniterative reconstruction method that interpolates a missing point as a shift-invariant linear combination of neighboring samples (with the set of neighbors defined by the kernel shape) using kernel weights previously estimated from autocalibration (ACS) data. SPIRiT [7] is an iterative reconstruction method that uses the constraint that every point in k-space (regardless of whether that point was sampled or missing in the original acquisition) can be predicted as a shift-invariant linear combination of neighboring samples. As before, the set of neighbors is defined by the kernel shape, and the kernel weights are estimated using ACS data. The ESPIRiT [8], SAKE [31], LORAKS [32], and AC-LORAKS [11] methods are all based on structured matrices, in which convolution-structured (i.e., Hankel or Toeplitz) matrices are formed from k-space data such that

multiplying the structured matrix with a vector is equivalent to a convolution with the k-space data. In all of these methods, the kernel shape influences the way that the structured matrices are constructed. Each method uses structured matrices in slightly different ways. ESPIRiT is a sensitivity-map estimation method that forms a structured matrix from ACS data, and obtains sensitivity maps from the eigenvectors of this matrix. SAKE is a reconstruction method that recovers missing k-space samples by enforcing a constraint that a structured matrix formed from the reconstructed k-space data is expected to have low-rank characteristics due to parallel imaging and image support constraints. Similarly, LORAKS uses similar low-rank matrix constraints to recover missing k-space samples, but constructs the structured matrix and enforces the low-rank constraints differently than SAKE does. For the results shown in this work, we utilized the version of LORAKS based on the "**S**-matrix", which simultaneously imposes parallel imaging, image support, and smooth-phase constraints. AC-LORAKS is an autocalibrated version of LORAKS, in which the nullspace of the LORAKS matrix is estimated from ACS data. This allows image reconstruction to be reformulated as a simple least-squares problem instead of a more complicated low-rank matrix recovery problem, greatly improving computation speed. The CNN-based method we considered is the U-Net approach that has been used as a benchmark in the FastMRI challenges [94]. This approach uses a CNN to remove the artifacts from coil-combined images formed through simple zero-filled reconstructions, where the convolution filters are trained using thousands of reference datasets. It should be noted that the U-Net is based on image-domain convolution instead of k-space convolution (which all of the other methods are based on).

We assessed the performance of GRAPPA, SPIRiT, SAKE, LORAKS, and AC-LORAKS by reconstructing retrospectively-undersampled data using the ten different kernel supports shown in Fig. 4.1, comprising five rectangular (square) kernels with different side lengths and five ellipsoidal (circular) kernels with different radii. Two datasets were used: 12-channel T2-weighted brain data acquired with a 256×187 matrix size; and 4-channel T1-weighted brain data acquired with a 258×256 matrix size. These datasets are shown in Fig. 4.2. Retrospective undersampling was performed for several different sampling patterns as shown in Fig. 4.2. The sampling patterns always include 24 fully-sampled lines at the center of k-space to be used as ACS data. Outside

of this region, we used conventional uniform undersampling with acceleration factors of $R = 2, 3$, and 4. In order to test both kernel shapes with other undersampling schemes, the T2-weighted dataset was accelerated using a partial Fourier acquisition scheme with the same number of k-space lines used for the T1-weighted dataset. The lines outside the ACS region were distributed proportionally such that only 6/8 of the k-space includes fully-sampled lines. Reconstruction results were evaluated using normalized root-mean-squared error (NRMSE), and we also evaluated the amount of computation time and the amount of RAM used in each case. These results were obtained on a computer with an Intel Xeon E5-1603 2.8 GHz quad core CPU processor and 32GB of RAM.
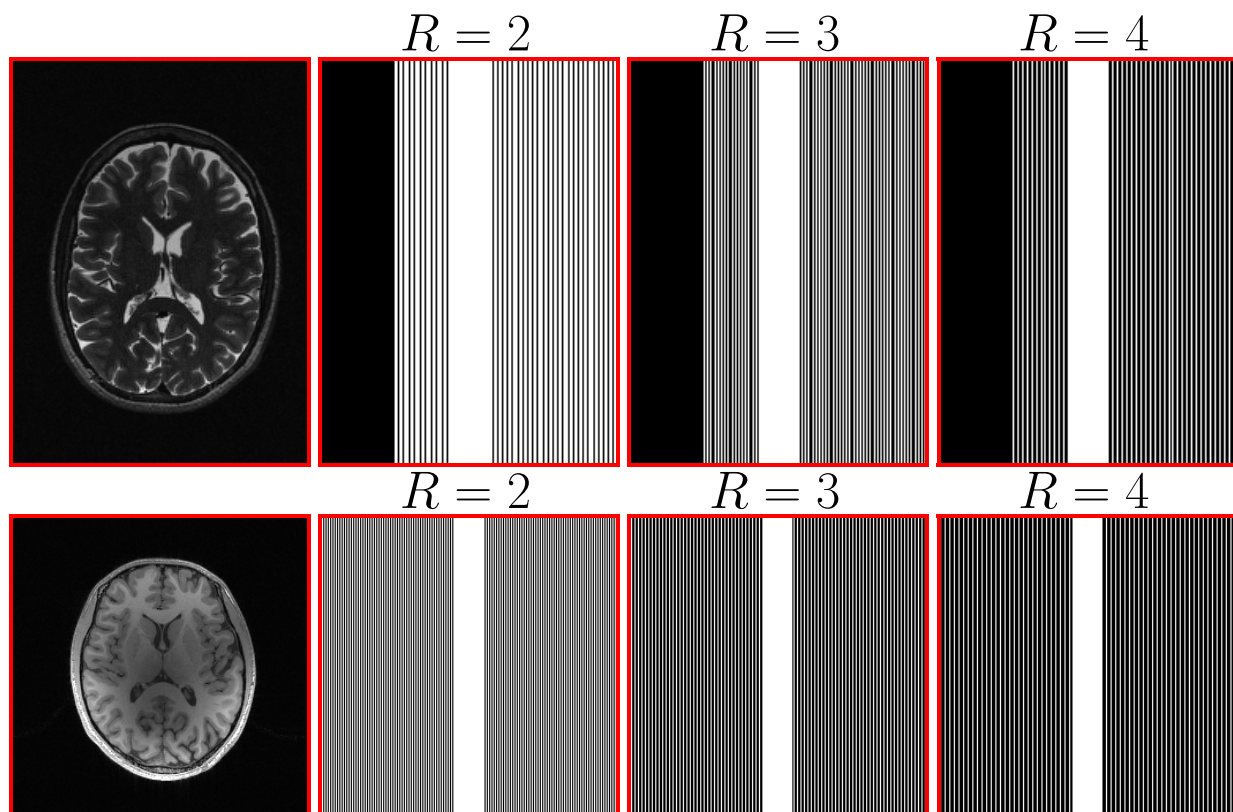
Figure 4.2: The T2-weighted (first row) and the T1-weighted (second row) datasets with the corresponding sampling patterns used in our experiments. We show results after coil combination (root sum-of-squares).

ESPIRiT was evaluated using the same two datasets, using the same ten kernel supports and the same 24 lines of ACS data described previously. The quality of the estimated sensitivity maps

was assessed by calculating the size of the residual error (quantified as NRMSE) after projecting the fully-sampled data onto the subspace spanned by the estimated sensitivity maps [8].

The U-Net was evaluated with knee data from the 2019 FastMRI challenge [94]. We compared the original U-Net (which used 3×3 rectangular kernels) against a modified U-Net (which used 3×3 ellipsoidal kernels, which was achieved using 3×3 rectangular kernels but forcing the values in the corners to be zero). Both the original and modified U-Nets were trained for 27 epochs from the same random initialization using $3,474$ single-channel training examples. We used the undersampling scheme used in [94]. For validation, we used $7,135$ single-channel datasets that were not part of the training set.

4.4   Results

Figure 4.3 shows the NRMSE results for the T2-weighted dataset obtained for GRAPPA, SPIRiT, ESPIRiT, SAKE, LORAKS, and AC-LORAKS, when using the kernel shapes and kernel sizes shown in Fig. 4.1, with the corresponding results for the T1-weighted dataset shown in Supporting Information Fig. S1. For both datasets it can be observed that, in most cases, the NRMSE achieved by both kernel shapes was quite similar, which can likely be attributed to the fact that both shapes have practically the same base spatial resolution. It should also be noted that for GRAPPA, the 3×3 rectangular kernel has significantly better performance than the 3×3 ellipsoidal kernel. This is expected since, for this small kernel size, the ellipsoidal kernel only uses two neighboring k-space samples (one on each side) to interpolate each missing sample, which is much smaller than the six neighboring k-space samples used by the rectangular kernel. For SPIRiT we observed a similar NRMSE for both kernel shapes comparable to the values obtained for the other methods, however, for the T1-weighted dataset, both kernel shapes exhibited a substantial decrease in performance when the kernel size was 11×11. The same phenomenon was observed for the T2-weighted dataset but only for the rectangular shape. This poor performance can be attributed to the fact that these large kernels have a large number of parameters that were difficult to estimate reliably based on the relatively small number of ACS lines.
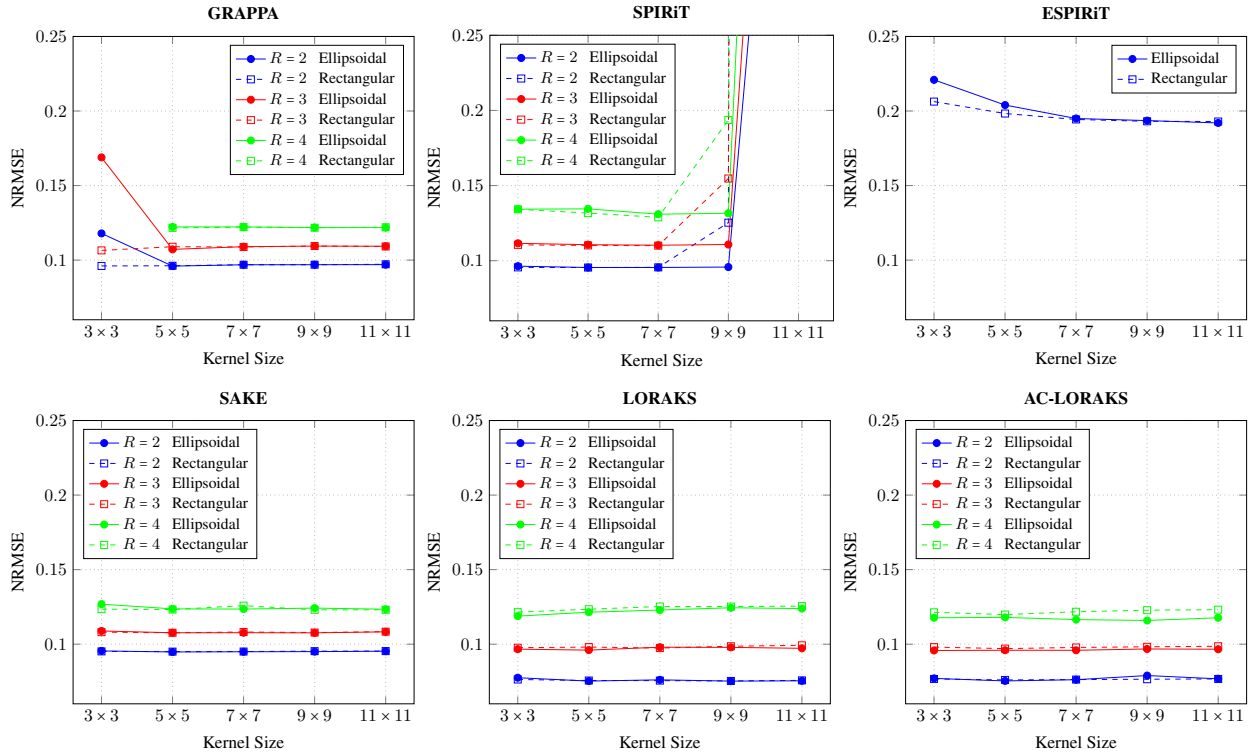
Figure 4.3: NRMSE for different reconstruction methods for the T2-weighted dataset as a function of kernel size and kernel shape.

Figure 4.5 and Supporting Information Fig. S2 respectively show corresponding computation times for the T2-weighted and T1-weighted datasets. We observe that the elliptical kernels shapes constitently yielded faster computation times than the rectangular kernel shapes, with generally bigger differences for larger kernel sizes. This matches the theoretical expectations described previously.

Supporting Information Figs. S3 and S4 respectively show the corresponding memory usage for the T2-weighted and T1-weighted datasets. Analogous results are shown in Supporting Information Fig. 4 for the T1-weighted dataset. In all cases, the ellipsoidal kernels used less memory than the rectangular kernels, as expected based on previous theoretical arguments.

Figure 4.9 shows the histogram of NRMSE values obtained from the validation set of images for CNN-based reconstruction. As can be seen, there is negligible difference in NRMSE performance for the two different kernel shapes, despite the fact that the network based on ellipsoidal kernels had substantially fewer parameters (i.e., $\approx 4 \times 10^6$) than the network based on rectangular kernels (i.e.,
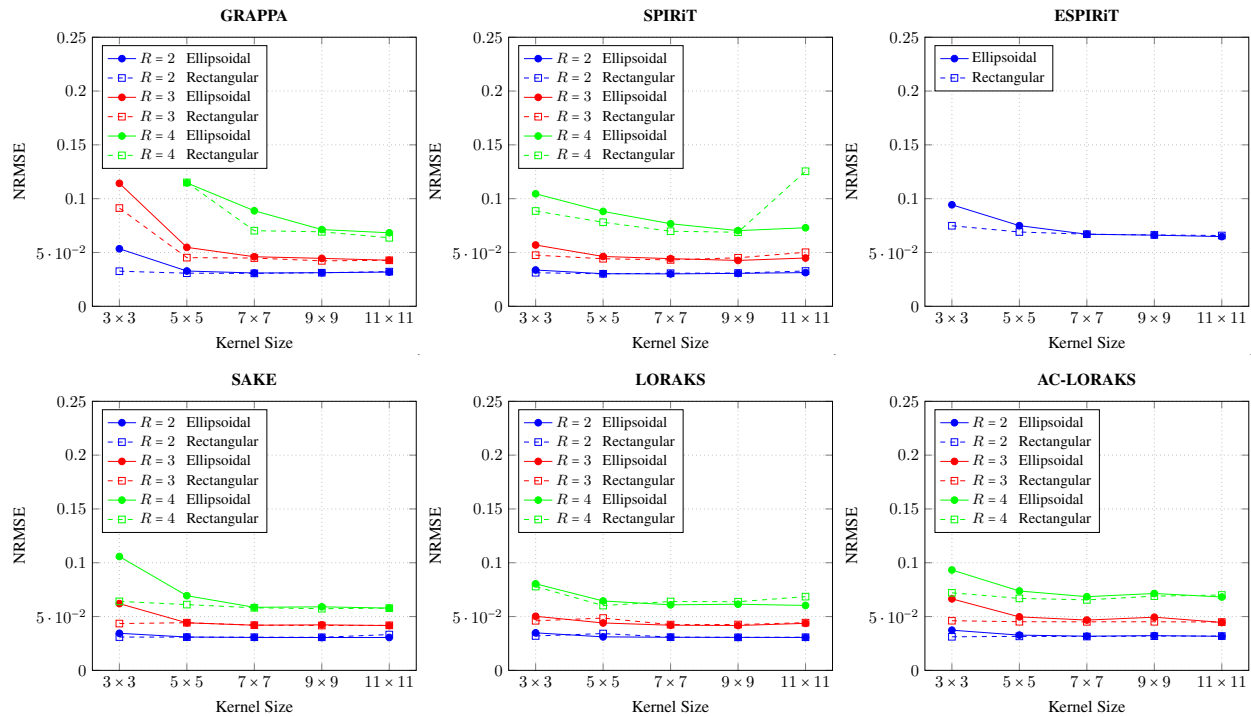
Figure 4.4: NRMSE for different reconstruction methods for the T1-weighted dataset as a function of kernel size and kernel shape.

$\approx 7 \times 10^6$). This difference could be potentially leveraged to reduce the complexity of the training process, although for simplicity, we utilized a naive implementation that did not (and would not be expected to) demonstrate computational benefits.
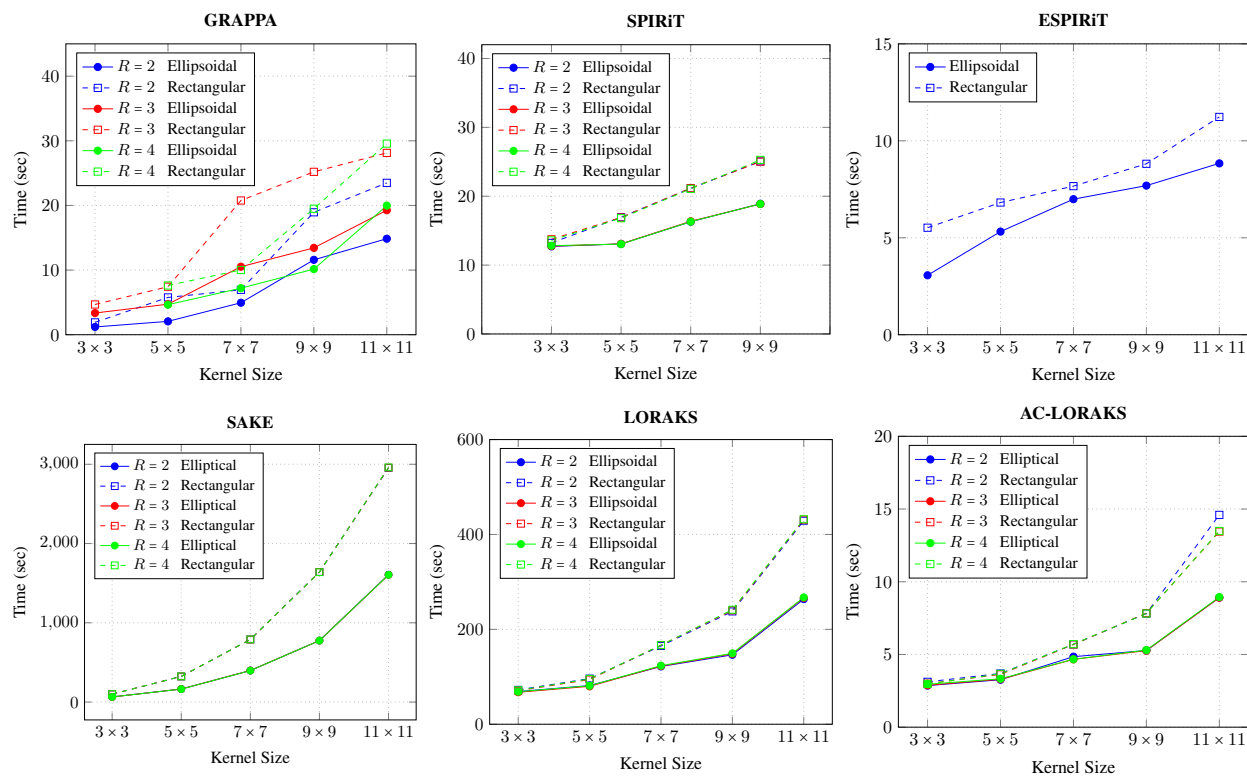
Figure 4.5: Computation times for different reconstruction methods for the T2-weighted dataset as a function of kernel size and kernel shape.
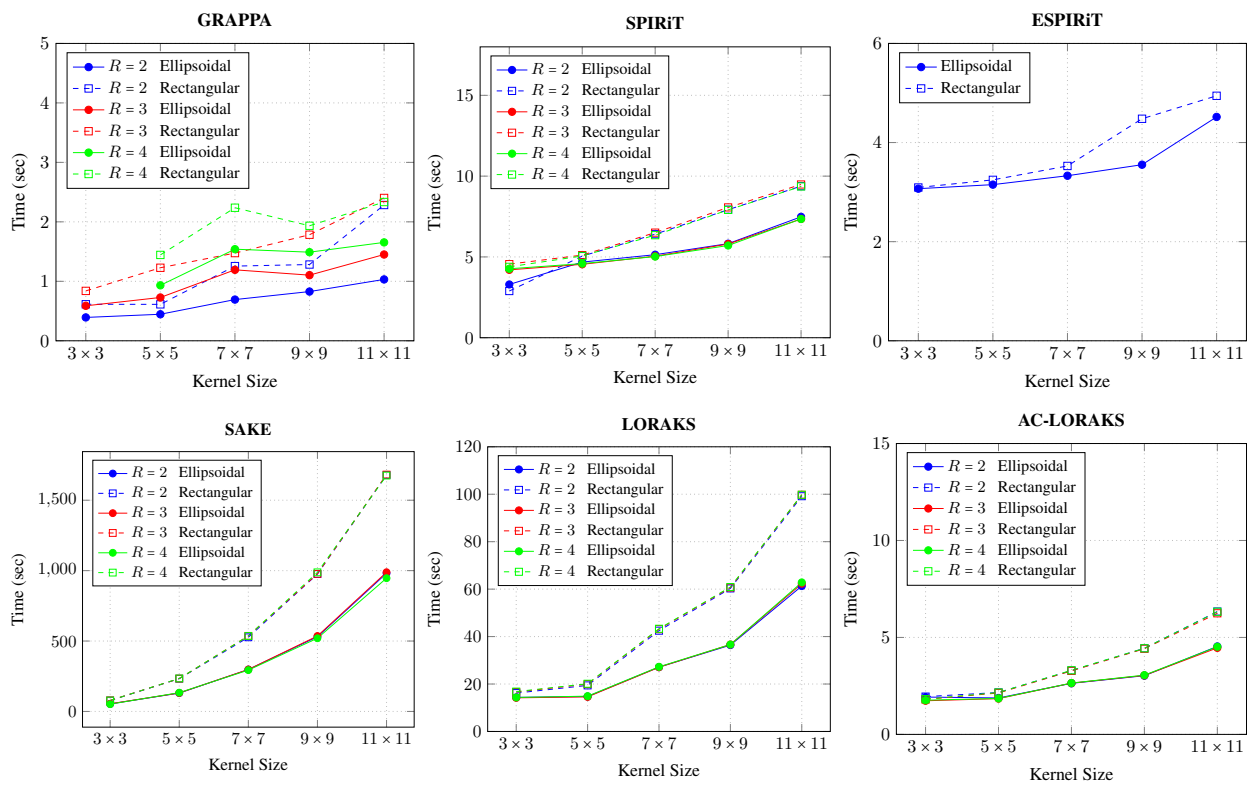
Figure 4.6: Computation times for different reconstruction methods for the T1-weighted dataset as a function of kernel size and kernel shape.
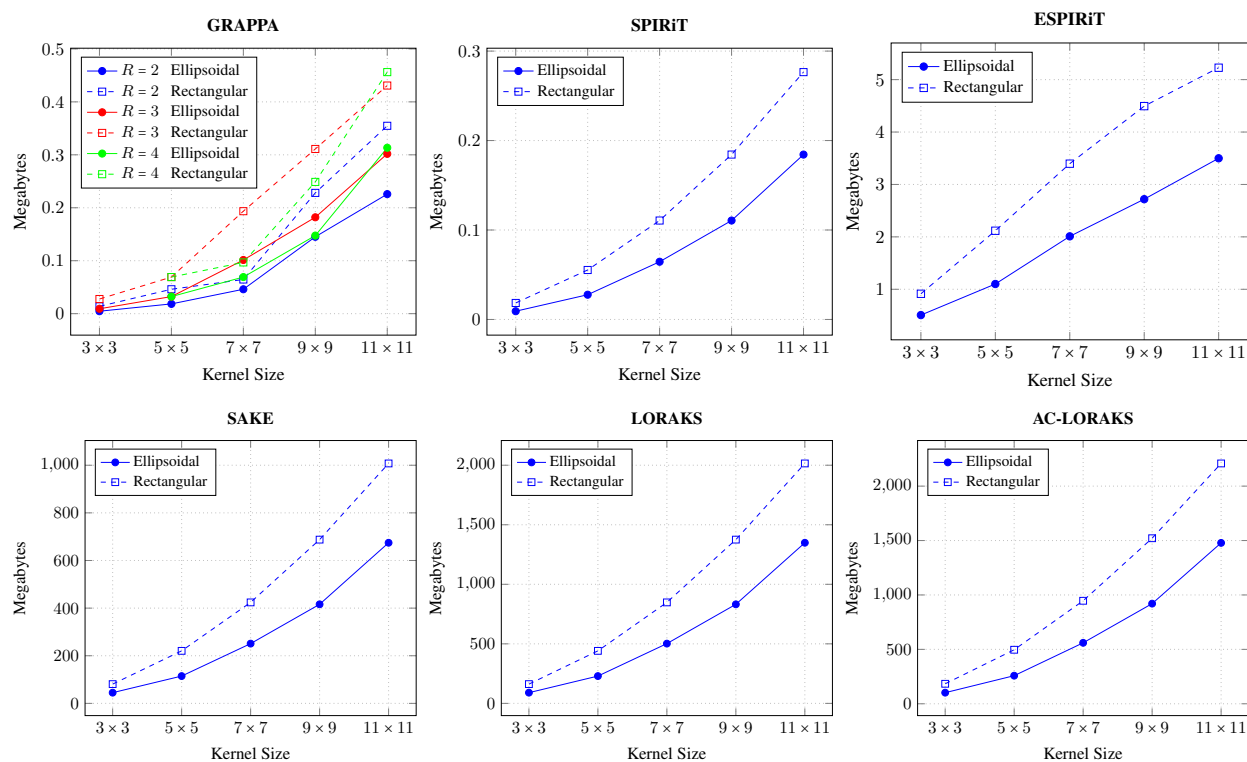
Figure 4.7: Memory usage for different reconstruction methods for the T2-weighted dataset as a function of kernel size and kernel shape.
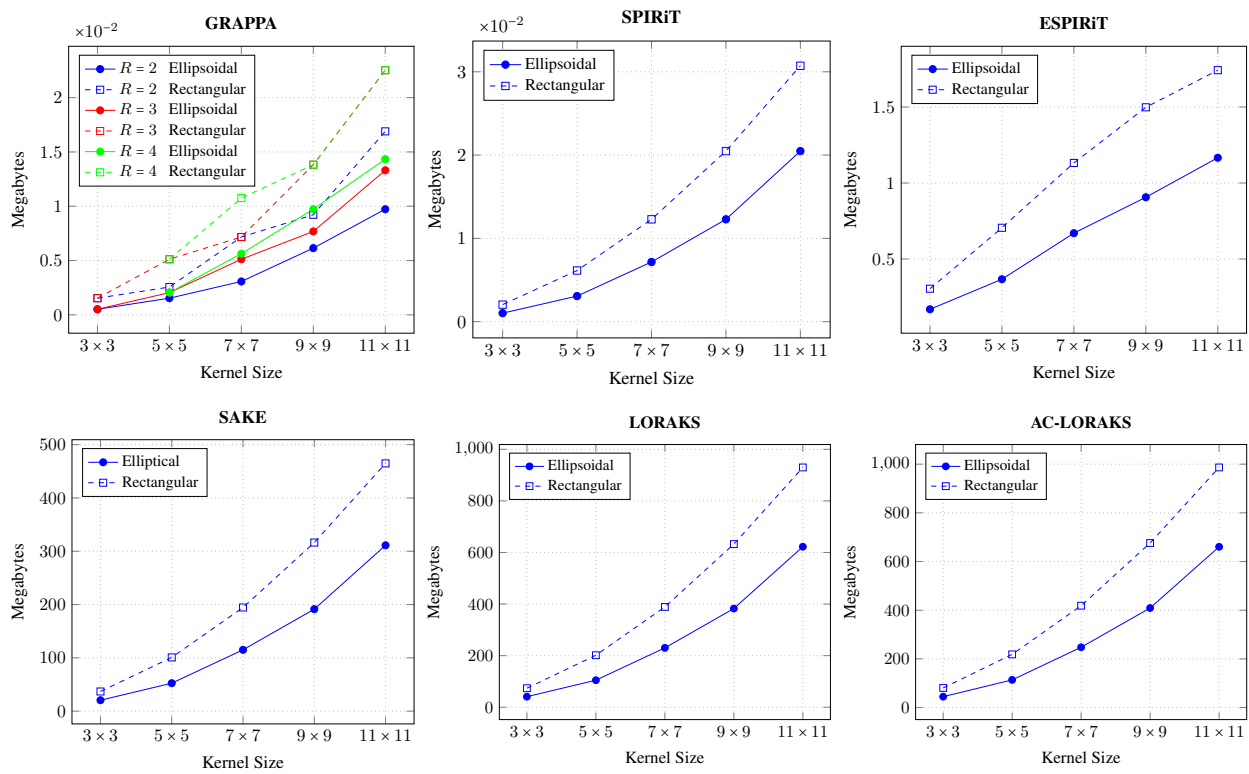
Figure 4.8: Memory usage for different reconstruction methods for the T1-weighted dataset as a function of kernel size and kernel shape.
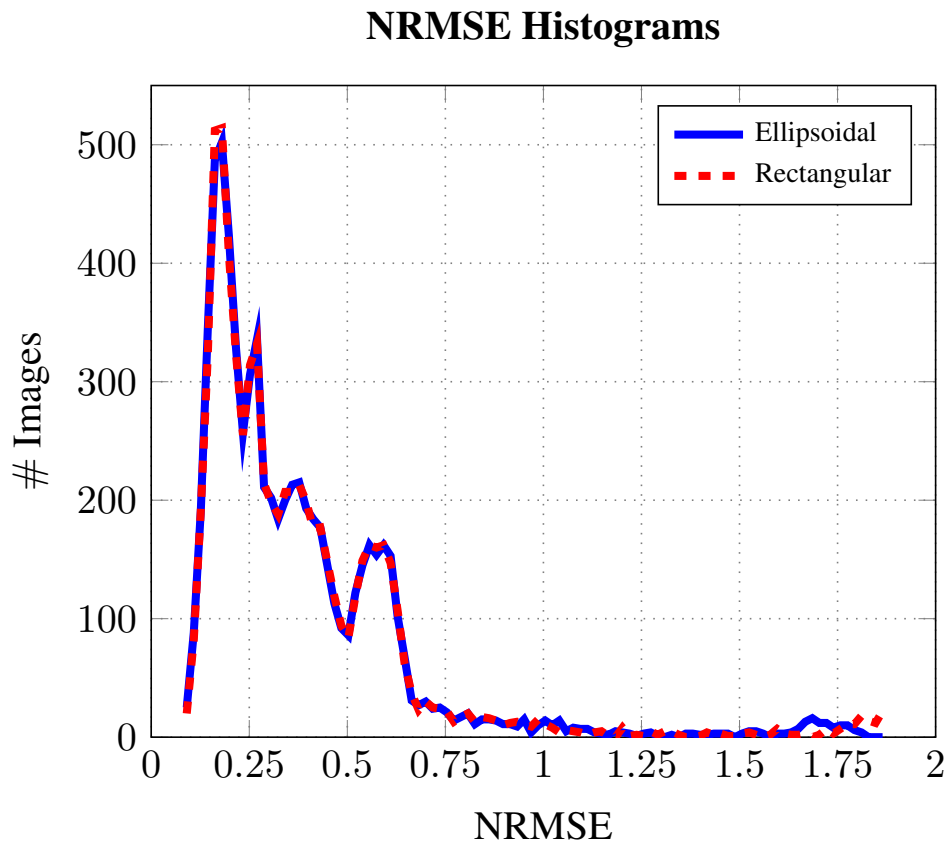
NRMSE Histograms

Figure 4.9: Histograms of the NRMSE values obtained on the validation set by the CNN-based reconstruction for the different kernel shapes.

Discussion and Conclusions

To the best of our knowledge, this work represents the first systematic evaluation of kernel shapes for convolution-based MRI reconstruction methods. Although ellipsoidal kernels are not very popular in the modern literature, our results suggest that ellipsoidal kernels often offer advantages (in computation time, memory usage, and the number of model parameters) over more-common rectangular kernels, with largely similar NRMSE performance. We expect that these new insights may be valuable for improving the efficiency of MRI reconstruction in the future.

In our evaluations, we have focused only on square-shaped and circle-shaped kernels, without taking full advantage of the degrees of freedom offered by rectangles and ellipses (with different sizes along different axes). We anticipate that additional improvements might be obtained by leveraging this untapped degree of freedom. It should also be noted that the ellipsoidal kernel shape is not necessarily optimal, and that other shapes (perhaps learned directly from empirical data) might achieve even better results. Exploration of such ideas is a potentially interesting direction for further research, though is beyond the scope of the present work.

Chapter 5

New Theory and Faster Computations for Subspace-Based Sensitivity Map Estimation in Multichannel MRI

5.1 Introduction

Most modern magnetic resonance imaging (MRI) experiments are performed using multichannel phased-array receiver coils [4–7, 9, 95–106]. For acquisitions involving Q receiver channels, the measurements obtained in a multichannel MRI experiment are usually modeled using the Fourier transform as

$$d_q(\mathbf{k}_m) = \int c_q(\mathbf{x})\rho(\mathbf{x})e^{-i2\pi \mathbf{k}_m^T \mathbf{x}}d\mathbf{x} + \eta_{qm} \tag{5.1}$$

for $m = 1, \ldots, M$ and $q = 1, \ldots, Q$. In this expression, M represents the total number of measured k-space positions $\mathbf{k}_m \in \mathbb{R}^D$, where D is the dimension of the image;[1] $d_q(\mathbf{k}_m)$ represents the complex-valued k-space data measured from the qth receiver channel at the mth position in k-space; $\rho(\mathbf{x})$ is the complex-valued underlying MR image (reflecting the state of the excited magnetization at the time of data acquisition) as a function of the spatial position $\mathbf{x} \in \mathbb{R}^D$; $c_q(\mathbf{x})$ is a complex-valued sensitivity map describing the sensitivity of the qth receiver channel to excited magnetization

---

[1]For simplicity, we have adopted the use of "spatial" notation throughout this paper, which is consistent with typical use cases in 2D or 3D imaging. In these cases, we would have either $D = 2$ with $\mathbf{x} = [x, y]^T$ and $\mathbf{k} = [k_x, k_y]^T$ or $D = 3$ with $\mathbf{x} = [x, y, z]^T$ and $\mathbf{k} = [k_x, k_y, k_z]^T$. However, it should be noted that our approach is also compatible with other scenarios such as spatiotemporal imaging with time-varying sensitivities, which, for example, might use $D = 3$ with $\mathbf{x} = [x, y, f]^T$ and $\mathbf{k} = [k_x, k_y, t]^T$.

at spatial position $\mathbf{x}$; and η_{qm} represents complex-valued measurement noise. We will also use $\rho_q(\mathbf{x}) \triangleq c_q(\mathbf{x})\rho(\mathbf{x})$ to denote the qth sensitivity-weighted image.

This paper is focused on the problem of estimating the sensitivity profiles $c_q(\mathbf{x})$ from a set of calibration measurements (i.e., a set of k-space data samples that are sampled at the Nyquist rate). While prior knowledge of the coil sensitivity maps is not always required, there are many tasks that can benefit from estimated sensitivity maps, including coil combination (in which the underlying MRI image $\rho(\mathbf{x})$ is estimated from the multichannel images $\rho_q(\mathbf{x})$ [95, 107, 108]) and accelerated image reconstruction (in which the underlying MRI image $\rho(\mathbf{x})$ is estimated from multichannel k-space data $d_q(\mathbf{k}_m)$ that is sampled below the Nyquist rate [4, 6, 102–106]).

Over the years, many different sensitivity map estimation methods have been proposed [4, 6, 8, 66, 107–116]. Among these, subspace-based approaches [8, 113–115] have proven to be particularly powerful and popular, with the subspace-based ESPIRiT method [8] rising to become the most widely-used sensitivity map estimation method in the modern literature because of its simplicity (with few tuning parameters) and excellent performance.

The mathematical principles underlying subspace-based sensitivity map estimation are nontrivial and are not always easy to understand. In the first part of this work, we present a novel theoretical derivation of subspace-based sensitivity map estimation that is based on a nullspace/linear predictability perspective [13], and makes use of concepts from the recent literature on structured low-rank matrix modeling [11–13, 31, 32, 36, 37, 117]. This approach is distinct from and complementary to existing theoretical explanations, and we personally find it to be more intuitive than the alternatives.

In the second part of this work, we introduce several new computational approaches that can substantially reduce the time and memory required for subspace-based sensitivity map estimation. This can be important because existing subspace-based methods can be computationally demanding, with substantial memory requirements and slow computation times. While slow computations can already be problematic when considering individual datasets, they can be especially limiting in modern machine learning contexts where it may be of interest to calculate sensitivity maps for every image within a database containing hundreds or thousands of datasets.

Our computational contributions include the following:

- We leverage our previous observation that using ellipsoidal convolution kernels instead of rectangular convolution kernels can lead to substantial reductions in memory usage and computational complexity, without sacrificing accuracy [16].

- We introduce a fast FFT-based approach that allows us to rapidly calculate the nullspace vectors associated with a multichannel convolution-structured matrix, without ever having to directly construct the actual convolution-structured matrix (which would normally require substantial memory and time to assemble). This approach can be viewed as a multichannel generalization of an approach that was previously developed for the single-channel case [75]. In addition to being useful for subspace-based sensitivity map estimation, this type of approach may also be useful for the broad class of structured low-rank modeling methods for multichannel MRI image reconstruction [11–13, 31, 32, 36, 37, 117]. Note that a similar FFT-based approach was proposed for multichannel MRI reconstruction in Ref. [118]. Compared to that approach, our implementation uses substantially fewer FFTs, which is enabled by making use of substantial redundancies that are present in the convolution-structured matrix.

- Noting that sensitivity maps arise from Maxwell's equations and must be smooth, we observe that estimating sensitivity maps at the nominal desired spatial resolution can be computationally wasteful. Instead, we propose a method that first estimates the sensitivity maps on a low-resolution spatial grid, and then uses a fast FFT-based spatial interpolation procedure to obtain sensitivity maps at the desired spatial resolution.

- Subspace-based sensitivity map estimation methods often require assembling a large number of small-scale matrices (e.g., one matrix for every spatial location in the image) which in total require substantial memory. We propose a memory-efficient approach based on signal processing techniques to calculate these matrices.

- Subspace-based sensitivity map estimation methods and coil combination methods often require computing a large number of small-scale singular value decompositions (SVDs) (e.g.,

one SVD for every spatial location in the image) [8, 107]. We introduce an efficient itera-
tive approach (based on classical power iteration [119]) to calculate many small-scale partial
SVDs simultaneously. While this approach is directly useful for subspace-based sensitivity
map estimation, we expect that it may also have some utility for a range of locally low-rank
modeling and denoising methods that also require calculating partial SVDs over local image
patches for every spatial location in the image [120–124].

Combined together, these algorithmic improvements — which we collectively call PISCO (Power
iteration over simultaneous efficiently-assembled patches, Interpolation, ellipSoidal kernels, and
FFT-based COnvolution) — enable a substantial (roughly 100-fold in some cases) computational
acceleration for subspace-based sensitivity map estimation.

5.2   Characteristics of the Inverse Problem

Before describing our novel theory and methods, we will first review some important characteristics
of the sensitivity map estimation problem.

As mentioned in the previous section, the problem we consider in this paper is the estimation of
the sensitivity maps $c_q(\mathbf{x})$ for $q = 1, \ldots, Q$ from a set of Nyquist-sampled calibration measurements.
We will assume that these calibration measurements consist of a collection of M k-space samples
from each of the Q channels, for a total of MQ observations. Our goal will be to calculate the
values of the sensitivity maps on a discrete voxel grid of N spatial locations $\mathbf{x}_n$, for $n = 1, \ldots, N$,
resulting in a total of NQ unknowns. Most practical scenarios will involve substantially more spatial
locations than k-space samples (i.e., $NQ \gg MQ$), such that the inverse problem is underdetermined
(with fewer measurements than unknowns) unless additional assumptions are made about the
characteristics of $c_q(\mathbf{x})$. To overcome this issue, the literature on sensitivity map estimation typically
makes the assumption that the sensitivity maps $c_q(\mathbf{x})$ are spatially smooth, which can be justified
based on physical principles (e.g., Maxwell's equations).

A more pernicious issue is that, while our goal is to estimate the unknown sensitivity maps
$c_q(\mathbf{x})$, the image $\rho(\mathbf{x})$ that appears in Eq. (5.1) is also generally an unknown. This means that

Eq. (5.1) can be viewed as a bilinear inverse problem, which is problematic because bilinear inverse problems have well-known scaling ambiguities. In particular, given one set of estimated sensitivity maps $c_q(\mathbf{x})$ for $q = 1, \ldots, Q$ and a corresponding estimated image $\rho(\mathbf{x})$ that together provide a good fit to the measured data, it is straightforward to observe that another equally-good fit to the measured data can be obtained by combining a rescaled version of the sensitivity maps $\alpha(\mathbf{x})c_q(\mathbf{x})$ for $q = 1, \ldots, Q$ together with a rescaled version of the image $\frac{1}{\alpha(\mathbf{x})}\rho(\mathbf{x})$. The complex-valued scaling function $\alpha(\mathbf{x})$ must be nonzero for every point $\mathbf{x}$ within the field of view (FOV), but is otherwise completely unconstrained, and can have arbitrary spatial variations. As such, unless additional strong assumptions are made, there will be infinitely many ways of choosing $\alpha(\mathbf{x})$ that all lead to sensitivity map estimates that provide equally-good fits to the measured data. Practically, this means that both the magnitude and phase of the estimated sensitivity maps are inherently ambiguous, and we can only hope for sensitivity map estimates that are unique up to these ambiguities.

Because of the inherent ambiguity, it is common for sensitivity map estimation methods to search for one set of estimated sensitivity maps $c_q(\mathbf{x})$ that fit the data well, and subsequently rescale them with a heuristically-chosen scaling function $\alpha(\mathbf{x})$ that can be designed to endow the scaled sensitivity maps (and the corresponding scaled image) with properties that are deemed to be useful.

Note also that the sensitivity map values $c_q(\mathbf{x})$ will have no impact on measured data at spatial locations where $\rho(\mathbf{x}) = 0$. As a result, the values of the sensitivity maps are completely unidentifiable at spatial locations outside the support of $\rho(\mathbf{x})$, leading to additional ambiguity.

5.3 Nullspace/Linear Predictability Theory for Subspace-Based Sensitivity Map Estimation

In this section, we present a novel derivation of subspace-based sensitivity map estimation from a nullspace/linear predictability perspective [13]. We will start with a quick summary of the relevant linear predictability relationships that are known to exist for multichannel MRI, before making novel links to sensitivity map estimation.

5.3.1   Summary of Linear Predictability for Multichannel MRI

Define the spatial support of $\rho(\mathbf{x})$ as the set of points $\Omega = \{\mathbf{x} \in \mathbb{R}^D : |\rho(\mathbf{x})| > 0\}$. In what follows, we will assume (without loss of generality) that the spatial coordinate system has been normalized such that the support Ω is completely contained within a hypercube Γ with sides of length one, defined as $\Gamma \triangleq \{\mathbf{x} \in \mathbb{R}^D : \|\mathbf{x}\|_\infty < \frac{1}{2}\}$. We will refer to Γ as the FOV. This assumption allows us to equivalently represent the continuous images $\rho(\mathbf{x})$ and $\rho_q(\mathbf{x})$ using infinite Fourier series representations as

$$\rho(\mathbf{x}) = \mathbb{1}_\Gamma(\mathbf{x}) \sum_{\mathbf{n} \in \mathbb{Z}^D} s[\mathbf{n}] e^{i2\pi \mathbf{n}^T \mathbf{x}} \tag{5.2}$$

and

$$\rho_q(\mathbf{x}) = \mathbb{1}_\Gamma(\mathbf{x}) \sum_{\mathbf{n} \in \mathbb{Z}^D} s_q[\mathbf{n}] e^{i2\pi \mathbf{n}^T \mathbf{x}} \tag{5.3}$$

for $q = 1, \ldots, Q$, where $s[\mathbf{n}]$ and $s_q[\mathbf{n}]$ respectively represent samples of the Fourier transforms of $\rho(\mathbf{x})$ and $\rho_q(\mathbf{x})$ on the rectilinear Nyquist grid (taking $\Delta k = 1$ along each dimension, as enabled by our assumptions about Γ), and $\mathbb{1}_\Gamma(\mathbf{x})$ is the indicator function for the FOV, with

$$\mathbb{1}_\Gamma(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \Gamma \\ 0, & \text{else.} \end{cases} \tag{5.4}$$

A basic assumption of the linear predictability formalism [13] is that the multichannel collection of sequences $\{s_q[\mathbf{n}]\}_{q=1}^Q$ will be autoregressive, satisfying multiple multichannel shift-invariant linear predictability relationships [5,7–9,11–13,31,32,36,113–115,117,125]. Specifically, there should exist multiple different Q-channel finite impulse response (FIR) filters $h_r[\mathbf{n}, q]$ that each satisfy

$$\sum_{q=1}^Q \sum_{\mathbf{m} \in \Lambda} h_r[\mathbf{m}, q] s_q[\mathbf{n} - \mathbf{m}] \approx 0 \text{ for } \forall \mathbf{n} \in \mathbb{Z}^D \tag{5.5}$$

for $r = 1, \ldots, R$, where R is the number of different filters and $\Lambda \subset \mathbb{Z}^D$ is a finite index set describing the support of each FIR filter.

There are many different situations in which the existence of such filters $h_r[\mathbf{n}, q]$ has been
theoretically proven [13]. We review a few of these below:[2]

- **Limited Image Support [13, 32, 125].** In many situations, the actual spatial support Ω of
  the original image $\rho(\mathbf{x})$ will be smaller than the FOV Γ, such that the complement of Ω in Γ
  (denoted by $\Omega^{\complement} \triangleq \Gamma \setminus \Omega$), is a measurable nonempty set. In such cases, it is possible to define
  nonzero functions $f(\mathbf{x})$ that have their supports contained within $\Omega^{\complement}$. Since $\rho(\mathbf{x})$ and $f(\mathbf{x})$
  have nonoverlapping support, we must have $f(\mathbf{x})\rho(\mathbf{x}) = 0$ for $\forall \mathbf{x} \in \Gamma$. If we let $\tilde{f}[\mathbf{n}]$ denote
  the Nyquist-sampled Fourier series representation of $f(\mathbf{x})$, then taking the Fourier transform
  of $f(\mathbf{x})\rho(\mathbf{x})$ and applying the Fourier convolution theorem yields that

$$\sum_{\mathbf{m} \in \mathbb{Z}^D} \tilde{f}[\mathbf{m}]s[\mathbf{n} - \mathbf{m}] = 0 \text{ for } \forall \mathbf{n} \in \mathbb{Z}^D. \tag{5.6}$$

If $\Omega^{\complement}$ is large enough, then it becomes possible to choose relatively smooth functions $f(\mathbf{x})$ that
satisfy the conditions specified above. If we assume that these functions are smooth enough
that they satisfy $\tilde{f}[\mathbf{n}] \approx 0$ for $\forall \mathbf{n} \notin \Lambda$ (i.e., the functions $f(\mathbf{x})$ are approximately bandlimited
to the frequency range defined by Λ),[3] then we can approximate $\tilde{f}[\mathbf{n}]$ as an FIR filter, and
obtain

$$\sum_{\mathbf{m} \in \Lambda} \tilde{f}[\mathbf{m}]s[\mathbf{n} - \mathbf{m}] \approx 0 \text{ for } \forall \mathbf{n} \in \mathbb{Z}^D. \tag{5.7}$$

Note also that if we have $f(\mathbf{x})\rho(\mathbf{x}) \approx 0$ for $\forall \mathbf{x} \in \Gamma$, then the relationship between $\rho(\mathbf{x})$ and
$\rho_q(\mathbf{x})$ implies we must also have that $f(\mathbf{x})\rho_q(\mathbf{x}) \approx 0$ for $\forall \mathbf{x} \in \Gamma$ for $q = 1, \ldots, Q$. This implies
that the multichannel FIR filter

$$h_r[\mathbf{n}, q] = \begin{cases} \alpha_q \tilde{f}[\mathbf{n}], & \mathbf{n} \in \Lambda \\ 0, & \text{else} \end{cases} \tag{5.8}$$

---

[2]Note that shift-invariant linear predictability relationships can also be proven for images that possess smooth
phase [13, 32, 126] and sparsity characteristics [13, 36, 37, 70, 117, 127], although these relationships are not directly
relevant for sensitivity map estimation and we do not discuss them further.

[3]While we can often obtain very accurate approximations of this form, this must always be an approximation,
and will never be exact. Specifically, since $f(\mathbf{x})$ is assumed to have finite support, then $\tilde{f}[\mathbf{n}]$ cannot have exactly-
limited support by the Fourier uncertainty principle [128]. While this paper describes this approximation informally
to simplify the exposition, more formal treatments exist in the literature [13].

will satisfy Eq. (5.5), for any choices of $\alpha_q \in \mathbb{C}$ for $q = 1, \ldots, Q$. It is easy to obtain many different $h_r[\mathbf{n}, q]$ filters from this relationship because of the flexibility offered by choosing different functions $f(\mathbf{x})$ and different coefficients α_q. Using similar arguments, additional channel-specific filters can also be obtained in common situations where some sensitivity-weighted images $\rho_q(\mathbf{x})$ have smaller support than the underlying image $\rho(\mathbf{x})$.

- **Smooth Sensitivity Maps [13,113,115,117].** As mentioned previously, it is often assumed that the sensitivity maps $c_q(\mathbf{x})$ are spatially smooth. Linear predictability relationships can be derived if we make the assumption that they are smooth enough that they can be treated as approximately bandlimited to the frequency range defined by Λ, such that

$$c_q(\mathbf{x}) \approx \mathbb{1}_\Gamma(\mathbf{x}) \sum_{\mathbf{n} \in \Lambda} \tilde{c}_q[\mathbf{n}] e^{i 2\pi \mathbf{n}^T \mathbf{x}} \tag{5.9}$$

for appropriate Fourier coefficients $\tilde{c}_q[\mathbf{n}]$. In this case, it can be observed that $c_i(\mathbf{x})\rho_j(\mathbf{x}) - c_j(\mathbf{x})\rho_i(\mathbf{x}) = 0$ for $\forall \mathbf{x} \in \Gamma$ for any choices of i and j, which is sometimes called the "cross-relation" in the multichannel blind deconvolution literature [129–131]. Taking the Fourier transform of the cross-relation and applying the Fourier convolution theorem leads to

$$\sum_{\mathbf{m} \in \Lambda} \tilde{c}_i[\mathbf{m}] s_j[\mathbf{n} - \mathbf{m}] - \sum_{\mathbf{m} \in \Lambda} \tilde{c}_j[\mathbf{m}] s_i[\mathbf{n} - \mathbf{m}] \approx 0 \text{ for } \forall \mathbf{n} \in \mathbb{Z}^D. \tag{5.10}$$

This implies that the multichannel FIR filter

$$h_r[\mathbf{n}, q] = \begin{cases} -\tilde{c}_j[\mathbf{n}], & \mathbf{n} \in \Lambda, q = i \\ \tilde{c}_i[\mathbf{n}], & \mathbf{n} \in \Lambda, q = j \\ 0, & \text{else} \end{cases} \tag{5.11}$$

will satisfy Eq. (5.5) for any choices of i and j with $i \neq j$. As before, it is easy to obtain many different $h_r[\mathbf{n}, q]$ filters using different choices of i and j.

- **Compressible Receiver Arrays [46, 47, 132].** It is frequently observed that the different

    channels in a multichannel receiver array possess linear dependencies, such that there exist

    multiple distinct choices of linear combination weights w_q that will each satisfy

$$\sum_{q=1}^{Q} w_q \rho_q(\mathbf{x}) \approx 0 \text{ for } \forall \mathbf{x} \in \Gamma. \tag{5.12}$$

    This observation is often used for dimensionality reduction ("coil compression") [46, 47, 132].

    However, this observation also implies that a relationship in the form of Eq. (5.5) will hold if

    $h_r[\mathbf{n}, q]$ is chosen as

$$h_r[\mathbf{n}, q] = \begin{cases} w_q, & \mathbf{n} = \mathbf{0} \\ 0, & \text{else.} \end{cases} \tag{5.13}$$

    Since there often exist multiple distinct choices of w_q that satisfy Eq. (5.12), there can also

    be many different filters $h_r[\mathbf{n}, q]$ that arise from this kind of relationship.

The preceding arguments all support the existence of multiple filters $h_r[\mathbf{n}, q]$ that satisfy Eq. (5.5).

In addition, if we were given full information about the support Ω and the sensitivity maps $c_q(\mathbf{x})$,

then we would be able to calculate filters $h_r[\mathbf{n}, q]$ directly from Eqs. (5.8), (5.11), and (5.13). How-

ever, in practice, we do not have information about Ω or $c_q(\mathbf{x})$ (these are quantities that we are

interested in estimating!), and we must instead resort to other methods to obtain the FIR filters

$h_r[\mathbf{n}, q]$.

Towards this end, an important observation from the literature on structured low-rank matrix

modeling [11–13, 31, 32, 36, 37, 117] is that the convolution appearing in Eq. (5.5) can be represented

in matrix form as $\mathbf{Ch} \approx \mathbf{0}$, where the matrix $\mathbf{C}$ is given by

$$\mathbf{C} \triangleq \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \dots & \mathbf{C}_Q \end{bmatrix}; \tag{5.14}$$

the matrix $\mathbf{C}_q \in \mathbb{C}^{P \times |\Lambda|}$ is a convolution-structured matrix (i.e., a Hankel or Toeplitz matrix)

formed from Fourier samples $s_q[\mathbf{n}]$, with each row of $\mathbf{C}_q$ corresponding to one value of $\mathbf{n}$ from

Eq. (5.5); P represents the number of different $\mathbf{n}$ values that are used in the construction of each

$\mathbf{C}_q$ matrix; and $\mathbf{h} \in \mathbb{C}^{Q|\Lambda|}$ is the vector of $h_r[\mathbf{n}, q]$ samples. See Ref. [12] for a detailed description of the matrix construction we use.

Even though the filter coefficients $\mathbf{h}$ may be *a priori* unknown, all filters $h_r[\mathbf{n}, q]$ that satisfy Eq. (5.5) must appear as approximate nullspace vectors of an appropriately-constructed $\mathbf{C}$ matrix. Importantly, it is possible to directly form such a $\mathbf{C}$ matrix from Nyquist-sampled calibration data, and its approximate nullspace can be identified by applying the SVD. It is therefore possible to learn a set of R linearly independent filters $h_r[\mathbf{n}, q]$ (with R equal to the dimension of the approximate nullspace of $\mathbf{C}$) in a data-driven way, without requiring any additional prior information [11]. Note that this nullspace-based method of computation will not result in filters that directly match the forms given in Eqs. (5.8), (5.11), or (5.13). Instead, the obtained nullspace vectors are likely to be linear combinations of such filters. In addition, the full set of approximate nullspace vectors for an appropriately-constructed $\mathbf{C}$ matrix is expected to form a basis for the set of all $h_r[\mathbf{n}, q]$ filters that are consistent with the FIR filter support Λ that was used in the construction of $\mathbf{C}$.

5.3.2   Linear Predictability and Sensitivity Map Estimation

In what follows, we assume that we have access to R filters $h_r[\mathbf{n}, q]$ that each satisfy Eq. (5.5), which can, e.g., be obtained from the nullspace of a structured calibration matrix as described in the previous subsection.

Observe that the linear prediction relationship from Eq. (5.5) can be rewritten in the image domain as

$$\sum_{q=1}^{Q} h_r(\mathbf{x}, q)\rho_q(\mathbf{x}) \approx 0 \text{ for } \forall \mathbf{x} \in \Gamma, \tag{5.15}$$

where

$$h_r(\mathbf{x}, q) \triangleq \sum_{\mathbf{n} \in \Lambda} h_r[\mathbf{n}, q]e^{i2\pi\mathbf{n}^T\mathbf{x}}. \tag{5.16}$$

Using the relationship between $\rho(\mathbf{x})$ and $\rho_q(\mathbf{x})$, Eq. (5.15) can be further simplified to

$$\rho(\mathbf{x}) \sum_{q=1}^{Q} h_r(\mathbf{x}, q)c_q(\mathbf{x}) \approx 0 \text{ for } \forall \mathbf{x} \in \Gamma. \tag{5.17}$$

Since $|\rho(\mathbf{x})| > 0$ for $\forall \mathbf{x} \in \Omega$ (by the definition of Ω), this implies that the summation

$$\sum_{q=1}^{Q} h_r(\mathbf{x}, q) c_q(\mathbf{x}) \approx 0 \text{ for } \forall \mathbf{x} \in \Omega, \tag{5.18}$$

demonstrating that the sensitivity maps $c_q(\mathbf{x})$ and the spatial-domain representation of the filters $h_r(\mathbf{x}, q)$ are strongly related to each other, and that must hold for all $\mathbf{x} \in \Omega$. On the other hand, the summation in Eq. (5.18) could result in arbitrary nonzero values for $\mathbf{x} \in \Omega^{\complement}$, since the fact that $\rho(\mathbf{x}) = 0$ at these spatial locations will still ensure that Eq. (5.17) is satisfied regardless of the properties of $h_r(\mathbf{x}, q)$ and $c_q(\mathbf{x})$.

Since we have R filters $h_r[\mathbf{n}, q]$ that each satisfy Eq. (5.5), the relationship from Eq. (5.18) can be expressed simultaneously for all R filters in matrix form as

$$\mathbf{H}(\mathbf{x})\mathbf{c}(\mathbf{x}) \approx \mathbf{0} \text{ for } \forall \mathbf{x} \in \Omega, \tag{5.19}$$

where $\mathbf{H}(\mathbf{x}) \in \mathbb{C}^{R \times Q}$ has entries $[\mathbf{H}(\mathbf{x})]_{rq} = h_r(\mathbf{x}, q)$, and $\mathbf{c}(\mathbf{x}) \in \mathbb{C}^Q$ is the vector of $c_q(\mathbf{x})$ values. Importantly, this demonstrates that the sensitivity map values at location $\mathbf{x} \in \Omega$ will be one of the approximate nullspace vectors of the corresponding matrix $\mathbf{H}(\mathbf{x})$.

Ideally, the matrix $\mathbf{H}(\mathbf{x})$ will only have a single approximate nullspace vector for each spatial position $\mathbf{x} \in \Omega$, since that would allow us to estimate $c_q(\mathbf{x})$ for $\mathbf{x} \in \Omega$ by calculating a basis for the approximate nullspace of $\mathbf{H}(\mathbf{x})$ (which can, e.g., be done using the SVD). Note that singular vectors associated with distinct singular values are always unique up to scaling ambiguities, which would allow us to obtain unique sensitivity map values $c_q(\mathbf{x})$ up to the same type of scaling ambiguities that are inherent to sensitivity map estimation (see previous discussion in Sec. 5.2).

Importantly, the ability to estimate unique sensitivity maps $c_q(\mathbf{x})$ (up to scaling ambiguities) for $\mathbf{x} \in \Omega$ is predicated on the matrix $\mathbf{H}(\mathbf{x})$ being approximately rank $Q - 1$. A necessary condition for this to be true is that $R > Q$. However, having $R > Q$ is not sufficient to ensure that $\mathbf{H}(\mathbf{x})$ possesses the desired rank characteristics. For example, filters associated with limited image support (Eq. (5.8)) will satisfy $h_r(\mathbf{x}, q) \approx 0$ for $\forall \mathbf{x} \in \Omega$, and will therefore not contribute significantly to

the rank of $\mathbf{H}(\mathbf{x})$. In practice however, we have observed empirically[4] that if we use the entire collection of filters from the approximate nullspace of $\mathbf{C}$, then the $\mathbf{H}(\mathbf{x})$ matrices always have an approximate rank of $Q-1$ for $\mathbf{x} \in \Omega$ under the assumptions made in this paper.[5] Interestingly, when $\mathbf{x} \in \Omega^{\complement}$ (i.e., the situation where the sensitivity maps are unidentifiable) and we use the entire collection of filters from the approximate nullspace of $\mathbf{C}$, we empirically observe that the matrix $\mathbf{H}(\mathbf{x})$ always has rank Q rather than $Q-1$, which provides a simple mechanism for identifying both Ω and $\Omega^{\complement}$.

In what follows, we will use $\mathbf{G}(\mathbf{x}) \in \mathbb{C}^{Q \times Q}$ to denote the positive semidefinite matrix formed from $\mathbf{H}(\mathbf{x})$ as $\mathbf{G}(\mathbf{x}) \triangleq (\mathbf{H}(\mathbf{x}))^{H}\mathbf{H}(\mathbf{x})$. The $\mathbf{G}(\mathbf{x})$ matrix is useful because it can be much smaller than $\mathbf{H}(\mathbf{x})$ in the common case that $R \gg Q$, while the two matrices share many characteristics. Specifically, the singular values and right singular vectors of $\mathbf{H}(\mathbf{x})$ have a one-to-one correspondence with the sorted squared eigenvalues and eigenvectors of $\mathbf{G}(\mathbf{x})$.

Empirical demonstrations of the nullspace characteristics of $\mathbf{H}(\mathbf{x})$ are presented in the next subsection.

5.3.3   Empirical Demonstration of Nullspace Characteristics

Throughout this paper, we will use the following three multichannel MRI datasets to illustrate theoretical principles and to validate our proposed methods:

- **Brain MPRAGE.** We used an MPRAGE sequence on a 3T scanner to acquire a T1-weighted human brain image using a 32-channel receiver array, as depicted in Fig. 5.1(a).

- **Brain TSE.** We used a turbo spin-echo sequence on a 3T scanner to acquire a T2-weighted human brain image using a 32-channel receiver array, as depicted in Fig. 5.1(b).

---

[4]Our experience is based on identifying the approximate nullspace by applying a heuristic thresholding rule to the SVD. Specifically, if σ_n denotes the nth singular value of $\mathbf{C}$ (with singular values indexed in order from largest to smallest), then our thresholding rule includes the nth right singular vector within the approximate nullspace whenever $\sigma_n < 0.05\sigma_1$. The threshold we chose of $0.05\sigma_1$ was determined heuristically and is unlikely to be optimal, although worked well for the datasets considered in this work.

[5]Note that different rank characteristics may be observed under different assumptions. For example, if the image support Ω is larger than the FOV Γ and if aliasing occurs within the FOV, then we empirically observe multiple sets of nullspace vectors at the spatial locations where the image support aliases onto itself. This is expected since such aliased voxels will be a linear combination of multiple sets of sensitivity-weighted images with distinct sensitivity maps. This enables the computation of multiple sensitivity maps, similar to what has been described previously in work on ESPIRiT [8], although we do not belabor this point as this scenario is not our primary interest.

(a) Brain MPRAGE   (b) Brain TSE   (c) Knee TSE
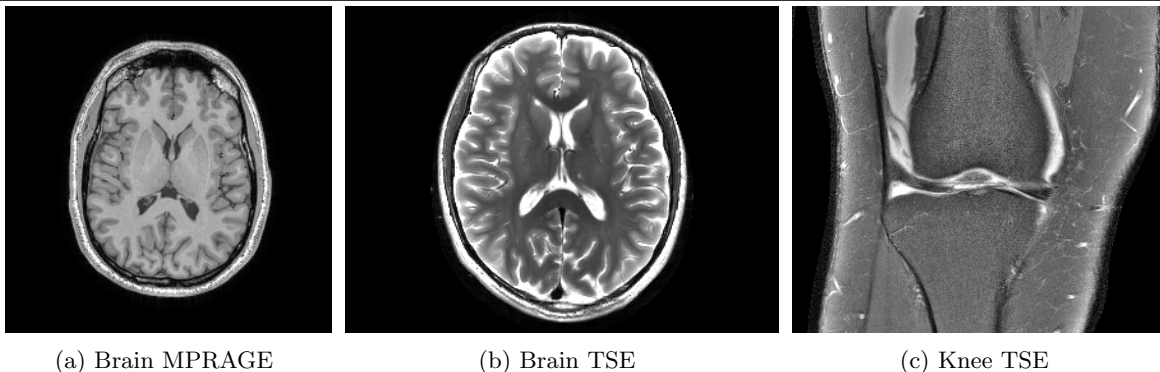
Figure 5.1: Depiction of the three datasets we use for illustration and validation.

(a) Brain MPRAGE   (b) Brain TSE   (c) Knee TSE
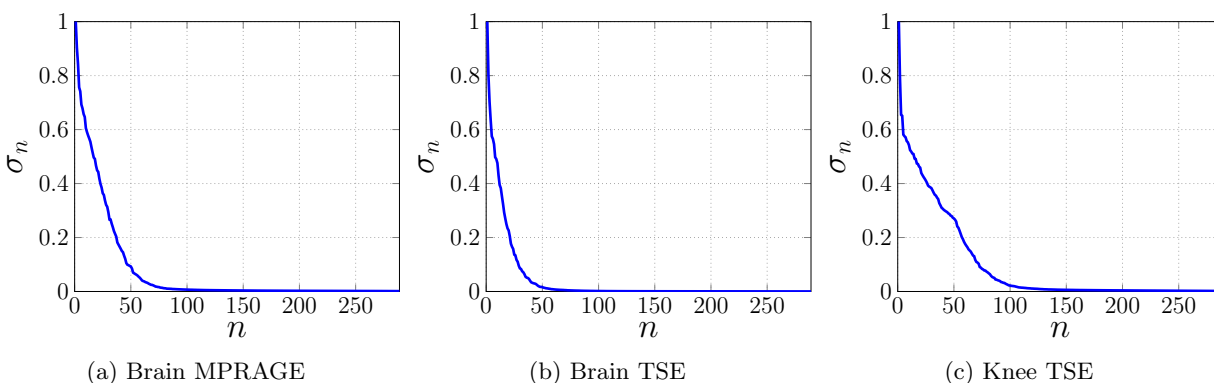
Figure 5.2: Normalized singular value curves for the $\mathbf{C}$ matrices corresponding to the three datasets from Fig. 5.1.

- **Knee TSE.** We used a 15-channel knee dataset from the fastMRI database [94], as depicted in Fig. 5.1(c). This image was acquired using a fat-suppressed turbo spin-echo sequence on a 3T scanner, producing a proton-density weighted image.

To illustrate the low-rank characteristics of the $\mathbf{C}$ matrix as described in Sec. 5.3.1, Fig. 5.2 plots the singular values of $\mathbf{C}$ for all three datasets. In each case, the $\mathbf{C}$ matrix was calculated assuming a 7×7 rectangular FIR filter support (i.e., $\Lambda = \{\mathbf{n} \in \mathbb{Z}^2 : \|\mathbf{n}\|_\infty \leq 3\}$), and the central 24×24 k-space samples from the Nyquist grid. As expected from theory, the singular values decay rapidly in each case, and all three $\mathbf{C}$ matrices are approximately low-rank with a substantial approximate nullspace.
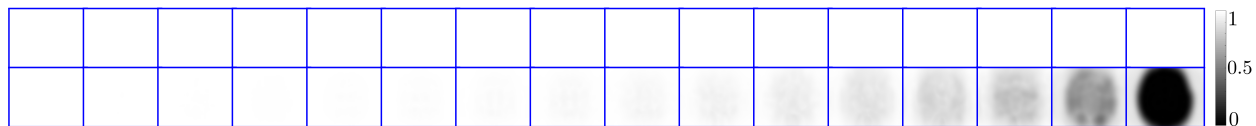
Figure 5.3: Spatial maps of the 32 eigenvalues of the $\mathbf{G}(\mathbf{x})$ matrices for the Brain TSE dataset. The eigenvalues have been normalized by $|\Lambda|$ and sorted from largest to smallest, with the largest in the top right and the smallest in the bottom right.
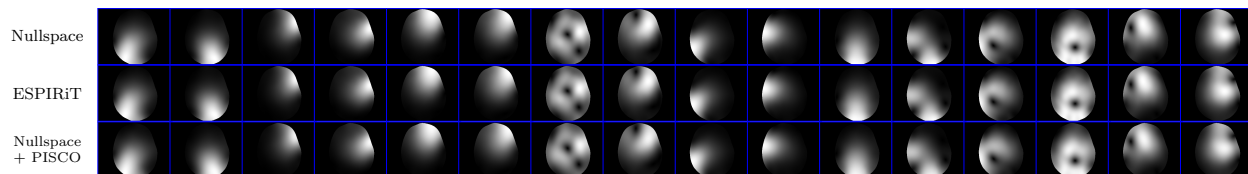
Figure 5.4: Magnitude of sensitivity maps for 16 representative coils estimated using the nullspace-based method (first row), ESPIRiT (second row), and the nullspace-based method combined with the PISCO techniques (third row).

To illustrate the characteristics of the $\mathbf{H}(\mathbf{x})$ and $\mathbf{G}(\mathbf{x})$ matrices described in Sec. 5.3.2, Fig. 5.3 shows spatial maps of the $Q = 32$ eigenvalues of $\mathbf{G}(\mathbf{x})$ for the Brain TSE dataset. As expected, the smallest eigenvalue is approximately zero within the support of the image (i.e., $\mathbf{H}(\mathbf{x})$ and $\mathbf{G}(\mathbf{x})$ are approximately rank $= Q - 1$ for $\mathbf{x} \in \Omega$), while the matrices have full rank (i.e., rank $= Q$) outside the image support.

5.3.4   Naive Algorithm for Nullspace-Based Sensitivity Map Estimation

Based on the theoretical framework described in the previous subsections, a naive nullspace-based algorithm for calulating sensitivity maps would proceed as follows:

1. Choose an FIR filter support set Λ, and construct the corresponding $\mathbf{C}$ matrix from calibration data.

2. Calculate a set of R approximate nullspace vectors of the $\mathbf{C}$ matrix using the SVD. (In our implementation, we will compute the SVD of $\mathbf{C}^H\mathbf{C}$ rather than the SVD of $\mathbf{C}$, since the two matrices have the same nullspace while $\mathbf{C}^H\mathbf{C}$ is generally much smaller.)

3. Using the nullspace vectors from Step (2), calculate the $\mathbf{H}(\mathbf{x})$ matrices for each spatial position $\mathbf{x}$, which can be simplified into the $\mathbf{G}(\mathbf{x})$ matrices as described in the previous subsection.

4. Calculate the eigenvalue decomposition of $\mathbf{G}(\mathbf{x})$ for each spatial location $\mathbf{x}$, and set $c_q(\mathbf{x})$ equal to the qth entry of the eigenvector associated with the smallest eigenvalue of $\mathbf{G}(\mathbf{x})$.

5. (Optional) If desired, choose a scaling function $\alpha(\mathbf{x})$ that has useful properties, and use it to normalize the sensitivity maps.

6. (Optional) If desired, apply thresholding to the spatial map of the smallest eigenvalue of $\mathbf{G}(\mathbf{x})$ to obtain a binary mask for the image support Ω.

As we will demonstrate in the sequel, this naive algorithm works well although can require substantial computation and memory usage, while a modified algorithm that uses our proposed PISCO approach leads to substantial computational accelerations and memory savings.

5.3.5 Memory Requirements of the Nullspace Algorithm and the Direct Calculation of $\mathbf{G}(\mathbf{x})$

One of the limiting factors in the previously described nullspace-based algorithm occurs in Step (3), where it is necessary to calculate the R spatial-domain multichannel filters $h_r(\mathbf{x}, q)$ in order to construct the $\mathbf{H}(\mathbf{x})$ matrices for each spatial location $\mathbf{x}$. This can be computationally demanding in terms of memory usage when R is big, which is commonly observed in practice. To overcome this limitation, we propose a computational procedure to directly calculate the matrices $\mathbf{G}(\mathbf{x})$ for each spatial location $\mathbf{x}$ without constructing $\mathbf{H}(\mathbf{x})$ first. This procedure correspond to the first of the proposed PISCO techniques. Before describing how this task is carried out, we introduce the matrix $\mathbf{V} \in \mathbb{C}^{Q|\Lambda| \times R}$ which columns correspond to the approximate nullspace vectors of $\mathbf{C}$ or, equivalently, to the Fourier-domain multichannel filters satisfying Eq. (5.5). Specifically, $\mathbf{V}$ has the form

$$\mathbf{V} \triangleq \begin{bmatrix} \mathbf{h}_1^{(1)} & \mathbf{h}_2^{(1)} & \dots & \mathbf{h}_R^{(1)} \\ \mathbf{h}_1^{(2)} & \mathbf{h}_2^{(2)} & \dots & \mathbf{h}_R^{(2)} \\ \vdots & \ddots & \dots & \vdots \\ \mathbf{h}_1^{(Q)} & \mathbf{h}_2^{(Q)} & \dots & \mathbf{h}_R^{(Q)} \end{bmatrix}, \tag{5.20}$$

where $\mathbf{h}_r^{(q)} \in \mathbb{C}^{|\Lambda|}$ is the vectorized version of $h_r[\mathbf{n}, q], \mathbf{n} \in \Lambda$. Now we describe our efficient procedure to calculate $\mathbf{G}(\mathbf{x})$. The entries of this matrix are given by $[\mathbf{G}(\mathbf{x})]_{pq} = \sum_{r=1}^R \bar{h}_r(\mathbf{x}, p) h_r(\mathbf{x}, q)$, where $\bar{h}_r$ corresponds to the complex conjugate version of h_r. This expression can also be calculated using the vectorized Fourier-domain filters as follows

$$[\mathbf{G}(\mathbf{x})]_{pq} = \mathbf{f}^H(\mathbf{x}) \left(\sum_{r=1}^R \mathbf{h}_r^{(q)} (\mathbf{h}_r^{(p)})^H \right) \mathbf{f}(\mathbf{x}), \tag{5.21}$$

where $\mathbf{f}(\mathbf{x}) \triangleq [e^{-i2\pi \mathbf{n}_1^T \mathbf{x}} \ e^{-i2\pi \mathbf{n}_2^T \mathbf{x}} \ \cdots \ e^{-i2\pi \mathbf{n}_{|\Lambda|}^T \mathbf{x}}]^T$ is the vector of complex sinusoids associated to the frequency components $\{\mathbf{n}_s\}_{s=1}^{|\Lambda|}$ within Λ. Equation (5.21) is important since the matrix $\mathbf{W}_{qp} \triangleq \sum_{r=1}^R \mathbf{h}_r^{(q)} (\mathbf{h}_r^{(p)})^H \in \mathbb{C}^{|\Lambda| \times |\Lambda|}$ can be easily extracted from the projection matrix $\mathbf{V}\mathbf{V}^H$, which is not expensive to calculate. Then, if we denote the sth column of $\mathbf{W}_{qp}$ by $\mathbf{w}_{qp}^{(s)}$, we have that $\mathbf{W}_{qp} = \sum_{s=1}^{|\Lambda|} \mathbf{w}_{qp}^{(s)} \mathbf{e}_s^H$, where $\mathbf{e}_s$ corresponds to the canonical vector with only zeros except for the sth entry which is equal to one. This allows us to write

$$\begin{aligned}[\mathbf{G}(\mathbf{x})]_{pq} &= \mathbf{f}^H(\mathbf{x}) \left(\sum_{s=1}^{|\Lambda|} \mathbf{w}_{qp}^{(s)} \mathbf{e}_s^H \right) \mathbf{f}(\mathbf{x}) \\ &= \sum_{s=1}^{|\Lambda|} w_q^{(s)}(\mathbf{x}, p) e^{-i2\pi \mathbf{n}_s^T \mathbf{x}}, \end{aligned} \tag{5.22}$$

where $w_q^{(s)}(\mathbf{x}, p)$ corresponds to the spatial-domain filter associated to $\mathbf{w}_{qp}^{(s)}$. Alternatively, we can perform the previous calculation in the Fourier domain. Denoting by $w_q^{(s)}[\mathbf{n}, p]$ the Fourier representation of $w_q^{(s)}(\mathbf{x}, p)$ (which nonzero components are the entries of $\mathbf{w}_{qp}^{(s)}$) and by $\mathcal{F}^{-1}$ the inverse Fourier transform operator, we have that $[\mathbf{G}(\mathbf{x})]_{pq}$ corresponds to $\mathcal{F}^{-1} \left(\sum_{s=1}^{|\Lambda|} w_q^{(s)}[\mathbf{n} + \mathbf{n}_s, p] \right)$ at location $\mathbf{x}$. Therefore, in order to calculate the matrices $\mathbf{G}(\mathbf{x})$ for each location $\mathbf{x}$, we only need to store Q spatial-domain multichannel filters (i.e., $w_q^{(s)}(\mathbf{x}, p)$ for $\forall q \in \{1, \ldots, Q\}$), which is usually substantially smaller than R. It should also be noted that a more efficient calculation of $\mathbf{V}\mathbf{V}^H$ is possible using the row-space vectors of $\mathbf{C}$, which are usually considerably less than the approximate nullspace vectors. In fact, we have that $\mathbf{V}\mathbf{V}^H = \mathbf{I}_{Q|\Lambda|} - \mathbf{U}\mathbf{U}^H$, where $\mathbf{U}$ is the matrix which columns correspond to the row-space vectors, and $\mathbf{I}_{Q|\Lambda|}$ corresponds to the $Q|\Lambda| \times Q|\Lambda|$ identity

matrix. In the following we consider this approach to calculate the projection matrix $\mathbf{V}\mathbf{V}^H$, given that calculating $\mathbf{U}\mathbf{U}^H$ is faster than calculating $\mathbf{V}\mathbf{V}^H$.

In the following experiments the proposed nullspace-based algorithm is implemented considering the previously described approach to calculate the $\mathbf{G}(\mathbf{x})$ matrices in Step (3). Otherwise, the memory requirements become restrictive making the nullspace-based algorithm unfeasible.

5.3.6   Comparisons of the Nullspace Algorithm and ESPIRiT

Before moving on to describe the rest of our novel PISCO computational acceleration techniques, we will first spend a little time to describe how PISCO compares to the popular ESPIRiT approach for sensitivity map estimation [8], which is also subspace-based. While the two approaches are derived using very different arguments, the resulting approaches end up having quite similar structure. Specifically, in our nullspace-based approach, we use the nullspace vectors of the $\mathbf{C}$ matrix to identify matrices $\mathbf{G}(\mathbf{x})$ that satisfy $\mathbf{G}(\mathbf{x})\mathbf{c}(\mathbf{x}) \approx \mathbf{0}$, where $\mathbf{c}(\mathbf{x}) \in \mathbb{C}^Q$ is the vector of $c_q(\mathbf{x})$ sensitivity map values. This causes us to estimate sensitivity maps using the eigenvector associated with the smallest eigenvalue of $\mathbf{G}(\mathbf{x})$. In contrast, ESPIRiT extracts the rowspace vectors of the $\mathbf{C}$ matrix, and uses these to construct matrices $\mathbf{B}(\mathbf{x}) \in \mathbb{C}^{Q \times Q}$ that are expected to satisfy $\mathbf{B}(\mathbf{x})\mathbf{c}(\mathbf{x}) \approx \mathbf{c}(\mathbf{x})$. As a result, the sensitivity maps in ESPIRiT can be estimated using the eigenvector associated with the largest eigenvalue of $\mathbf{B}(\mathbf{x})$. Interestingly, it can be derived (with a substantial amount of tedious-but-straightforward derivations that are too involved to reproduce here) that the $\mathbf{B}(\mathbf{x})$ matrix used by ESPIRiT will satisfy

$$\mathbf{B}(\mathbf{x}) = \mathbf{I}_Q - \frac{\mathbf{G}(\mathbf{x})}{|\Lambda|}. \tag{5.23}$$

In addition to theoretical similarities between ESPIRiT and the nullspace-based approach, the two approaches also empirically produce sensitivity map estimates that are nearly identical to one another. Results are illustrated in Fig. 5.4, where we show representative sensitivity maps estimated from the Brain TSE dataset. For both methods we chose an FIR filter support Λ of dimensions 7×7 and a calibration data set of dimensions 24×24 to calculate $\mathbf{C}$. We calculated the singular
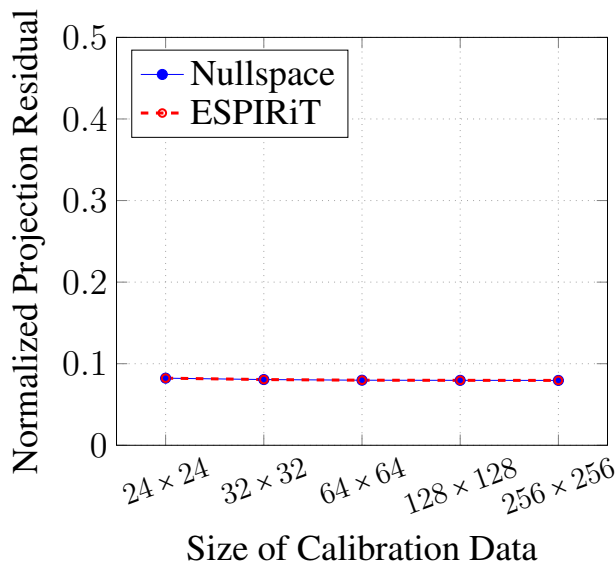
Figure 5.5: Quantitative comparison between the nullspace-based approach and ESPIRiT for the Brain TSE data in terms of (a) accuracy (as measured through the normalized projection residual) and (b) speed (as measured through the total computation time).

vectors of $\mathbf{C}^H\mathbf{C}$ corresponding to singular values which were greater than the 5% of the greatest singular value. These singular vectors are supposed to approximate a basis for the rowspace of $\mathbf{C}^H\mathbf{C}$, and can be used to calculate the matrices $\mathbf{B}(\mathbf{x})$ for each spatial location. Analogously, we calculated the singular vectors of $\mathbf{C}^H\mathbf{C}$ corresponding to singular values smaller than the 5% of the greatest singular value (i.e., approximate nullspace vectors), and use them to directly calculate the matrices $\mathbf{G}(\mathbf{x})$ (i.e., using the method described in Sect. 5.3.5). Then, we estimated the sensitivity maps at each spatial location by calculating the first eigenvector of $\mathbf{B}(\mathbf{x})$ in ESPIRiT, and the last eigenvector of $\mathbf{G}(\mathbf{x})$ in the nullspace-based algorithm. Next, the estimated sensitivity maps for both methods were normalized in magnitude using a sum-of-squares approach, and referenced in phase with respect to the phase of the sensitivity map estimate corresponding to the first channel. Finally, both sets of sensitivity maps were masked by identifying locations where the smallest eigenvalue of $\mathbf{G}(\mathbf{x})$ was smaller than the 8% of the greatest eigenvalue. From the results in Fig. 5.4 it can be observed that the visual differences between the two approaches are negligible. Figure 5.4 also shows that the results are negligibly different from those obtained when combining the nullspace-based approach with PISCO (which leads to dramatic improvements in computation speed and

Table 5.1: Normalized Projection Residual with 32×32 calibration data.

Dataset	ESPIRiT	ESPIRiT +PISCO	Nullspace	Nullspace +PISCO
Brain MPRAGE	0.136	0.140	0.136	0.140
Brain TSE	0.081	0.087	0.081	0.087
Knee TSE	0.292	0.293	0.292	0.293

Table 5.2: Computation Time (secs.) with 32×32 calibration data.

Dataset	ESPIRiT	ESPIRiT +PISCO	Nullspace	Nullspace +PISCO
Brain MPRAGE	44.3	1.0	NA	1.0
Brain TSE	99.2	1.1	NA	1.1
Knee TSE	66.2	0.3	NA	0.3

memory usage), although we will save that discussion for later in the paper after we have discussed the remaining details of PISCO.

To evaluate the quality of a sensitivity map estimate in a more quantitative way, we have also computed a "normalized projection residual" to quantify how well the estimated sensitivity maps match the fully-sampled data. Specifically, we first take the fully-sampled multichannel data and project it onto the subspace spanned by the sensitivity maps. If the sensitivity map estimation procedure worked well, then the projection onto the sensitivity maps should capture all of the true signal, with only noise remaining in the residual after projection [8]. Our "normalized projection residual" metric is obtained by taking the ℓ_2-norm of the difference between the original multichannel data and the projected data, and normalizing by the ℓ_2-norm of the original multichannel data. A smaller value of this metric is indicative of better sensitivity map estimation performance. Figure 5.5(a) shows plots of this metric for the Brain TSE data, for both ESPIRiT and the nullspace-based approach, using a 7×7 FIR filter support Λ and different calibration data sizes. As can be seen, both methods have very similar residual values. Table 5.1 also shows the normalized projection residual for all three datasets for both ESPIRiT and the nullspace-based approach, this time using 32×32 calibration data while still using a 7×7 FIR filter support. As before, the differences between ESPIRiT and the nullspace-based approach are negligible. (This table also shows PISCO results, although we will defer a detailed discussion of these results until after we have described PISCO).

Table 5.3: Memory Usage (GB) with 32×32 calibration data.

Dataset	ESPIRiT	ESPIRiT +PISCO	Nullspace	Nullspace +PISCO
Brain MPRAGE	1.01	0.55	25.33	0.55
Brain TSE	1.71	1.44	68.20	1.44
Knee TSE	1.91	0.43	18.89	0.43

To evaluate the speed of different algorithmic approaches, we measured the computation time employed by both approaches using a MATLAB in-house implementation on a computer with an Intel Xeon E5-1603 2.8GHz quad core CPU and 32GB of RAM. Table 5.2 shows the computation time for all three datasets for ESPIRiT, the nullspace-based approach, and the combination of these two algorithms with the PISCO techniques. We have not included the time corresponding to the nullspace-based approach since its implementation makes use of the procedure to directly calculate the matrices $\mathbf{G}(\mathbf{x})$ described in Sect. 5.3.5. Therefore, comparing its computation time against ESPIRiT would be an unfair comparison given that this type of techniques is not used in the original ESPIRiT description. However, as we will discuss later in the paper (and as is already evident from Table 5.2), there is no computational speed difference between ESPIRiT and the nullspace-based method when these are combined with all the PISCO techniques.

We also evaluated the memory usage for the nullspace-based algorithm and ESPIRiT, by measuring the RAM memory associated to the most expensive variables needed in each algorithm. These correspond to the spatial-domain multichannel filters needed to naively calculate the matrices $\mathbf{G}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ in the nullspace-based algorithm and ESPIRiT, respectively[6]. We also added the RAM memory associated to the convolutional matrix $\mathbf{C}$, which is also considerable. These results are reported in Table 5.3 from where it can be observed how massive the memory requirements are for the nullspace-based algorithm in comparison to ESPIRiT (if a naive implementation is used). However, we can also observe that the memory associated to the considered variables is reduced drastically for both algorithms when combined with the PISCO techniques, making them similar in terms of memory usage. The detailed explanation of these memory savings will be given after our description of the rest of the PISCO techniques in the next section.

---

[6]It should be noted that in this assessment we are not considering the memory savings due to the procedure described in Sect. 5.3.5. This type of memory savings are only considered in the reported PISCO results.

Overall, these results demonstrate strong theoretical, empirical, and computational similarities between ESPIRiT and the nullspace-based algorithm. Of course, another contribution of this paper is to introduce PISCO, which will enable substantial improvements in computational complexity for the nullspace method, ESPIRiT, and a number of related techniques. The remaining details of PISCO and a numerical evaluation of its performance will be presented in the next section.

5.4   Proposed Computational Methods

Now we describe the remaining of our computational contributions included in PISCO, noticing that one of them has already been described in Sect. 5.3.5. As we explain next, each of these computational methods provides important improvements either in memory usage and/or computational complexity when incorporated into the naive algorithm described in Section 5.3.4. Notably, these efficiency improvements do not compromise estimation quality and, given the structural similarities between ESPIRiT and the nullspace-based method, it is also straightforward to incorporate PISCO into ESPIRiT. In the following, besides providing a theoretical description of the computational methods in PISCO, we empirically assess efficiency improvements when including these methods individually into the nullspace-based algorithm. In all the cases we use the Brain TSE dataset and consider different sizes for the calibration data as in Fig. 5.5. It should be mentioned that one of the PISCO techniques was introduced in Sect. 5.3.5, which makes the nullspace-based algorithm feasible in terms of memory usage. In the following experiments, the reported results regarding accuracy and computation time of our baseline nullspace-based method were obtained using the procedure described in Sect. 5.3.5 for the calculation of the $\mathbf{G}(\mathbf{x})$ matrices. However, unless it is specified, the reported results regarding memory usage were not necessarily obtained using the procedure described in Sect. 5.3.5.

5.4.1   Ellipsoidal versus Rectangular Convolution Kernels

In previous work we have shown that considering an ellipsoidal support for the FIR filters in the construction of $\mathbf{C}$ offers significant efficiency improvements in comparison to using a rectangular
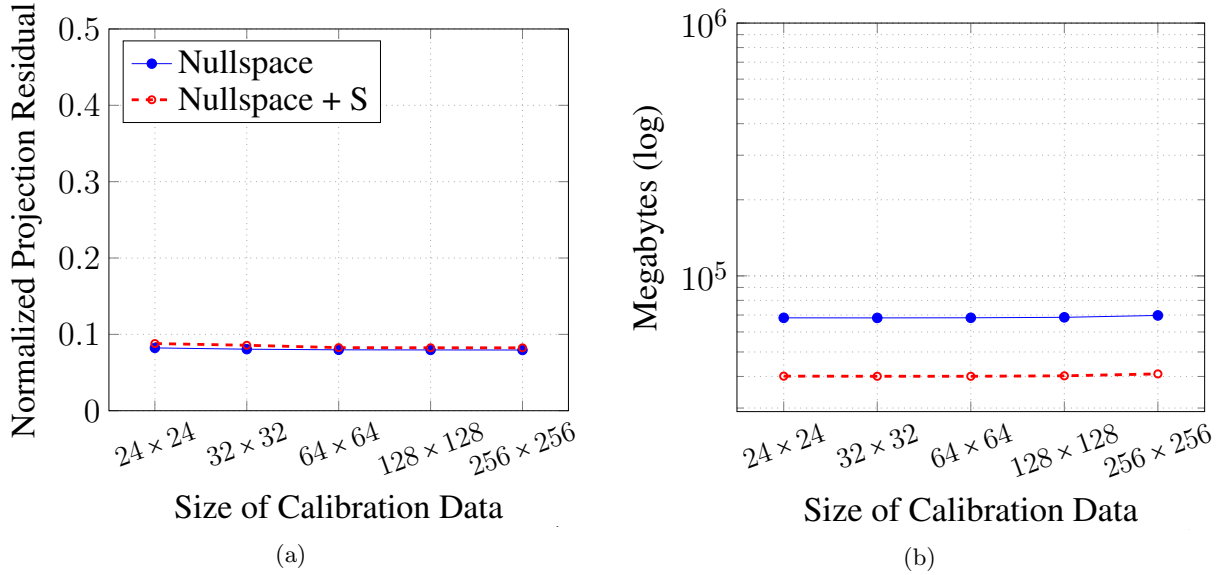
Figure 5.6: Assessment of the nullspace-based algorithm using filters with a rectangular (Nullspace) and an ellipsoidal support (Nullspace + S).

support [16]. Specifically, if we consider the ellipsoidal neighborhood $\Lambda_e = \{\mathbf{n} \in \mathbb{Z}^2 : \|\mathbf{n}\|_{\ell_2} \leq \tau\}$ in Eq. (5.7), instead of the rectangular neighborhood $\Lambda_r = \{\mathbf{n} \in \mathbb{Z}^2 : \|\mathbf{n}\|_{\infty} \leq \tau\}$, the size of $\mathbf{C}$ is considerably reduced given that $|\Lambda_e| < |\Lambda_r|$. Then, memory usage and computation time should be reduced when using the ellipsoidal support instead of the rectangular support, given that $\mathbf{C}$ has smaller dimensions and therefore it is faster to construct. In Fig. 5.6 we assess the nullspace-based algorithm when using both types of support. In order to have a fair comparison we considered an ellipsoidal support Λ_e with $\tau = 3$, and the rectangular support Λ_r with size $(2\tau + 1) \times (2\tau + 1)$. We can see from Fig. 5.6(a) that both types of support obtained similar accuracy, however, there are considerable gains in memory usage when using the ellipsoidal support as we see in Fig. 5.6(b). In this case we are reporting the RAM memory associated to the spatial-domain multichannel filters needed to calculate the $\mathbf{G}(\mathbf{x})$ matrices in addition to the RAM memory associated to $\mathbf{C}$. The computation time for both types of support was similar in this case, given that the difference in size for both supports is not big enough. We should expect bigger differences in memory usage and computational complexity for higher values of τ as shown in [16]. We have decided to only consider $\tau = 3$ for the results in Fig. 5.6 and in the following sections, given that the estimation quality varied negligibly for higher values of τ.

5.4.2 FFT-Based Calculation of Nullspace Vectors

One important step in the nullspace-based method and ESPIRiT is the calculation of nullspace and rowspace vectors of $\mathbf{C}$, respectively. As indicated in the Step (2) of the nullspace-based algorithm in Section 5.3.4, it is convenient to work with $\mathbf{C}^H\mathbf{C}$ instead of $\mathbf{C}$, since $\mathbf{C}^H\mathbf{C}$ has the same nullspace vectors and usually much smaller dimensions. Even though the small dimensions of $\mathbf{C}^H\mathbf{C}$ permit a fast calculation of the SVD, we still need to calculate $\mathbf{C}$ in prior, which is a demanding task in terms of memory usage and computation time. Following previous work [75, 118], we propose to calculate $\mathbf{C}^H\mathbf{C}$ directly without calculating $\mathbf{C}$ first. In [75] the authors proposed an FFT-based approach to calculate $\mathbf{C}^H\mathbf{C}$ directly in the single-channel case (i.e., $Q = 1$), which was extended to the multichannel case in [118]. However, the multichannel approach proposed in [118] can be demanding in terms of computation time. In the following we make the observation that it is possible to rapidly calculate $\mathbf{C}^H\mathbf{C}$ in the multichannel case by leveraging on its structure and signal processing techniques.

In the proposed approach the idea is to calculate each column of $\mathbf{C}^H\mathbf{C}$ separately. Let $\mathbf{e}_i \in \mathbb{C}^{Q|\Lambda|}$ be the ith canonical vector. Then, if we write $\mathbf{e}_i$ using a stack of Q subvectors $\{\mathbf{e}_q^{(i)}\}_{q=1}^Q$ of size $|\Lambda|$, we have that all but the $q(i) \triangleq \lfloor i/|\Lambda| \rfloor + 1$ subvector has a nonzero components. Then, the ith column of $\mathbf{C}^H\mathbf{C}$ would be given by

$$\mathbf{C}^H\mathbf{C}\mathbf{e}_i = \begin{bmatrix} \mathbf{C}_1^H \\ \vdots \\ \mathbf{C}_Q^H \end{bmatrix} \mathbf{C}_{q(i)}\mathbf{e}_{q(i)}^{(i)}, \tag{5.24}$$

which reduces to calculating Q products of the form $\mathbf{C}_p^H\mathbf{C}_q\mathbf{f}$ where $p,q \in \{1,\ldots,Q\}$ and $\mathbf{f} \in \mathbb{C}^{|\Lambda|}$ is a canonical vector. Given the convolutional structure of the matrices $\{\mathbf{C}_q\}_{q=1}^Q$, it is possible to approximate these products efficiently using the FFT [74, 75]. It can be theoretically shown that this procedure to calculate approximations of the columns of $\mathbf{C}_p^H\mathbf{C}_q$ is equivalent to performing a sliding-window that extracts a series of patches of size $|\Lambda|$ from

$$\mathcal{F}^{-1}(\overline{\mathcal{F}(\mathcal{Z}(\tilde{\rho}_p))} \odot \mathcal{F}(\mathcal{Z}(\tilde{\rho}_q))),$$
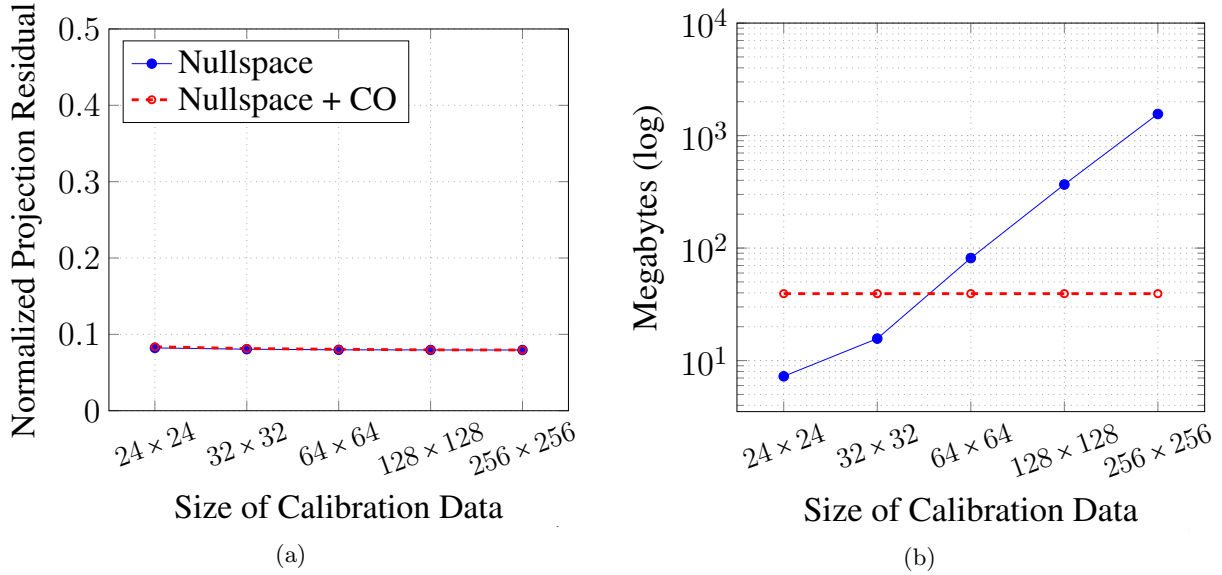
Figure 5.7: Assessment of the nullspace-based algorithm calculating the nullspace vectors in Step (2) with (Nullspace + CO) and without (Nullspace) the FFT-based approach.

where each patch is centered at a different frequency component in Λ. In the last expression $\tilde{\rho}_p \triangleq \mathcal{F}(\rho_p)$; $\mathcal{Z}(\tilde{\rho}_p)$ correspond to a zero-padded version of $\tilde{\rho}_p$; and $\odot$ denotes the element-wise product. It should be noted that this sliding-window approach has been previously proposed in [75] for the single-channel. To the best of our knowledge this is the first time that this approach has been extended to the multichannel case.

In Fig. 5.7 we show how the performance and memory usage of the nullspace-based algorithm changes when calculating $\mathbf{C}^H\mathbf{C}$ directly instead of calculating $\mathbf{C}$ first. In this case we only report the RAM memory associated to $\mathbf{C}$ and $\mathbf{C}^H\mathbf{C}$ given that all the other variables remain the same. We observe that there are dramatic improvements in memory usage for big calibration data sizes, which correspond to the cases when the dimensions of $\mathbf{C}^H\mathbf{C}$ are considerably smaller than the dimensions of $\mathbf{C}$. It should be noted that there are cases where the memory usage is worse for the proposed approach. These cases correspond to calibration data sizes where the dimensions of $\mathbf{C}^H\mathbf{C}$ are higher than the dimensions of $\mathbf{C}$. However, in these cases the memory consumption by both matrices is not considerable. Notably, the improvements provided by the proposed approach negligibly affect estimation accuracy and computation time (not reported).

5.4.3 Smoothness-Based Interpolation

In the nullspace-based algorithm proposed in Sect. 5.3.4 Step (4) corresponds to a location-wise estimation where the vector $\mathbf{c}(\mathbf{x})$ is estimated for every location $\mathbf{x}$ using the eigenvalue decomposition of $\mathbf{G}(\mathbf{x})$. This can be expensive in terms of computation time given that usually the number of locations is big. In this work we make the observation that it is possible to reduce the number of locations where the previously described process is performed, and estimate $\mathbf{c}(\mathbf{x})$ for the rest of locations using a fast interpolation procedure. In fact, sensitivity maps arise from Maxwell's equations and therefore they possess smooth spatial characteristics. Based on this principle, we propose to estimate $\mathbf{c}(\mathbf{x})$ for a low-resolution spatial grid using the eigenvalue decomposition of $\mathbf{G}(\mathbf{x})$, and then use an FFT-based interpolation procedure to obtain $\mathbf{c}(\mathbf{x})$ for the rest of locations considered at the desired spatial resolution. However, for this approach to be successful, it is necessary to obtain smooth sensitivity map estimates for the low-resolution spatial grid. This is not usually the case when conventional computational methods are used to calculate the eigenvalue decomposition of $\mathbf{G}(\mathbf{x})$, since the phase of the resulting sensitivity map estimates commonly exhibits a random structure. To solve this issue, we reference the phase of the low-resolution spatial grid sensitivity map estimates to the phase of a low-resolution image obtained from the calibration data. To obtain this low-resolution image we propose to use the low-resolution spatial grid sensitivity map estimates (with random phase) and a filtered version of the calibration data to obtain a SENSE-based reconstructed image [6]. We have observed that filtering the calibration data with a Gaussian window produces a smooth SENSE-reconstructed image with adequate phase characteristics to be used as reference. However, the Gaussian window might not be necessarily optimal, and other windows could achieve even better results. In Fig. 5.8 we show how the performance and computation time of the nullspace-based algorithm change when using the proposed FFT-based interpolation approach in Step (4). We can see that there are negligible variations in accuracy and dramatic improvements in computation time.
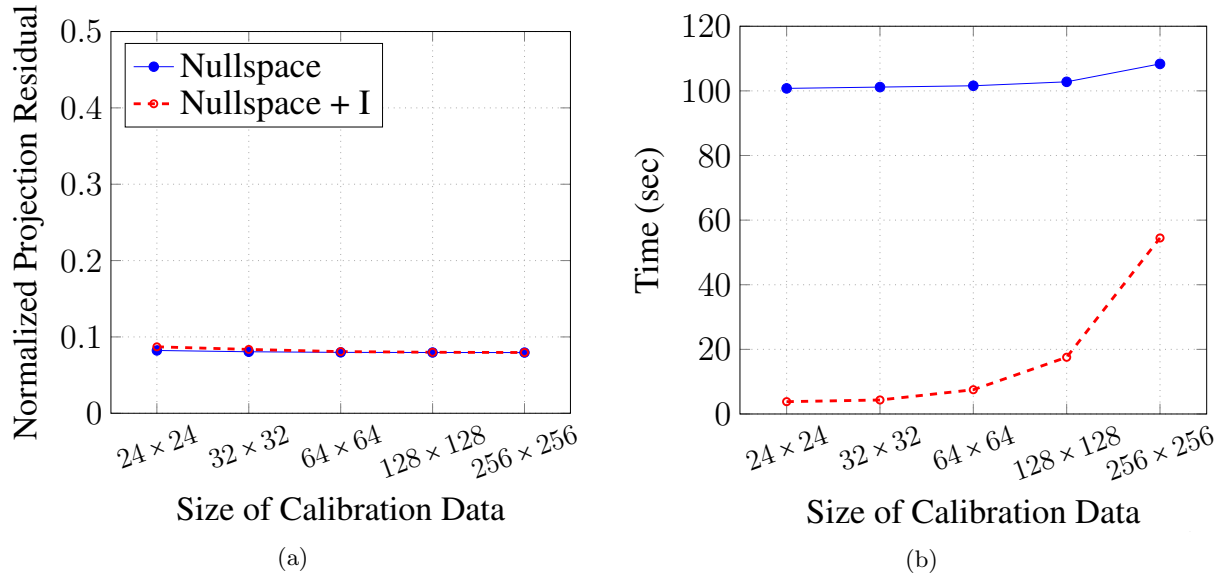
Figure 5.8: Assessment of the nullspace-based algorithm using (Nullspace + I) and not using (Nullspace) the FFT-based interpolation approach in Step (4).

5.4.4   Power Iteration

One of the most expensive steps in the nullspace-algorithm given in Sec. 5.3.4 corresponds to Step (4), where $\mathbf{c}(\mathbf{x})$ is estimated by calculating the eigenvector corresponding to the smallest eigenvalue of $\mathbf{G}(\mathbf{x})$. One possible approach to do this is to calculate the SVD of each $\mathbf{G}(\mathbf{x})$ for every location $\mathbf{x}$, and then extract the last eigenvector of each matrix. However, this process can be computationally expensive when the number of locations is big. We make the observation that, given that the nullspace of $\mathbf{G}(\mathbf{x})$ tends to be unidimensional, it would be enough to only calculate one vector in this subspace instead of the whole SVD. For this purpose, we propose to use an iterative algorithm based on Power Iteration [119]. Remarkably, the steps of this algorithm can be easily implemented in parallel for all the locations, which highly improves the computation time. Before introducing the proposed algorithm, we review how to use Power Iteration to estimate $\mathbf{c}(\mathbf{x})$ for a particular location $\mathbf{x}$, and then we show how we can perform the same procedure for multiple locations at the same time. The power iteration procedure for a particular location is summarized in Algorithm 1, where η corresponds to an upper bound for the magnitude of the eigenvalues of $\mathbf{G}(\mathbf{x})$. Even though this process is faster than SVD, repeating it for all the considered locations can still be slow. We

---

Algorithm 1 Power Iteration procedure to estimate $\mathbf{c}(\mathbf{x})$ for a particular location $\mathbf{x}$

  Input: $\mathbf{G}(\mathbf{x})$
  Output: $\mathbf{c}(\mathbf{x})$

  Initialization: Initialize $\mathbf{c}(\mathbf{x})$ randomly

  for $k \leftarrow 1$ to K **do**
    $\mathbf{c}(\mathbf{x}) \leftarrow (\eta \mathbf{I}_Q - \mathbf{G}(\mathbf{x}))\mathbf{c}(\mathbf{x})$
    $\mathbf{c}(\mathbf{x}) \leftarrow \mathbf{c}(\mathbf{x})/\|\mathbf{c}(\mathbf{x})\|_2$
  end for

---

propose a computational method to execute the steps in Algorithm 1 for all the locations at the same time. For this purpose we first define the vectors $\{\mathbf{g}_{pq}\}_{p,q=1}^Q$, where $\mathbf{g}_{pq}$ is a vector which contains the (p,q) entry of every matrix $\mathbf{G}(\mathbf{x})$. Then, we define the vectors $\{\tilde{\mathbf{g}}_{pq}\}_{p,q=1}^Q$ such that

$$
\tilde{\mathbf{g}}_{pq} = \begin{cases} \eta_{\max}\mathbf{1} - \mathbf{g}_{pq}, & \text{if } p = q \\ -\mathbf{g}_{pq}, & \text{if } p \neq q \end{cases}
$$

where $\eta_{\max}$ corresponds to an upper bound for the eigenvalues of all the matrices $\mathbf{G}(\mathbf{x})$ and $\mathbf{1}$ is a vector with all entries equal to one. Given the structure of $\mathbf{G}(\mathbf{x})$, a simple choice for this bound corresponds to $\eta_{\max} = |\Lambda|$. Then, we also define the vectors $\{\mathbf{c}_q\}_{q=1}^Q$ where $\mathbf{c}_q$ correspond to the vectorized version of $c_q(\mathbf{x})$ considering every location $\mathbf{x}$. It should be noted that estimating $\mathbf{c}(\mathbf{x})$ for every location, as indicated in Step (4) of the nullspace-based algorithm, is equivalent to estimating the vectors $\{\mathbf{c}_q\}_{q=1}^Q$. Using these definitions, we propose Algorithm 2, which is equivalent to execute the steps in Algorithm 1 for all the locations in an all-at-once fashion. Notably, Algorithm 2 can be efficiently implemented using programming languages like MATLAB or Python, where vectorization techniques can be used to rapidly calculate summations and reduce the number of for-loops. In our in-house implementation we only consider the outer for-loop associated to the variable k. The inner for-loop and summations are all performed using vectorization techniques.

A fast convergence of the previous algorithms depends on having a big difference in magnitude between the last two eigenvalues of the matrices $\mathbf{G}(\mathbf{x})$. In practice we have observed that this is usually the case (e.g., see Fig. 5.3), therefore, only a few iterations are needed for convergence. However, there are cases where the last two eigenvalues are close to each other, which corresponds

---

Algorithm 2 All-at-once PowerIteration-based Algorithm

Input: $\{\tilde{\mathbf{g}}_{pq}\}_{p,q=1}^{Q}$
Output: Estimated sensitivity map vectors $\{\mathbf{c}_q\}_{q=1}^{Q}$

Initialization: Initialize $\{\mathbf{c}_q\}_{q=1}^{Q}$ randomly

for $k \leftarrow 1$ to K **do**
 $\mathbf{c}_p \leftarrow \sum_{q=1}^{Q} \tilde{\mathbf{g}}_{pq} \odot \mathbf{c}_q$ for all $p \in \{1 \ldots, Q\}$
 Compute the vector $\boldsymbol{\alpha}$ with entries

$$[\boldsymbol{\alpha}]_m = 1/\sqrt{\sum_{p=1}^{Q} |[\mathbf{c}_p]_m|^2}$$

 $\mathbf{c}_p \leftarrow \boldsymbol{\alpha} \odot \mathbf{c}_p$ for all $p \in \{1 \ldots, Q\}$
end for

---

to situations where more than one sensitivity map is required to describe a specific location. It should also be mentioned that there are other more advanced algorithms besides Power Iteration which can also be used to compute only one eigenvector of the matrices $\mathbf{G}(\mathbf{x})$, such as the Lanczos algorithm, which theoretically has better convergence that Power Iteration. However, we have decided to use Power Iteration since its steps can be easily implemented in all-at-once fashion as we previously described, while algorithms like Lanczos include normalization steps which are more involved to implement for multiple matrices at the same time. In Fig. 5.9 we show how the performance and computation time of the nullspace-based algorithm change when using the proposed PowerIteration-based approach instead of a location-wise SVD decomposition approach in Step (4). We can see that there are negligible variations in accuracy and significant improvements in computation time. For each case we selected $K = 10$ as the number of iterations in Algorithm 2. The same number of iterations is used in the experiments in the following subsection where all the PISCO techniques are used at the same time.
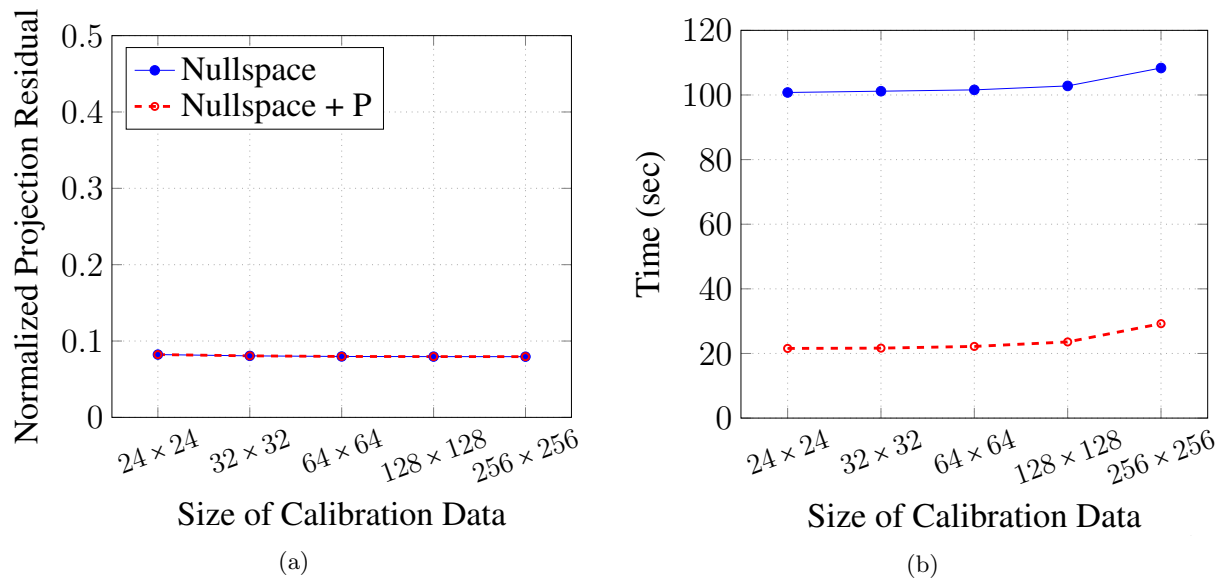
Figure 5.9: Comparison between using the location-wise SVD approach (Nullspace) and the all-at-once PowerIteration-based approach (Nullspace + P) in Step (4) of the nullspace-based algorithm.
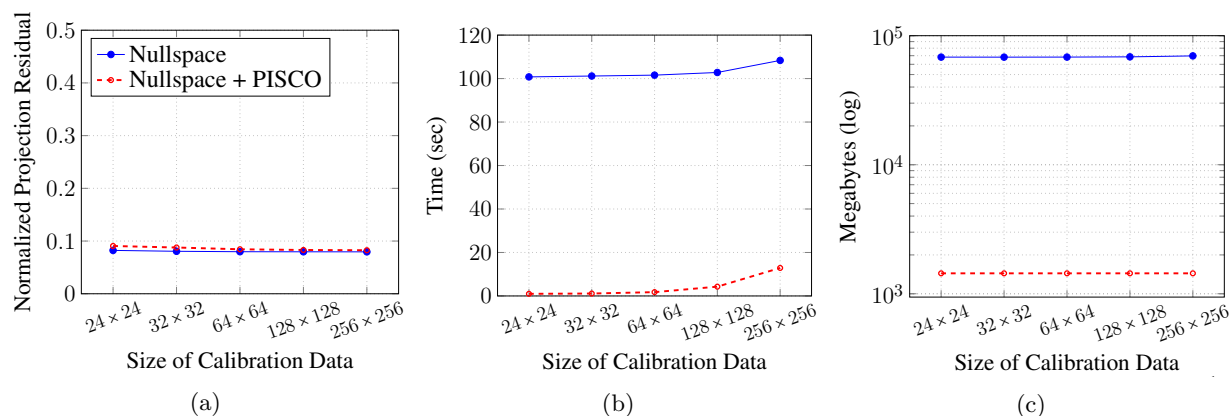
Figure 5.10: Comparison between the nullspace-based algorithm with (Nullspace + PISCO) and without (Nullspace) the PISCO techniques.

5.4.5   Combined Approach

Now that we have introduced each of the PISCO computational techniques, we empirically show the efficiency improvements that are achieved when they are all integrated at the same time to the nullspace-based estimation algorithm. In Fig. 5.10 we compare the nullspace-based algorithm with its enhanced version after integrating all the PISCO techniques. Remarkably, it can be observed that massive improvements in computation time and memory usage are achieved while maintaining estimation accuracy. For the memory usage analysis of the naive nullspace-based algorithm we have considered the RAM memory associated to the spatial-domain multichannel filters needed to calculate the $\mathbf{G}(\mathbf{x})$ matrices, in addition to the RAM memory associated to $\mathbf{C}$. As expected, this memory usage is huge for each of the considered calibration data sizes. Notably, the memory usage is reduced drastically when using PISCO. In this case, we are considering the RAM memory associated to the $\mathbf{G}(\mathbf{x})$ matrices (which are calculated efficiently using the method in Sect. 5.3.5), in addition to the RAM memory associated to $\mathbf{C}^H\mathbf{C}$ (which is calculated efficiently using the method in Sect. 5.4.2). In Fig. 5.11 we show the estimated sensitivity maps for four representative channels of the three considered datasets when using the nullspace-based algorithm and its enhanced version using PISCO. For both algorithms we used a calibration data size equal to 32×32. For the three datasets it can be observed that there are strong similarities between the sensitivity maps estimated with each of the methods. In addition, we report estimation accuracy, computation times, and memory usage in Tables I, II, and III, respectively. For the sake of completeness, we also report the results of integrating the PISCO techniques into ESPIRiT, which is straightforward given the strong similarities in structure between the nullspace-based algorithm and ESPIRiT. From these tables it can be observed that the nullspace-based algorithm and ESPIRiT provide similar results when including PISCO, which is expected given the theoretical analysis provided in Sect. 5.3.6. We obtained that in all the cases PISCO offers massive improvements in computation time and memory usage without compromising estimation accuracy.
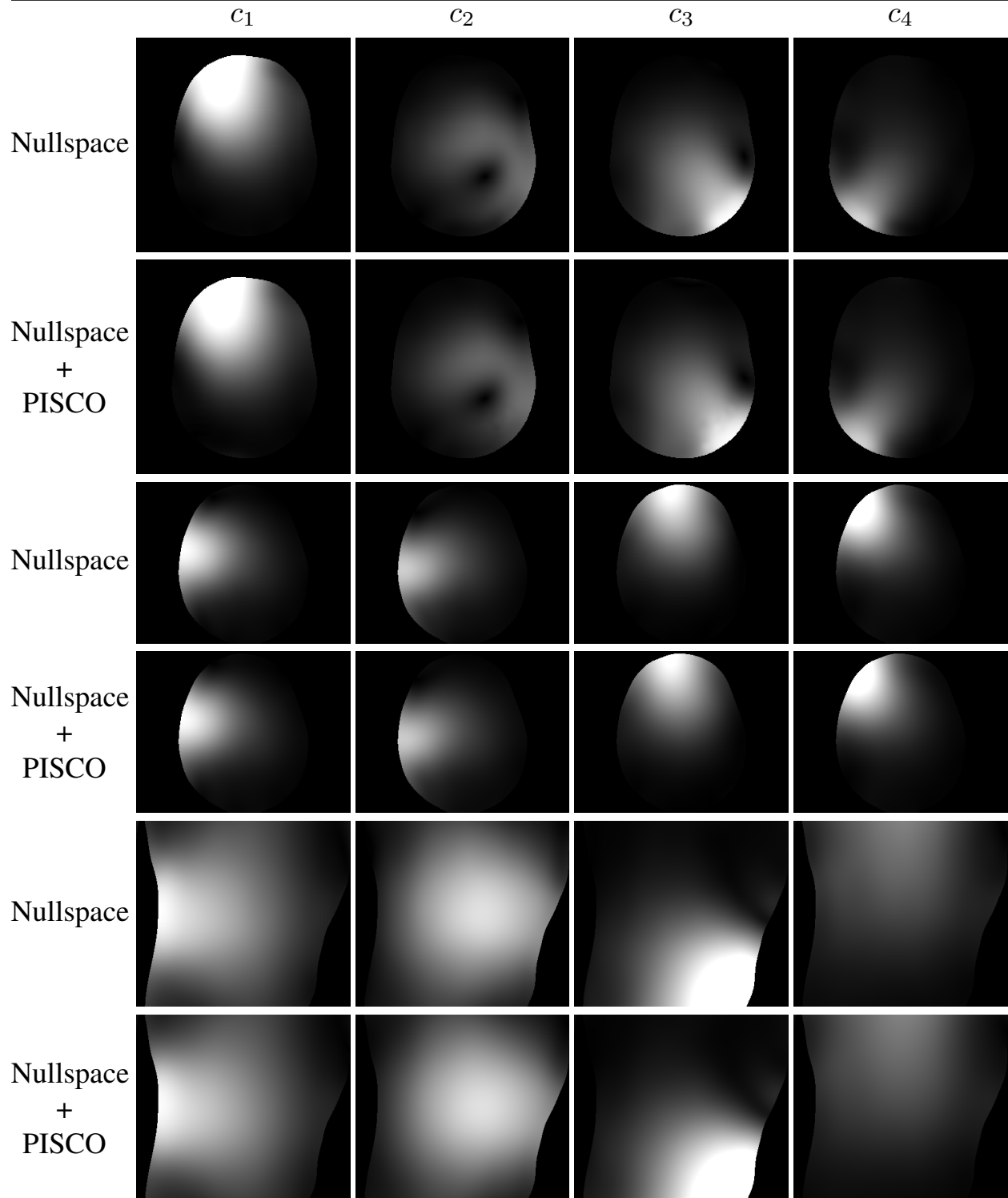
Figure 5.11: Estimated sensitivity maps using the nullspace method with and without PISCO for the three considered datasets. We show the magnitude of the sensitivity maps of four representative coils.

5.5   Discussion and Conclusions

In this work we have proposed a novel theoretical framework for the subspace-based sensitivity map estimation problem. In comparison to the existent literature, our results offer an alternative and complementary theoretical description that in our opinion provides a more intuitive mathematical explanation. Using this framework, we have proposed a sensitivity map estimation method that relies on nullspace and structured low-rank matrix modeling theory. We showed that the proposed nullspace-based method possesses strong theoretical and structural similarities with ESPIRiT. In the second part of this work we proposed a set of powerful computational methods that allow an efficient implementation of the proposed nullspace-based method, and that can be straightforwardly integrated into existent methods like ESPIRiT. Remarkably, we showed that our proposed computational techniques produce massive improvements in memory usage and computation time in comparison to naive implementations of subspace-based methods, while maintaining estimation accuracy. As an example, roughly a 100-fold acceleration can be achieved when integrating our techniques to ESPIRiT.

For the results shown in this paper we used an in-house MATLAB implementation for all the considered methods. It should be mentioned that even bigger gains in performance could be obtained with more efficient implementations and other programming languages like C. As an example, the BART software package [133] provides a C-based implementation of ESPIRiT which is expected to be much faster than MATLAB implementations. Notably, using our PISCO-based MATLAB implementation we were able to estimate sensitivity maps much faster than the C-based implementation of ESPIRiT provided in BART. For instance, using the Brain TSE dataset and a calibration data of size 24×24, our PISCO-based MATLAB implementation completed in ~ 1 second while the BART C-based implementation completed in ~ 19 seconds. It should be mentioned that both implementations obtained similar estimation accuracy. Therefore, substantial efficiency improvements should be expected if the computational methods of this paper are implemented using more powerful programming languages. Moreover, the automatic selection of the different parameters in the proposed computational methods should also contribute to improving the obtained results.

For instance, approaches based on the SURE estimator, as the one proposed in [134], could be combined with the proposed methodology.

Finally, it should be highlighted that the computational methods developed in this work can also be applied beyond the sensitivity map estimation context. For instance, many accelerated MRI reconstruction methods are based on the construction of convolutional structured matrices as the ones considered in this work (e.g., SPIRiT [7], PRUNO [9], SAKE [31], LORAKS [32], ALOHA [36]). Therefore, these methods could be benefited from the the proposed computational method described in Sect. 5.4.2, which efficiently calculates this type of matrices using FFT-based convolutions.

This page intentionally left blank

Chapter 6

Conclusion

During the last five decades MRI has been constantly evolving in terms of performance and efficiency, providing researchers and clinicians with an outstanding imaging technology capable of generating high-quality images of the living tissue in a safe and noninvasive manner. Even though MRI is an omnipresent resource in modern medicine, there are still several challenges that have limited MRI from reaching its full potential. Particularly, the time needed for the data-acquisition process can be restrictively long, which affects scanning throughput, patient comfort, and accessibility to MRI technology. To address this problem, widely used approaches have consisted of accelerating the data-acquisition process by considering advanced hardware and/or the acquisition of only a fraction of the data. In the case of partial acquisition, the accelerated data need to be efficiently reconstructed before the generation of the corresponding images. In this work we made significant contributions in this area. Firstly, we provided powerful algorithms able to successfully reconstruct accelerated MRI data in challenging acquisition scenarios. Secondly, we provided computational techniques which allow substantial efficiency improvements for reconstruction algorithms without compromising image-quality performance. A more detailed description of our main contributions is given in the following:

1. We have theoretically and empirically studied reconstruction algorithms based on structured low-rank modeling, which are capable of leveraging on constraints learned from reference datasets in order to improve reconstruction performance. These algorithms account for the

missing data samples by automatically identifying linear predictability relationships embedded in the data, and complement the reconstruction task with in prior knowledge learned from correlated datasets. We empirically showed that the use of reference data constraints in the Fourier domain offers performance advantages over constraints in the spatial domain. We showed that, in the case of accelerated data acquired using an uniform undersampling scheme, the use of a nonconvex cost-function leads to better reconstruction performance than the use of a convex cost-function. Notably, we show that these developments allow the reconstruction of highly accelerate data in important MRI applications such as Echo Planar Imaging acquisition.

2. We have developed a robust reconstruction method based on structured low-rank modeling which is able to account for imperfections in the reference data. Additionally, this reconstruction method is able to learn linear predictability relationships between the reference data and the accelerated data, which produces important improvements in reconstruction performance. We have empirically shown that this method is able to successfully reconstruct data in high acceleration regimes.

3. By making a systematic comparison we showed that the efficiency of MRI reconstruction methods based on linear predictability principles can be substantially improved by a simple modification in the underlying model used for prediction. When predicting missing data samples, we showed that using an ellipsoidal neighborhood of samples instead of using a rectangular neighborhood of samples has a negligible effect on reconstruction performance. However, the computational time and memory usage are substantially improved. We showed this effect empirically by implementing both neighborhood shapes in several conventional and state-of-the-art reconstruction methods.

4. We provided a novel theoretical description of the parallel imaging MRI sensitivity map estimation problem based on linear predictability and structured low-rank modeling theory.

These theoretical results provide an intuitive mathematical perspective to study subspace-based sensitivity map estimation methods. In addition, by relying on advanced signal processing techniques, we developed and tested powerful computational techniques which allow massive efficiency improvements when integrated to conventional sensitivity map estimation methods. For instance, we showed that approximately a 100-fold acceleration in computational time can be achieved without compromising estimation accuracy. Therefore, the overall reconstruction time of MRI reconstruction methods that estimate sensitivity maps in prior should also be substantially reduced.

Even though MRI can be considered as one the most remarkable technologies in medicine, there are still many areas where MRI can be improved, as the time needed for the data-acquisition. In this work we have developed theoretical and computational tools that have a direct impact on reducing the acquisition time, and we have empirically shown how they allow substantial improvements in performance and efficiency. However, there are still exciting research lines that we have not explored in this work. For example, in the first part of this work, we have focused the analysis of our proposed reconstruction methods on performance by following a proof-of-principles approach, and we have not necessarily provided the most efficient computational implementation. Therefore, the exploration of a more efficient implementation appears as a compelling research line with the potential of allowing the translation of our developed technology to clinical settings. Consequently, given that our proposed methods offer substantial advantages over conventional methods, we expect that their use in clinical applications will open new exciting research opportunities for the MRI community. There are also promising research lines related to the methods developed in the second part of this work. We have focused the assessment of the proposed computational techniques in specific contexts related to MRI reconstruction, however, these techniques can be straightforwardly translated to other important medical imaging applications such as computed tomography. Moreover, the proposed computational techniques can also be used to improve the efficiency of computational imaging applications in areas beyond medical imaging.

Bibliography

[1] Q.-S. Xiang and F. Q. Ye, "Correction for geometric distortion and N/2 ghosting in EPI by phase labeling for additional coordinate encoding (PLACE)," *Magn. Reson. Med.*, vol. 57, pp. 731–741, 2007.

[2] W. S. Hoge, F. R. Preiswerk, J. R. Polimeni, S. Yengul, P. A. Ciris, and B. Madore, "Ultrasound monitoring of a respiratory phantom for the development and validation of segmented EPI reconstruction methods," in *Proc. Int. Soc. Magn. Reson. Med.*, 2017, p. 1307.

[3] Z.-P. Liang, F. Boada, T. Constable, E. M. Haacke, P. C. Lauterbur, and M. R. Smith, "Constrained reconstruction methods in MR imaging," *Rev. Magn. Reson. Med.*, vol. 4, pp. 67–185, 1992.

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

[5] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Reson. Med.*, vol. 47, pp. 1202–1210, 2002.

[6] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, pp. 952–962, 1999.

[7] M. Lustig and J. M. Pauly, "SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitary k-space," *Magn. Reson. Med.*, vol. 65, pp. 457–471, 2010.

[8] M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "ESPIRiT – an eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magn. Reson. Med.*, vol. 71, pp. 990–1001, 2014.

[9] J. Zhang, C. Liu, and M. E. Moseley, "Parallel reconstruction using null operations," *Magn. Reson. Med.*, vol. 66, pp. 1241–1253, 2011.

[10] J. P. Haldar, "Low-rank modeling of local k-space neighborhoods (LORAKS): Implementation and examples for reproducible research," University of Southern California, Los Angeles, CA, Tech. Rep. USC-SIPI-414, April 2014, http://sipi.usc.edu/reports/abstracts.php?rid=sipi-414.

[11] ——, "Autocalibrated LORAKS for fast constrained MRI reconstruction," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2015, pp. 910–913.

[12] J. P. Haldar and J. Zhuo, "P-LORAKS: Low-rank modeling of local k-space neighborhoods with parallel imaging data," *Magn. Reson. Med.*, vol. 75, pp. 1499–1514, 2015.

[13] J. P. Haldar and K. Setsompop, "Linear predictability in magnetic resonance imaging reconstruction: Leveraging shift-invariant Fourier structure for faster and better imaging," *IEEE Signal Process. Mag.*, vol. 37, pp. 69–82, 2020.

[14] R. A. Lobos, T. H. Kim, W. S. Hoge, and J. P. Haldar, "Navigator-free EPI ghost correction with structured low-rank matrix models: New theory and methods," *IEEE Trans. Med. Imag.*, vol. 37, no. 11, pp. 2390–2402, 2018.

[15] R. A. Lobos, W. S. Hoge, A. Javed, C. Liao, K. Setsompop, K. S. Nayak, and J. P. Haldar, "Robust autocalibrated structured low-rank EPI ghost correction," *Magn. Reson. Med.*, vol. 85, no. 6, pp. 3403–3419, 2021.

[16] R. A. Lobos and J. P. Haldar, "On the shape of convolution kernels in MRI reconstruction: Rectangles versus ellipsoids," *Magn. Reson. Med.*, vol. 87, no. 6, pp. 2989–2996, 2022.

[17] P. Mansfield, "Multi-planar image formation using NMR spin echoes," *J. Phys. C: Solid State Phys.*, vol. 10, p. L55, 1977.

[18] M. A. Bernstein, K. F. King, and X. J. Zhou, *Handbook of MRI Pulse Sequences*. Burlington: Elsevier Academic Press, 2004.

[19] H. Bruder, H. Fischer, H.-E. Reinfelder, and F. Schmitt, "Image reconstruction for echo planar imaging with nonequidistant k-space sampling," *Magn. Reson. Med.*, vol. 23, pp. 311–323, 1992.

[20] X. Hu and T. H. Le, "Artifact reduction in EPI with phase-encoded reference scan," *Magn. Reson. Med.*, vol. 36, pp. 166–171, 1996.

[21] N.-k. Chen and A. M. Wyrwicz, "Removal of EPI Nyquist ghost artifacts with two-dimensional phase correction," *Magn. Reson. Med.*, vol. 51, pp. 1247–1253, 2004.

[22] M. H. Buonocore and L. Gao, "Ghost artifact reduction for echo planar imaging using image phase correction," *Magn. Reson. Med.*, vol. 38, pp. 89–100, 1997.

[23] S. Skare, D. Clayton, R. Newbould, M. Moseley, and R. Bammer, "A fast and robust minimum entropy based non-interactive Nyquist ghost correction algorithm," in *Proc. Int. Soc. Magn. Reson. Med.*, 2006, p. 2349.

[24] D. Xu, K. F. King, Y. Zur, and R. S. Hinks, "Robust 2D phase correction for echo planar imaging under a tight field-of-view," *Magn. Reson. Med.*, vol. 64, pp. 1800–1813, 2010.

[25] U. Yarach, M.-H. In, I. Chatnuntawech, B. Bilgic, F. Godenschweger, H. Mattern, A. Sciarra, and O. Speck, "Model-based iterative reconstruction for single-shot EPI at 7T," *Magn. Reson. Med.*, vol. 78, pp. 2250–2264, 2017.

[26] J. D. Ianni, E. B. Welch, and W. A. Grissom, "Ghost reduction in echo-planar imaging by joint reconstruction of images and line-to-line delays and phase errors," *Magn. Reson. Med.*, vol. 79, pp. 3114–3121, 2018.

[27] W. S. Hoge, H. Tan, and R. A. Kraft, "Robust EPI Nyquist ghost elimination via spatial and temporal encoding," *Magn. Reson. Med.*, vol. 64, pp. 1781–1791, 2010.

[28] W. S. Hoge and J. R. Polimeni, "Dual-polarity GRAPPA for simultaneous reconstruction and ghost correction of echo planar imaging data," *Magn. Reson. Med.*, vol. 76, pp. 32–44, 2016.

[29] H.-C. Chang and N.-K. Chen, "Joint correction of Nyquist artifact and minuscule motion-induced aliasing artifact in interleaved diffusion weighted EPI data using a composite two-dimensional phase correction procedure," *Magn. Reson. Imag.*, vol. 34, pp. 974–979, 2016.

[30] V. B. Xie, M. Lyu, Y. Liu, Y. Feng, and E. X. Wu, "Robust EPI Nyquist ghost removal by incorporating phase error correction with sensitivity encoding (PEC-SENSE)," *Magn. Reson. Med.*, vol. 79, pp. 943–951, 2018.

[31] P. J. Shin, P. E. Z. Larson, M. A. Ohliger, M. Elad, J. M. Pauly, D. B. Vigneron, and M. Lustig, "Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion," *Magn. Reson. Med.*, vol. 72, pp. 959–970, 2014.

[32] J. P. Haldar, "Low-Rank Modeling of Local-Space Neighborhoods (LORAKS) for Constrained MRI," *IEEE Trans. Med. Imag.*, vol. 33, pp. 668–681, 2014.

[33] ——, "Low-rank modeling of local k-space neighborhoods: From phase and support constraints to structured sparsity," in *Wavelets and Sparsity XVI, Proc. SPIE 9597*, 2015, p. 959710.

[34] J. P. Haldar and T. H. Kim, "Computational imaging with LORAKS: Reconstructing linearly predictable signals using low-rank matrix regularization." in *Proc. Asilomar Conf. Sig. Sys. Comp.*, 2017.

[35] T. H. Kim, K. Setsompop, and J. P. Haldar, "LORAKS makes better SENSE: Phase-constrained partial Fourier SENSE reconstruction without phase calibration," *Magn. Reson. Med.*, vol. 77, pp. 1021–1035, 2016.

[36] K. H. Jin, D. Lee, and J. C. Ye, "A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank Hankel matrix," *IEEE Trans. Comput. Imaging*, vol. 2, pp. 480–495, 2016.

[37] G. Ongie and M. Jacob, "Off-the-grid recovery of piecewise constant images from few Fourier samples," *SIAM J. Imaging Sci.*, vol. 9, pp. 1004–1041, 2016.

[38] P. Kellman and E. R. McVeigh, "Phased array ghost elimination," *NMR Biomed.*, vol. 19, pp. 352–361, 2006.

[39] J. Lee, K. H. Jin, and J. C. Ye, "Reference-free EPI Nyquist ghost correction using annihilating filter-based low rank Hankel matrix for K-space interpolation," *Magn. Reson. Med.*, vol. 76, pp. 1775–1789, 2016.

[40] M. Mani, M. Jacob, D. Kelley, and V. Magnotta, "Multi-shot sensitivity-encoded diffusion data recovery using structured low-rank matrix completion (MUSSELS)," *Magn. Reson. Med.*, vol. 78, pp. 494–507, 2017.

[41] M. Mani, V. Magnotta, D. Kelley, and M. Jacob, "Comprehensive reconstruction of multi-shot multi-channel diffusion data using MUSSELS," in *Proc. IEEE Eng. Med. Bio. Conf.*, 2016, pp. 1107–1110.

[42] M. Lyu, M. Barth, V. B. Xie, Y. Liu, Y. Feng, and E. X. Wu, "Robust 2D Nyquist ghost correction for simultaneous multislice (SMS) EPI using phase error correction SENSE and virtual coil SAKE," in *Proc. Int. Soc. Magn. Reson. Med.*, 2017, p. 515.

[43] R. A. Lobos, T. H. Kim, W. S. Hoge, and J. P. Haldar, "Navigator-free EPI ghost correction using low-rank matrix modeling: Theoretical insights and practical improvements," in *Proc. Int. Soc. Magn. Reson. Med.*, 2017, p. 0449.

[44] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, vol. 52, pp. 471–501, 2010.

[45] D. G. Luenberger, *Optimization by Vector Space Methods.* New York: John Wiley & Sons, 1969.

[46] M. Buehrer, K. P. Pruessmann, P. Boesiger, and S. Kozerke, "Array compression for MRI with large coil arrays," *Magn. Reson. Med.*, vol. 57, pp. 1131–1139, 2007.

[47] F. Huang, S. Vijayakumar, Y. Li, S. Hertel, and G. R. Duensing, "A software channel compression technique for faster reconstruction with many channels," *Magn. Reson. Imag.*, vol. 26, pp. 133–141, 2008.

[48] S. B. Reeder, E. Atalar, A. Z. Faranesh, and E. R. McVeigh, "Referenceless interleaved echo-planar imaging," *Magn. Reson. Med.*, vol. 41, pp. 87–94, 1999.

[49] S. M. Grieve, A. M. Blamire, and P. Styles, "Elimination of Nyquist ghosting caused by read-out to phase-encode gradient cross-terms in EPI," *Magn. Reson. Med.*, vol. 47, pp. 337–343, 2002.

[50] N.-K. Chen, A. V. Avram, and A. W. Song, "Two-dimensional phase cycled reconstruction for inherent correction of echo-planar imaging Nyquist artifacts," *Magn. Reson. Med.*, vol. 66, pp. 1057–1066, 2011.

[51] W. S. Hoge and J. R. Polimeni, "Artifact correction in accelerated-segmented EPI data via dual-polarity GRAPPA," in *Proc. Int. Soc. Magn. Reson. Med.*, 2017, p. 449.

[52] E. J. Candès, C. A. Sing-Long, and J. D. Trzasko, "Unbiased risk estimates for singular value thresholding and spectral estimators," *IEEE Trans. Signal Process.*, vol. 61, pp. 4643–4657, 2013.

[53] T. H. Kim and J. P. Haldar, "SMS-LORAKS: Calibrationless simultaneous multislice MRI using low-rank matrix modeling," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2015, pp. 323–326.

[54] B. A. Poser and K. Setsompop, "Pulse sequences and parallel imaging for high spatiotemporal resolution MRI at ultra-high field," *NeuroImage*, 2017, Early View.

[55] T. H. Kim, B. Bilgic, D. Polak, K. Setsompop, and J. P. Haldar, "Wave-LORAKS for faster wave-CAIPI MRI," in *Proc. Int. Soc. Magn. Reson. Med.*, 2017, p. 1037.

[56] W. Jiang, P. E. Z. Larson, and M. Lustig, "Simultaneous estimation of auto-calibration data and gradient delays in non-Cartesian parallel MRI using low-rank constraints," in *Proc. Int. Soc. Magn. Reson. Med.*, 2016, p. 939.

[57] M. Mani, S. Poddar, V. Magnotta, and M. Jacob, "Trajectory correction of radial data using MUSSELS," in *Proc. Int. Soc. Magn. Reson. Med.*, 2017, p. 1197.

[58] J. P. Haldar, "Calibrationless partial Fourier reconstruction of MR images with slowly-varying phase: A rank-deficient matrix recovery approach," in *ISMRM Workshop on Data Sampling & Image Reconstruction*, Sedona, 2013.

[59] M. K. Stelhing, R. Turner, and P. Mansfield, "Echo-planar imaging: Magnetic resonance imaging in a fraction of a second," *Science*, vol. 254, pp. 43–50, 1991.

[60] M. Lyu, M. Barth, V. B. Xie, Y. Liu, X. Ma, Y. Feng, and E. X. Wu, "Robust SENSE reconstruction of simultaneous multislice EPI with low-rank enhanced coil sensitivity calibration and slice-dependent 2D Nyquist ghost correction," *Magn. Reson. Med.*, vol. 80, pp. 1376–1390, 2018.

[61] Y. Liu, M. Lyu, M. Barth, Z. Yi, A. T. Leong, F. Chen, Y. Feng, and E. X. Wu, "PEC-GRAPPA reconstruction of simultaneous multislice EPI with slice-dependent 2D Nyquist ghost correction," *Magn. Reson. Med.*, vol. 81, pp. 1924–1934, 2019.

[62] L. Ying and J. Sheng, "Joint image reconstruction and sensitivity estimation in SENSE (JSENSE)," *Magn. Reson. Med.*, vol. 57, pp. 1196–1202, 2007.

[63] D. Sheltraw, B. Inglis, V. Deshpande, and M. Trumpis, "Simultaneous reduction of two common autocalibration errors GRAPPA EPI time series data," *Preprint*, 2012, arXiv:1208.0972.

[64] C. A. Baron and C. Beaulieu, "Motion robust GRAPPA for echo-planar imaging," *Magn. Reson. Med.*, vol. 75, no. 3, pp. 1166–1174, 2016.

[65] S. L. Talagala, J. E. Sarlls, S. Liu, and S. J. Inati, "Improvement of temporal signal-to-noise ratio of GRAPPA accelerated echo planar imaging using a FLASH based calibration scan," *Magn. Reson. Med.*, vol. 75, no. 6, pp. 2362–2371, 2016.

[66] H. C. M. Holme, S. Rosenzweig, F. Ong, R. N. Wilke, M. Lustig, and M. Uecker, "ENLIVE: An efficient nonlinear method for calibrationless and robust parallel imaging," *Sci. Rep.*, vol. 9, no. 1, pp. 1–13, 2019.

[67] B. Bilgic, T. H. Kim, C. Liao, M. K. Manhard, L. L. Wald, J. P. Haldar, and K. Setsompop, "Improving parallel imaging by jointly reconstructing multi-contrast data," *Magn. Reson. Med.*, vol. 80, no. 2, pp. 619–632, 2018.

[68] R. A. Lobos, A. Javed, K. S. Nayak, W. S. Hoge, and J. P. Haldar, "Robust autocalibrated LORAKS for EPI ghost correction," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2018, pp. 663–666.

[69] ——, "Robust autocalibrated LORAKS for improved EPI ghost correction with structured low-rank matrix models," in *Proc. Int. Soc. Magn. Reson. Med.*, 2018, p. 3533.

[70] Z.-P. Liang, E. M. Haacke, and C. W. Thomas, "High-resolution inversion of finite Fourier transform data through a localised polynomial approximation," *Inverse Probl.*, vol. 5, pp. 831–847, 1989.

[71] J. R. Polimeni, H. Bhat, T. Witzel, T. Benner, T. Feiweier, S. J. Inati, V. Renvall, K. Heberlein, and L. L. Wald, "Reducing sensitivity losses due to respiration and motion in accelerated echo planar imaging by reordering the autocalibration data acquisition," *Magn. Reson. Med.*, vol. 75, no. 2, pp. 665–679, 2016.

[72] W. S. Hoge, K. Setsompop, and J. R. Polimeni, "Dual-polarity slice-GRAPPA for concurrent ghost correction and slice separation in simultaneous multi-slice EPI," *Magn. Reson. Med.*, vol. 80, no. 4, pp. 1364–1375, 2018.

[73] C. Liao, M. K. Manhard, B. Bilgic, Q. Tian, Q. Fan, S. Han, F. Wang, D. J. Park, T. Witzel, J. Zhong *et al.*, "Phase-matched virtual coil reconstruction for highly accelerated diffusion echo-planar imaging," *NeuroImage*, vol. 194, pp. 291–302, 2019.

[74] T. H. Kim and J. P. Haldar, "LORAKS software version 2.0: Faster implementation and enhanced capabilities," University of Southern California, Los Angeles, CA, Tech. Rep. USC-SIPI-443, May 2018. [Online]. Available: http://sipi.usc.edu/reports/abstracts.php?rid=sipi-443

[75] G. Ongie and M. Jacob, "A fast algorithm for convolutional structured low-rank matrix recovery," *IEEE Trans. Comput. Imaging*, vol. 3, pp. 535–550, 2017.

[76] D. Le Bihan, C. Poupon, A. Amadon, and F. Lethimonnier, "Artifacts and pitfalls in diffusion MRI," *J. Magn. Reson. Imag.*, vol. 24, no. 3, pp. 478–488, 2006.

[77] A. Javed and K. S. Nayak, "Single-shot EPI for ASL-CMR," *Magn. Reson. Med.*, vol. 84, pp. 738–750, 2020.

[78] C. Bhushan, J. P. Haldar, S. Choi, A. A. Joshi, D. W. Shattuck, and R. M. Leahy, "Co-registration and distortion correction of diffusion and anatomical images based on inverse contrast normalization," *Neuroimage*, vol. 115, pp. 269–280, 2015.

[79] M. Buehrer, P. Boesiger, and S. Kozerke, "Virtual body coil calibration for phased-array imaging," in *Proc. Int. Soc. Magn. Reson. Med.*, 2009, p. 760.

[80] T. H. Kim and J. P. Haldar, "The Fourier radial error spectrum plot: A more nuanced quantitative evaluation of image reconstruction quality," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2018, pp. 61–64.

[81] M. Mani, M. Jacob, G. McKinnon, B. Yang, B. Rutt, A. Kerr, and V. Magnotta, "SMS-MUSSELS: A navigator-free reconstruction for slice-accelerated multi-shot diffusion imaging," in *Proc. Int. Soc. Magn. Reson. Med.*, 2019, p. 0233.

[82] B. Bilgic, C. Liao, M. K. Manhard, Q. Tian, I. Chatnuntawech, S. S. Iyer, S. F. Cauley, T. Feiweier, S. Giri, a. Y. H. Y. Hu, J. R. Polimeni, L. L. Wald, and K. Setsompop, "Robust high-quality multi-shot EPI with low-rank prior and machine learning," in *Proc. Int. Soc. Magn. Reson. Med.*, 2019, p. 1250.

[83] Y. Chang, D. Liang, and L. Ying, "Nonlinear GRAPPA: A kernel approach to parallel MRI reconstruction," *Magn. Reson. Med.*, vol. 68, pp. 730–740, 2011.

[84] J. Lyu, U. Nakarmi, D. Liang, J. Sheng, and L. Ying, "KerNL: Kernel-based nonlinear approach to parallel MRI reconstruction," *IEEE Trans. Med. Imag.*, vol. 38, pp. 312–321, 2019.

[85] M. Akçakaya, S. Moeller, S. Weingärtner, and K. Uğurbil, "Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging," *Magn. Reson. Med.*, vol. 81, no. 1, pp. 439–453, 2019.

[86] T. H. Kim, P. Garg, and J. P. Haldar, "LORAKI: Autocalibrated recurrent neural networks for autoregressive reconstruction in k-space," *Preprint*, 2019, arXiv:1904.09390.

[87] F. Knoll, K. Hammernik, C. Zhang, S. Moeller, T. Pock, D. K. Sodickson, and M. Akcakaya, "Deep-learning methods for parallel magnetic resonance image reconstruction: A survey of the current approaches, trends, and issues," *IEEE Signal Process. Mag.*, vol. 37, pp. 128–140, 2020.

[88] D. Liang, J. Cheng, Z. Ke, and L. Ying, "Deep magnetic resonance image reconstruction: Inverse problems meet neural networks," *IEEE Signal Process. Mag.*, vol. 37, pp. 141–151, 2020.

[89] E. N. Yeh, C. A. McKenzie, M. A. Ohliger, and D. K. Sodickson, "Parallel magnetic resonance imaging with adaptive radius in k-space (PARS): Constrained image reconstruction using k-space locality in radiofrequency coil encoded data," *Magn. Reson. Med.*, vol. 53, no. 6, pp. 1383–1392, 2005.

[90] R. A. Lobos and J. P. Haldar, "Improving the performance of accelerated image reconstruction in k-space: The importance of kernel shape," in *Proc. Int. Soc. Magn. Reson. Med.*, 2019, p. 2407.

[91] M. A. Bernstein, S. B. Fain, and S. J. Riederer, "Effect of windowing and zero-filled reconstruction of MRI data on spatial resolution and acquisition strategy," *Magn. Reson. Med.*, vol. 14, no. 3, pp. 270–280, 2001.

[92] J. A. Bankson and S. M. Wright, "Generalized partially parallel imaging with spatial filters," in *Proc. Int. Soc. Magn. Reson. Med.*, 2001, p. 1795.

[93] J. Wang, B. Zhang, K. Zhong, and Y. Zhuo, "Image domain based fast GRAPPA reconstruction and relative SNR degradation factor," in *Proc. Int. Soc. Magn. Reson. Med.*, 2005, p. 2428.

[94] J. Zbontar, F. Knoll, A. Sriram, T. Murrell, Z. Huang, M. J. Muckley, A. Defazio, R. Stern, P. Johnson, M. Bruno *et al.*, "fastmri: An open dataset and benchmarks for accelerated mri," *arXiv preprint arXiv:1811.08839*, 2018.

[95] P. B. Roemer, W. A. Edelstein, C. E. Hayes, S. P. Souza, and O. M. Mueller, "The NMR phased array," *Magn. Reson. Med.*, vol. 16, pp. 192–225, 1990.

[96] J. S. Hyde, A. Jesmanowicz, W. Froncisz, J. B. Kneeland, and T. M. Grist, "Parallel image acquisition from noninteracting local coils," *J. Magn. Reson.*, vol. 70, pp. 512–517, 1986.

[97] J. W. Carlson, "An algorithm for NMR imaging reconstruction based on multiple RF receiver coils," *J. Magn. Reson.*, vol. 74, pp. 376–380, 1987.

[98] M. Hutchinson and U. Raff, "Fast MRI data acquisition using multiple detectors," *Magn. Reson. Med.*, vol. 6, pp. 87–91, 1988.

[99] J. R. Kelton, R. L. Magin, and S. M. Wright, "An algorithm for rapid image acquisition using multiple receiver coils," in *Proc. Int. Soc. Magn. Reson. Med.*, Amsterdam, 1989, p. 1172.

[100] D. Kwiat, S. Einav, and G. Navon, "A decoupled coil detector array for fast image acquisition in magnetic-resonance-imaging," *Med. Phys.*, vol. 18, pp. 251–265, 1991.

[101] J. B. Ra and C. Y. Rim, "Fast imaging using subencoding data sets from multiple detectors," *Magn. Reson. Med.*, vol. 30, pp. 142–145, 1993.

[102] K. P. Pruessmann, M. Weiger, P. Börnert, and P. Boesiger, "Advances in sensitivity encoding with arbitrary k-space trajectories," *Magn. Reson. Med.*, vol. 46, pp. 638–651, 2001.

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

[104] M. Blaimer, F. Breuer, M. Mueller, R. M. Heidemann, M. A. Griswold, and P. M. Jakob, "SMASH, SENSE, PILS, GRAPPA: how to choose the optimal method," *Topics in Magnetic Resonance Imaging*, vol. 15, no. 4, pp. 223–236, 2004.

[105] L. Ying and Z.-P. Liang, "Parallel MRI using phased array coils," *IEEE Signal Process. Mag.*, vol. 27, no. 4, pp. 90–98, 2010.

[106] A. Deshmane, V. Gulani, M. A. Griswold, and N. Seiberlich, "Parallel mr imaging," *J. Magn. Reson.*, vol. 36, no. 1, pp. 55–72, 2012.

[107] D. O. Walsh, A. F. Gmitro, and M. W. Marcellin, "Adaptive reconstruction of phased array MR imagery," *Magn. Reson. Med.*, vol. 43, no. 5, pp. 682–690, 2000.

[108] M. Bydder, D. J. Larkman, and J. V. Hajnal, "Combination of signals from array coils using image-based estimation of coil sensitivity profiles," *Magn. Reson. Med.*, vol. 47, pp. 539–548, 2002.

[109] S. L. Keeling and R. Bammer, "A variational approach to magnetic resonance coil sensitivity estimation," *Applied Mathematics and Computation*, vol. 158, no. 2, pp. 359–388, 2004.

[110] L. Ying and J. Sheng, "Joint image reconstruction and sensitivity estimation in SENSE (JSENSE)," *Magn. Reson. Med.*, vol. 57, no. 6, pp. 1196–1202, 2007.

[111] M. Uecker, T. Hohage, K. T. Block, and J. Frahm, "Image reconstruction by regularized nonlinear inversion – joint estimation of coil sensitivites and content," *Magn. Reson. Med.*, vol. 60, pp. 674–682, 2008.

[112] M. J. Allison, S. Ramani, and J. A. Fessler, "Accelerated regularized estimation of MR coil sensitivities using augmented Lagrangian methods," *IEEE Trans. Med. Imag.*, vol. 32, no. 3, pp. 556–564, 2012.

[113] R. L. Morrison, M. Jacob, and M. N. Do, "Multichannel estimation of coil sensitivities in parallel MRI," in *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro.* IEEE, 2007, pp. 117–120.

[114] D. G. Gungor, R. Ahmad, and L. C. Potter, "A subspace approach to blind coil sensitivity estimation in parallel MRI," *Journal of Cardiovascular Magnetic Resonance*, vol. 16, no. 1, pp. 1–2, 2014.

[115] H. She, R.-R. Chen, D. Liang, Y. Chang, and L. Ying, "Image reconstruction from phased-array data based on multichannel blind deconvolution," *Magn. Reson. Imag.*, vol. 33, pp. 1106–1113, 2015.

[116] X. Peng, B. P. Sutton, F. Lam, and Z.-P. Liang, "DeepSENSE: Learning coil sensitivity functions for SENSE reconstruction using deep learning," *Magn. Reson. Med.*, vol. 87, no. 4, pp. 1894–1902, 2022.

[117] M. Jacob, M. P. Mani, and J. C. Ye, "Structured low-rank algorithms: Theory, magnetic resonance applications, and links to machine learning," *IEEE Signal Process. Mag.*, vol. 37, pp. 54–68, 2020.

[118] S. Zhao, L. C. Potter, and R. Ahmad, "High-dimensional fast convolutional framework (HICU) for calibrationless MRI," *Magn. Reson. Med.*, vol. 86, no. 3, pp. 1212–1225, 2021.

[119] G. Golub and C. van Loan, *Matrix Computations*, 3rd ed. London: The Johns Hopkins University Press, 1996.

[120] J. Trzasko and A. Manduca, "Local versus global low-rank promotion in dynamic MRI series reconstruction," in *Proc. Int. Soc. Magn. Reson. Med.*, 2011, p. 4371.

[121] J. D. Trzasko and A. Manduca, "CLEAR: Calibration-free parallel imaging using locally low-rank encouraging reconstruction," in *Proc. Int. Soc. Magn. Reson. Med.*, 2012, p. 517.

[122] J. V. Manjón, P. Coupé, L. Concha, A. Buades, D. L. Collins, and M. Robles, "Diffusion weighted image denoising using overcomplete local PCA," vol. 8, p. e73021, 2013.

[123] J. Veraart, D. S. Novikov, D. Christiaens, B. Ades-Aron, J. Sijbers, and E. Fieremans, "De-noising of diffusion MRI using random matrix theory," *NeuroImage*, vol. 142, pp. 394–406, 2016.

[124] L. Cordero-Grande, D. Christiaens, J. Hutter, A. N. Price, and J. V. Hajnal, "Complex diffusion-weighted image estimation via matrix recovery under general noise models," *NeuroImage*, vol. 200, pp. 391–404, 2019.

[125] K. F. Cheung and R. J. Marks II, "Imaging sampling below the Nyquist density without aliasing," vol. 7, pp. 92–105, 1990.

[126] F. Huang, W. Lin, and Y. Li, "Partial Fourier reconstruction through data fitting and convolution in k-space," *Magn. Reson. Med.*, vol. 62, pp. 1261–1269, 2009.

[127] Z.-P. Liang, "Constrained image reconstruction from incomplete and noisy data: A new parametric approach," Ph.D. dissertation, Case Western Reserve University, Cleveland, OH, USA, 1989.

[128] R. N. Bracewell, *The Fourier Transform and its Applications*.   Boston: McGraw-Hill, 2000.

[129] M. İ. Gürelli and C. L. Nikias, "EVAM: An eigenvector-based algorithm for multichannel blind deconvolution of input colored signals," *IEEE Trans. Signal Process.*, vol. 43, pp. 134–149, 1995.

[130] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Process.*, vol. 43, pp. 2982–2993, 1995.

[131] G. Harikumar and Y. Bresler, "Perfect blind restoration of images blurred by multiple filters: theory and efficient algorithms," *IEEE Trans. Image Process.*, vol. 8, pp. 202–219, 1999.

[132] D. Kim, S. F. Cauley, K. S. Nayak, R. M. Leahy, and J. P. Haldar, "Region-optimized virtual (ROVir) coils: Localization and/or suppression of spatial regions using sensor-domain beamforming," *Magn. Reson. Med.*, vol. 86, pp. 197–212, 2021.

[133] M. Uecker, F. Ong, J. I. Tamir, D. Bahri, P. Virtue, J. Y. Cheng, T. Zhang, and M. Lustig, "Berkeley advanced reconstruction toolbox," in *Proc. Intl. Soc. Mag. Reson. Med*, vol. 23, no. 2486, 2015.

[134] S. Iyer, F. Ong, K. Setsompop, M. Doneva, and M. Lustig, "SURE-based automatic parameter selection for ESPIRiT calibration," *Magn. Reson. Med.*, vol. 84, no. 6, pp. 3423–3437, 2020.