Accelerated Dynamic MR Imaging Using Linear And Non-Linear Machine Learning-Based Image Reconstruction Models

Yanis Djebra

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2023IPPAT011

Thèse de doctorat

Gordon Center for Medical Imaging

MASSACHUSETTS GENERAL HOSPITAL

HARVARD MEDICAL SCHOOL

TELECOM Paris

IP PARIS

Accelerated Dynamic MR Imaging Using Linear And Non-Linear Machine Learning-Based Image Reconstruction Models

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Télécom Paris & Gordon Center for Medical Imaging
(Massachusetts General Hospital, Harvard Medical School)

École doctorale n°626 de l'Institut Polytechnique de Paris (ED IP Paris)
Spécialité de doctorat: Signal, Images, Automatique et Robotique

Thèse présentée et soutenue à Palaiseau, le 11 mai 2023, par

YANIS DJEBRA

Nadjia Kachenoura Directrice de Recherches, INSERM (LIB, Sorbonne Université)	Présidente / Examinatrice
Loïc Boussel Professeur des universités – Praticien hospitalier, INSA Lyon (CREATIS)	Rapporteur
Lotfi Chaari Professeur Associé, INP-Toulouse (IRIT)	Rapporteur
Marc Normandin Professeur Associé, Harvard Medical School (GCMI)	Examinateur
Chao Ma Professeur Assistant, Harvard Medical School (GCMI)	Examinateur
Isabelle Bloch Professeure, Sorbonne Université (LIP6), et Télécom Paris	Directrice de thèse
Georges El Fakhri Professeur, Harvard Medical School (GCMI)	Co-Directeur de thèse

Acknowledgments

بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ

I would like to take this moment to extend my deepest appreciation to everyone who supported and guided me on this journey that was my PhD thesis. First and foremost, I extend my heartfelt gratitude to both my advisors, Dr. Isabelle Bloch at Telecom Paris and Sorbonne Université in France, and Dr. Georges El Fakhri at the Gordon Center for Medical Imaging (GCMI) in Boston, who allowed this PhD to be concretized and flourished.

Naturally, I would like to thank my mentor, Dr. Chao Ma, as well as Dr. Paul Kyu Han, and Dr. Thibault Marin whom I also closely worked with. Your consistent guidance, expertise, and unwavering support played an invaluable role in the development and completion of my work. I appreciate your dedication, patience, and the wisdom you shared with me in the fields of MR and optimization techniques.

To the esteemed members of the jury, thank you for your time and constructive feedback. Your insights not only enriched this research but also my growth as a scientist. I am grateful for the opportunity to learn from your extensive knowledge and experience.

The majority of my work was conducted in Boston, and to all the faculty and staff at the GCMI, I cannot express enough thanks for your unwavering support and camaraderie. This journey would not have been as fruitful or enjoyable without your contribution. I consider myself fortunate to have been part of such an enriching environment and look forward to maintaining these connections in my future endeavors.

Finally, a big thank you to my family and friends. Your ongoing encouragements have been instrumental in my journey. Your support has been not just a source of strength during my PhD studies, but throughout my life. This accomplishment is not just my own, but a testament to the incredible support system I have been fortunate enough to be part of. Thank you.

Résumé

L'imagerie par Résonance Magnétique (IRM) dynamique est d'une grande valeur pour le diagnostic médical grâce à sa polyvalence en termes de contraste, sa haute résolution spatiale, son rapport signal/bruit élevé et permet l'obtention non invasives d'images multi-planaires. Elle peut être particulièrement utile pour l'imagerie du cerveau, du cœur, de la colonne vertébrale et des articulations, ainsi que pour la détection d'anomalies. De plus, la disponibilité croissante de machines de Tomographie par Émission de Positrons (TEP) / IRM permet l'acquisition simultanée de données de TEP et pour de meilleures reconstructions et des informations complémentaires. Cependant, un défi majeur en IRM dynamique est la reconstruction d'images à partir de données d'espace-k échantillonnées en dessous de la fréquence de Nyquist. De nombreuses méthodes ont été proposées pour l'imagerie IRM sous-échantillonnée, notamment l'imagerie parallèle et le compressed sensing, mais ces méthodes restent limitées en imagerie dynamique du au taux de sous-échantillonnage intrinsèquement élevé.

Le premier objectif de cette thèse est de montrer le potentiel et l'utilité du modèle de sous-espace linéaire pour l'imagerie IRM sous respiration libre. Un tel modèle peut théoriquement capturer des mouvements respiratoires et cardiaques réguliers. Cependant, lorsqu'il s'agit de scans prolongés, des mouvements irréguliers peuvent survenir, tels qu'une respiration erratique ou un mouvement global causé par l'inconfort du patient. Une première question se pose donc naturellement : un tel modèle peut-il capturer ces types de mouvement et, si oui, peut-il reconstruire des images à partir d'un scan IRM dynamique présentant ces mouvements ? Nous démontrons dans cette thèse comment le modèle de sous-espace peut efficacement reconstruire des images sans d'artefacts à partir de données d'espace-k fortement sous-échantillonnées. Une première application est présentée où nous reconstruisons des images IRM dynamiques avec haute résolution spatiale et temporelle et les utilisons pour corriger le mouvement des données TEP simultanément acquise, capturant ainsi des mouvement complexes tels qu'une respiration erratique ou un mouvement global. Une deuxième application sur la cartographie T1 cardiaque est présentée. Des données d'espace-k sous-échantillonnées ont été acquises à l'aide d'une séquence de type inversion-récupération et sous respiration libre, et des images IRM 3D dynamiques du cœur entier ont été reconstruites en utilisant le modèle de sous-espace linéaire.

Le deuxième objectif de cette thèse est de comprendre les limites du modèle de sous-espace linéaire et de développer un nouveau modèle de reconstruction d'imagerie IRM dynamique qui

pallie ces limitations. Plus spécifiquement, le modèle de sous-espace suppose que les données de haute dimension résident dans un sous-espace linéaire de basse dimension qui capture les corrélations spatiotemporelles des images dynamiques. Cette méthode repose donc sur un modèle de réduction de dimension linéaire et ne prend pas en compte les caractéristiques intrinsèquement non linéaires du signal, ce qui peut montrer ses limites avec des taux de sous-échantillonnage plus élevés. Des modèles basés sur l'apprentissage de variétés ont donc été explorés pour la reconstruction d'images en IRM dynamique et visent à apprendre la structure intrinsèque du signal en résolvant des problèmes de réduction de dimensionnalité non linéaires. Nous présentons dans cette thèse une stratégie alternative pour la reconstruction d'images IRM basée sur l'apprentissage de variétés. La méthode proposée apprend la structure de variété via un alignement linéaire des espaces tangents (LTSA) et peut être interprétée comme une généralisation non linéaire du modèle de sous-espace. Des validations ont été effectuées sur des études de simulation numérique ainsi que sur des expériences d'imagerie cardiaque 2D et 3D in vivo, démontrant des performances améliorées par rapport à l'état-de-l'art.

Les deux premiers objectifs présentent respectivement des modèles linéaires et non linéaires, mais ces méthodes utilisent des techniques d'optimisation linéaire conventionnelles pour résoudre le problème de reconstruction. En revanche, l'utilisation de réseaux de neurones profonds pour l'optimisation peut procurer une meilleure puissance de représentation non linéaire. Cet axe de recherche a donc été exploré à la fin de cette thèse. Des premiers résultats sur les approches basées sur l'apprentissage profond sont présentés dans cette thèse et l'état-de-l'art est discuté. Le dernier chapitre présente les conclusions, discute des contributions de l'auteur et détaille les perspectives de recherche potentielles ouvertes par le travail effectué dans cette thèse.

Abstract

Dynamic Magnetic Resonance (MR) imaging is of high value in medical diagnosis thanks to its contrast versatility, high spatial resolution, and high Signal-to-Noise Ratio (SNR), and allows for non-invasive multi-planar images of the body. It can be particularly useful for imaging the brain, heart, spine, and joints, as well as for detecting abnormalities. In addition, MR images can now be simultaneously acquired with Positron Emission Tomography (PET) data with the increasing availability of hybrid PET/MR. This can aid the PET data reconstruction using techniques such as MR-based PET data motion correction or get complementary information such as relaxation times T1 that can prove useful for diagnosis or derived parameters calculations (e.g., membrane potential calculations). However, a key challenge in dynamic MRI is reconstructing high-dimensional images from sparse k-space data sampled below the Nyquist sampling rate. Many methods have been proposed for accelerated imaging with sparse sampling, including parallel imaging, compressed sensing.

The first objective of this thesis is to show the potential and usefulness of free-breathing MR imaging using a linear subspace model. It has been proved that such a model can in principle capture regular respiratory and cardiac motion. However, when dealing with lengthy scans, irregular motion patterns can occur, such as erratic breathing or bulk motion caused by patient discomfort. A first question thus naturally arises: can such a model capture irregular types of motion and, if so, can it reconstruct images from a dynamic MR scan presenting bulk motion and irregular respiratory motion? To answer this question, we demonstrate how the subspace model can efficiently reconstruct artifact-free images from highly undersampled k-space data with various motion patterns. A first application is presented where we reconstruct high-resolution, high frame-rate dynamic MR images from a PET/MR scanner. Reconstructed dynamic MR images were then used to correct motion in the PET reconstruction by estimating phase-to-phase nonrigid motion fields able to capture complex motion patterns such as irregular respiratory patterns and bulk motion. A second application on cardiac T1 mapping is presented. Undersampled k-space data were acquired using a free-breathing, ECG-gated inversion recovery sequence, and dynamic 3D MR images of the whole heart were reconstructed leveraging the linear subspace model.

The second objective of this thesis is to understand the limits of the linear subspace model and develop a novel dynamic MR reconstruction scheme that palliates these limitations. More specifically, the subspace model assumes that high-dimensional data reside in a low-dimensional linear

subspace that captures the spatiotemporal correlations of dynamic MR images. This dimensionality reduction significantly reduces the degrees of freedom of the underlying signals and therefore allows for undersampled signal reconstruction. However, this model relies on a linear dimensionality reduction model of highly dimensional data and does not account for intrinsic non-linear features of the signal, which may show its limits with higher undersampling rates. Manifold learning-based models have therefore been explored for image reconstruction in dynamic MRI and aim at learning the intrinsic low-dimensional structure of the input data that are embedded in a high-dimensional signal space by solving non-linear dimensionality reduction problems. We present in this thesis an alternative strategy for manifold learning-based image reconstruction in dynamic MRI. The proposed method is closely related to machine learning methods that learn manifolds via linear tangent space alignment (LTSA) and can be interpreted as a non-linear generalization of the subspace model. Validation on numerical simulation studies as well as in vivo 2D and 3D cardiac imaging experiments were performed. The proposed method offered improved performances in terms of noise and undersampling artifacts compared to the state-of-the-art techniques.

Finally, research on deep learning-based approaches for MR image reconstruction is ongoing to further improve reconstruction performances. The two first objectives present respectively linear and non-linear based approaches for MR image reconstruction. However, even though the LTSA method is based on a non-linear model, both methods actually use conventional linear optimization techniques to solve the reconstruction problem. Using deep neural networks for optimization has shown promising results in many fields, as they procure non-linear and extensive representation power. Early results are presented in this thesis and state-of-the-art techniques are discussed. A last chapter then presents conclusions, discusses the author's contributions, and considers the potential research perspectives that have been opened up by the work presented in this thesis.

Contents

Introduction

Magnetic resonance imaging (MRI) is a widely used medical imaging technique that enables non-invasive visualization of the structure and function of the human body. It is based on the detection of nuclear magnetic resonance (NMR) signals emanating from the hydrogen atoms present in water and fat molecules within the body. By applying a strong, static magnetic field and pulsed radiofrequency (RF) excitation, MR signals can be selectively detected from different tissues and used to generate detailed images of the body. This modality has several advantages over other imaging methods such as X-ray, CT, and ultrasound. First, it does not expose the patient to ionizing radiation, which makes it safer for repeated examinations and for imaging pregnant women or children. In addition, MR imaging has excellent soft tissue contrast, high spatial resolution, and relatively straightforward full 3D-imaging. These features yield one of the main advantages of MRI: the differentiation between several types of tissues and visualization of subtle abnormalities that may not be visible with other imaging modalities. Finally, MR imaging is highly versatile and can be used for both anatomical and functional imaging. Using various types of sequences, we can obtain virtually an infinite number of images with different contrast between the tissues. With adequate MR sequences, we can even obtain information such as blood flow and tissue perfusion. Techniques such as functional MRI (fMRI), which measures brain activity by detecting changes in blood flow, and diffusion MRI, which examines the movement of water molecules to study the integrity of white matter in the brain (and other organs), can then be real assets to assess neurological disorders or tumor evolution over time.

In addition, dynamic MRI, which refers to the acquisition of MR images over time in a dynamic way, is becoming more and more common and allows for the visualization of dynamic processes within the body such as cardiac contractions or dynamic contrast enhanced (DCE) imaging of the liver. With the increasing availability of hybrid PET/MR, MR imaging can now be simultaneously acquired with PET data. This can aid the PET data reconstruction using techniques such as MR-based PET data motion correction, and get complimentary information such as relaxation times T1 or T2 that can prove useful for diagnosis or for derived parameters calculations. However, MRI – and dynamic MRI mostly, have some inherent limitations. The imaged objects in dynamic MRI move, and can produce motion artifacts. Then, to be able to correctly capture motion, one needs a high temporal resolution, which can be obtained but at the cost of a decreased spatial resolution. The compromise is therefore to be able to acquire k-space samples at a temporal resolution high

enough to capture the motion of the moving objects (e.g., heart beating or respiratory motion). This compromise is translated as undersampled MR measurement, especially for dynamic MRI, leading to an ill-posed optimization problem when it comes to recover the true image behind the measurement. Applying a simple Fourier transform is therefore not sufficient anymore, as undersampled MR data result in images corrupted with aliasing artifacts and noise.

A solution to this problem is to use advanced image reconstruction schemes, using mathematical models, or leveraging prior information on the data. These reconstruction techniques are iterative (unlike the Fourier transform reconstruction) and optimize some well-defined cost function to recover the image from the undersampled measurements. Many techniques have been developed through the years, such as Parallel imaging (PI) techniques [43, 121, 146], Compressed Sensing (CS) [88], or more recently, subspace/manifold based techniques that model the data as lying close to a lower dimensional manifold [118, 170, 171].

This PhD thesis explores reconstruction techniques for MRI. We used pre-existing models for innovative applications, understood the limitations of such models, and contributed to the MR image reconstruction field with a new reconstruction model overcoming these limitations. The first objective of this thesis is to show the potential and usefulness of free-breathing MR imaging using a linear subspace model. Indeed, it has been proved that such a model can in principle capture regular respiratory and cardiac motion. However, when dealing with lengthy scans, irregular motion patterns can occur, such as erratic breathing or bulk motion caused by patient discomfort. A first question thus naturally arises: can such a model capture irregular types of motion and, if so, can it reconstruct images from a dynamic MR scan presenting bulk motion and irregular respiratory motion? To answer this question, we demonstrate that the subspace model can efficiently reconstruct artifact-free images from highly undersampled k-space data with various motion patterns. A first application is presented where we reconstruct high-resolution, high frame-rate dynamic MR images from a PET/MR scanner. Reconstructed dynamic MR images were then used to correct motion in the PET reconstruction by estimating phase-to-phase nonrigid motion fields able to capture complex motion patterns such as irregular respiratory patterns and bulk motion. A second application on cardiac T1 mapping is presented. Undersampled k-space data were acquired using a free-breathing, ECG-gated inversion recovery sequence, and dynamic 3D MR images of the whole heart were reconstructed leveraging the linear subspace model.

The second objective of this thesis is to understand the limits of the linear subspace model and develop a novel dynamic MR reconstruction scheme that palliates these limitations. More specifically, the subspace model assumes that high-dimensional data reside in a low-dimensional linear subspace that captures the spatiotemporal correlations of dynamic MR images. This dimensionality reduction significantly reduces the degrees of freedom of the underlying signals and therefore allows for undersampled signal reconstruction. However, this model relies on a linear dimensionality reduction model of highly dimensional data and does not account for intrinsic non-linear features of the signal, which may show its limits with higher undersampling rates. Manifold learning-based models have therefore been explored for image reconstruction in dynamic MRI and aim at learning

the intrinsic low-dimensional structure of the input data that are embedded in a high-dimensional signal space by solving non-linear dimensionality reduction problems. We present in this thesis an alternative strategy for manifold learning-based image reconstruction in dynamic MRI. The proposed method is closely related to machine learning methods that learn manifolds via linear tangent space alignment (LTSA) and can be interpreted as a non-linear generalization of the subspace model. Validation on numerical simulation studies as well as in vivo 2D and 3D cardiac imaging experiments were performed. The proposed method offered improved performances in terms of noise and undersampling artifacts compared to the state-of-the-art techniques.

Finally, research on deep learning-based approaches for MR image reconstruction is ongoing to further improve reconstruction performances. The two first objectives present respectively linear and non-linear based approaches for MR image reconstruction. However, even though the LTSA method is based on a non-linear model, both methods actually use conventional linear optimization techniques to solve the reconstruction problem. Using deep neural networks for optimization has shown promising results in many fields as they procure non-linear and extensive representation power. General concepts are discussed, a state-of-the-art technique is discussed, and early results are presented.

This manuscript is organized as follows. In Chapter 1, we present the key concepts of NMR: where the signal originates, how it can be used for imaging, and how images can be reconstructed from that signal. In Chapter 2, we present the subspace model [170] and how we can leverage spatiotemporal correlation to reconstruct artifact-free images with high spatial and temporal resolution. Two contributions are then presented. First, we show that applying the subspace model to reconstruct high quality dynamic MR images can aid PET applications with PET data motion correction. Secondly, we demonstrate that the subspace model manages to reconstruct 4D volumes (3D + time) from highly undersampled k-space data acquired in a cardiac scan with free-breathing. We also show that the model successfully captures motion and dynamic contrast changes, allowing for 3D cardiac T_1 mapping. Chapter 3 then presents the main contribution of this PhD thesis, that is, a novel image reconstruction method based on linear tangent space alignment developed to overcome the limitations of the linear subspace model. We also provide mathematical connections between this model and state-of-the-art subspace models, and discuss ongoing research to further improve the developed method. Chapter 4 presents current investigations on deep learning-based MR image reconstruction techniques and discussions on future directions. Finally, Chapter 5 concludes on the work done during this PhD thesis, and gives a global outlook and perspectives for the future.

Chapter 1

Background on MRI

This chapter briefly summarizes the physics behind Nuclear Magnetic Resonance (NMR) and how this phenomenon is used for Magnetic Resonance Imaging (MRI). Section 1.1 presents the NMR phenomenon with a spin-based point of view along the Bloch equations that mathematically describe the macroscopic changes in magnetization. Section 1.2 explains how the NMR signal can be used for imaging with spatial encoding, and defines the notion of k-space and MR signal sampling, which lays ground to the principle of MR imaging. Section 1.3 details the basic and advanced image reconstruction techniques in MR imaging in order to recover images from undersampled measurements. Finally, Section 1.4 discusses the limitations of conventional reconstruction techniques along with the challenges in dynamic MRI.

1.1 Nuclear Magnetic Resonance

Historically, the term NMR was first used to describe the phenomenon of stimulated absorption and emission of energy from nuclei placed within a magnetic field. This phenomenon discovery is credited to Isidor Isaac Rabi in the 1930s, who received the Nobel Prize in Physics in 1944 for that discovery. Felix Bloch and Edward Mills Purcell later discovered NMR in liquid and solid materials. NMR was then used for imaging, which gave birth to NMR imaging, later re-branded as Magnetic Resonance Imaging (MRI) in the 1980s due to the negative connotation of the word "nuclear". Due to the highly complicated nature and history of the MR phenomenon, we only give here a brief and compact overview of MR physics for the reader's convenience. In-depth explanation of MR physics can be found in [13, 19, 82, 110, 144].

1.1.1 Introduction to the NMR phenomenon

Magnetic resonance relies on the resonance phenomenon of nuclei possessing a magnetic moment when undergoing a strong magnetic field (e.g., hydrogen protons ^{1}H in most MRI studies, but also ^{19}F, ^{31}P, etc.). Each of these nuclei possesses an intrinsic magnetic moment $\vec{\boldsymbol{\mu}}$ called "spin" aligned

along an axis with random orientation under normal conditions. Due to the randomness of the spin orientation, the macroscopic resulting magnetic moment $\vec{\boldsymbol{M}}$ of a body full of these nuclei is null, i.e., $\sum \vec{\boldsymbol{\mu}} = \vec{\boldsymbol{M}} = \vec{0}$. However, in the presence of a strong magnetic field (denoted as B_0 or $\vec{\boldsymbol{B}}_0$), all the magnetic spins align along its direction either in a parallel or antiparallel fashion. In average, more spins will tend to align in a parallel fashion, as it represents a lower energy level. This leads to an observable macroscopic magnetic moment $\vec{\boldsymbol{M}} \neq \vec{0}$ with a component along the $\vec{\boldsymbol{B}}_0$ axis (see Figure 1.1).

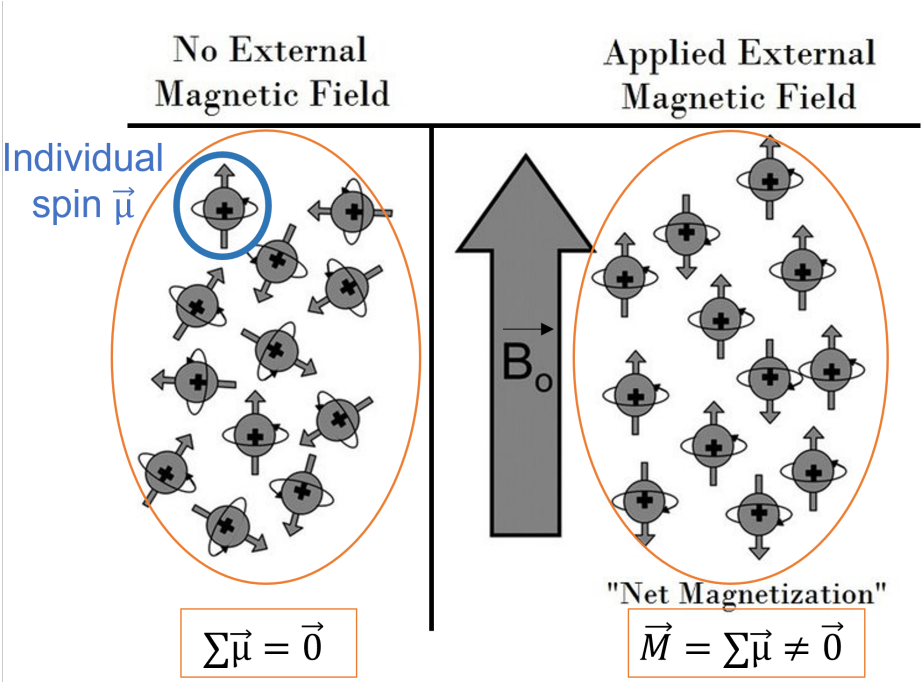


Figure 1.1: Magnetic moment alignment of protons in presence of an external magnetic field [120].

Then, a key concept of NMR is what we call "precession". This can can roughly be explained with a simple analogy, illustrated in Figure 1.2: a spinning gyroscope exhibits angular momentum that makes it resist potential changes in orientation, preventing it from aligning with Earth's gravitational field. Therefore, instead of falling on its side, the gyroscope will rotate on a plane perpendicular to the field it undergoes at a frequency and direction correlated to its spinning velocity, mass, shape and external gravitational field. Similarly, a nucleus precession is described by its gyromagnetic ratio γ, a nuclei-specific quantity related to its spin, mass and size, and the external magnetic field strength $|\vec{\boldsymbol{B}}_0|$. This relationship is mathematically represented by the so-called Larmor frequency $f_0 = \gamma B_0$ (or similarly $w_0 = \frac{\gamma}{2\pi} B_0$ in rad.s^{-1}), i.e., the resonant frequency at which the magnetic spin precesses along $\vec{\boldsymbol{B}}_0$. For example, in the case of a hydrogen nuclei ^{1}H, the gyromagnetic ratio is equal to 42.58 MHz.T^{-1}.

Note that even though individual spins precess on a plan perpendicular to $\vec{\boldsymbol{B}}_0$, the vector sum of the spin angular moment $\vec{\boldsymbol{M}}$ does not have a component on the transversal plane. Despite the quantum behavior of the individual spin angular moment, once summed and averaged over many

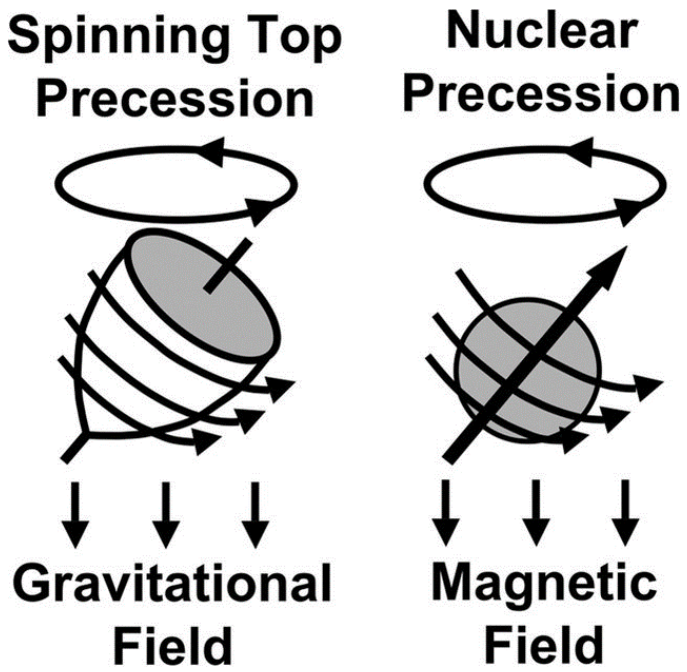


Figure 1.2: Precession of a spinning gyroscope and nuclear precession. They are both similar in the sense that an external force induces precession due to their inherent spinning energy [120].

of them, the macroscopic resultant ($\vec{\boldsymbol{M}}$) follows the rules of classical physics. As such, $\vec{\boldsymbol{M}}$ does not precess and is aligned with the external field $\vec{\boldsymbol{B}}_0$ when at equilibrium. However, when the equilibrium is disrupted with external energy input, we will see in Section 1.1.2 that $\vec{\boldsymbol{M}}$ can also precess. Upon defining the z axis as $\vec{\boldsymbol{z}} = \frac{\vec{\boldsymbol{B}}_0}{|\vec{\boldsymbol{B}}_0|}$, the global magnetic moment (or magnetization) of an excited body is then commonly expressed as a combination of two terms: a longitudinal ($\vec{\boldsymbol{M}}_z$) and a transversal ($\vec{\boldsymbol{M}}_{xy}$) term:

$$\vec{\boldsymbol{M}} = \vec{\boldsymbol{M}}_z + \vec{\boldsymbol{M}}_{xy} \tag{1.1}$$

where $\vec{\boldsymbol{M}}_{xy} = \vec{0}$ at equilibrium.

1.1.2 Radio-frequency pulse excitation

In practice, measuring the magnetization $\vec{\boldsymbol{M}}$ is difficult at equilibrium due to its negligible order of magnitude compared to the external magnetic field $\vec{\boldsymbol{B}}_0$. However, it is possible to indirectly measure it by "flipping" the magnetization to the transversal plane using a radio frequency (RF) pulse. To do so, we apply an oscillating magnetic field $\vec{\boldsymbol{B}}_1$, oscillating at the Larmor frequency ω_0 (or equivalently f_0) in a plane perpendicular to $\vec{\boldsymbol{B}}_0$. This will make the spins resonate and change their energy level, which as a resultant will create phase coherence of individual spins and thus transversal magnetization $\vec{\boldsymbol{M}}_{xy} \neq \vec{0}$.

As long as the RF pulse $\vec{\boldsymbol{B}}_1$ is applied, the macroscopic magnetization $\vec{\boldsymbol{M}}$ will rotate along an axis perpendicular to $\vec{\boldsymbol{B}}_0$ and form an angle θ with $\vec{\boldsymbol{B}}_0$ called "flip angle". $\vec{\boldsymbol{M}}$ is then out of equilibrium due to the external energy absorption caused by $\vec{\boldsymbol{B}}_1$ and will therefore try to get back

to equilibrium. Doing so, it will start precessing around $\vec{B}_0$ (i.e., in the transverse plane), as shown in Figure 1.3.

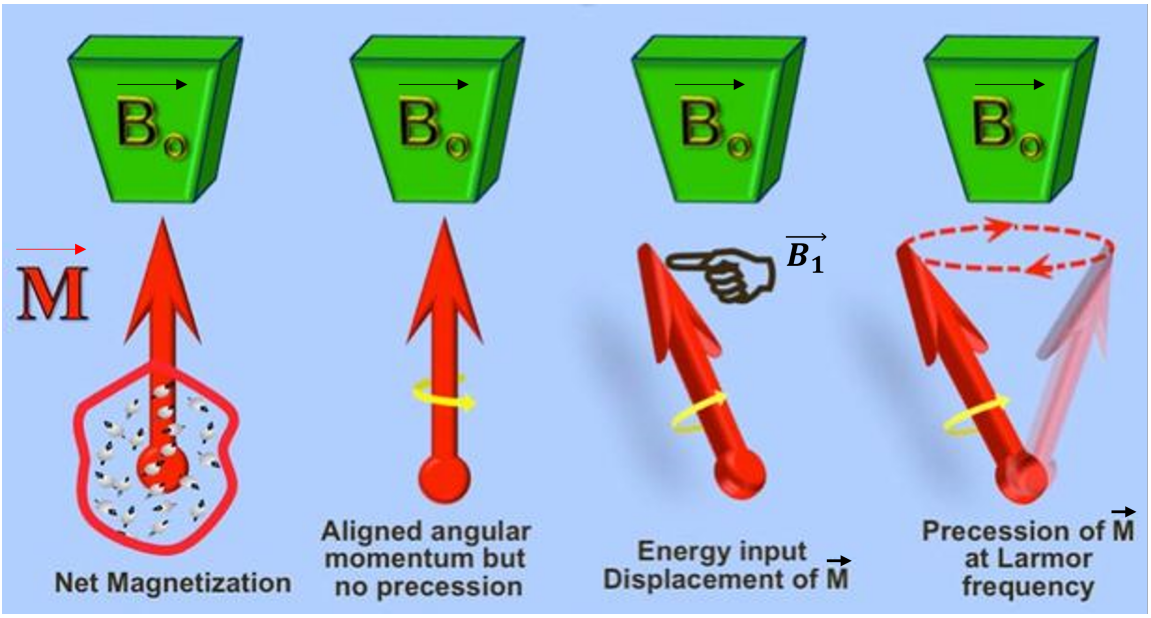


Figure 1.3: Net Magnetization $\vec{M}$ can be tipped out of equilibrium with an oscillating magnetic field $\vec{B}_1$, rotating at the Larmor. Once tipped off, $\vec{M}$ precesses until reaching equilibrium back. Courtesy of Allen D. Elster, MRIquestions.com

1.1.3 Bloch equations and free-induction decay

The NMR phenomenon was mathematically described by Felix Bloch in 1946 [9], where he presented the so-called Bloch equations. Instead of the more common quantum point of view of the individual spins, Bloch hypothesized that the system — potentially composed of millions of spins, could actually be seen as a single macroscopic vector: the magnetization $\vec{M}$. This vector behaves like a regular vector in the sense of classical physics (i.e., unlike quantum physics) and has three spatial components. Bloch equations then tell us that $\vec{M}$ is actually a function of time $\vec{M}(t) = (M_x(t), M_y(t), M_z(t))$ directly related to the gyromagnetic ratio γ, the external magnetic field the nuclei undergoes $\vec{B}(t) = \vec{B}_0 + \vec{B}_1(t)$, and characteristic constants of the excited body called relaxation times (T_1 and T_2) that represent respectively the longitudinal regrowth and transversal decay of the magnetization:

$$\frac{d}{dt}\begin{pmatrix} M_x \\ M_y \\ M_z \end{pmatrix} = \begin{pmatrix} -\frac{1}{T_2} & \gamma B_z & -\gamma B_y \\ -\gamma B_z & -\frac{1}{T_2} & \gamma B_x \\ \gamma B_y & -\gamma B_x & -\frac{1}{T_1} \end{pmatrix} \begin{pmatrix} M_x \\ M_y \\ M_z \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{M_0}{T_1} \end{pmatrix} \tag{1.2}$$

These equations allow us to predict the behavior of the NMR signal, and have simple close form solutions in some trivial cases. Take the example of a sample with homogeneous magnetic field $\vec{B}_0$

to which is applied an RF pulse with a flip angle $\theta = 90$ degrees. Denote M_0 the original maximum value of the magnetization $\vec{\boldsymbol{M}}$, a quantity proportional to the proton density. The solution of the Bloch equations then predicts the following magnetization over time:

$$M_x(t) = M_0 \exp\left(-\frac{t}{T_2}\right)\sin(\omega t) \tag{1.3}$$

$$M_y(t) = M_0 \exp\left(-\frac{t}{T_2}\right)\cos(\omega t) \tag{1.4}$$

$$M_z(t) = M_0(1 - \exp\left(-\frac{t}{T_1}\right)) \tag{1.5}$$

This means that upon applying a 90 degrees RF pulse, the magnetization $\vec{\boldsymbol{M}}$ is flipped and gradually returns to equilibrium by precessing on a plane perpendicular to $\vec{\boldsymbol{B}}_0$ (Equations (1.3) and (1.4)) while growing back its transverse component following a negative exponential trend with time constant T_1. Note that the magnitude of the transverse magnetization $M_{xy}(t) = M_0 \exp\left(-\frac{t}{T_2}\right)$ decays in an exponential trend with time constant T_2. This is illustrated in Figure 1.4.

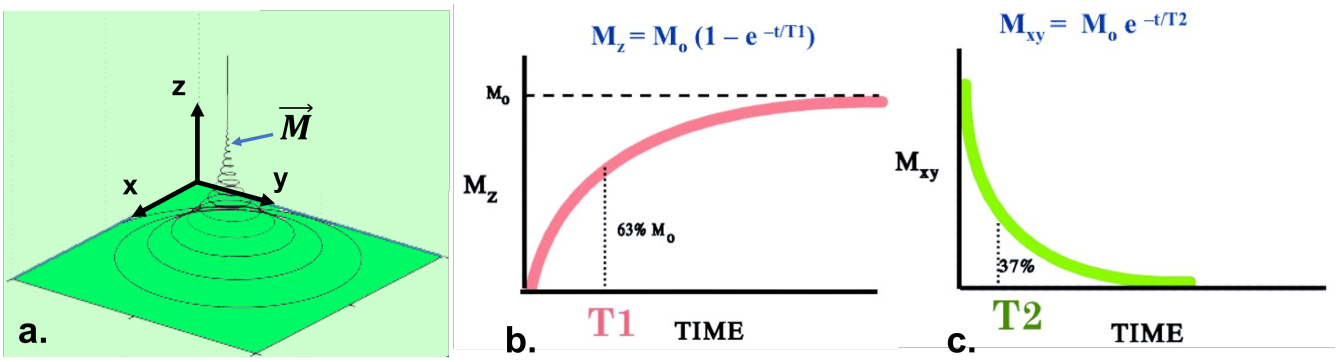


Figure 1.4: Visualization of $\vec{\boldsymbol{M}}$ during precession with a 90 degrees flip angle. **(a.)** 3D Plot of the trajectory of $\vec{\boldsymbol{M}}$ along time during precession. **(b.)** Plot of the longitudinal magnetization decay $\vec{\boldsymbol{M}}_{xy}$ along time. **(b.)** Plot of the transversal magnetization regrowth of $\vec{\boldsymbol{M}}_z$ along time. Figures modified from MRIquestions.com

T_1 is also called the spin-lattice relaxation time or sometimes thermal relaxation time, as this time constant is related to the energy transfers from the spins to their environment (i.e., the lattice). Similarly, T_2 is also called the spin-spin relaxation time and is not always related to energy transfer like T_1 as it can occur with spin dephasing caused by local field homogeneity and the neighboring spins. In short, most phenomenon inducing T_1 relaxation will also cause T_2 relaxation, while T_2 relaxation may also appear without T_1 relaxation.

As a side note, it should be noted that the magnetization signal usually decays faster than the expected T_1 and T_2, yielding shorter apparent T_1 and T_2. This effect can be mostly caused by B_0 field inhomogeneity but also potential magnetic susceptibility differences within the sample inducing fields distortions. In the case of T_2, the shorter apparent relaxation time, denoted by T_2^* (T_2 "star"),

is related to T_2 in the following way:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_{2_{in}}} \tag{1.6}$$

where $T_{2_{in}}$ is the contribution of all the potential inhomogeneities to the transverse relaxation.

The decaying signal in the transverse plane (Equations (1.3) and (1.4)) has exponential envelope (related to T_1 and T_2), and sinusoidal carrier wave: this is what we actually measure in MRI, and more globally in NMR. This signal is called free-induction decay (FID) and corresponds to the transverse magnetization signal induced by an RF pulse excitation. Most tissues and/or liquids in our body possess distinguishable relaxation times T_1 and T_2, which therefore allow differentiation of their signals.

1.2 Magnetic Resonance Imaging

Traditional MRI mostly focuses on the magnetic resonance of hydrogen nuclei (i.e., protons) as they are the most prominent in the human body and therefore yield the highest sensitivity. Even though other nuclei can be excited by changing the frequency of the excitation pulse for example, we will focus here on traditional ^{1}H MRI. A brief RF pulse (typically less than a few milliseconds) rotates $\vec{\boldsymbol{M}}$ in the transverse plane by a degree θ. The magnetization then starts precessing, creating an alternate magnetic field in the transverse plane that produces an alternate voltage in presence of a coil sensor (induction).

1.2.1 From magnetization to voltage

Coil receivers are used to capture the oscillating magnetic signal and transform it into an electrical signal as per the Faraday Law. The electromagnetic signal produced in the coil receiver is then derived from the time-variation of the magnetic flux going through its surface. This will form an alternate voltage signal $\tilde{s}(t)$ with a carrier wave frequency being the Larmor frequency ω_0. For example, in the case of a 1 Tesla magnetic field B_0, the electrical signal will oscillate around 42.58 MHz within a relatively small bandwidth. Therefore, instead of directly measuring the incoming signal from the coil(s), we actually perform complex demodulation of the real-valued signal $\tilde{s}(t)$ to re-center it on zero. Define a small spatial variation in the magnetic field $\boldsymbol{\delta}\vec{\boldsymbol{B}}$. In the case of a FID, and upon defining $\omega = \gamma\big(B_0 + \delta B(x,y,z,t)\big) = \omega_0 + \delta\omega$, the demodulated signal $s(t)$ is then:

$$s(t) \propto \int_{\mathcal{V}} M(x,y,z)\omega_0 \, \exp\Big(-\frac{t}{T_2}\Big) \exp\Big(-i\delta\omega t\Big) dx\, dy\, dz \tag{1.7}$$

where $\mathcal{V}$ is the volume of the excited sample. An important note from Equation (1.7) is that both M and ω_0 are linearly dependent on the external magnetic field strength $B = B_0$, which means that the voltage is proportional to B^2 and therefore the sensitivity of the signal too. Having higher values for B is therefore favorable, at the expense of potential other issues such as field inhomogeneity.

However, the signal captured is not localized at this point: it comes from the whole sample within the external magnetic field and does not contain spatial information. Forming an image out of the NMR signal therefore requires some sort of "spatial encoding" to discriminate the location of the signal.

1.2.2 Spatial encoding of the NMR signal

In 1973, Paul Lauterbur and Peter Mansfield found a way to get spatial encoding in the NMR signal and therefore produce an image out of it. They were later awarded the 2003 Nobel Prize for their contribution in what is nowadays known as Magnetic Resonance Imaging. In a nutshell, in addition to the magnetic field $\vec{\boldsymbol{B}}_0$, they used a spatially-varying magnetic field. Therefore, as the magnetic field is different between two spatial locations, the spatial gradient field $\vec{\boldsymbol{G}}(x, y, z, t) = \nabla \vec{\boldsymbol{B}}(x, y, z, t)$ is now non-zero in magnitude. Because of this gradient, the magnetic field $\vec{\boldsymbol{B}}(x, y, z, t)$ perceived by the spins throughout the sample will be different based on their spatial location, and their precessing frequency $\omega(x, y, z, t) = \gamma B(x, y, z, t)$ will also be different: their spatial location is encoded by the gradient (Figure 1.5).

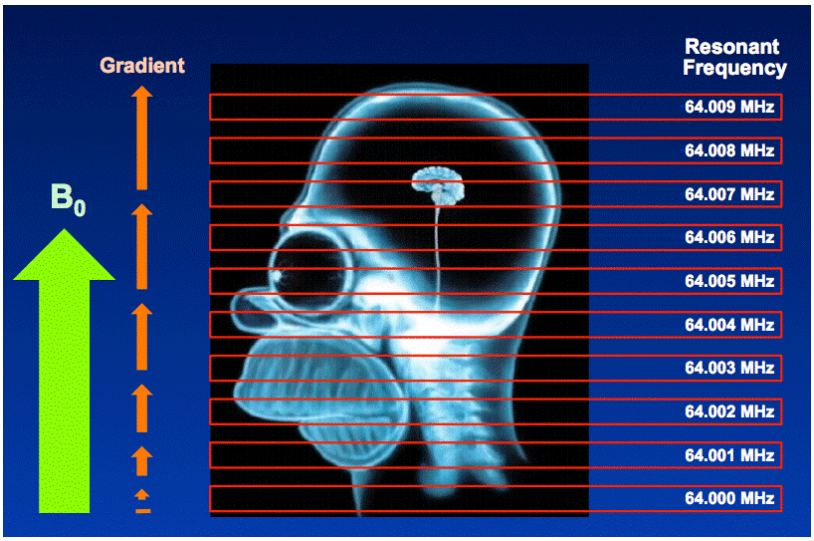


Figure 1.5: Spatial encoding with magnetic field gradient. The spatial location is encoded by the varying magnetic field strength due to a different resonating frequency at different spatial locations in the sample. Courtesy of Allen D. Elster, MRIquestions.com

Without loss of generality, we define the small spatial variation in the magnetic field $\delta B(x, y, z)$ varying according to a linear magnetic gradient in the spatial directions:

$$\boldsymbol{\delta}\vec{\boldsymbol{B}}(x, y, z) = \begin{bmatrix} G_x x \\ G_y y \\ G_z z \end{bmatrix} \tag{1.8}$$

At a location (x, y, z), $\delta\omega$ defined in Equation (1.7) can therefore be expressed as $\delta\omega = \gamma(G_x x +$

$G_y y + G_z z$). Equation (1.7) can then be written:

$$s(t) \propto \int_{\mathcal{V}} \rho(x, y, z, t) \exp\Big(- i\gamma(G_x x + G_y y + G_z z)t\Big) dx\, dy\, dz, \tag{1.9}$$

where $\rho(x, y, z, t) = \omega_0 M(x, y, z) \exp\Big(- \frac{t}{T_2}\Big)$ in the case of a FID.

In addition, we introduce a spatial location vector $\mathbf{r} = (x, y, z)$ and another vector denoted $\mathbf{k}$ related to the gyromagnetic ratio and the gradient field:

$$\mathbf{k}(t) = \frac{\gamma}{2\pi} \begin{bmatrix} G_x t \\ G_y t \\ G_z t \end{bmatrix}, \tag{1.10}$$

Equation (1.9) can finally be further simplified to:

$$s(\mathbf{k}, t) \propto \int_{\mathcal{V}} \rho(\mathbf{r}, t) \exp\big(- i2\pi \mathbf{k} \cdot \mathbf{r}\big) d\mathbf{r} \tag{1.11}$$

One can notice that Equation (1.11) links $s(\mathbf{k}, t)$ as being the Fourier transform (denoted $\mathcal{F}$) of the quantity $\rho(\mathbf{r}, t)$, i.e.,

$$s(\mathbf{k}, t) \propto \mathcal{F}\{\rho(\mathbf{r}, t)\} \tag{1.12}$$

An important conclusion from this relationship is that the vector $\mathbf{k}$ represents a spatial frequency in a space dual to the image space represented by $\mathbf{r}$. The consequence of this conclusion is that the signal retrieved from the coil receiver is actually in the spatial frequency domain and not in the image domain. The outputted MR signal is therefore commonly called k-space as it maps the spatial frequency space ($\mathbf{k}$) rather than the original image space ($\mathbf{r}$). As $\mathbf{k}$ and $\mathbf{r}$ are dual variables linked by the Fourier transform, controlling the trajectory of $\mathbf{k}$ using gradients (Equation (1.10)) also gives us some control on the spatial location of the signal, hence the denomination "Fourier encoding" for this kind of spatial encoding.

Slice, phase, and frequency encoding

In MRI, k-space sample locations are commonly controlled using gradients. Applying gradients will move the vector $\mathbf{k}$ to another location in the spatial frequency space (i.e., k-space) and a signal can be measured after RF pulse for different $\mathbf{k}$-space locations. The spatial encoding is commonly done in three step: slice selection, frequency encoding and phase encoding.

During the slice selection, a linear gradient G_z is applied. This means that spins in the sample will have different precession frequency, linearly dependent on their position along the z axis. An RF pulse with limited bandwidth will be able to excite a specific slice with specific slice thickness related to the gradient strength. For simplicity, let us assume we excite a 2D slice, i.e., after slice selection the spin excitation is purely in-plane in the x and y directions. Frequency and phase

encoding then commonly refer to the steps where gradients are applied along the x and y directions respectively (G_x and G_y). These gradients will encode the spatial location of the MR signal: voxels within the sample will have different phases and frequencies in the x and y directions due to the different apparent magnetic field.

MR sequences: Gradients and k-space location

One can also conceptually understand that applying these two gradients is the same as moving in the k-space of a value equal to the strength of the gradient multiplied by the elapsed time (assuming G_x and G_y are constant), as indicated by Equation (1.10). Coincidentally, if the gradients are not constant, the term $G_i t$, $i = [x, y, z]$ is replaced by $\int_0^t G_i(t')\,dt'$ in Equation (1.10). For example, applying a negative gradient for a duration Δt along the x axis is equivalent to move in the k-space of a value $\frac{\gamma}{2\pi} G_x t$ along the x axis in the spatial frequency domain (denoted k_x). The data samples are then acquired after an excitation pulse, mapping the k-space using gradients according to an MR sequence (Figure 1.6).

The aim in MRI is then to acquire MR signal samples at as many spatial frequencies as possible (k_x and k_y) to fill the k-space and obtain an image after Fourier transform. Note that the reconstructed image properties (such as resolution and field-of-view (FOV)) will depend on the k-space measurements such as the maximum frequency value k_{max} and the sampling interval Δk.

Two important sequence parameters controlling the contrast and acquisition speed are the echo time (TE) and the repetition time (TR) (see Figure 1.6). TE is defined as the time between the middle of the RF pulse and the middle of the echo signal. For multiple echos, we commonly use TE1, TE2, etc. When the RF spin echo and gradient echo are not coincident in time, TE refers to the time of the gradient spin echo [4]. TR is defined as the period of time between the beginning of a pulse sequence and the beginning of the succeeding pulse sequence [4]. Note that in the current and previous sections, the concept of the MR signal is explained using a simple excitation case with a single RF pulse: the free-induction decay. However, using multiple RF pulses or combining them with gradients is also possible. In particular, widely used MR imaging techniques such as Gradient-echo (GRE) and Spin-echo (SE) can be produced using respectively an RF pulse with gradient rephasing and 2 RF pulses (2^{nd} pulse one for spin rephasing) [62].

1.2.3 Signal and noise model in dynamic MRI

Equation (1.11) in the previous section presents the relationship between the acquired signal and the imaged object, but does not take into account potential noise from the MR signal. The noise of the MR signal is governed by a Rician distribution [44]. It is however commonly approximated to Gaussian noise when the signal-to-noise ratio (SNR) is above 2 [44]. In addition, for simplicity purpose, we will consider the two quantities to be equal instead of proportional. The signal model

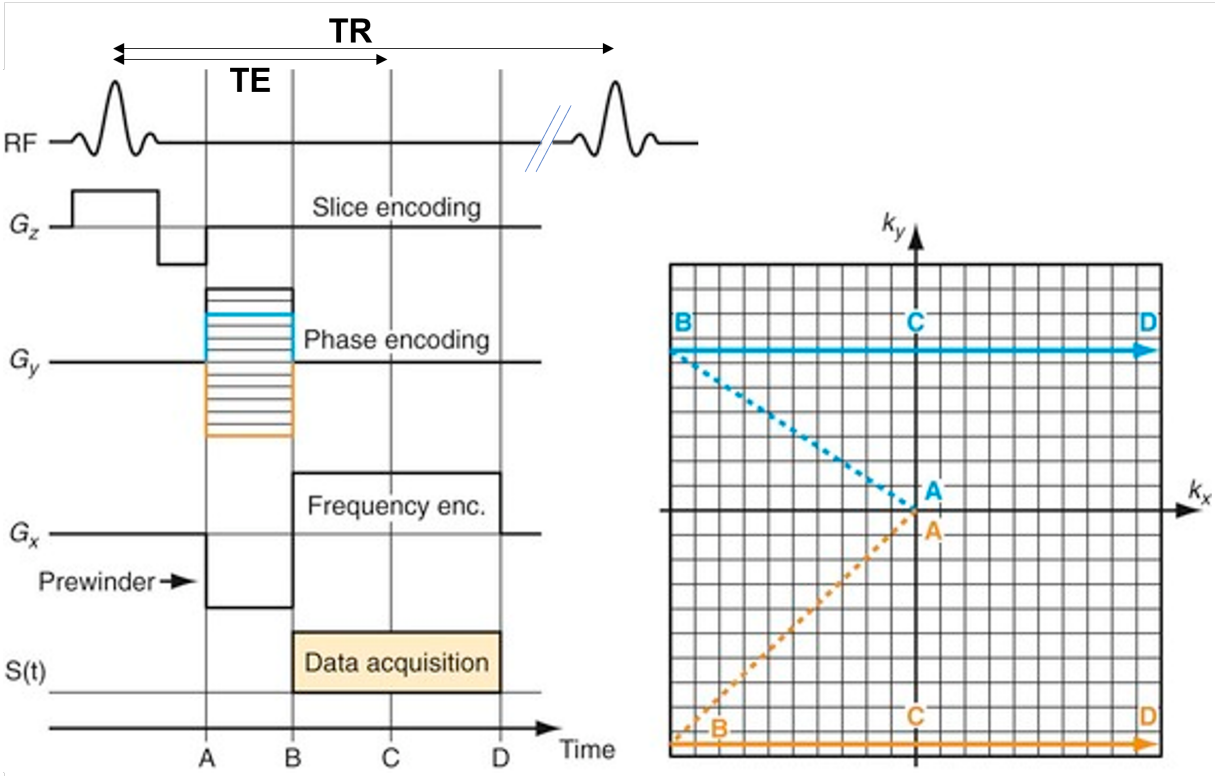


Figure 1.6: Basic MR sequence for 2D coverage of the k-space. On the left, the MR sequence with RF pulse and gradients application. On the right, the equivalent k-space trajectory for time-points A, B, C and D. The prewinding gradient is required to start the acquisition with negative spatial frequencies. The operation is then repeated for several phase-encoding gradient intensity (blue line) to cover the whole k-space. Courtesy of radiologykey.com

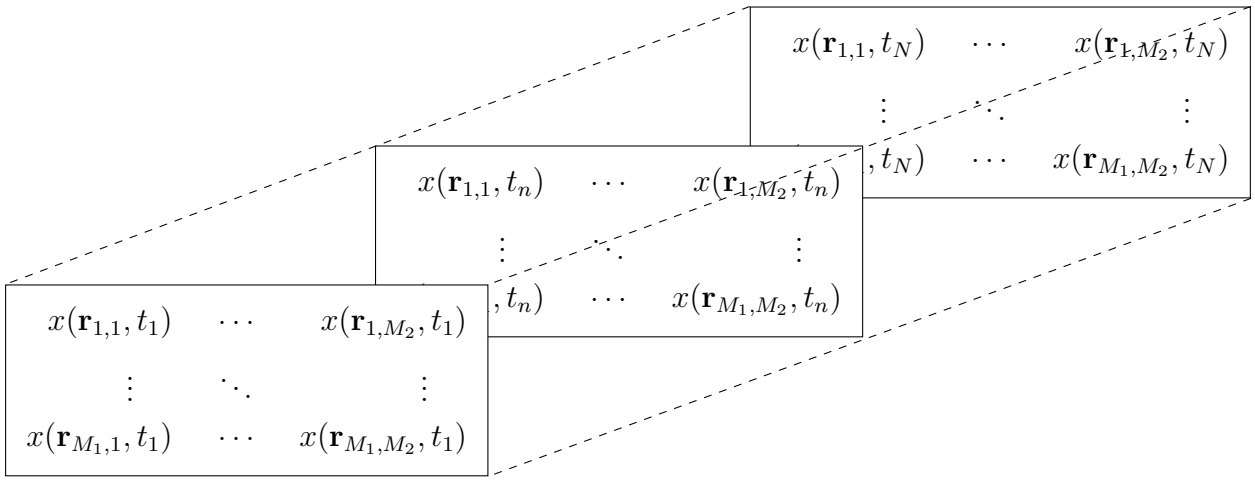
of the measured data at k-space location $\mathbf{k}$ and time t can then be written as:

$$s(\mathbf{k}, t) = \int_{\mathcal{V}} x(\mathbf{r}, t) e^{-i2\pi \mathbf{k}\cdot\mathbf{r}} d\mathbf{r} + \varepsilon(\mathbf{k}, t), \tag{1.13}$$

where $x(\mathbf{r}, t)$ denotes the image intensity at location $\mathbf{r}$ and time t in volume $\mathcal{V}$ and $\varepsilon(\mathbf{k}, t)$ is the measurement noise (assumed to be identical and independent Gaussian noise). Note that in the rest of this manuscript, the imaged object at a location $\mathbf{r}$ and time t will be denoted by $x(\mathbf{r}, t)$ and not $\rho(\mathbf{r}, t)$. This choice is purely practical and is motivated by the usual notations in optimization and image reconstruction, where the unknown matrix to recover is commonly denoted by $\mathbf{X}$.

The MR signal in Equation (1.13) represents a continuous signal in the k-space. Just like the input image space, the k-space is infinite and only a specific part of it is interesting to the user: its center, where most of the energy is. However, in terms of signal processing, we would need to cover the whole k-space for all frequencies at an infinite precision to obtain a perfect representation of the imaged object. To remedy this problem, we actually use a discretized representation of the imaged object (denoted by $\mathbf{X}$) and discretized k-space measurement (denoted by $\mathbf{s}$). In the discretized case,

the dynamic image series formed by $x(\mathbf{r}, t)$ is represented as follows:



where M_1 and M_2 are the vertical and horizontal sizes of the image, and N is the number of temporal frames. For mathematical practicality, the discretized images are vectorized, i.e., an image of size $M_1 \times M_2$ is represented as a vector $M \times 1$ where $M = M_1 M_2$, and the image vectors at different time points are concatenated along rows. This 2D matrix representing the image series is commonly called the Casorati matrix:

$$\mathbf{X} = \begin{bmatrix} x(\mathbf{r}_1, t_1) & \dots & x(\mathbf{r}_1, t_N) \\ \vdots & \ddots & \vdots \\ x(\mathbf{r}_M, t_1) & \dots & x(\mathbf{r}_M, t_N) \end{bmatrix}, \tag{1.14}$$

where $\mathbf{X} \in \mathbb{C}^{M \times N}$, M is the number of voxels, and N is the number of temporal frames.

The discretized MR signal is obtained from the continuous NMR signal sampled throughout time, aiming at filling the k-space "grid" defined by the spatial frequency step Δk and maximum spatial frequency k_{max}. The image is then recovered using the one-to-one mapping between the k-space samples and imaged object: the Fourier transform (see Figure 1.7).

Accordingly, using the relationship in Equation (1.12), the signal model in Equation (1.13) can be expressed in the discretized case as:

$$\mathbf{S} = \mathbf{F}_s \mathbf{X} + \mathbf{E}, \tag{1.15}$$

where $\mathbf{S} \in \mathbb{C}^{P \times N}$ is a matrix containing the measured k-space data at P frequency locations for the N frames, $\mathbf{F}_s$ denotes the spatial domain Discrete Fourier Transform (DFT), and $\mathbf{E}$ contains the noise of the measured data. Using this relationship, the image $\mathbf{X}$ can be obtained from the measured k-space data using an inverse DFT: this process is called image reconstruction.

Equation (1.15) represents a case where we fully sample the k-space, i.e., we acquire all the k-space locations needed at a rate above the Nyquist rate to be able to directly recover $\mathbf{X}$ from

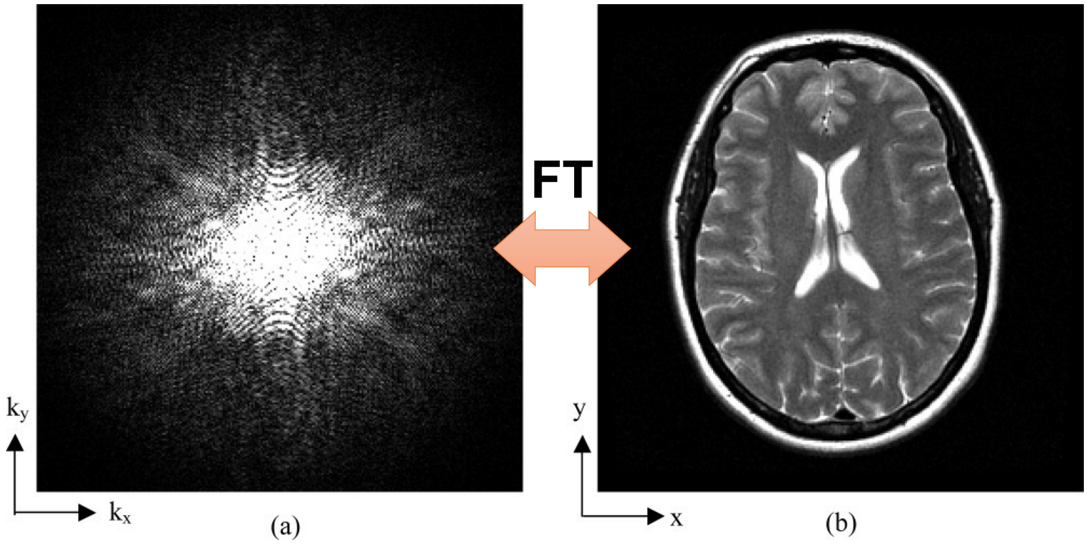


Figure 1.7: **(a)** Discretized 2D MR signal outputted from the scanner. The raw data shown here are denoted as k-space measurement. **(b)** Corresponding image data in the image space. The mapping between the two spaces is the Fourier transform (FT). As k-space measurements are complex numbers, both **(a)** and **(b)** are shown in terms of magnitude. Original figure from [111].

$\mathbf{S}$ with a DFT. However, in practice, scans are limited by the acquisition time and acquiring all the k-space locations can prove to be challenging. The acquired k-space data is therefore often undersampled. In such a case, the discretized signal model is then:

$$\mathbf{s} = \boldsymbol{\Omega}\left(\mathbf{F}_s\mathbf{X}\right) + \boldsymbol{\varepsilon}, \tag{1.16}$$

where $\mathbf{s} \in \mathbb{C}^{K\times 1}$ contains the vectorized undersampled k-space measurements, $\boldsymbol{\Omega}$ denotes a sampling mask operator in the (k,t)-space, and $\boldsymbol{\varepsilon}$ is a vector containing the noise of the measured data. Image reconstruction can be performed solving Equation (1.16) using the least-square method:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \left\|\boldsymbol{\Omega}\left(\mathbf{F}_s\mathbf{X}\right) - \mathbf{s}\right\|_2^2, \tag{1.17}$$

The signal model presented in Equation (1.16) is in the case of single-coiled measured data. If multiple coils are available for image acquisition, a data vector $\mathbf{s}$ is acquired from each of them, forming a multicoil array $[\mathbf{s}_1, \ldots, \mathbf{s}_{n_{coil}}]$, where n_{coil} is the number of coil receivers. In such a case, images can be reconstructed separately for each coil data $\mathbf{s}_i, i = 1, \ldots, n_{coil}$, forming n_{coil} reconstructed images $\mathbf{X}_i$. A magnitude dynamic image series $\mathbf{X}$ can then be generated by combining the reconstructed dynamic images of each coil using the sum-of-square (SoS) method:

$$\mathbf{X} = \sqrt{\sum_{i=1}^{n_{coil}} |\mathbf{X}_i|^2}, \tag{1.18}$$

where $|\bullet|$ is the element-wise absolute value and $\bullet^2$ is the element-wise square function.

1.2.4 Sampling of the MR signal

When performing an MR scan, various k-space locations are sampled throughout time. As explained in the previous section, the k-space locations correspond to spatial frequencies: the center of the k-space represents the low frequency components while the edges represent the higher frequency components. As such, most of the energy of the signal is located in the center of the k-space, which determines the overall contrast and basic shapes of the image, while the edges of the k-space encode the image details and sharpness [99, 111]. It is therefore crucial to acquire samples in the center of the k-space (k-space without its center would only show edges [29, 99]), without neglecting the edges to obtain reasonable image details. To do so, many sampling trajectories are possible, each of them with pros and cons.

Cartesian k-space sampling

Cartesian sampling refers to the sampling of k-space points on a Cartesian grid. This is done by acquiring data samples at a fixed phase encoding gradient G_y while varying the frequency encoding gradient G_x. Data points are sampled regularly along the frequency direction (also called readout direction). In general, one k-space line along the readout direction is acquired per excitation pulse. An example of k-space measurements sampled on a Cartesian grid is shown in Figure 1.8. However, depending on the imaged object and the desired resolution/FOV, acquiring one readout line per TR might be too slow. In such cases, some fast MR imaging techniques exist such as Echo planar imaging (EPI) to cover larger k-space zones with a single RF pulse. One of the main advantages of Cartesian sampling is that the data collection is uniform, there is therefore no need to compensate for its density when performing image reconstruction. Additionally, efficient reconstruction algorithms such as the Fast Fourier Transform (FFT) can be used.

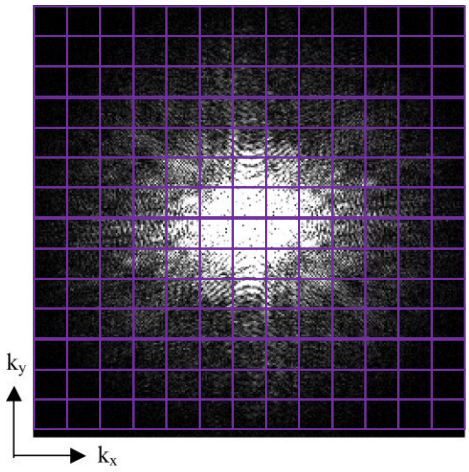


Figure 1.8: Discretized k-space acquired using Cartesian sampling trajectory. Every purple square in the k-space represents one measurement on the Cartesian grid.

Non-Cartesian k-space sampling

Despite the efficiency of the Cartesian sampling for image reconstruction using the FFT, other sampling trajectories such as radial or spiral are getting more popular by the day. Indeed, as explained above, most of the signal energy is located in the center of the k-space, and that location also encodes the basic shapes and contrast of the image while the edges of the k-space have lower energy and are more prone to noise. Despite these facts, we still want to acquire signal at both locations. Yet, it could be beneficial to acquire data in a way such that the center of the k-space is more sampled than the edges. Non-uniform sampling of the k-space therefore seems to be of interest as classical Cartesian sampling does not offer this flexibility.

Most common non-Cartesian sampling trajectories are radial and spiral (see Figure 1.9). With these trajectories, both frequency and phase encoding gradients are turned on, creating a spiral or radial trajectory. Some of the main advantages with these techniques are their relatively low sensitivity to motion artifacts, incoherently distributed aliasing artifacts, and higher SNR due to the heavily sampled center of the k-space at regular rate [29, 111]. However, due to their non-uniform sample density, reconstruction is computationally heavier than with Cartesian sampling (the FFT algorithm cannot be utilized here) and due to gradients delays and distortions, the exact position of the k-space samples may be different than expected, creating artifacts.

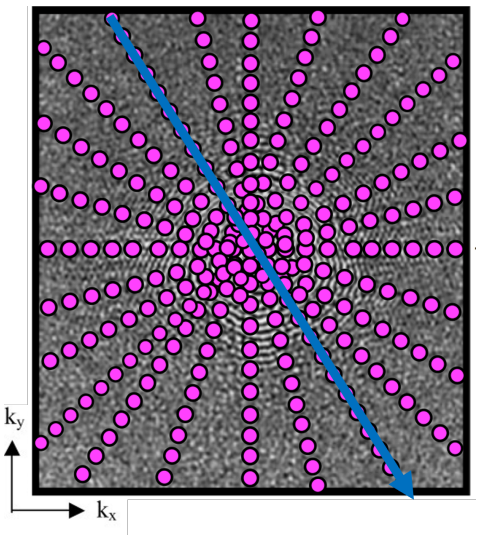


Figure 1.9: Example of a non-Cartesian sampling trajectory. Here, k-space measurements are acquired using a radial sampling trajectory. Every pink dot in the k-space represents one measurement along the radial trajectory indicated by the blue line. Using this sampling trajectory, the center of the k-space is denser than the edges (i.e., more samples in the center). Courtesy of Allen D. Elster, MRIquestions.com

1.3 Magnetic Resonance image reconstruction

Applying the Fourier transform to the measured k-space data, along some potential advanced optimization schemes following Equations (1.15) and (1.16), is called image reconstruction. We present

here some state-of-the-art techniques for MR image reconstruction.

1.3.1 Non-iterative image reconstruction techniques

Cartesian k-space sampling: Fast Fourier Transform (FFT)

As presented in Section 1.2.4, k-space samples acquired following a Cartesian trajectory can be reconstructed using a Fast Fourier transform, especially fast for matrix sizes being a power of 2. When the k-space is fully sampled, i.e., above the Nyquist rate, the image array is obtained with a simple FFT. With the current available CPU machines, applying the FFT to a 256×256 images is in the order of the millisecond (or less), without considering potential acceleration from GPUs.

However, when the k-space is not fully sampled, i.e., there is a sampling mask Ω as in Equation (1.16), applying the FFT on the undersampled k-space will produce aliasing artifacts. These artifacts are mostly dependent on the k-space trajectory and on which samples are missing. In the case of Cartesian trajectory, the undersampling is mostly applied on the phase direction (generally y axis) as it corresponds to the bottleneck in terms of acquisition time (the frequency encoding direction is almost "free" in terms of acquisition time). In such a case, the artifacts will correspond to aliasing along the phase-encoding direction, as shown in Figure 1.10c. As fewer samples are acquired in the phase-encoding direction, the FOV will be smaller in that direction. Note that zero-filled image reconstruction can also be performed to get an image of same FOV as the fully sampled, yet still with heavy aliasing artifacts.

Non-Cartesian k-space sampling: Gridding and Non-Uniform FFT (NuFFT)

In the case of non-Cartesian k-space trajectory, we cannot rely on the FFT for the reconstruction of the image as the samples are not uniformly distributed anymore. There are mainly three alternatives [113]: the first one is using specific trajectories (e.g., spiral) with specific angular criteria to use pre-existing algorithms such as the projection reconstruction from X-ray CT. Another technique, a little more common, is to re-sample the non-uniform samples $\mathbf{s}$ on a Cartesian grid using interpolation kernels to produce new uniformly distributed samples $\tilde{\mathbf{s}}$. The FFT can then be applied on $\tilde{\mathbf{s}}$ to reconstruct images as in Section 1.3.1 – Cartesian case. This process is called gridding, and can be seen as a pre-processing step on the original input data that produces a new input easier to reconstruct with more conventional algorithms (i.e., the FFT). Lastly, the third and most popular option is to use a fast Fourier transform algorithm adapted to non-uniform sampling patterns. Notably, due to the growing interest and need to perform fast Fourier transform on arbitrarily distributed samples, many fast interpolation algorithms have been developed in the late 1990s and early 2000s [32, 41]. These algorithms, commonly referred to as Non-uniform FFT (NuFFT), rely on one step of interpolation (gridding) and one step of FFT. Despite their flexibility, these algorithms require more computation time than the FFT (possibly 10-50 times) due to the interpolation step, which is one of the main drawbacks when dealing with high dimensional data.

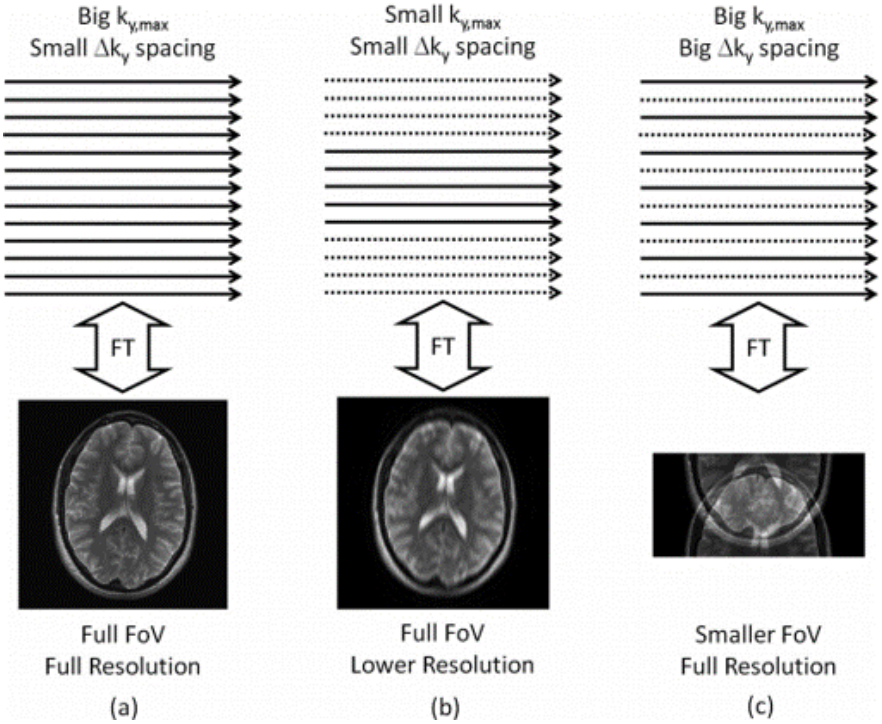


Figure 1.10: (a) full-resolution image with samples covering the k-space at the Nyquist rate. **(b)**: Undersampling of the edges of the k-space. Reducing $k_{y,max}$ maintains the FOV but decreases the image resolution. **(c)** Regular undersampling in the phase-encoding direction. Increasing Δk_y while holding $k_{y,max}$ constant maintains image resolution but decreases the FOV, resulting in spatial aliasing artifacts in the corresponding image [28].

Figure 1.11 shows the NuFFT reconstructed image from undersampled k-space data with radial trajectory. Due to the radial pattern of the sampling trajectory, the aliasing occurring is different from the Cartesian case, causing streaks along the angles of the missing spoke angles (i.e., missing projections).

Note that in the case of a non-iterative reconstruction scheme, applying the gridding method or the NuFFT is strictly equivalent. However, in the case of iterative reconstruction schemes (see Section 1.3.2) a difference arises. When using the gridding method, the input is $\tilde{\mathbf{s}}$ and $\mathbf{F}_s$ in Equation (1.16) corresponds to the FFT. Yet, when reconstructing using the NuFFT, the input data vector $\mathbf{s}$ is not pre-processed. More precisely, the input of the reconstruction scheme is the original input data $\mathbf{s}$ and the NuFFT is applied as long as we iterate (i.e., $\mathbf{F}_s$ in Equation (1.16) corresponds to the NuFFT).

1.3.2 Iterative image reconstruction techniques

Despite numerous advantages such as versatility and spatial resolution, MR imaging is inherently a modality with slow imaging speed. The acquired dynamic MR data are therefore often highly undersampled. A key challenge in dynamic MRI is then reconstructing high-dimensional images from sparse (k, t)-space data sampled below the Nyquist sampling rate. Henceforth, several MR image reconstruction schemes based on different types of assumptions have been developed as a

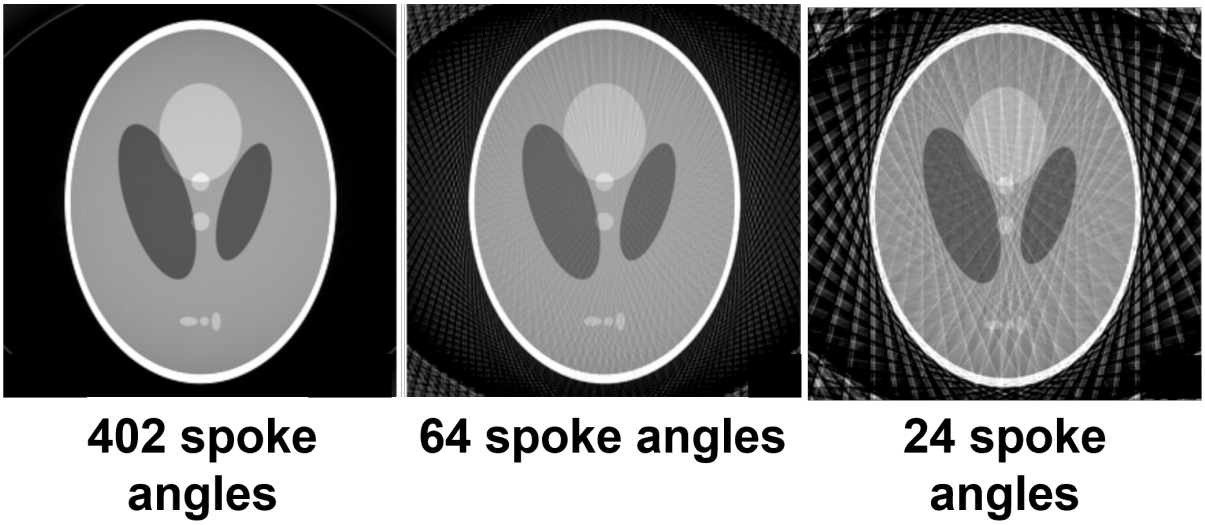


Figure 1.11: Reconstructed images of a Shepp-Logan phantom with matrix size 256×256. The Nyquist rate for radial sampling is $256 \times \frac{\Pi}{2} = 402$ spokes, hence the artifact-free image reconstructed on the left. When decreasing the number of spokes acquired, streaking artifacts appear.

mean to counter the slow-imaging. As a result, advanced constrained/model-based reconstruction techniques are often used to recover the dynamic MR images from the highly undersampled (k, t)-space data. We show here some of the state of the art iterative MR image reconstruction techniques used to deal with undersampling.

Parallel imaging (PI)

Parallel imaging has emerged as a powerful tool to reduce scanning time and improve patient comfort, and became widely available in industrial MR scanners in the early 2000s. This technique aims at using the spatial sensitivity information of the coil receivers to combine them during reconstruction, instead of using the sum-of-square technique to combine independently reconstructed images.

In parallel imaging, the k-space measurements are sampled at a frequency rate lower than the Nyquist rate, skipping the acquisition of certain lines in the k-space (e.g., in the phase encoding direction). With the conventional non-iterative reconstruction method, this undersampled pattern leads to artifacts as shown in Section 1.3.1. The gist of PI is then to rely on the numerous coil receivers to "unfold" the overlapping signal causing aliasing.

One of the main advantages of parallel imaging is its versatility as it can be used to nearly all MR applications and only requires multi-coil receivers for image acquisition with no need of higher gradient system performances [8] to increase imaging by a factor R of 1.5 to 3 in most cases [38]. However, parallel imaging can be challenging in practice as it tends to cause a decrease in signal-to-noise ratio (SNR) and, when needed, the estimation of the coil sensitivity map may not always be straightforward. PI techniques can be categorized in two: the techniques applied in the image domain, i.e., untangling the aliasing signal in the image; and the techniques applied in the k-space

domain, i.e., estimating the missing data-points causing aliasing.

SENSE The Sensitivity Encoding (SENSE) [121, 122] is one of the most common PI techniques applied in the image domain [28]. As shown in Section 1.3.1, missing k-space lines cause aliasing and reduced FOV in the image domain. With n_{coil} coils covering the whole imaged object, we can reconstruct n_{coil} individual coil images. For a given pixel of the full-FOV reconstructed image, the n_{coil} will contribute to its signal in a way that is weighted by the coils sensitivities in that region. In other words, assuming a reduction factor R, the n_{coil} individual coil images are a R-way "folded" version of the original image, and we rely on a known coil sensitivity to "unfold" the images and recover the underlying full-FOV signal [8, 28]. For simplicity, let us consider $R = n_{coil} = 2$ in the Cartesian case, i.e., 1 k-space line over 2 is missing in the phase encoding direction. Denote the reduced-FOV and aliased images as $\tilde{\mathbf{X}}_1$ and $\tilde{\mathbf{X}}_2$, and their underlying non-aliased reduced-FOV version $\mathbf{X}_1$ and $\mathbf{X}_2$. We can then write:

$$\begin{aligned} \tilde{\mathbf{X}}_1 &= \mathbf{X}_1 \circ \mathbf{S}_1 + \mathbf{X}_2 \circ \mathbf{S}_1 \\ \tilde{\mathbf{X}}_2 &= \mathbf{X}_1 \circ \mathbf{S}_2 + \mathbf{X}_2 \circ \mathbf{S}_2 \end{aligned} \tag{1.19}$$

where $\circ$ represents the element-wise product and $\mathbf{S}_1$ and $\mathbf{S}_2$ are the coil sensitivity maps of the coils 1 and 2. The principle of this technique is illustrated in Figure 1.12.

Assuming the sensitivity maps are known, the above two equations with two unknowns $\mathbf{X}_1$ and $\mathbf{X}_2$ can easily be solved with linear algebra to form the full-FOV image $\mathbf{X}$. This can be generalized to an arbitrary undersampling factor R and coil receivers n_{coil}, where we obtain n_{coil} equations with R unknowns (due to the R-way "folding"). Therefore, in theory, as long as $R \leq n_{coil}$ the system of equations can be inverted to recover the full-FOV image.

Despite being first theorized for the Cartesian case in [121], the SENSE method is generalized to arbitrary trajectories along with conjugate gradient (CG) methods to solve the optimization problem more efficiently [122]. The coil sensitivity matrix is then incorporated in the reconstruction formulation:

$$\mathbf{s}_{mc} = \mathcal{A}_{SENSE}(\mathbf{X}) + \boldsymbol{\varepsilon} \tag{1.20}$$

where $\mathbf{s}_{mc}$ is the vectorized multi-coil k-space data and $\mathcal{A}_{SENSE}$, often called generalized encoding matrix (GEM), is an operator composed of the sampling mask, the Fourier transform (as in Equation (1.16)) and the coil sensitivity maps for the SENSE model.

A least square estimate of the above equation can be obtained by solving the normal equation [125] and using a CG iterative scheme. This allows for better memory usage / processing time for large linear systems, which is particularly suited to for the FFT, and does not require strong provisions to ensure convergence [122]. This technique, now widely used in clinical settings as well, assumes that coil sensitivity maps are known and estimated during a pre-scan, which is not always feasible. In such a case, the coil sensitivity maps can be estimated from low-resolution images or

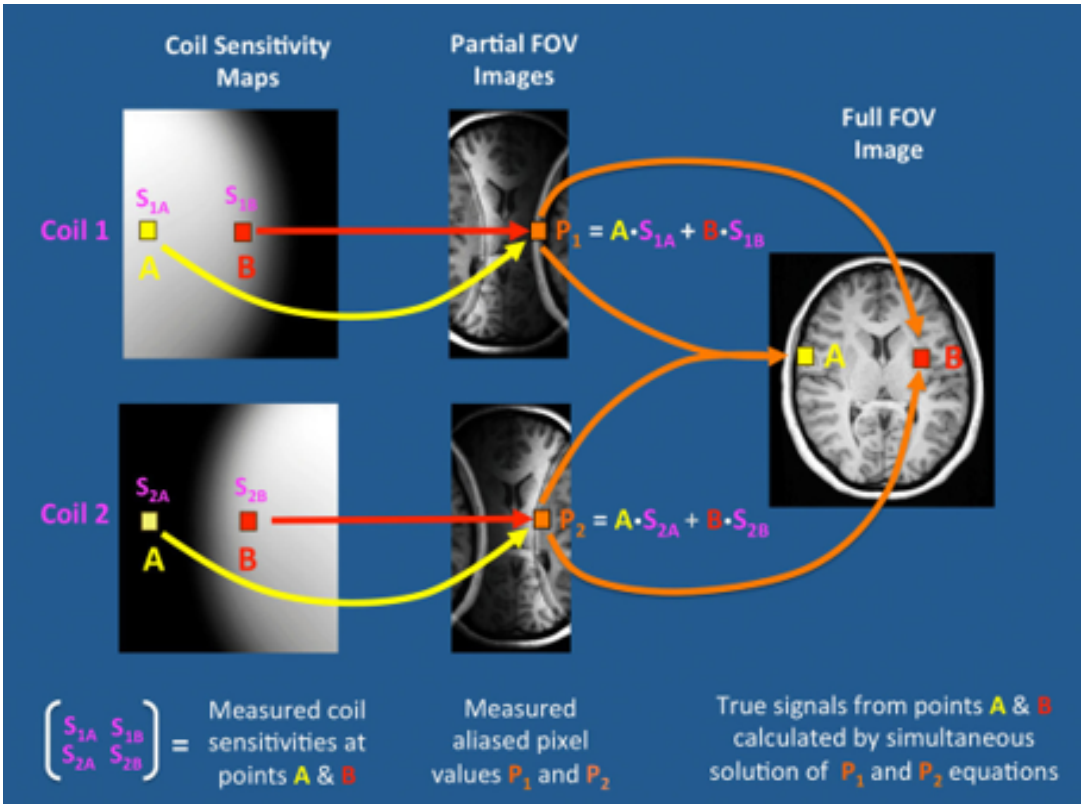


Figure 1.12: Principle of SENSE-like techniques. Known (or estimated) coil sensitivity maps are used to untangle the aliased signals in the image space. S_{iA} and S_{iB} denote the sensitivity of the i^{th} coil for the voxels A and B. P_i represents the voxel signal intensity in the aliased image $\tilde{\mathbf{X}}_i$ for the i^{th} coil. [29]

time-averaged images for dynamic MRI. However, note that inaccurate coil sensitivity maps can cause artifact in the full-FOV reconstructed image [28, 164].

GRAPPA While the SENSE method focuses on removing the aliasing signal in the image domain, the Generalized Autocalibrating Partially Parallel Acquisition (GRAPPA) (and its non-Cartesian adaptations) aims at processing directly the input k-space data to recover the missing information [15, 43, 56, 140]. Unlike SENSE, the GRAPPA method reconstructs full-FOV individual coil images that are then combined using the sum-of-square to produce the original image $\mathbf{X}$. One can think of GRAPPA as the dual method of SENSE: if the original image multiplied by a sensitivity matrix yields an aliasing-free individual coil image, then, equivalently, the original k-space convoluted with some weight-matrix should yield the non-aliasing k-space data. Another way to see it is to consider that each coil weighs the image signal intensity based on its sensitivity (i.e. the closer to the coil, the better). This translates into a smearing of information in the k-space in coil-dependent fashion [28, 164] and if that smearing pattern can be estimated for each coil (through weights), missing k-space samples can be estimated from their neighbors.

Differently from SENSE, the GRAPPA method does not require knowing the coil sensitivity before reconstruction: it uses a technique called autocalibration to determine the relationship between the data from the several coil receivers. This autocalibration is done using a set of fully

sampled k-space lines acquired in the center of the k-space called the autocalibration signal (ACS). The GRAPPA method is then typically formulated as a least-squares optimization problem, where the objective is to minimize the difference between the estimated k-space data and the ACS. Note that in this method, missing k-space lines are zero-filled. The principle of missing lines estimation is illustrated in Figure 1.13.

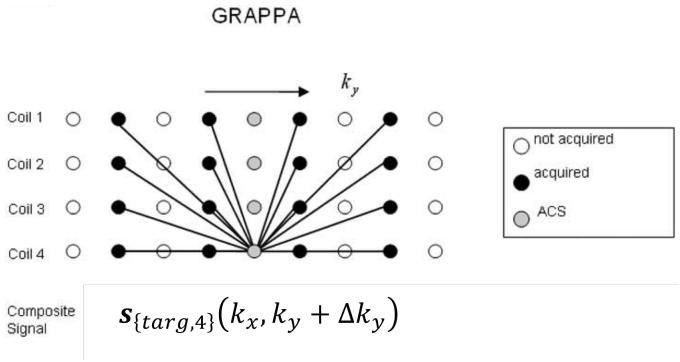


Figure 1.13: Missing k-space lines estimation using GRAPPA. Each dot represents a line in the k-space. GRAPPA uses multiple lines from all coils to fit one line in a target coil (here coil 4). This procedure is repeated for every coil, resulting in fully-sampled k-space measurements estimated for each coil. The FFT-reconstructed images can then be combined using a sum-of-squares reconstruction [8].

For simplicity, take the case of a 2D Cartesian dataset with an R reduction factor, i.e., only one k-space line over R is acquired in the phase-encoding direction. The GRAPPA formulation then fits the target points from a coil i (i.e., the missing samples $s_{\text{targ},i}$) to the source points from all the coils (i.e., acquired samples $s_{\text{src},j}, j = 1, \ldots, n_{coil}$):

$$s_{\text{targ},i}\left(k_x, k_y + \Delta k_y\right) = \sum_{j=1}^{n_{coil}} \sum_{\tau_x} \sum_{\tau_y} w\left(i, j, \tau_x, \tau_y\right) \cdot s_{\text{src},j}\left(k_x + \tau_x, k_y + \tau_y\right) \tag{1.21}$$

where τ_x and τ_y represent the position along x and y in the kernel formed by the weights w, $w\left(i, j, \tau_x, \tau_y\right)$ represents the value of the kernel for a target sample in coil i, source sample in coil j, and at the indexes τ_x, τ_y. Note that τ_x and τ_y depend on the kernel size (e.g., 2×3) that is chosen before reconstruction.

Thanks to the ACS, the target samples can be known for the center of the k-space, i.e., $s_{\text{targ},i}$ is known and only the weights w are unknown. We can then form a weight matrix $\mathbf{W}$ of size $n_{coil} \times n_{coil} n_w$ from the n_w kernel weights for the n_{coil} coils sources and n_{coil} targets to write Equation (1.21) in matrix notation:

$$\mathbf{s}_{ACS,\text{targ}} = \mathbf{W}\, \mathbf{s}_{ACS,\text{src}} \tag{1.22}$$

where $\mathbf{s}_{ACS,\text{targ}}$ and $\mathbf{s}_{ACS,\text{src}}$ are vectors containing the target and the source samples.

Note that the weights are shift-invariant, which means that the same set of weights is used for the whole k-space, and that the relationship in Equation (1.22) must be satisfied for all the samples in the k-space. The above problem can efficiently be solved using a least-square method and, if

ill-conditioned, a regularization term can also be added on $\mathbf{W}$. One of the main advantages of this technique is that it does not require the coil sensitivity to be known before reconstruction and can be directly estimated using the acquired data, making it usually more robust than SENSE in inhomogeneous areas.

Compressed Sensing (CS) The compressed sensing theory for MRI application has been described in [88, 89] and aims at recovering signals and images from data sampled below the theoretical Nyquist frequency. Well-known transform-based compression algorithms such as JPEG are already extensively used to compress images. Natural images in particular can often be compressed without loss of information. In MR image reconstruction, CS exploits the MR data intrinsic property of being sparse in an appropriate transformed domain. Lustig *et al.* therefore present the MR signal as being compressible [88], making it possible from the acquisition itself to only acquire ("sense") the relevant ("compressed") components of the data. This new concept of "Compressed Sensing" was well described in [88]: *"Since the images we intend to acquire will be compressible, with most transform coefficients negligible or unimportant, is it really necessary to acquire all that data in the first place? Can we not simply measure the compressed information directly from a small number of measurements, and still reconstruct the same image, which would arise from the fully sampled set? Furthermore, since MRI measures Fourier coefficients [...] the question is whether it is possible to do the above by measuring only a subset of k-space."*

To successfully apply CS, three key conditions have to be met [89]. The most fundamental one is that the data must be sparse in some transformed domain [36, 89]. As explained above, this condition is already met for MR images. In addition, the undersampling artifacts must be incoherent in the sparsifying transformed domain. For MR data, this means that the artifacts created by the k-space undersampling should not have a consistent pattern (i.e., streaks or obvious aliasing) and can be obtained using pseudo-random sampling. A way to meet this condition is to acquire k-space data for random locations along the acquisition. Finally, the sparsity of the image in the transformed domain must be enforced by some non-linear reconstruction scheme. In practice, this can be done by adding a 1-norm regularization term on the domain-transformed data.

To summarize, in CS applications, the image model from Equation (1.16) is fitted to the undersampled k-space data with an additional term, sometimes denoted as prior. The optimization problem to solve can be written as:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \left\| \mathbf{\Omega}\left(\mathbf{F}_s \mathbf{X}\right) - \mathbf{s} \right\|_2^2 + \lambda \left\| \mathcal{D}(\mathbf{X}) \right\|_1 , \tag{1.23}$$

where $\mathcal{D}$ is a sparsifying transform operator such as finite difference, λ is a parameter weighing the CS regularization, and $\|\bullet\|_1$ is the l_1-norm defined as the sum of the absolute value of entries elements. This can be solved using algorithms such as ISTA/FISTA [5] or the alternating direction methods of multipliers (ADMM) algorithm [12]. Note that compressed sensing can be used along other models, such as SENSE. In such a case, the coil sensitivity matrix is added to the forward

model in Equation (1.23) and the multi-coil data is used.

1.4 Limitations for dynamic MRI

Despite advanced reconstruction schemes such as SENSE/GRAPPA or CS presented in the previous section, reconstructing artifact-free images is still a challenge for dynamic MRI. In theory, parallel imaging can offer a reduction factor as high as the number of coil receivers, however, in practice, this kind of technique only offers a reduction factor of 1.5 to 3 without noticeable artifacts [38] and is not always fitted for dynamic imaging with inhomogeneous regions such as lungs. CS also allows for higher undersampling of the k-space, but the actual reduction factor it allows is highly dependent on the application (i.e., DCE-liver, MRA, cardiac, etc.) and a too strong CS regularization will tend to blur the image, rendering them non-diagnostic [64].

Overall, the difficulty with dynamic MRI is that it requires high spatial resolution to provide detailed images of the anatomy being imaged (just as static MRI) but also high temporal resolution to capture rapid changes in the image over time. The main limitation is the trade-off between temporal and spatial resolution. Increasing the temporal resolution typically requires sacrificing spatial resolution: the shorter the temporal resolution, the smaller the time window for k-space line sampling within one frame. Dynamic MR image reconstruction is therefore a highly underdetermined problem due to the extra time dimension of the dynamic series. Another limitation of dynamic MRI is motion artifact. Motion during the imaging process can cause blurring and distortion in the reconstructed images, particularly in fast-moving or highly deformable tissue such as the heart or lungs. Motion correction techniques [39, 153, 163] can be used to mitigate these artifacts, but they may not always be effective, particularly in cases of large or complex motion. Finally, conventional MR image reconstruction techniques may also be limited by the length of the acquisition time. In dynamic MRI, the longer the acquisition time, the more detailed the images will be, but this can also lead to increased motion artifact, and increased burden for the patient. Finding the optimal balance between acquisition time and image quality is often a challenge in dynamic MRI.

To address these challenges, dynamic MRI reconstruction algorithms often rely on additional constraints or prior knowledge about the image sequence to help guide the reconstruction process. This can include assumptions about the smoothness/pseudo-periodicity of the dynamic image series, or assumptions about an underlying low-dimensional space where the k-space samples lie. These types of technique will be discussed in Chapter 2 where a linear low-dimensional model for dynamic MRI is presented with applications to PET motion correction, and in Chapter 3 where a novel non-linear low-dimensional model based on linear tangent space alignment is derived and presented with application to dynamic MRI.

Chapter 2

Subspace model-based image reconstruction for dynamic MRI

Chapter 1 presented background knowledge on the NMR phenomenon and how it is used in MR imaging. Notions such as k-space sampling strategies (e.g., Cartesian and radial trajectories) were detailed, and state-of-the-art iterative and non-iterative reconstruction techniques were presented to reconstruct images from undersampled k-space data. However, these techniques show limitations with high undersampling rates, which is especially bound to happen in dynamic MRI due to the limited sampling time-window per temporal frame. Using a model that takes into account spatiotemporal correlations in the MR signal can therefore be particularly useful to overcome these limitations. We present in this chapter a method that models the dynamic k-space data as high-dimensional data lying close or onto a low-dimensional linear subspace and demonstrate how it can reconstruct MR images with fewer unknowns by fixing a set of temporal basis functions. The mathematical formulation of the subspace model is first shown along an efficient MR acquisition scheme to maximize the potency of the model by estimating temporal basis from the undersampled k-space data. Two applications leveraging the subspace model for MR image reconstruction are then presented, respectively for MR-based PET data motion correction and free-breathing 3D cardiac T_1 mapping.

2.1 Subspace model

2.1.1 Mathematical formulation

We denote the dynamic image series as $x(\mathbf{r}, t)$ and its matrix representation $\mathbf{X} \in \mathbb{C}^{M \times N}$ as described in Section 1.2.3:

$$\mathbf{X} = \begin{bmatrix} x(\mathbf{r}_1, t_1) & \dots & x(\mathbf{r}_1, t_N) \\ \vdots & \ddots & \vdots \\ x(\mathbf{r}_M, t_1) & \dots & x(\mathbf{r}_M, t_N) \end{bmatrix}. \tag{2.1}$$

The temporal signal at the m-th voxel of the dynamic images is denoted as $\mathbf{z}_m = \left[x(\mathbf{r}_m, t_1), \ldots, x(\mathbf{r}_m, t_N)\right]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}$, where 'T' represents the matrix transpose operation. In the subspace model, we assume that $\mathbf{z}_m$ resides in a low-dimensional subspace and can therefore be expressed as:

$$\mathbf{z}_m = \mathbf{\Phi}\boldsymbol{\theta}_m, m = 1, \ldots, M, \tag{2.2}$$

where $\mathbf{\Phi} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_D] \in \mathbb{C}^{N\times D}$ concatenates along columns the temporal bases (i.e., $\boldsymbol{\phi}_d = [\phi_d(t_1), \ldots, \phi_d(t_N)]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}$) that span a D-dimensional subspace ($D << M$ or N), and $\boldsymbol{\theta}_m \in \mathbb{C}^{D\times 1}$ concatenates the corresponding spatial coefficients (or equivalently coordinates) of $\mathbf{z}_m$ in this subspace.

The subspace model in Equation (2.2) can be written more concisely in the form of a low-rank decomposition of the Casorati matrix $\mathbf{X}$ presented in Equation (2.1):

$$\mathbf{X} = \mathbf{\Theta}\mathbf{\Phi}^{\mathrm{T}}, \tag{2.3}$$

where $\mathbf{\Theta} \in \mathbb{C}^{M\times D}$ concatenates along rows the spatial coefficients $\boldsymbol{\theta}_m^T$ of all M voxels with respect to the subspace basis, $\mathbf{\Phi}$. However, this decomposition may not be always possible in an exact way, and an approximation is then searched for.

The optimal $\mathbf{\Theta}$ and $\mathbf{\Phi}$ then minimize the low-rank approximation error of $\mathbf{X}$:

$$\arg\min_{\mathbf{\Theta},\mathbf{\Phi}} \left\|\mathbf{X} - \mathbf{\Theta}\mathbf{\Phi}^{\mathrm{T}}\right\|_F^2 \tag{2.4}$$

The above optimization problem can be solved by singular value decomposition (SVD), which provides the best rank-D linear approximation of the Casorati matrix $\mathbf{X}$:

$$\mathbf{\Theta}\mathbf{\Phi}^{\mathrm{T}} = \mathbf{U}_D\mathbf{\Sigma}_D\mathbf{V}_D^{\mathrm{H}}, \tag{2.5}$$

where 'H' represents the matrix Hermitian transpose operation, $\mathbf{\Sigma}_D = \mathrm{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_D\}$ are the D largest singular values of $\mathbf{X}$, and $\mathbf{U}_D$ and $\mathbf{V}_D$ are the corresponding left and right singular vectors. The singular value decomposition in Equation (2.5) actually provides a way of verifying the assumption of low-rankness implied in the subspace model by plotting the singular value distribution of $\mathbf{X}$, in addition to an ideal cut-off rank D.

In practice, $\mathbf{X}$ is the matrix we want to recover in the image reconstruction process, therefore, directly performing an SVD on $\mathbf{X}$ to verify the low-rankness assumption is usually not possible. However, workarounds exist, such as performing the SVD on a first estimate $\hat{\mathbf{X}}$ provided from another reconstruction technique. Another more practical way is to perform the SVD on a set of k-space data forming a basis spanning the underlying spatiotemporal subspace [22] (see details on "training" k-space samples defined in Section 2.1.2). The temporal basis functions $[\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_D]$ can then be estimated from the acquired data and fixed before image reconstruction (Sections 2.1.2 and 2.1.3). This only leaves $\mathbf{\Theta}$ to be estimated during image reconstruction, i.e., $M \times D$ unknowns

instead of the original $M \times N$ unknowns ($N >> D$).

Therefore, the subspace model significantly reduces the degrees of freedom of the underlying spatiotemporal distribution and enables special (k, t)-space sampling schemes and image reconstruction methods that recover dynamic images from undersampled data with high spatial resolution and high frame rate [21, 22, 81, 83, 114, 171]. In practice, the dimension of the subspace just needs to be sufficiently high to capture the temporal changes of the signals in dynamic MRI [21, 22, 171].

2.1.2 Data acquisition scheme

The data acquisition scheme of the subspace-based methods usually relies on two interleaved sets of k-space data sampled within one acquisition [81]. In each frame, a set of data is acquired at a high temporal rate (to enable high temporal resolution) but covering only limited k-space for the determination of the temporal bases. This set is referred to as "training" k-space data. Note that the term "training" k-space data (or "navigator data" in some manuscripts) is commonly used in the literature related to subspace-based approaches for samples acquired with high temporal rate for temporal basis estimation and is different from the "training" datasets used in deep learning applications. The other set, namely the "imaging" set, is composed of data acquired in the entire k-space sparsely sampled throughout the acquisition (to enable high spatial resolution), i.e., covering the entire k-space but at a low temporal rate. This "imaging" set is then fitted to the imaging model (i.e., subspace model) in the image reconstruction process.

To further illustrate the data acquisition scheme of subspace-based image reconstruction methods, take the case of an acquisition with radial trajectory in the k_x-k_y plane. Denote n_t the number of k-space lines per frame for temporal basis estimation and n_{im} the number of k-space lines per frame for image reconstruction. A total of $(n_t + n_{im})$ k-space lines are acquired each frame, resulting in $K = N(n_t + n_{im})$ acquired k-space lines in the acquisition. In such a case, the scheme described above commonly translates to $n_t = 2$ k-space lines acquired each frame in 2D imaging and $n_t = 3$ in 3D imaging: one line along k_x, one line along k_y and one line along k_z in 3D imaging; to span the spatiotemporal subspace. Another n_{im} lines at random spoke angles are acquired, sparsely covering the rest of the k-space. These two steps are repeated N times to produce the undersampled k-space data vector $s \in \mathbb{C}^{K \times 1}$. An example of such an acquisition in the 2D case is shown in Figure 2.1.

Many k-space trajectories, such as Cartesian, spiral, and radial trajectories, can be used [89]. In the work presented in this manuscript, we used random radial trajectories for the 2D imaging experiments and random stack-of-stars trajectories for the 3D imaging experiments because of their superior performance in dynamic MRI applications.

2.1.3 Temporal basis estimation

In the subspace model, a temporal basis spanning the underlying linear subspace is estimated prior to image reconstruction. The approach presented below is based on the assumption that the training

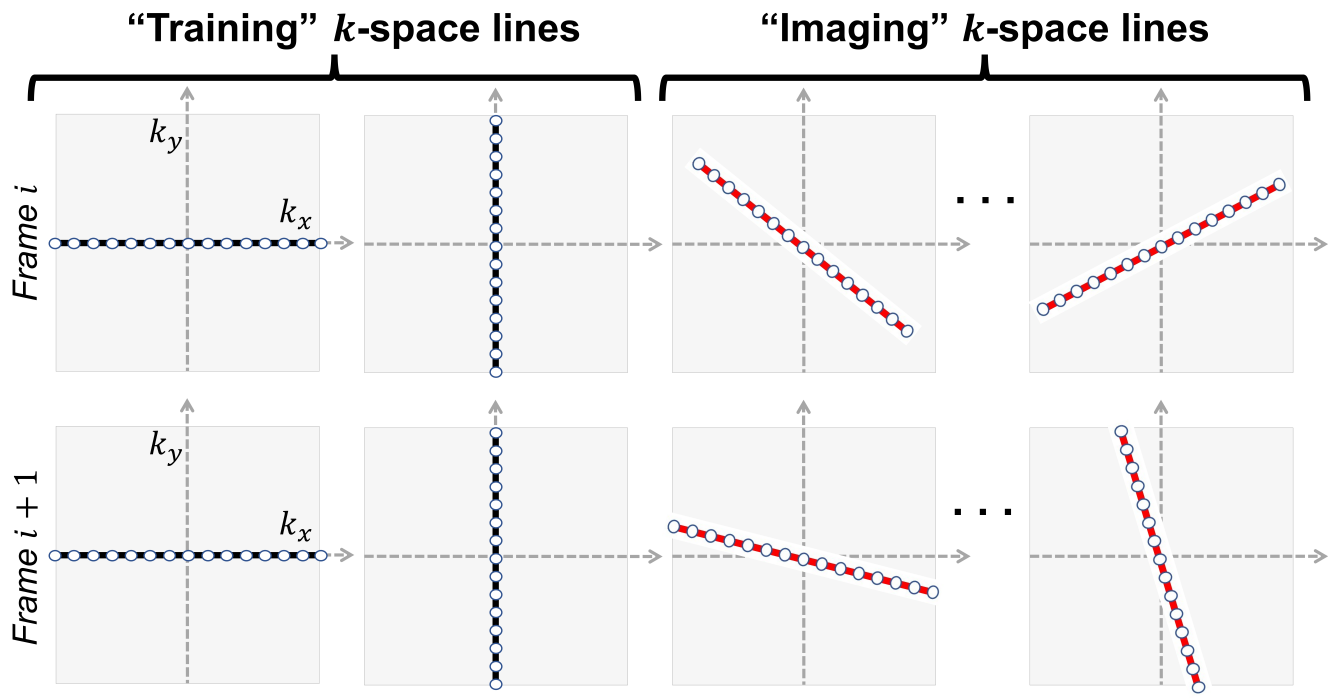


Figure 2.1: Example of an efficient 2D data acquisition scheme for temporal basis estimation and subspace modeling. The white dots represent individual k-space measurements along the readout direction. In this example, training lines for temporal basis estimation are acquired at the beginning of each frame along the k_x and k_y directions (black lines). For the remaining time of each frame, k-space lines along random spoke angles are then acquired (red lines).

data and the imaging data share the same temporal subspace, which has been proved to be valid [22].

Suppose there are in total N temporal frames with, in each frame, n_t k-space lines acquired as training data, as described in Section 2.1.2, and denote R_o the readout dimension of the measured k-space data. A Casorati matrix $\mathbf{\Lambda}$ of size $(n_t R_o \times N)$ can be formed by concatenating the n_t training k-space lines of each frame along rows. In the case of imaging with n_{coil} phased-array coils, with $n_{coil} > 1$, the training lines from each coil are additionally concatenated along columns, forming a multi-coil Casorati matrix of size $(n_{coil} n_t R_o \times N)$. The temporal basis of the subspace, $\mathbf{\Phi} \in \mathbb{C}^{N \times D}$, is then estimated by calculating the first D right singular vectors of $\mathbf{\Lambda}$. See [22] for an in-depth analysis of the subspace estimation using "training" k-space data acquired in this way.

2.1.4 Image reconstruction

Once an estimation of $\mathbf{\Phi}$ is obtained, the image reconstruction problem is reduced to the determination of the spatial coefficients matrix $\mathbf{\Theta}$. We solve this problem by fitting the subspace model – also referred to as Low-Rank (LR) model – to the undersampled (k,t)-space data with additional sparsity constraints [171]:

$$\mathbf{\Theta} = \arg\min_{\mathbf{\Theta}} \left\| \Omega \left(\boldsymbol{F_s} \mathbf{\Theta} \mathbf{\Phi}^{\mathrm{T}} \right) - \mathbf{s} \right\|_2^2 + \lambda \left\| T(\mathbf{\Theta} \mathbf{\Phi}^{\mathrm{T}}) \right\|_1 + \mu \left\| \mathbf{\Theta} \right\|_F^2, \tag{2.6}$$

where $\mathbf{s}$ is the measured k-space data, $\boldsymbol{F_s}$ is the Fourier transform operator in the spatial domain, *i.e.* Non-Uniform FFT (NuFFT) operator [32] for the stack-of-stars trajectory, Ω is the sparse sampling operator in the (k,t)-space, T is a sparsifying transform, e.g., the finite difference operator along the spatial and/or temporal directions, $\|.\|_F$ is the Frobenius norm and the scalar variables λ and μ are regularization parameters. The first term in Equation (2.6) is a data fidelity term, fitting the subspace model to the measured k-space data $\mathbf{s}$. The second term is a complementary sparsity penalty promoting the piecewise smoothness of the reconstructed images $\mathbf{X} = \boldsymbol{\Theta}\boldsymbol{\Phi}^{\mathrm{T}}$ along the spatial and temporal dimensions. The third term is an l_2-norm regularization term introduced to improve the numerical stability and favoring minimal norm solutions for $\boldsymbol{\Theta}$.

We solve the optimization problem in Equation (2.6) using the Alternating Direction Methods of Multipliers (ADMM) algorithm [12], which leads to solving the following three sub-optimization problem in an alternative fashion:

$$\mathbf{z}^{(k+1)} = \mathcal{S}_{\frac{\lambda}{\rho}}\left(T\left(\boldsymbol{\Theta}^{(k)}\,\boldsymbol{\Phi}^{\mathrm{T}}\right) + \boldsymbol{\eta}^{(k)}\right), \tag{2.7}$$

$$\boldsymbol{\Theta}^{(k+1)} = \arg\min_{\boldsymbol{\Theta}} \left\|\Omega\left(\boldsymbol{F_s}\boldsymbol{\Theta}\boldsymbol{\Phi}^{\mathrm{T}}\right) - \mathbf{s}\right\|_2^2 + \frac{\rho}{2}\left\|T\left(\boldsymbol{\Theta}\boldsymbol{\Phi}^{\mathrm{T}}\right) - \mathbf{z}^{(k+1)} + \boldsymbol{\eta}^{(k)}\right\|_F^2 + \mu\left\|\boldsymbol{\Theta}\right\|_F, \tag{2.8}$$

$$\boldsymbol{\eta}^{(k+1)} = \boldsymbol{\eta}^{(k)} + \left(T\left(\boldsymbol{\Theta}^{(k+1)}\,\boldsymbol{\Phi}^{\mathrm{T}}\right) - \mathbf{z}^{(k+1)}\right), \tag{2.9}$$

where $\mathbf{z}$ is the split variable, $\boldsymbol{\eta}$ is the dual variable, and ρ is a scalar relaxation parameter. The $\mathbf{z}$ update (Equation (2.7)) is a soft thresholding operation and the $\boldsymbol{\Theta}$ update (Equation (2.8)) is a convex-optimization problem, which is solved using the conjugate gradient algorithm. In the rest of this manuscript, the subspace model will also be denoted Low-Rank (LR) model.

2.2 Application to MR-based motion correction for PET-MR

2.2.1 Introduction

Motion, including physiological motion (*i.e.*, cardiac and respiratory motions) and involuntary bulk motion, is a major source of image quality degradation in Positron Emission Tomography (PET), which can result in spatial blurring artifacts and mismatch between emission and attenuation maps, altering quantification of tracer concentration and deteriorating the diagnostic value of PET images [85, 109, 133]. The conventional way to handle motion in PET is the gating method, which bins PET list-mode data to different cardiac and/or respiratory motion phases followed by reconstructions of images of each phase. However, gating results in increased noise levels due to the reduced number of events in each motion phase. To address the limitations of the gating method, many PET motion correction methods have been developed, which consist of two consecutive steps: motion field estimation and motion correction by either applying the estimated motion fields to the gated images or modeling it within motion-compensated PET image reconstruction [126].

PET motion correction methods can be divided into two major subcategories, depending on how

the motion field is estimated: PET-based methods and Magnetic Resonance (MR)-based methods. In the PET-based motion correction methods, the measured emission data are first assigned to specific motion phases based on surrogate signals [66], *e.g.*, electrocardiogram (EKG), respiratory bellow, optical tracking, etc. [35, 103, 168], or the PET-data themselves [69, 86, 147], *e.g.*, center of mass, time-of-flight information, frame-by-frame images, etc. Motion fields are then estimated by registering the reconstructed image of each phase to a reference phase [27]. However, the accuracy of the motion fields estimated by the PET-based methods is limited by low signal-to-noise ratio (SNR), especially in the case of dual gating, and the overall lack of anatomical structural information of PET images [109, 117].

The increasing availability of hybrid PET/MR systems provides a unique opportunity for mitigating effects of motion in PET using MR-based motion correction. Because of its excellent soft-tissue contrast, high spatial resolution, and high SNR, MR provides more accurate estimation of motion fields than the PET-based methods. MR-based PET motion correction methods have been successfully applied to compensate for respiratory and cardiac motion in various applications involving both static and dynamic PET imaging [14, 37, 59, 73, 115, 117]. One major limitation of the MR-based motion correction methods is that the conventional non-iterative MR imaging methods are unable to resolve cardiac or respiratory motion in real time due to the slow imaging speed. Binning-based MR imaging methods [31, 42, 104, 127, 130] are often used to address this issue, where MR k-space data are grouped into different motion phases based on surrogate signals (e.g., EKG), navigator signals, or k-space data alone, and images of each motion phase are then reconstructed for the estimation of motion fields. However, the binning-based MR imaging methods suffer from three noticeable limitations. First, they assume pseudo periodic motion, which does not hold well in the case of arrhythmia and irregular respiratory motion. Secondly, they rely on either surrogate signals or navigator signals acquired along a single direction to assign k-space data to specific motion phases, which cannot reliably capture involuntary bulk motion. Thirdly, their performance is limited by the inherent trade-off between the number of motion phases (and thus the accuracy of motion field measurement) and data acquisition time.

In this work, we propose a real-time MR imaging method for PET motion corrections in PET/MR. High resolution real-time MR imaging is achieved by a subspace-based imaging method, which takes advantage of a unique property of high-dimensional dynamic MR signals [80]. This model takes advantages of the spatial-temporal correlations of dynamic MR images, significantly reduces the number of unknowns of the underlying spatiotemporal signal, and makes it possible to recover high resolution, high frame-rate dynamic MR images from highly undersampled k-space data [22, 171]. For PET motion correction, the reconstructed real-time MR images are used to determine motion phases and estimate motion fields. PET list-mode data are binned into sinograms accordingly and ordered-subset expectation-maximization (OSEM) reconstruction [61] is performed integrating the estimated displacement in the system matrix for motion correction. We demonstrate the performance of the proposed method by carrying out *in vivo* ^{18}F-FDG PET/MR imaging experiments using a 3T simultaneous PET/MR scanner.

2.2.2 Methods

Simulation Study: Irregular respiratory motion

The key assumption of the subspace-based imaging method is the low-rank property of dynamic MR signals. We performed a simulation study to investigate this property in the case of regular and irregular respiratory motion. Two phantoms (shown in Figure 2.2) were generated frame by frame using the XCAT software [139] to simulate regular and irregular respiratory motion. Respiratory and cardiac cycles were divided into respectively 30 and 40 phases and 3D volumes were computed for each respiratory and cardiac phase combination (i.e. 1200 volumes). A 4D (3D space + time) phantom was then built by selecting and concatenating frame by frame the 3D volumes based on simulated EKG and respiratory signals. Both the breathing frequency and diaphragm expansion were varied while keeping a constant heart rate in the simulation of the irregular respiratory motion. Each phantom contained 6 respiratory cycles. The contrast was designed to simulate a Balanced Steady-State Free Precession (bSSFP) signal for several compartments such as fat, muscles, etc. using T_1 and T_2 values from the literature [10].

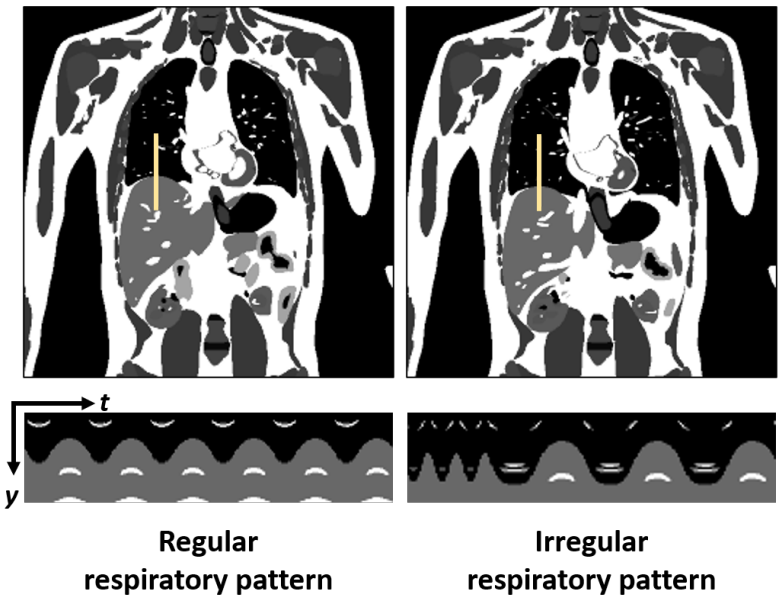


Figure 2.2: Simulated MR phantom for regular (left) and irregular (right) respiratory patterns. The plots shown in the bottom are 1D time-profiles of the liver along the yellow vertical line in the phantom image. These profiles show the frequency and amplitude variations in both cases.

A singular value decomposition was then performed to investigate the effect of an irregular respiratory pattern on the rank, and its corresponding approximation error with low-rank truncation. The result, shown in Figure 2.3, presents plots of the normalized singular values calculated using the phantom with regular (solid lines) and irregular (dashed line) respiratory pattern. In both cases, an approximation error smaller than 3% is achieved with a rank up to 9. The decay of the calculated singular values from both phantoms is therefore very similar, indicating that the breathing pattern does not substantially affect the rank of the data.

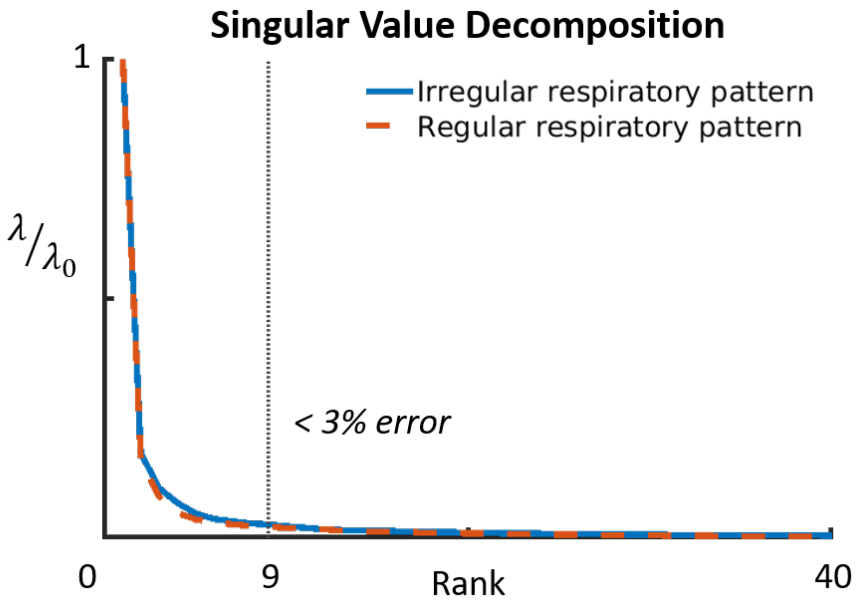


Figure 2.3: Normalized singular values decay for irregular respiratory pattern (plain blue line) and regular respiratory pattern (dashed orange line). An approximation error inferior to 3% can be achieved with a rank of 9 or higher, proving the signal indeed shows low-rank characteristics.

PET/MR imaging experiment

An ^{18}F-FDG PET/MR scan was performed on one healthy subject under a study protocol approved by our local Institutional Review Board (IRB). PET and MR data were simultaneously acquired 30 minutes after ^{18}F-FDG injection (around 10 mCi) using a 3T PET/MR scanner (Siemens Biograph mMR, Siemens Healthcare, Erlangen, Germany).

Two 5-minute MR acquisitions were performed using a spoiled gradient-recalled echo (GRE) sequence with stack-of-stars radial sampling trajectories in the coronal plane. The imaging parameters are as follows: image size $= 384 \times 384 \times 32$, voxel size $= 1.9 \times 1.9 \times 5\,\text{mm}^3$, TR/TE $= 3/1.6$ ms, and flip angle $= 7$ degrees. The (k, t)-space data were acquired using a random sampling pattern shown in Figure 2.4. A total of 35 k-space spokes were sampled in each frame, resulting in a frame rate of 9.5 volumes per second. For each frame, the first three spokes were respectively acquired along the k_x, k_y and k_z directions across the center of the k-space to estimate the temporal basis of the subspace model, as described in Section 2.1.3. The remaining 32 spokes were along a random angle in the k_x-k_y plane for every k_z.

During the first 5-minute acquisition, the subject was instructed to move once to assess the effect of both respiratory and bulk motion. During the second 5-minute acquisition, the subject was instructed to simulate an irregular respiratory pattern including both deep and shallow breaths. The vendor-provided two-point Dixon sequence was performed with breath-holding to obtain attenuation coefficients.

For comparison, we reconstructed MR images using the same data by a binning-based method, known as XD-GRASP [31]. The respiratory motion signal used for binning was processed the

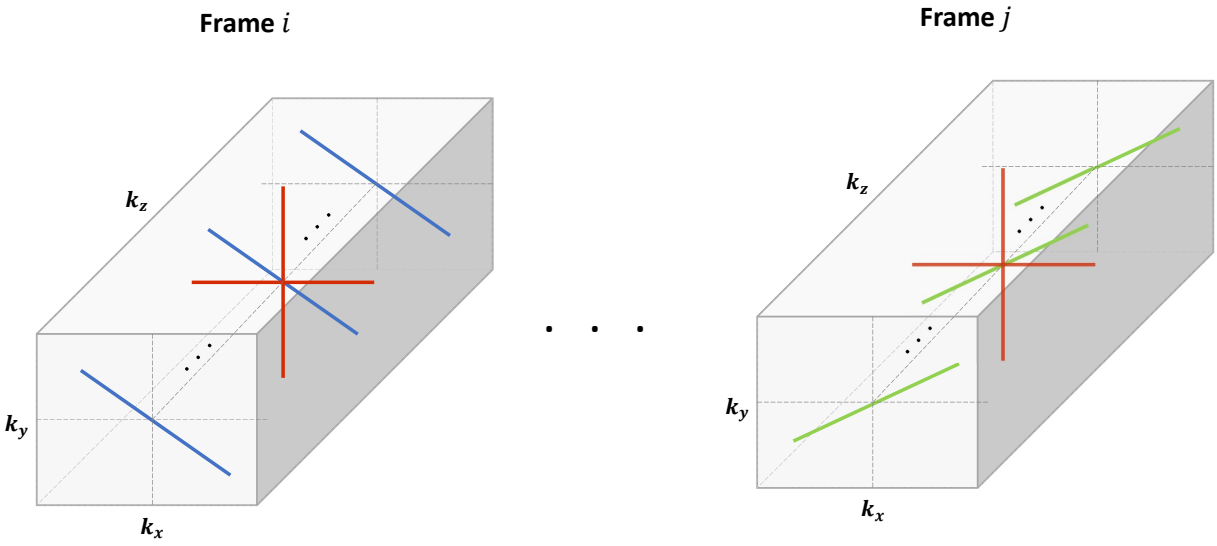


Figure 2.4: MR Sampling scheme used for the proposed method. 35 lines per frame are acquired: 1 line along k_z and 34 lines in the k_x-k_y plane. For a better visibility, only the k_x-k_y in-plane acquired lines are shown in the figure. Two training lines along k_x and k_y at $k_z = 0$ are consistently acquired through the whole acquisition to estimate the temporal basis V_t (red lines). A random angle is chosen every frame and is consistently acquired every k_z for imaging.

same way the authors of [31] did, i.e., the Fourier transform of the data at the center of the k-space at each frame was sorted into a 2D matrix, with data from each coil concatenated along the first dimension. A Principal Component Analysis (PCA) was then applied on this matrix and the component with the highest peak in the respiratory frequency range $[0.1, 0.5]$ Hz was selected as the binning signal. The k-space data were then grouped into balanced bins, i.e., each bin containing the same number of spokes. The actual number of bins is experiment dependent and will be described later in Section 2.2.3. The XD-GRASP reconstruction was performed the same way as in [31], i.e., solving the same optimization problem with sparsity penalty. Note that in this study, the sparsifying transform T in Equation (2.6) for both the subspace model and XD-GRASP was the finite difference operator along the spatial and temporal directions.

Motion estimation

The reconstructed real-time MR images were first binned into a small number of phases corresponding to different body positions (respiratory and bulk motion phases). Binning was performed in three steps. The first step consists in visually determining the bulk motion phases from the MR images and discarding time frames corresponding to the transition between bulk motion phases. In the second step, a bin is assigned to each (real-time) frame by tracking the tip of the right lobe of the liver over time while ensuring balanced bins (*i.e.*, all bins should contain a similar number of frames). Finally, a combined MR image is formed for each bin by averaging all real-time images in a bin. Volumetric image registration was then performed between all bins and a reference bin using the multiscale B-spline registration algorithm described in [24].

PET reconstruction

The acquired list-mode PET events were first rearranged into B sinograms $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_B)$ following the binning determined from MR images and discarding PET list-mode events occurring during bulk motion transitions.

PET reconstruction was performed using the OSEM algorithm [61] (an iterative statistical algorithm maximizing the log-likelihood of the measured data given the model of the data) and integrating the estimated motion fields in the forward model [84, 116]. Let $\mathbf{x}$ denote the PET image to reconstruct arranged in vector form. The system matrix, denoted by $\mathbf{F}$, is decomposed as $\mathbf{F} = \mathbf{N}\,\mathbf{A}\,\mathbf{G}\,\mathbf{M}$, where:

- $\mathbf{M} = [\mathbf{M}_1, \ldots, \mathbf{M}_B]^\top$ is a stack of deformation operators estimated using the procedure described in Section 2.2.2,
- $\mathbf{G} = \mathrm{diag}(\overline{\boldsymbol{G}}, \ldots, \overline{\boldsymbol{G}})$ is a block-diagonal geometrical projection matrix constructed by repeating the static projection matrix $\overline{\boldsymbol{G}}$ implemented using Siddon's algorithm [145] (a fast implementation of the Radon transform specifically tailored for PET imaging),
- $\mathbf{A} = \mathrm{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_B)$ is a diagonal matrix with time-varying attenuation coefficients,
- $\mathbf{N} = \mathrm{diag}(\overline{\boldsymbol{S}}, \ldots, \overline{\boldsymbol{S}})$ is a diagonal matrix with detector sensitivity coefficients $\overline{\boldsymbol{S}}$ repeated for all bins.

With these notations, the motion-corrected OSEM update for a given subset l is given by:

$$\mathbf{x}^{(n+1)} = \frac{\mathbf{x}^{(n)}}{\mathbf{F}_l^\top \mathbb{1}} \mathbf{F}_l^\top \frac{\mathbf{y}_l}{\mathbf{F}_l \mathbf{x}^{(n)} + \mathbf{s}_l}, \tag{2.10}$$

where $\mathbf{F}_l$ is the system matrix for the l-th subset, $\mathbb{1}$ is a vector full of 1's, and $\mathbf{s}_l$ is the combined additive correction sinogram for subset l including randoms and scatter. Correction sinograms were constructed as follows. Random coincidences were estimated using the delayed window method [72]. Scatter was estimated using the single scatter simulation algorithm [161] from an initial reconstruction performed without motion correction. Scatter was estimated separately for each bulk motion phase. Attenuation coefficients were obtained from a vendor-provided Dixon sequence during breath-holding. The attenuation map was deformed to each bin and forward projected to calculate sinogram-domain attenuation coefficients.

The OSEM used 12 subsets and 5 iterations. This motion-corrected reconstruction is denoted by MC in the rest of this work. For comparison, two other reconstruction methods were considered: a traditional OSEM without motion correction (NMC) and a gated reconstruction where only list-mode events occurring in a given motion phase are reconstructed without motion correction (Gated). Note that no time-of-flight information was available and that the scanner's point spread function was not included in the imaging model. Both reference methods used 4 iterations to account for the difference in convergence speed, aiming to match the noise level in MC and NMC reconstructions.

The noise level was not matched for the Gated method in order to keep gated reconstructions sharp and better compare with motion correction methods (in practice, gated reconstruction would typically use fewer subsets and/or iterations).

Quantitative analysis

In order to compare PET reconstructions, two evaluation measures were used: the contrast-to-noise ratio (CNR) and target-to-background ratio (TBR). The contrast-to-noise ratio is defined as:

$$\mathrm{CNR}(x, \mathcal{R}_1) = \frac{\bar{x}_{\mathcal{R}_1} - \bar{x}_{\mathcal{R}_0}}{\sigma_0}, \tag{2.11}$$

where σ_0 is the standard deviation in the background region $\mathcal{R}_0$ and $\bar{x}_{\mathcal{R}_1}$ is the average activity of image x in region $\mathcal{R}_1$. The background region $\mathcal{R}_0$ was a spherical region with a diameter of 14 mm, which was placed in a low activity region (lung) and defined large enough to estimate the noise level. With the same notations, the target-to-background ratio is given by:

$$\mathrm{TBR}(x, \mathcal{R}_1) = \frac{\bar{x}_{\mathcal{R}_1}}{\bar{x}_{\mathcal{R}_2}}, \tag{2.12}$$

where $\mathcal{R}_2$ is a small spherical region with diameter 4 mm located in the liver. Both measures were evaluated in a small spherical region with diameter 4 mm located in the kidney ($\mathcal{R}_1$ shown in Figure 2.5). The region was selected in the kidney with the highest activity, and its size was defined to cover the high activity region. Note that all regions of interest were located at the same position for all reconstruction methods. In the absence of ground truth, the contrast-to-noise and target-to-background ratios were used as indicators of image quality.

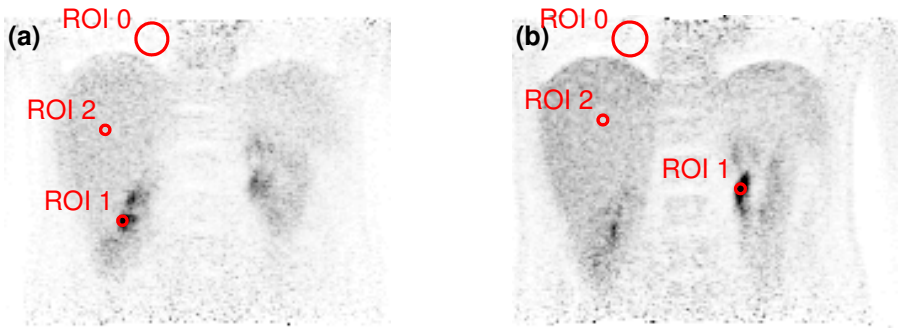


Figure 2.5: Regions of interest used for quantitative analysis: (a) bulk motion experiment, (b) irregular respiratory motion experiment. ROI 0 ($\mathcal{R}_0$) in the lung, ROI 1 ($\mathcal{R}_1$) in the kidney, ROI 2 ($\mathcal{R}_2$) in the liver. For each experiment, a small region in the kidney with the highest activity was selected to assess the effects of different PET reconstruction methods on contrast-to-noise ratio.

2.2.3 Results

Correction of bulk motion

In this experiment the subject was instructed to move after around 2.5 minutes in the 5-minute acquisition. Images reconstructed by XD-GRASP and the proposed method at the end-inhalation and end-exhalation phases are shown in Figure 2.6. The k-space data were grouped into 6 balanced bins as in [31] in the XD-GRASP reconstruction. The images obtained by XD-GRASP method show noticeable blurring artifacts largely because the bulk motion was not detected from the navigator signal. More specifically, Figure 2.7(a) shows the navigator signal obtained from the training line along the k_z direction in each frame as in [31]. Since the bulk motion of the subject was along the x direction (left to right), the navigator signal only recorded abnormal changes during the bulk motion, *i.e.*, the red region in Figure 2.7(a), but did not contain sufficient information to indicate what type of motion occurred. Therefore, six motion bins were chosen in XD-GRASP while the k-space data acquired in the red region of Figure 2.7(a) were discarded, resulting in blurring artifacts.

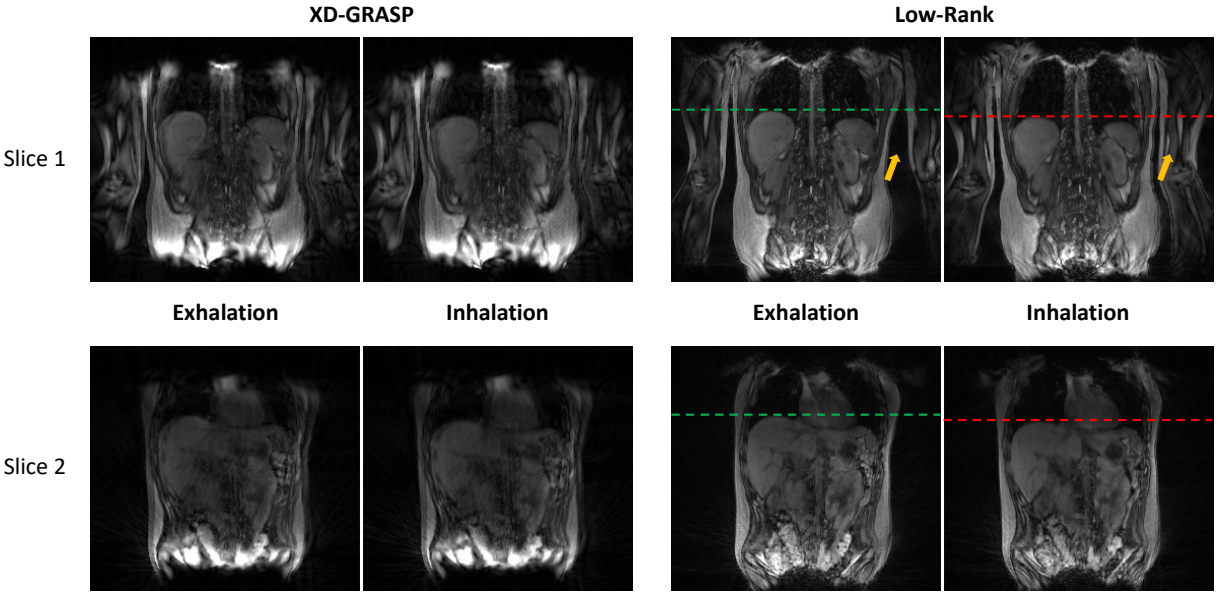


Figure 2.6: Representative MR images obtained with the XD-GRASP and the proposed method. The red dashed-line indicates the top of the liver position for end-inhalation and the green dashed-line indicates the top of the liver position for end-exhalation. Note that, besides respiratory motion, bulk motion indicated by the yellow arrows is clearly seen in the proposed low-rank based image reconstruction. The images obtained by XD-GRASP show blurring artifacts largely because the bulk motion was not detected from the navigator signal (see Figure 2.7 for more details).

The images obtained by the proposed method shown in Figure 2.6 successfully capture both respiratory motion (as indicated by the red and green dashed lines) and bulk motion (as indicated by the yellow arrows). Figure 2.7(b) to (d) show the temporal basis functions of the PS model estimated from the three training lines, where, intuitively, the first component (Figure 2.7(b)) shows respiratory patterns and the second component (Figure 2.7(c)) indicates bulk motion. To further demonstrate the real-time capabilities of the proposed method, Figure 2.8 shows images at multiple time frames along with a 1D profile through the liver along time. The images from

before and after bulk motion demonstrate the ability to capture both respiratory and bulk motion. The yellow overlay emphasizes the body displacement between Stage 1 and Stage 2. The profile plot shows the respiratory motion, captured for both bulk motion phases. The transition portion between the two bulk motion phases corresponds to the frames that were discarded in the PET reconstruction.

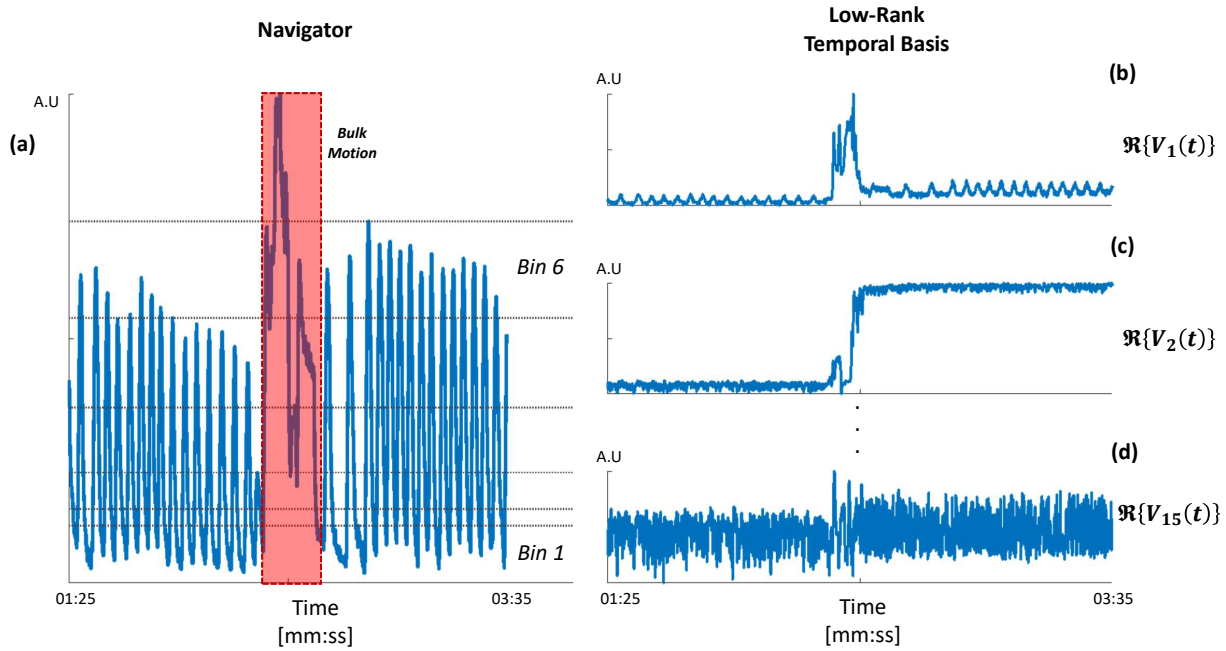


Figure 2.7: (a) Plot of the processed navigator along time, which was used to bin the k-space data in XD-GRASP. (b) to (d) Real part of the temporal basis $\mathbf{\Phi}^{\mathrm{T}}$ for the components 1, 2, and 15 of the PS model, respectively.

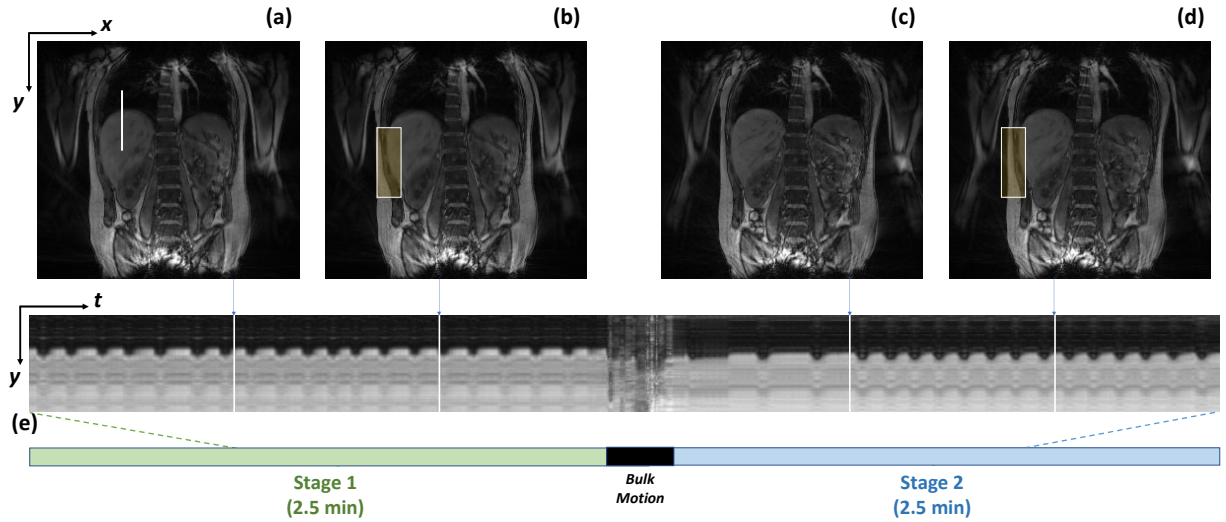


Figure 2.8: Dynamic MR images reconstructed by the proposed method. The real-time profile (e) is plotted for a part of the experiment where the bulk motion happens, and two representative images (a) & (b) and (c) & (d) are shown for each body position. The white line in (a) shows where the time profile was taken. A yellow box at the edge of the patient has been drawn for the first body position (b) and the same box was also drawn for the second body position (d) at the same coordinates (regarding the image). One can clearly see that the body of the subject moved to the right of the image during the bulk motion, and that the proposed method managed to catch that motion.

To account for the two body positions and for respiratory motion in PET reconstruction, real-time MR images obtained by the proposed method were grouped into 12 bins (6 bins for each body position) for motion field estimation. Frames in the transition between the two bulk motion phases were excluded (a total of 15 seconds were discarded). Motion was estimated between all bins and the bin corresponding to the end-exhalation, which was used as reference bin. Estimated motion fields are shown in Figure 2.9. The top left image shows motion caused by respiration, mostly visible as a vertical displacement near the liver (shown with more details in the inset image). The left column images show the bulk motion, which is mostly lateral. Finally, the bottom right figure shows a combination of respiratory motion between end-inhalation and end-exhalation and bulk motion.

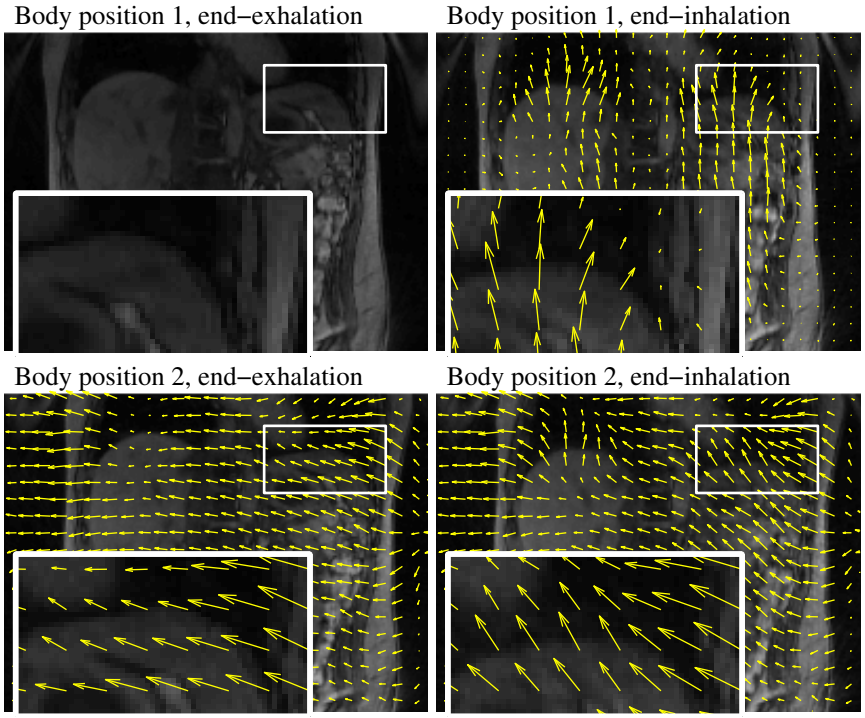


Figure 2.9: Estimated motion field between bins from different bulk motion phases. The top left panel corresponds to the bin used as reference (Body position 1 and the end of exhalation). The top right panel shows the same body position at the end of inhalation; the overlaid motion field exhibits mostly vertical displacement near the liver, corresponding to respiratory motion. The bottom row shows the estimated motion field at the end of exhalation and inhalation for the second body position (after bulk motion). Motion fields demonstrate the lateral displacement between body positions.

Reconstructed PET images are shown in Figure 2.10. Figures 2.10(a) and 2.10(b) show coronal and axial slices using different reconstruction methods. Motion caused blurring artifacts are clearly visible on the NMC reconstruction primarily in the lateral direction, corresponding to bulk motion but also in the vertical direction due to respiratory motion. The gated reconstruction, which uses one sixth of the PET counts at a single body position, shows sharper features but is severely corrupted by noise. The motion-corrected reconstruction using XD-GRASP for motion estimation (denoted by MC-XDG) still shows noticeable motion artifacts because the XD-GRASP reconstruction failed

to capture the bulk motion. The proposed method compensates both respiratory and bulk motion, significantly reducing motion-blur, while exhibiting a low noise level. Figure 2.10(c) shows line profiles through the kidney. Without motion correction, the activity peak is lowered by motion. Instead, two distinct peaks are visible, which correspond to the two bulk motion phases. The two peaks are also visible for the XD-GRASP based motion correction. The gated reconstructions and the motion-corrected reconstructions by the proposed method both preserve the peak activity. However, the gated reconstructions exhibit a high level of noise due to the reduced amount of data used for reconstruction.

Evaluation measures are reported in Table 2.1. The table shows that motion corrected reconstruction leads to the highest CNR: an 83% improvement was observed over reconstruction without motion correction, 47% over XD-GRASP-based motion correction and 198% over gated reconstruction. The noise level in NMC and MC is similar (within 15%) but the contrast is substantially improved by motion correction, while the noise level in the gated reconstruction is three times higher leading to the low CNR. This is expected since gated reconstructions were not tuned to reduce noise but to increase contrast. One could expect a lower CNR for gated reconstruction with a tuned number of subsets and iterations. For the TBR, the gated reconstruction achieves the highest ratio, because gated reconstructions favor high contrast (at the expense of high noise). The proposed motion compensation method approaches the gated TBR (25% decrease) and outperforms NMC (20% increase).

Table 2.1: Contrast-to-noise ratio (CNR) and target-to-background ratio (TBR) for kidney region of interest. See Figure 2.5 for a view of the regions of interest.

	NMC	Gated	MC-XDG	MC-LR
CNR	42.07	25.88	52.49	77.26
TBR	6.05	9.57	6.84	7.25

Correction of irregular respiratory motion

The second experiment was designed to evaluate the performance of the proposed method in the case of irregular respiratory motion. The subject was instructed to alternate between slow deep and fast shallow breaths throughout the 5 minutes PET/MR acquisition.

MR images obtained by the proposed method are shown in Figure 2.11. The top row shows images at different frames: two at the end of inhalation and two at the end of exhalation taken from different breathing patterns (deep/shallow), respectively. The full extent of the respiratory motion is captured and the images are artifacts-free. Figures 2.11(e) and 2.11(f) show 1D profiles of the image through the liver changing over time. Both the images and the plot in Figure 2.11 clearly show the breathing patterns, alternating between deep slow breaths and fast shallow ones. Based on the reconstructed real-time MR images, 12 bins were determined through analysis of the liver displacement in the MR images and were consequently used for motion field estimation and motion corrected PET reconstruction.

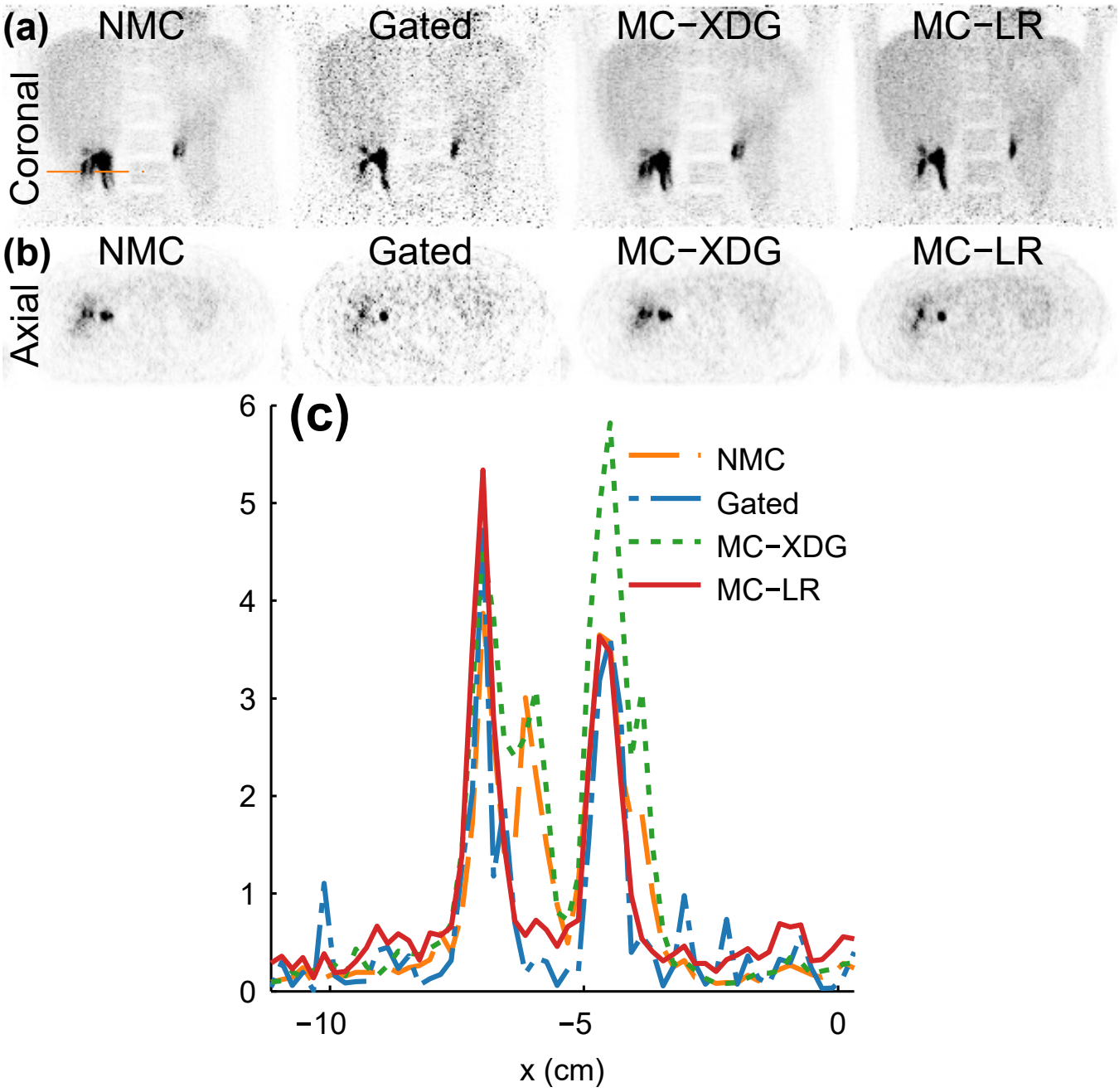


Figure 2.10: PET reconstructions for the bulk motion experiment using three different methods: reconstruction without motion correction (NMC), reconstruction from PET data corresponding to a single respiratory phase and body position (Gated), motion correction using motion estimated from the XD-GRASP MR reconstructions (MC-XDG) and proposed motion-corrected reconstruction from low-rank MR reconstruction (MC-LR). Profile plots through the right kidney (along the orange line drawn on the NMC image) are shown in (c).

Corresponding PET reconstructions are shown in Figure 2.12. Images reconstructed without motion correction (NMC) exhibit blurring artifacts. This is particularly visible on the left kidney (see the green line on the gated coronal image) where the bright spot visible on other images is elongated in the vertical direction, due to the large amplitude of the respiratory motion. The gated

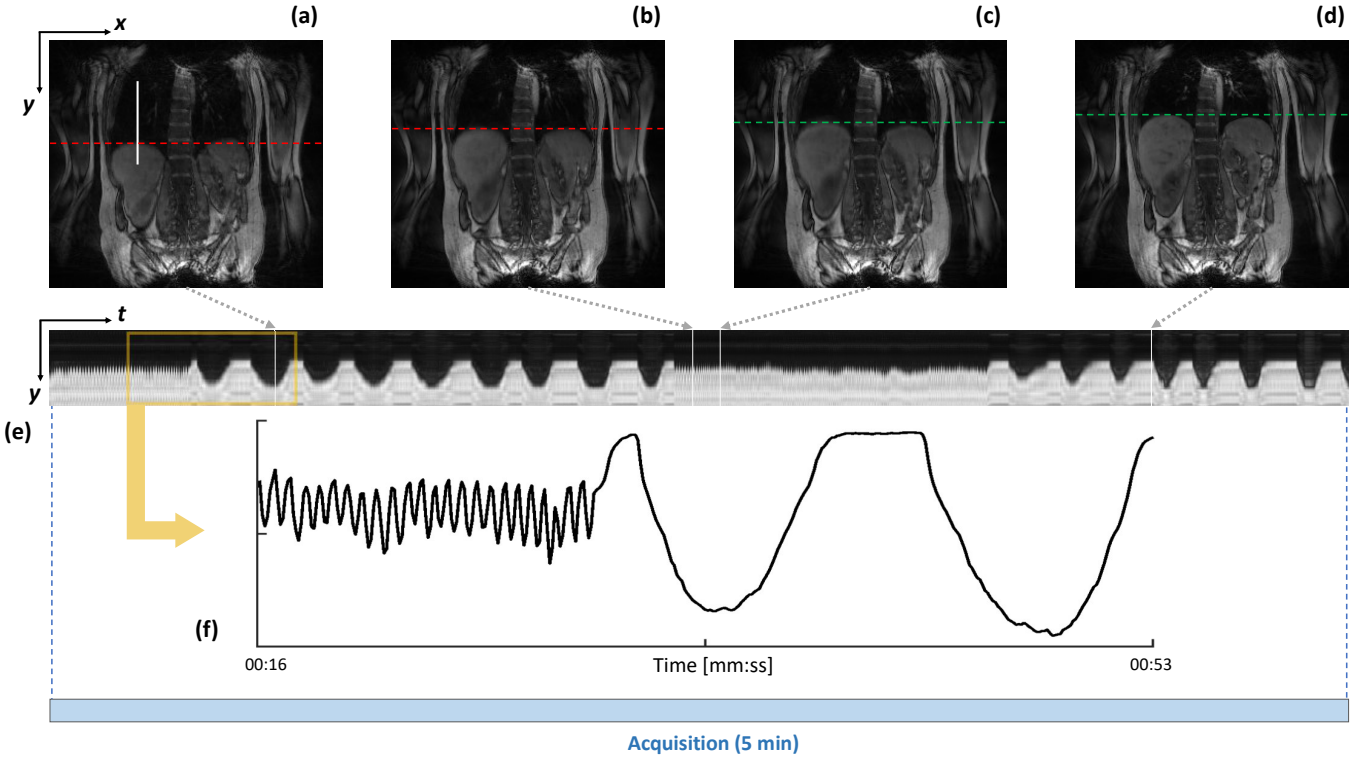


Figure 2.11: Reconstructed MR images in the case of irregular respiratory motion. The real-time profile (e) is plotted for a few minutes of the experiment, where the subject changes their breathing pattern from slow and deep to fast and shallow respiration. Images (a) and (b) show representative images at end-inhalation during deep breath and shallow breath, respectively; (c) and (d) at end-exhalation. The red dashed-line indicates the top of the liver position for end-inhalation and the green dashed-line indicates the top of the liver position for end-exhalation for each breathing pattern.

reconstruction uses one sixth of the total number of counts and therefore is degraded by noise, despite resulting in a sharper image. The motion corrected reconstructions using XD-GRASP and the proposed method for motion estimation both produce sharper images with higher SNR, while the proposed motion correction method results in the best image quality, in terms of noise and contrast. Corresponding line profiles are plotted in Figure 2.12(c). The NMC peak is elongated along the y-axis, due to the large extent of the mostly vertical respiratory motion. The Gated line profile is sharper near its peak but has a large noise level. The proposed MC method results in a good compromise between sharpness and low noise. Contrast-to-noise and target-to-background ratios (defined in Equation (2.11) and Equation (2.12) respectively) are reported in Table 2.2 (regions of interest are shown in Figure 2.5(b)). These measures show the superior performance of the proposed motion correction method. The improvement in CNR is around 163% over NMC and over 200% over gated reconstruction. As for the bulk motion experiment, the gated reconstructions were tuned to maximize contrast rather than reduce noise. The CNR for the gated reconstructions could be improved with further tuning (i.e. by decreasing the number of subsets and iterations). The TBR for the proposed method is within 15% of the gated TBR and around 95% larger than NMC.

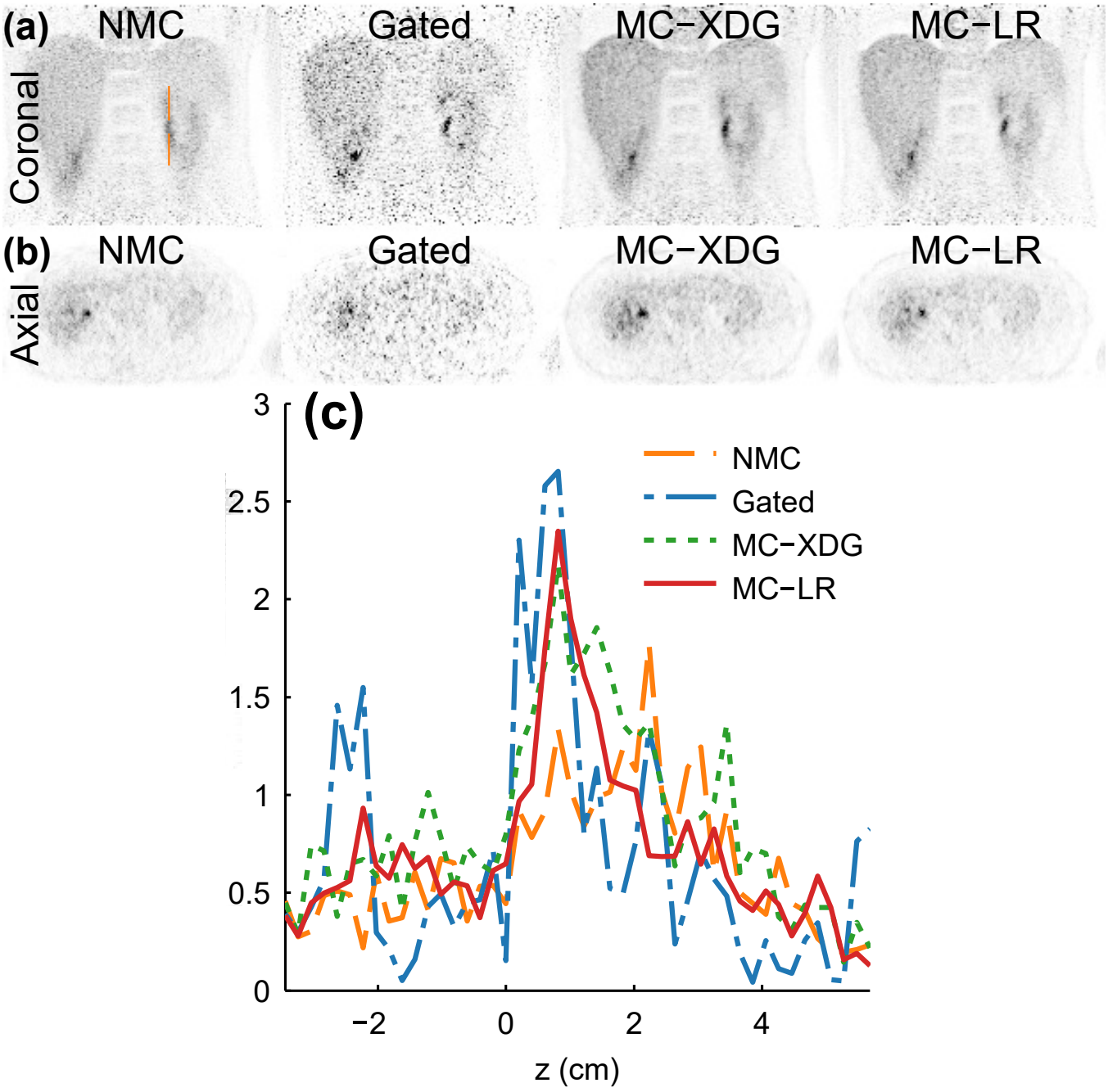


Figure 2.12: PET reconstructions for the irregular motion experiment using three different methods: reconstruction without motion correction (NMC), reconstruction from PET data corresponding to a single respiratory phase (Gated) and proposed motion-corrected reconstruction (MC-LR). (c) shows profile plots through the left kidney (the profile line is shown on the NMC coronal PET image).

Table 2.2: Contrast-to-noise ratio and target-to-background ratio for kidney region of interest. Regions are shown in Figure 2.5.

	NMC	Gated	MC-XDG	MC-LR
CNR	17.52	15.28	37.39	46.18
TBR	3.92	9.00	6.33	7.64

2.2.4 Discussion

We have demonstrated the performance of the proposed MR-based motion correction for PET in two challenging cases: bulk motion and irregular respiratory motion. The proposed subspace-based MR imaging method allows for reconstruction of high-resolution 3D volumes at a rate of 9.5 volume/s, which enables accurate motion field estimation even in the case of irregular motions. Another important benefit of the proposed approach is the ability to perform informed binning for PET motion correction, rather than relying on navigators or external markers which offer limited information on the subject motion. With full real-time volumetric MR images, detecting motion becomes straightforward, and the process of determining an appropriate number of bins is greatly simplified.

The proposed method utilizes an MR acquisition which fully overlaps with the PET acquisition and provides real-time MR images for motion correction. The proposed method can still have benefits for other commonly used acquisition protocols. It is common in practice to reserve a first part of the PET acquisition to perform MR motion field measurements and use the remaining PET acquisition time to perform additional MR measurements (e.g. using T_1 or T_2 contrast sequences) that can be used for other diagnostic tasks [117]. The proposed method can advantageously replace the motion field measurement sequence, possibly reducing the acquisition time while preserving image quality. A gating signal (e.g. navigator or external marker) can then be used in subsequent MR sequences to select an appropriate bin for each PET frame. Another approach is to integrate contrast sequences into the motion field estimation sequence described in this work.

The study reported in this chapter has some limitations. First, the computation time for the low-rank reconstruction with sparsely sampled non-Cartesian k-space data could be a concern [165]. The current MATLAB (The MathWorks, Inc., Natick, Massachusetts, United States) implementation performs reconstruction of one slice and one coil in around one hour. However, progress has been made on that end and the low-rank reconstruction scheme has been implemented using a lower level programming language and parallel computing devices (i.e., GPU) to help achieve reasonable runtimes (about 5 minutes per coil for all slices). Secondly, the proposed method does not have sufficient temporal resolution to resolve the motion in the transition phase between the two bulk motion phases of experiment 1 (Figure 2.8(e)). The time-varying profile plot shows that the image quality in the transition is severely degraded. The corresponding list-mode data were excluded from the PET reconstruction. Since the duration of the bulk motion was short, only about 5% of the list-mode data were discarded and thus should not be a significant limitation. Thirdly, this study focuses on demonstrating the feasibility of using subspaced-based real-time MR for PET motion correction. We showed the performance of our method in two cases (bulk motion and irregular respiratory motion) from *in vivo* PET/MR experiments on a healthy subject. More subjects are needed to fully evaluate the performance of the proposed method in clinical settings. It is also worthwhile to explore incorporating advanced system modeling including time-of-flight information and point-spread function modeling into the proposed method. While advanced imaging models

are expected to improve low-count (e.g. gated) reconstructions and quantitative metrics, all reconstruction methods, including the proposed motion correction, would benefit from more accurate modeling.

2.3 Application to cardiac T_1 mapping

2.3.1 Introduction

Cardiac T_1 mapping is a powerful cardiovascular magnetic resonance imaging (MRI) technique that allows for quantitative assessment of tissue characteristics and underlying pathology of myocardium. Native (i.e., without the usage of exogenous contrast agent) myocardial T_1 is a well-recognized biomarker that characterizes alterations in the structure and intra-/extra-cellular components of myocardium, such as iron deposition, amyloid disease, Anderson-Fabry disease, and myocarditis [47, 98, 136]. Extracellular volume fraction (ECV), measured from pre- and post-contrast T_1 values, provides quantitative measurement of interstitial expansion and associated diseases, such as amyloidosis, fibrosis, or myocardial edema [47, 98]. ECV is an emerging biomarker for diffuse fibrosis (e.g., in heart failure, dilated cardiomyopathy, and amyloidosis) [98], which is challenging to detect using qualitative late gadolinium enhancement (LGE) methods.

Modified Look-Locker inversion-recovery (MOLLI) [97] is a widely used method for 2D cardiac T_1 mapping which utilizes adiabatic inversion pulses for magnetization preparation and performs electrocardiogram (ECG)-gated balanced steady-state free precession (bSSFP) acquisitions through multiple cardiac cycles with breath-holding. Although MOLLI produces myocardial T_1 maps with high precision [68], the method is limited to a single-slice imaging per breath-hold. Methods have been developed to extend conventional 2D MOLLI method to multi-slice 2D or 3D acquisitions with breath-holding by leveraging the state-of-the-art parallel imaging, simultaneous multi-slice acquisition, compressed sensing, and non-Cartesian sampling techniques [17, 160]. However, these methods suffer from limited through-plane resolution and coverage, often involving long or repetitive breath-holds to obtain volumetric T_1 maps of the heart which imposes significant burden on patients.

Various methods have been developed to overcome the limitations of breath-holding and allow for 3D cardiac T_1 mapping with free-breathing acquisitions. Respiratory and cardiac gating-based T_1 mapping methods have been developed to acquire interleaved multi-slice 2D [46, 63, 159] or segmented 3D k-space data [45, 106, 158] at end-diastole with free-breathing, where effects from respiratory motion were mitigated by prospectively tracking respiratory motion using navigators or self-navigation techniques. MR fingerprinting approaches have been combined with free-breathing ECG gated acquisitions for multi-parametric cardiac MRI [26]. However, most of these methods are limited by spatial coverage, resolution in slice direction, or imaging time due to the low data acquisition efficiency of gating. Recently, free-running (i.e., no cardiac or respiratory gating) continuous acquisition methods have been proposed for 2D or 3D cardiac T_1 mapping [23, 49, 65, 101, 123, 142]. Of note, Qi et al. have reported a free-running 3D whole-heart T_1 mapping method [123],

which uses translational respiratory motion correction and a patch-based low-rank tensor model to reconstruct 3D T_1 maps with isotropic resolution. The T_1 maps obtained by this method, however, were from 1.5T and the method may not translate well to 3T for reasons discussed below.

While 3D cardiac T_1 mapping methods developed up to date have been applied mostly at 1.5T, unique technical challenges arising from more severe B_0 and transmit B_1 (B_1^+) inhomogeneities need to be addressed at 3T. For instance, spoiled gradient-echo (SPGR) readout with small flip angle (FA) is often used in 3D cardiac imaging at 3T to avoid B_0 inhomogeneity-caused banding-artifacts associated with bSSFP readout [17, 23, 49, 142]. However, T_1 mapping with SPGR readout is known to be sensitive to errors in FA caused by imperfect RF pulses and B_1^+ inhomogeneities [33]. The latter is particularly problematic at 3T, where B_1^+ variation over the left ventricle with body-coil transmission is reported on the order of 30-60% [148], leading to bias in T_1 estimation. The robustness of cardiac T_1 mapping methods with SPGR acquisitions needs to be thoroughly investigated in the presence of B_1^+ inhomogeneity [141, 172].

In this work, we present a new cardiac T_1 mapping method for rapid 3D T_1 mapping of the heart at 3T. Here, we focus on the main contribution of the author of this PhD thesis, which was the MR image reconstruction of high-dimensional data with dynamic contrast (due to the inversion-recovery sequence). To do so, a sparse (k,t)-space sampling along a stack-of-stars trajectory was used to accelerate imaging and a subspace-based image reconstruction method was used to recover high frame-rate dynamic images from highly undersampled (k,t)-space data. These reconstructions were then used for T_1 fitting to obtain 3D T_1 maps of the heart, as presented in [54]. The details of the improvement in the acquisition protocol, T_1 fitting, and other aspect of the study can be found in the published article [54], also provided in Appendix A. The performance of the proposed method was characterized and validated through numerical simulations, phantom studies, and *in vivo* experiments on healthy human subjects (n=6). Preliminary accounts of this work have been presented previously in the form of conference abstracts [51, 52, 53].

2.3.2 Methods

Data Acquisition

The proposed ECG-gated cardiac T_1 mapping sequence is shown in Figure 2.13. A non-selective inversion pulse was applied every N+M heartbeats with two different inversion times (TI). This scheme is referred to as N-(M)-N-(M) protocol for simplicity, where N denotes number of cardiac cycles for acquisition and M denotes number of cardiac cycles for signal recovery. Data were acquired at end-diastole period using SPGR readout. A special data acquisition scheme was employed for sparse sampling (k,t)-space data along a stack-of-stars trajectory, as described in Section 2.1.2. A limited number of "training" data (e.g., along the k_x, k_y, and k_z directions across the center of the k-space) were acquired with high sampling rate to determine the temporal changes of the underlying signal. Data at all other k-space locations were sparsely-sampled over the entire (k,t)-space to ensure that a sufficient number of measurements were acquired at each k-space location

for subspace-based image reconstruction.

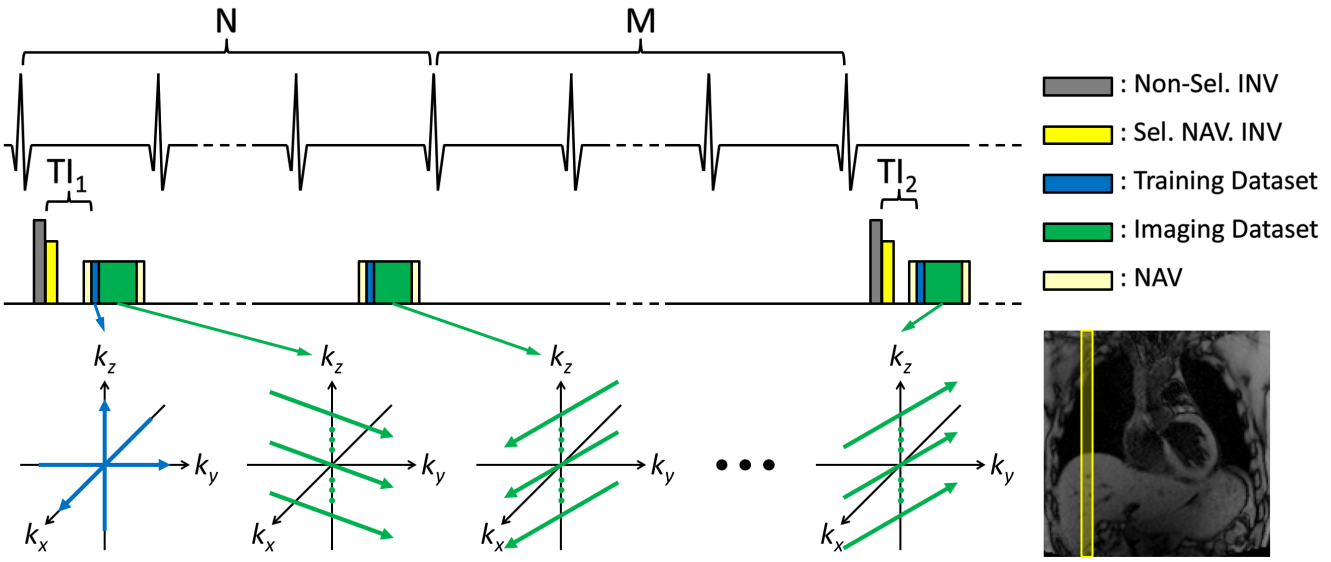


Figure 2.13: Schematic diagram of the proposed data acquisition scheme. N-(M)-N-(M) protocol is shown with non-selective inversion pulse applied every N+M heartbeats (where N denotes the number of cardiac cycles for acquisition and M denotes the number of cardiac cycles for signal recovery) with two different inversion times (TI). Data acquisition consists of "training" dataset acquiring a limited number of k-space lines with high sampling rate and "imaging" dataset sparsely-sampling all other k-space locations for subspace-based image reconstruction (see Section 2.1.2). To track respiratory motion, 1D respiratory navigator signals were acquired in the sagittal plane at the dome of the right hemidiaphragm after a spatially-selective inversion pulse was applied in the same sagittal plane to invert the magnetization signals back to the equilibrium state.

To track respiratory motion, 1D respiratory navigator signals were acquired in the sagittal plane at the dome of the right hemidiaphragm at the beginning and end of data acquisition of each cardiac cycle. A spatially-selective inversion pulse was applied right after the non-selective inversion pulse to invert the magnetization signals in the same sagittal plane back to the equilibrium state and, therefore, to mitigate the contrast changes caused by the non-selective inversion pulses in the navigator signals.

Image reconstruction

Image reconstruction of sparsely-sampled data was performed by solving a constrained optimization problem using the subspace model from Equation (2.6) with a sparsity constraint in the spatio-spectral domain [171]:

$$\boldsymbol{\Theta} = \arg\min_{\boldsymbol{\Theta}} \left\| \Omega \left(\boldsymbol{F_s} \boldsymbol{\Theta} \boldsymbol{\Phi}^{\mathrm{T}} \right) - \mathbf{s} \right\|_2^2 + \lambda \left\| \mathbf{F}_t (\boldsymbol{\Theta} \boldsymbol{\Phi}^{\mathrm{T}}) \right\|_1 + \mu \left\| \boldsymbol{\Theta} \right\|_F^2 , \tag{2.13}$$

where, as in Section 2.1.4, $\mathbf{s}$ is the measured k-space data, $\boldsymbol{F_s}$ is the Fourier transform operator in the spatial domain, i.e., Non-Uniform FFT (NuFFT) operator [32] for the stack-of-stars trajectory, Ω is the sparse sampling operator in the (k, t)-space, $\mathbf{F}_t$ is the Fourier transform operator in the

temporal domain, $\|.\|_F$ is the Frobenius norm and the scalar variables λ and μ are regularization parameters. Note that compared to the image reconstruction in Section 2.2.2, only the sparsifying transform is different.

In this work, $\mathbf{\Phi}$ was estimated separately from training data using singular value decomposition (SVD) as described in Section 2.1.3. The image reconstruction problem is then reduced to determining the spatial coefficients $\mathbf{\Theta}$ from the measured data. The optimization problem in Equation (2.13) was solved using an alternating direction method of multipliers (ADMM) [12]. For fast computation of F_s, 1D Fourier transform was applied along the k_z direction first and the NuFFT was applied in the k_x-k_y plane for slice-by-slice reconstruction. The dynamic images were reconstructed in a coil-by-coil fashion and then combined using the sum-of-squares method to form the final reconstruction. We implemented the image reconstruction algorithm in Python and utilized the SigPy package [108] to accelerate the computation using Graphical Processing Units (GPUs). Reconstructions were performed on four NVIDIA Tesla V100 SXM2 GPUs (parallelized over slice and coil dimensions) with reconstruction time around one minute for each slice and coil.

In vivo study

Six healthy volunteers (four males and two females; 32 ± 3 years) were recruited under a study protocol approved by our local Institutional Review Board (IRB). Written informed consent was obtained from all subjects before study participation. Imaging experiments were performed using a 3T MR scanner (MAGNETOM Trio, Siemens Healthcare, Erlangen, Germany) with a body-coil for transmission and spine and surface coils for reception. Imaging was performed using 10-(3)-10-(3) protocol with data sampling following a stack-of-stars trajectory. Two frames were acquired per cardiac cycle, each consisting of k-space spokes along the same angle in the k_x-k_y plane over all k_z encodings and 3 additional training lines at the center of the k-space along the k_x, k_y, and k_z direction, respectively (Figure 2.13). The spoke angle varied randomly from frame to frame following uniform random distribution. The other imaging parameters were: FOV=$308 \times 308 \times 180\text{mm}^3$, matrix-size=$160 \times 160 \times 40$, image orientation=short-axis view, FA=9°, TR/TE=3.4/1.7ms, and inversion delay times=100/180ms. The reconstructed images were then used for T_1 fitting. More details on the acquisition and T_1 and B_1 maps estimation can be found in Appendix A.

2.3.3 Results

In vivo study

The results from the *in vivo* study are shown in Figures 2.14. Figure 2.14 shows representative reconstructed images at various slice positions and inversion times from Subject 1 using the proposed method. Visually, no noticeable artifacts were shown in the reconstructed images from the proposed method across different slices and TI times. More result can be found in Appendix A.

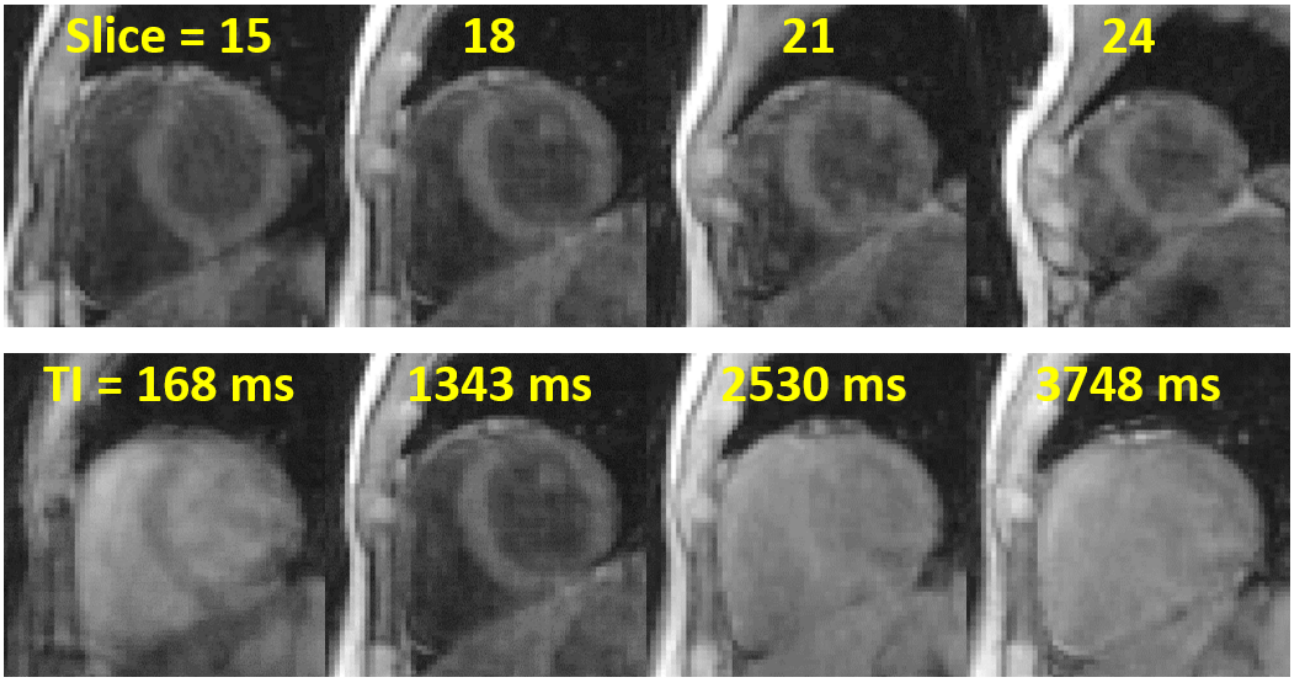


Figure 2.14: *in vivo* study results of reconstructed images from the proposed method. Representative reconstructed images from Subject 1 are shown at various slice positions for a fixed TI of 1343 ms (top row) and various TI times for slice position index of 18 (bottom row).

2.3.4 Discussion

In this work, we proposed a new cardiac T_1 mapping method for fast, high-resolution 3D T_1 mapping of the whole heart at 3T. The proposed method has several novel features compared to the existing cardiac T_1 mapping methods, among which acquisition protocol advances and simulation studies (presented in Appendix A). The novel feature presented in this manuscript is how the proposed method utilizes special (k,t)-space sampling scheme and subspace-based image reconstruction to recover dynamic images in clock time from the undersampled data, i.e., two 3D volumes for every cardiac cycle. This allows mitigating the effects of natural heart-rate variations on T_1 mapping by fitting the reconstructed dynamic signals to a signal dictionary generated with subject-specific timing of data acquisition recorded during the imaging experiment. The data acquisition advances along with the subspace-based image reconstruction with sparsity constraint allowed for accurate T_1 mapping in the presence of B_1^+ inhomogeneity at 3T with high spatial resolution (i.e., 1.9 mm in-plane and 4.5 mm through-plane) and large through-plane coverage (i.e., 40 slices) in a 10-min free-breathing acquisition.

The current work has several limitations. First, the method proposed in this work involves ECG-gated acquisition, which is susceptible to ECG mis-triggering and may suffer from image blurring due to cardiac motion. The former could be addressed by adaptive heartbeat rate prediction as in the double-gating technique [31]. The latter can be mitigated by retrospectively discarding k-space

data acquired outside the end-diastole window based on the recorded ECG signals. Free-running (i.e., no cardiac or respiratory gating)-based continuous acquisition scheme may be preferable over ECG-gated acquisition schemes for maximizing data acquisition efficiency. However, in the case of free-running continuous acquisition scheme, the inversion recovery rate is strongly coupled with B_1^+ and estimation of B_1^+ is expected to be critical for accurate and precise T_1 estimation in the presence of B_1^+ inhomogeneity at 3T. Secondly, the performance of the proposed was validated using *in vivo* experiments with a small number of healthy subjects (n=6). Studies with a larger number of healthy subjects and patients are needed to assess the accuracy and reproducibility of the proposed method and to evaluate its value in clinical applications. More detailed discussion on the estimation of the T_1 maps and B_1^+ inhomogeneity (e.g., joint T_1 B_1^+ estimation) can be found in Appendix A.

2.4 Conclusion

This chapter presented the subspace model (also denoted as low-rank model) and how it can be efficiently used for MR image reconstruction of highly undersampled k-space data. We first presented an application for PET/MR scanner where dynamic MR data are used for PET data motion correction. In this contribution, we proposed an MR-based method for PET motion correction using a subspace-based real-time MR imaging for motion field estimation. We demonstrated the feasibility of the proposed method using ^{18}F-FDG-PET/MR studies on a healthy subject. Our results show that the proposed method can capture and correct for normal and irregular respiratory motions as well as bulk body motion. The proposed method can be beneficial to a range of clinical applications where irregular motion patterns are expected. Then, we demonstrated the versatility of the subspace model by presenting how that same model can be used for free-breathing 3D cardiac T_1 mapping. A new method was developed and optimized for fast, high-resolution 3D T_1 mapping of the whole heart at 3T with advances in the acquisition side and the reconstruction side. The main contribution of the author in this PhD thesis was the image reconstruction, where we utilized the subspace model with sparsity constraint implemented on GPU to achieve 3D T_1 mapping of the whole-heart with high spatial resolution and through-plane coverage in 10 minutes with reasonable reconstruction time.

However, we do notice some limitations in this linear subspace model. For example, this model was not powerful enough to reconstruct images from 10 minutes worth of continuously sampled k-space data and we acquired ECG-triggered data to simplify the reconstruction process as cardiac motion would not come into play. Indeed, continuous free-breathing 3D cardiac scans induce a lot of dynamics in the data: respiratory motion, cardiac motion, and dynamic contrast on top of that. That is without mentioning the very high undersampling rate due the large 3D matrix size necessary for high spatial resolution, exponentially increasing the number of unknowns. All these reasons will cause the subspace model to fail to reconstruct artifacts-free images as it assumes linear dimensionality reduction is enough to model complex data. Then, such a model may not

handle the increasing complexity in the data (i.e., without ECG-triggering), as shown in the next chapter (Figure 3.9). To alleviate ourselves from these limitations, we developed a new MR image reconstruction method with more flexibility and better modeling of the underlying manifold (i.e., in a non-linear way). We present this model in Chapter 3 and demonstrate its added value to MR image reconstruction.

Chapter 3

Non-linear manifold learning-based model: Linear tangent space alignment

Chapter 2 presented a strategy using a linear subspace to represent the MR data, and reconstruct them with fewer unknowns by fixing a set of temporal basis functions. Despite offering promising results, this technique relies on a linear dimensionality reduction model of high-dimensional data, which may show its limits with higher undersampling rate. We derive and present here a novel and more versatile MR image reconstruction method relying on non-linear dimensionality reduction techniques through manifold-learning and tangent space alignment. The rest of this chapter is organized as follows. Section 3.1 introduces related work on dynamic MRI. Section 3.2 presents background information of dynamic MR, subspace model, and tangent spaces of a manifold. Section 3.3 then describes the LTSA model-based MR image reconstruction method, a novel MR image reconstruction scheme developed during the PhD thesis. Section 3.4 presents results and comparisons of the proposed method with state-of-the-art methods. Finally, Section 3.5 discusses the proposed method and presents how the LTSA model is a non-linear generalization of existing linear subspace models.

3.1 Related Work

A key challenge in dynamic Magnetic Resonance Imaging (MRI) is reconstructing high-dimensional images from sparse (k, t)-space data sampled below the Nyquist sampling rate. Many methods have been proposed for accelerated imaging with sparse sampling, including parallel imaging [43, 121, 146], compressed sensing [88, 89], and more recently, the subspace model based imaging methods [21, 22, 81, 83, 114, 171]. The subspace model, or the so-called partially separable (PS) model [81], assumes that high-dimensional data reside in a low-dimensional subspace which captures the spatiotemporal correlations of dynamic MR images. Mathematically, the subspace model leads to a low-rank (LR) model of the spatiotemporal signals, which significantly reduces the degrees of freedom of the underlying signals and enables reconstruction of dynamic MR images from highly undersampled (k, t)-space data. The subspace model has been extended to the local low-rank (LLR) model [151,

152], exploiting the local low-dimensional subspace structure, and to the low-rank tensor (LRT) model [55] leveraging the low-dimensional structure of multi-array data for accelerated dynamic MRI. The subspace model-based image reconstruction methods have been successfully applied to free-running multi-parametric cardiac MR, e.g., the multitasking method [23], and high-resolution MR Spectroscopic Imaging (MRSI) [74, 75, 92].

The dimensionality of dynamic MR images can be further reduced using nonlinear models. For example, deep neural networks (DNNs) have been studied for applications in MR image reconstruction for improved denoising and dealiasing [79, 87], often requiring a large database with labels for training. By contrast, manifold learning-based methods aim at learning the intrinsic low-dimensional structure of the input data that are embedded in a high-dimensional signal space without any training labels. Many methods have been proposed to solve the nonlinear dimensionality reduction problem, including locally linear embedding (LLE) [132], Isomap [149], graph Laplacian [6], etc. These methods can be interpreted by a kernel principal component analysis (kernel PCA) framework [48], where an implicit mapping between training data and a feature space is constructed while preserving the local geometry of the training data. Manifold learning-based methods have been applied to pattern recognition, data clustering and classification, showing superior performance over the conventional linear model-based methods such as principal component analysis (PCA) [48].

Manifold models have also been explored for image reconstruction in dynamic MRI [16, 102, 105, 118, 119, 143, 154]. Different from applications in data clustering and classification, where tasks are performed in the feature space, manifold learning based-image reconstruction requires solving an additional manifold reconstruction problem (or the so-called pre-imaging problem), which maps the learned low-dimensional representation in the feature space back to the signal space. This imposes a unique challenge for manifold model-based image reconstruction when the kernel matrix is, as often the case in practice, formed implicitly from training data. The SToRM method and its variations [2, 7, 102, 118, 119, 173] address this issue by adding a manifold regularization term to the image reconstruction problem, which penalizes the smoothness of the embedding manifold through a graph Laplacian matrix formed by the input data. The kernel-based low-rank (KLR) method [105] explicitly models dynamic MR images as low-dimensional manifolds for image reconstruction. Shetty et al., have recently proposed a machine-learning framework, called bi-linear modeling of data manifolds (BiLMDM), for image reconstruction in dynamic MRI [143]. While both the BiLMDM method and the proposed method represent dynamic images using a bi-linear model, the two methods are different from each other in terms of model construction and image reconstruction algorithm (see Section 3.5.3 for more detailed discussions on the difference between the proposed method and BiLMDM).

However, these models have some limitations. First, applying non-linear dimensionality reduction (i.e., finding the best low-dimensional feature space) is not enough in manifold learning methods, as one needs to be able to go back to the input space. In image reconstruction, this is called the pre-imaging problem. For example, in the case of PCA, the mapping is straightforward

as the model is linear, and we can just express the low-dimensional coordinates as a linear combination of basis function vectors. Methods based on kernel PCA [105] are limited by the use of generic kernels with invertible mappings between the signal space and the feature space, which may not be most effective in capturing the inherent low-dimensional structure of the underlying data. This type of method does not have data-driven non-linear mapping but "man-made" kernels for easier pre-imaging, which greatly reduces the representation power. Methods such as SToRM [118] astutely penalize the low-dimensional manifold from the input space to avoid solving the pre-imaging problem, but such methods do not manage to explicitly characterize the underlying manifold, which in turn may also decrease the representation power of the model.

This chapter presents an alternative strategy for manifold learning-based image reconstruction in dynamic MRI. The proposed method is closely related to machine learning methods that learn manifolds via linear tangent space alignment (LTSA) [170]. While most manifold learning methods focus on learning a low-dimensional representation of a manifold in a feature space, the LTSA methods solve both the nonlinear dimensionality reduction and manifold reconstruction problems explicitly. In this work, we propose a LTSA model-based framework for reconstructing dynamic MR images from highly undersampled (k, t)-space data. An efficient Alternating Direction Methods of Multipliers [11] (ADMM)-based algorithm is developed to solve the resultant image reconstruction problem with additional sparsity constraints. Numerical simulation studies as well as *in vivo* 2D and 3D cardiac imaging experiments are performed to validate the performance of the proposed method and compare with state-of-the-art image reconstruction methods.

3.2 Background

3.2.1 Subspace model of dynamic images

We use in this chapter the same notations for the signal model and the subspace model of the MR signal as in Section 2.1.1. As a reminder, the temporal signal at the m-th voxel of the dynamic images (denoted by $\mathbf{z}_m = \left[x(\mathbf{r}_m, t_1), \ldots, x(\mathbf{r}_m, t_N)\right]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}$) , can be expressed as:

$$\mathbf{z}_m = \mathbf{\Phi}\boldsymbol{\theta}_m, m = 1, \ldots, M, \tag{3.1}$$

or equivalently in the form of a low-rank decomposition of the Casorati matrix formed by the dynamic images:

$$\mathbf{X} = \mathbf{\Theta}\mathbf{\Phi}^{\mathrm{T}}, \tag{3.2}$$

where $\mathbf{\Theta} \in \mathbb{C}^{M\times D}$ concatenates along rows the spatial coefficients $\boldsymbol{\theta}_m^T$ of all M voxels with respect to the basis of the temporal subspace $\mathbf{\Phi} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_D] \in \mathbb{C}^{N\times D}$.

As detailed in Section 2.1.1, the subspace model assumes that the high-dimensional input data (i.e., dynamic images) reside in a low-dimensional subspace. Then, one just needs to use a sufficiently

high subspace dimension to capture the temporal changes of the signals in dynamic MRI. However, the subspace model is a linear model treating the data globally: it does not take advantage of the local structure of the data in the low-dimensional space, resulting in a potentially larger number of unknowns than actually needed. On the other side, manifold models leverage the intrinsic low-dimensional nonlinear manifold structure of the dynamic MRI signals for image reconstruction [16, 102, 105, 118, 119, 143, 154], leading to a smaller number of unknowns and thus reduced imaging time and better image reconstruction performance compared to the linear subspace model.

3.2.2 Tangent space of a manifold

A parametric manifold model of the temporal signal $\mathbf{z}_m$ is given by:

$$\mathbf{z}_m = f(\boldsymbol{\tau}_m), m = 1, \ldots, M, \tag{3.3}$$

where $\boldsymbol{\tau}_m \in \mathbb{C}^{D\times 1}$ represents a low-dimensional vector (or global coordinate) in the feature space, and f is a nonlinear function mapping a vector in the feature space to a data point in the signal space.

If f is smooth, the signal at a location $\boldsymbol{\tau}_m$ can be approximated by its neighbor at $\boldsymbol{\tau}$ using the first-order Taylor expansion:

$$f(\boldsymbol{\tau}_m) = f(\boldsymbol{\tau}) + \mathbf{J}_f(\boldsymbol{\tau})\big(\boldsymbol{\tau}_m - \boldsymbol{\tau}\big) + \mathcal{O}\big(\|\boldsymbol{\tau}_m - \boldsymbol{\tau}\|_2^2\big), \tag{3.4}$$

where $\mathbf{J}_f(\boldsymbol{\tau}) \in \mathbb{C}^{N\times D}$ is the Jacobian matrix of f. The tangent space $\mathcal{T}_{\boldsymbol{\tau}}$ of f is defined as the subspace spanned by the column vectors of $J_f(\boldsymbol{\tau})$. If the manifold is regular, $J_f(\boldsymbol{\tau})$ is full rank for any $\boldsymbol{\tau}$ [170].

3.3 Linear tangent space alignment

This work presents a linear tangent space alignment (LTSA) model-based image reconstruction method, which leverages the intrinsic low dimensional manifold structure of the underlying signals for accelerated dynamic MRI through tangent space alignment. Specifically, the dynamic images are first grouped into several neighborhoods along the time axis. The local geometry of the manifold is learned via a subspace approximation of the images in each neighborhood. The learned local coordinates can be aligned to the global coordinates of the manifold using linear transforms if the underlying manifold is smooth. In image reconstruction, the global coordinates along with the linear transform matrices are determined by fitting the LTSA model to the measured k-space data. The LTSA model and the formulation of the corresponding image reconstruction problem are discussed in detail below.

3.3.1 LTSA model of dynamic images

Given a set of dynamic images expressed in the form of a Casorati matrix $\mathbf{X}$ as in Equation (3.2), we assume that the temporal signals $\mathbf{z}_m = \left[x(\mathbf{r}_m, t_1), \ldots, x(\mathbf{r}_m, t_N)\right]^{\mathrm{T}}$ reside in a low-dimensional manifold:

$$\mathbf{z}_m = f(\boldsymbol{\tau}_m), m = 1, \ldots, M, \tag{3.5}$$

where $\boldsymbol{\tau}_m \in \mathbb{C}^{D\times 1}$ denotes the global coordinate of $\mathbf{z}_m$ in the feature space.

We propose to group the dynamic images into C neighborhoods based on a certain similarity metric:

$$\mathbf{X}_c = \begin{bmatrix} x(\mathbf{r}_1, t_{j_{c,1}}) & \ldots & x(\mathbf{r}_1, t_{j_{c,P_c}}) \\ \vdots & \ddots & \vdots \\ x(\mathbf{r}_M, t_{j_{c,1}}) & \ldots & x(\mathbf{r}_M, t_{j_{c,P_c}}) \end{bmatrix} \in \mathbb{C}^{M\times P_c}, \tag{3.6}$$

where $\{t_{j_{c,p}}\}_{p=1}^{P_c} \subset [1, \ldots, N]$ are the selected temporal frames in neighborhood c. The temporal signals in each neighborhood, denoted by $\mathbf{z}_m^{(c)} = \left[x(\mathbf{r}_m, t_{j_{c,1}}), \ldots, x(\mathbf{r}_m, t_{j_{c,P_c}})\right]^{\mathrm{T}} \in \mathbb{C}^{P_c\times 1}$ are given by

$$\mathbf{z}_m^{(c)} = \boldsymbol{\Pi}_c \mathbf{z}_m = \boldsymbol{\Pi}_c f(\boldsymbol{\tau}_m) \in \mathbb{C}^{P_c\times 1}, \tag{3.7}$$

where $\boldsymbol{\Pi}_c \in \mathbb{R}^{P_c\times N}$ is a selection operator determined by the indices $\{j_{c,p}\}_{p=1}^{P_c}$. Note that the neighborhoods may overlap, in which case the selection operator $\boldsymbol{\Pi}_c$ is normalized by the number of neighborhoods that a frame contributes to.

If the underlying manifold is smooth, the signals in the neighborhood c can be approximated using the first-order Taylor expansion:

$$\begin{aligned} \mathbf{z}_m^{(c)} &\approx \boldsymbol{\Pi}_c f(\bar{\boldsymbol{\tau}}) + \boldsymbol{\Pi}_c \mathbf{J}_f(\bar{\boldsymbol{\tau}})(\boldsymbol{\tau}_m - \bar{\boldsymbol{\tau}}) \\ &\approx \boldsymbol{\Pi}_c \mathbf{J}_f(\mathbf{0})\boldsymbol{\tau}_m. \end{aligned} \tag{3.8}$$

Here, we further assume that the centroid of the global coordinates is $\bar{\boldsymbol{\tau}} = \mathbf{0}$ and $f(\bar{\boldsymbol{\tau}}) = \mathbf{0}$ for the simplicity of expressions and without loss of generality. However, direct calculation of the Jacobian of the manifold $\mathbf{J}_f$ is difficult because the nonlinear mapping f is built implicitly.

If the manifold is also regular, Equation (3.8) indicates that the Casorati matrix in Equation (3.6) can be approximated by a low-rank matrix, which can be obtained by PCA:

$$\mathbf{X}_c \approx \boldsymbol{\Theta}_c \boldsymbol{\Phi}_c^{\mathrm{T}}, \tag{3.9}$$

where $\boldsymbol{\Phi}_c \in \mathbb{C}^{P_c\times D}$ concatenates along columns the temporal bases that span the subspace of $\{\mathbf{z}_m^{(c)}\}_{m=1}^{M}$, and $\boldsymbol{\Theta}_c \in \mathbb{C}^{M\times D}$ concatenates along rows the corresponding local coordinates. Here we assume that the rank of $\boldsymbol{\Pi}_c \mathbf{J}_f(\mathbf{0})$ is also D for the simplicity of expressions. Note that $\boldsymbol{\Phi}_c^{\mathrm{T}} \boldsymbol{\Phi}_c = \mathbf{I} \in \mathbb{C}^{D\times D}$.

Comparing Equation (3.8) and Equation (3.9), the local coordinates in Equation (3.9) can be

aligned with the global coordinates through a linear transform[a]:

$$\begin{aligned}\boldsymbol{\Theta}_c &= \mathbf{T}\mathbf{J}_f(\mathbf{0})^{\mathrm{T}}\boldsymbol{\Pi}_c^{\mathrm{T}}\boldsymbol{\Phi}_c \\ &= \mathbf{T}\mathbf{L}_c,\end{aligned} \tag{3.10}$$

where $\mathbf{T} = [\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_M]^{\mathrm{T}} \in \mathbb{C}^{M\times D}$ concatenates along rows the global coordinates of the manifold and $\mathbf{L}_c \in \mathbb{C}^{D\times D}$ represents a linear transform aligning the local coordinates with the global coordinates. The matrices of local coordinates $\boldsymbol{\Theta}_c$ of each neighborhood can be concatenated into a matrix $\boldsymbol{\Theta} = [\boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_C] \in \mathbb{C}^{M\times CD}$. Equation (3.10) indicates that $\boldsymbol{\Theta}$ has the following low-rank decomposition:

$$\boldsymbol{\Theta} = \mathbf{T}\mathbf{L}, \tag{3.11}$$

where $\mathbf{L} = [\mathbf{L}_1, \ldots, \mathbf{L}_C] \in \mathbb{C}^{D\times CD}$.

For image reconstruction applications, Equation (3.9) and Equation (3.10) lead to a linear tangent space alignment (LTSA) model that explicitly reconstructs dynamic images from feature vectors:

$$\mathbf{X}_c = \mathbf{T}\mathbf{L}_c\boldsymbol{\Phi}_c^{\mathrm{T}}, \tag{3.12}$$

and

$$\mathbf{X} = \sum_{c=1}^{C} \mathbf{T}\mathbf{L}_c\boldsymbol{\Phi}_c^{\mathrm{T}}\boldsymbol{\Pi}_c. \tag{3.13}$$

3.3.2 LTSA model-based MR image reconstruction

The image reconstruction problem is formulated as determining the global coordinates $\mathbf{T} = [\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_M]^{\mathrm{T}}$ and linear transform matrices lumped in $\mathbf{L} = [\mathbf{L}_1, \ldots, \mathbf{L}_C]$ by fitting the LTSA-model of dynamic images in Equation (3.13) to the measured (k, t)-space data with a sparsity constraint:

$$\begin{aligned}\arg\min_{\mathbf{T},\mathbf{L}} \left\| \boldsymbol{\Omega}\Big(\mathbf{F}_s \sum_{c=1}^{C} \mathbf{T}\mathbf{L}_c\boldsymbol{\Phi}_c^{\mathrm{T}}\boldsymbol{\Pi}_c\Big) - \mathbf{s} \right\|_2^2 + \frac{\mu_T}{2}\left\|\mathbf{T}\right\|_F^2 \\ + \frac{\mu_L}{2}\left\|\mathbf{L}\right\|_F^2 + \lambda\left\|\mathcal{D}(\mathbf{T}\mathbf{L})\right\|_1,\end{aligned} \tag{3.14}$$

where, as in Section 2.1.4, $\mathbf{s}$ is the measured k-space data, $\boldsymbol{F_s}$ is the Fourier transform operator in the spatial domain, i.e., Non-Uniform FFT (NuFFT) operator [32] for the stack-of-stars trajectory, Ω is the sparse sampling operator in the (k, t)-space, $\mathcal{D}$ is the finite difference operator along the spatial dimensions, $\|.\|_F$ is the Frobenius norm and the scalar variables λ and μ_T and μ_L are regularization parameters.

In Equation (3.14), the first term of the cost function is a data fidelity penalty; the second

[a]Note that stacking the M temporal signal $\mathbf{z}_m^{(c)}$ on the 2^{nd} dimension gives $\mathbf{X}_c^{\mathrm{T}} = \boldsymbol{\Phi}_c\boldsymbol{\Theta}_c^{\mathrm{T}}$. We transposed it for better readability. Without transposition, Equation 3.10 would be $\boldsymbol{\Theta}_c^{\mathrm{T}} = \mathbf{L}_c^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}$ with a more common expression of the linear transform operation on the left side of the global spatial coordinates $\mathbf{T}^{\mathrm{T}}$.

and the third ones are l_2-norm regularization terms introduced to improve the numerical stability of the image reconstruction; and the last term is a complementary sparsity penalty on the local coordinates to promote the piece-wise smoothness of the reconstructed images. Here we assume that the temporal bases of each neighborhood $\{\mathbf{\Phi}_c\}_{c=1}^{C}$ are pre-determined from simultaneously acquired "training" k-space data as described in Section 3.3.3.

The optimization problem in Equation (3.14) is a non-convex optimization problem but can be solved efficiently using the ADMM algorithm [11], which leads to solving the following sub-problems in an alternative fashion:

$$\mathbf{T}^{(k+1)} = \arg\min_{\mathbf{T}} \left\{ \left\| \mathbf{\Omega}\Big(\mathbf{F}_s \sum_{c=1}^{C} \mathbf{T}\mathbf{L}_c{}^{(k)} \mathbf{\Phi}_c^{\mathrm{T}} \mathbf{\Pi}_c\Big) - \mathbf{s} \right\|_2^2 + \frac{\rho}{2} \left\| \mathcal{D}\Big(\mathbf{T}\mathbf{L}^{(k)}\Big) - \mathbf{G}^{(k)} + \boldsymbol{\eta}^{(k)} \right\|_2^2 + \frac{\mu_T}{2} \|\mathbf{T}\|_F^2 \right\}, \tag{3.15}$$

$$\mathbf{L}^{(k+1)} = \arg\min_{\mathbf{L}} \left\{ \left\| \mathbf{\Omega}\Big(\mathbf{F}_s \sum_{c=1}^{C} \mathbf{T}^{(k+1)} \mathbf{L}_c \mathbf{\Phi}_c^{\mathrm{T}} \mathbf{\Pi}_c\Big) - \mathbf{s} \right\|_2^2 + \frac{\rho}{2} \left\| \mathcal{D}\Big(\mathbf{T}^{(k+1)}\mathbf{L}\Big) - \mathbf{G}^{(k)} + \boldsymbol{\eta}^{(k)} \right\|_2^2 + \frac{\mu_L}{2} \|\mathbf{L}\|_F^2 \right\}, \tag{3.16}$$

$$\mathbf{G}^{(k+1)} = \mathcal{S}_{\frac{\lambda}{\rho}}\Big(\mathcal{D}\Big(\mathbf{T}^{(k+1)}\mathbf{L}^{(k+1)}\Big) + \boldsymbol{\eta}^{(k)}\Big), \tag{3.17}$$

$$\boldsymbol{\eta}^{(k+1)} = \boldsymbol{\eta}^{(k)} + \mathcal{D}\Big(\mathbf{T}^{(k+1)}\mathbf{L}^{(k+1)}\Big) - \mathbf{G}^{(k+1)}, \tag{3.18}$$

where $\mathbf{G}$ is the augmented Lagrangian split variable, $\boldsymbol{\eta}$ is the dual variable, ρ is a relaxation parameter, and $\mathcal{S}_{\frac{\lambda}{\rho}}$ is a soft-thresholding operator with threshold $\frac{\lambda}{\rho}$. The optimization problem in Equation (3.15) and Equation (3.16) are convex and can be efficiently solved using the conjugate gradient method.

The LTSA-based image reconstruction problem in Equation (3.14) is formulated with an explicit low-rank constraint on the local coordinate matrix, i.e., $\mathbf{\Theta} = \mathbf{T}\mathbf{L}$. Alternatively, the image reconstruction problem can be formulated with a soft nuclear norm penalty on $\mathbf{\Theta}$ to promote the low rankness of $\mathbf{\Theta}$, which could be advantageous in terms of representation power. However, implementing the LTSA with a nuclear norm penalty requires performing SVD on $\mathbf{\Theta}$ in each iteration, which can be both computationally expensive and memory demanding in the case of 3D dynamic MRI.

3.3.3 Implementation details

Data acquisition scheme

The data acquisition scheme of the proposed method shares similarities with that of the subspace-based methods [81]. Many k-space trajectories, such as Cartesian, spiral, and radial trajectories, can be used [89]. In this work, we used random radial trajectories in the 2D imaging experiment and random stack-of-stars trajectories in the 3D imaging experiment because of their superior performance in dynamic MRI applications. Exploring optimal sampling trajectories for a specific application is beyond the scope of this work. For the 2D imaging experiment, two sets of k-space data were acquired in each temporal frame during the imaging experiment. The first set of k-space data consisted of two k-space lines acquired at the center of the k-space along the k_x- and k_y-axis. This dataset was used to estimate the temporal bases of the LTSA model as in the subspace-based methods [22, 81], which is often referred to as "training" k-space data in the literature. See [22] for an in-depth analysis of subspace estimation using "training" k-space data acquired as such. The second set of k-space data sparsely sampled the entire (k, t)-space (see Section 2.1.2). This dataset was fitted to the LTSA model in image reconstruction as in Equation (3.14). The data acquisition scheme of the 3D imaging experiment was similar. See Section 3.4 for more details.

Forming the neighborhoods

Given an initial reconstruction of the dynamic images, grouping the images into different neighborhoods can be solved using any clustering methods (e.g., k-means). However, this would require a high-quality initial image reconstruction and could be time-consuming when the dimension of the reconstructed dynamic images is large. In the current implementation, we chose to group (k, t)-space data into different neighborhoods based on respiratory motion. The motion phases can be estimated using respiratory surrogate signals such as navigators or by tracking the dome of the liver over time in the dynamic images reconstructed by a LR-based method. In this work, the latter strategy was implemented. The number of neighborhoods C was set as the maximum amplitude of the liver displacement in pixels. An example of how respiratory motion estimated through liver displacement can be used for neighborhood estimation is shown in Figure 3.1.

Estimation of temporal bases

In this work, temporal bases were estimated using the training data prior to image reconstruction in Equation (3.14). Suppose there were in total P_c temporal frames in neighborhood c, each had n_t k-space lines acquired as training data. A Casorati matrix of size $n_t r_o \times P_c$ was formed using these k-space lines, where r_o was the size of the readout dimension. In the case of imaging with phased-array coils, the training lines from each coil were concatenated along columns. The temporal bases of neighborhood c, i.e., $\mathbf{\Phi}_c$, were then estimated by calculating the first D right singular vectors of the Casorati matrix.

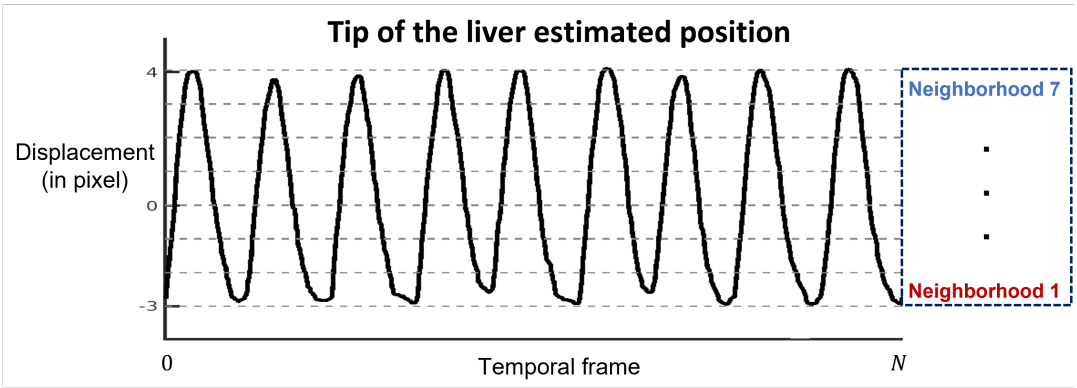


Figure 3.1: Estimated position of the liver throughout a 2D acquisition. The liver position is estimated with an erf fitting on the edge of the liver using an initial LR reconstruction. In this example, 7 neighborhoods would be used with the proposed LTSA method based on the maximum amplitude of the liver displacement in pixels, as described in Section 3.3.3.

Initialization

In the current implementation, the optimization problem in Equation (3.14) was solved by solving the sub-problems in Equation (3.15) to Equation (3.18) iteratively, which required a proper initialization of $\mathbf{L}_c$. One way to initialize $\mathbf{L}_c$ is to use Equation (3.11). More specifically, an initial image reconstruction was first performed using the LR-based method, followed by grouping the reconstructed images into different neighborhoods. The local coordinates, i.e., $\mathbf{\Theta}_c$, were then estimated by performing SVD on the Casorati matrix of each neighborhood as in Equation (3.9). Finally, an initial guess of $\mathbf{L}_c$ was obtained by performing SVD on the Casorati matrix formed by $\mathbf{\Theta}_c$ as in Equation (3.11). Another strategy was to initialize $\mathbf{L}_c$ with random matrices. According to our experience, the two initialization approaches led to similar image reconstruction performance. The latter initialization approach was used in this work due to its simple implementation.

CPU/GPU implementation

The proposed image reconstruction algorithm was implemented on CPUs using Matlab (The MathWorks, Inc., Natick, Massachusetts, United States) for reconstructions of 2D dynamic images. The typical computation time for reconstructing 2D dynamic images (e.g., those in Section 3.4.2 and Section 3.4.3) using a workstation (8 Intel Xeon 2.4 GHz CPUs) was 1.8 hours. The algorithm was further implemented on GPUs using Python with the CuPy [107], SigPy [108], and MRRT (https://github.com/mritools/mrrt.nufft) packages for accelerated image reconstructions of 3D dynamic images. The typical computation time for reconstructing 3D dynamic images (e.g., those in Section 3.4.4) using a GPU (NVIDIA Tesla V100 SXM2) was about 10 minutes.

Algorithm 1 LTSA model with sparsity constraint for MR data reconstruction

Input: Undersampled (k,t)-space data $\boldsymbol{s}$.
Output: Reconstructed dynamic MR image set $\boldsymbol{X}$.
A. Preparation of the data
Set the reconstruction rank D.
Define neighborhoods and set C as in Section 3.3.3.
Calculate temporal bases $\boldsymbol{\Phi}_c$ as in Section 3.3.3.
B. Reconstruction of $\boldsymbol{T}$ and $\{\boldsymbol{L}_c\}$
Set N_{loop}, λ, μ_T, μ_L, ρ.
Initialize $\{\boldsymbol{L}_c\}^{(0)}$ $c = 1, \ldots, C$ with random matrices.
for $k = 1$ to N_{loop} **do**
  Compute $\boldsymbol{T}^{(k)}$ by solving Equation (3.15).
  Compute ${\boldsymbol{L}_c}^{(k)}$ $c = 1, \ldots, C$ by solving Equation (3.16).
  Soft-threshold $\boldsymbol{G}^{(k)}$ as in Equation (3.17).
  Update the dual variable $\boldsymbol{\eta}^{(k)}$ from Equation (3.18).
end for
Form the dynamic set $\boldsymbol{X}$ using Equation (3.13).

3.4 Results

We performed both numerical simulation studies and *in vivo* MR cardiac imaging experiments using in-house developed dynamic MR sequences on healthy subjects at 3T to validate the performance of the proposed method. The results from the proposed method were compared with the low rank (LR) method, the temporal local low-rank (t-LLR) method, and the SToRM method. The image reconstruction error was evaluated using Normalized Root Mean Square Error (NRMSE) defined as

$$\text{NRMSE} = \frac{\|\mathbf{X}_{\text{ref}} - \mathbf{X}_{\text{recon}}\|_F}{\|\mathbf{X}_{\text{ref}}\|_F}, \tag{3.19}$$

where $\mathbf{X}_{\text{ref}}$ and $\mathbf{X}_{\text{recon}}$ denote the Casorati matrices formed by the reference and reconstructed images, respectively. The performance of the compared methods was also evaluated by calculating the Structural Similarity (SSIM) between the reconstructed images and the reference images, defined as:

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{3.20}$$

where μ_x is the pixel sample mean of $\mathbf{x}$, μ_y is the pixel sample mean of $\mathbf{y}$; σ_x^2 is the variance of $\mathbf{x}$, σ_y^2 is the variance of $\mathbf{y}$, σ_{xy} is the covariance of $\mathbf{x}$ and $\mathbf{y}$, $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$ are two variables to stabilize the division with weak denominator ($k_1 = 0.01$ and $k_2 = 0.03$ by default) and L is the dynamic range of the pixel-values.

For all the compared methods, dynamic images of each coil were reconstructed from undersampled complex-valued k-space data. Magnitude images were then generated by combining the reconstructed dynamic images of each coil using the sum-of-square method (described in Section 1.2.3)

for display and comparison purposes.

3.4.1 Compared methods

In the LR method (also denoted as subspace model, described in Chapter 2), dynamic images were reconstructed by fitting the subspace model in Equation (3.2) to the measured data, resulting in the following optimization problem [171]:

$$\arg\min_{\boldsymbol{\Theta}} \left\| \boldsymbol{\Omega}\left(\mathbf{F}_s \boldsymbol{\Theta} \boldsymbol{\Phi}^{\mathrm{T}}\right) - \mathbf{s} \right\|_2^2 + \frac{\mu}{2} \left\| \boldsymbol{\Theta} \right\|_F^2 + \lambda \left\| \mathcal{D}(\boldsymbol{\Theta}) \right\|_1 , \tag{3.21}$$

where $\boldsymbol{\Phi}$ is estimated from the "training" k-space data in the same way as in Section 3.3.3. The last term of Equation (3.21) is a sparsity penalty introduced to promote the piece-wise smoothness of the reconstructed images as in Equation (3.14). As a sidenote, images reconstructed with a sparsity constraint, i.e., $\lambda \neq 0$ in Equation (3.14) or Equation (3.21), are denoted using the prefix "l_1", e.g, l_1-LR and l_1-LTSA, in the rest of this section. We selected λ to balance the sparsity constraint term with the data fidelity term, setting a similar target ratio between the two terms for both the LR and LTSA methods.

In the t-LLR method, dynamic images were reconstructed by solving the following optimization problem:

$$\arg\min_{\{\boldsymbol{\Theta}_c\}_{c=1}^{C}} \left\| \boldsymbol{\Omega}\left(\mathbf{F}_s \sum_{c=1}^{C} \boldsymbol{\Theta}_c \boldsymbol{\Phi}_c^{\mathrm{T}} \boldsymbol{\Pi}_c\right) - \mathbf{s} \right\|_2^2 + \frac{\mu}{2} \sum_{c=1}^{C} \left\| \boldsymbol{\Theta}_c \right\|_F^2 . \tag{3.22}$$

Note that the temporal neighborhoods in the t-LLR method were defined in the same way as in the proposed LTSA method. Compared to the proposed LTSA method, the t-LLR method aims at recovering the local coordinates of each neighborhood independently without aligning them to the tangent space of the manifold.

The SToRM method with l_2-norm regularization was implemented using the code provided on GitHub as in [118]. The SToRM optimization problem is:

$$\arg\min_{\mathbf{X}} \left\| \boldsymbol{\Omega}\left(\mathbf{F}_s \mathbf{X}\right) - \mathbf{s} \right\|_2^2 + 2\lambda \, \mathrm{Tr}\left(\mathbf{X} \mathbf{L}_\mathbf{p} \mathbf{X}^H\right) \tag{3.23}$$

where Tr denotes the trace operator, and $\mathbf{L}_p$ is the Laplacian matrix of $\mathbf{X}$ (the calculation of $\mathbf{L}_p$ described in [118]).

Equation (3.21) describes a linear (global) subspace-based image reconstruction scheme, as described in Chapter 2. The resulting signal model of this method is a global low-rank decomposition of the dynamic image series $\mathbf{X} = \boldsymbol{\Theta}\boldsymbol{\Phi}^{\mathrm{T}}$ (see Equation 2.3) where we fix the temporal basis and recover the much smaller spatial coordinates, incidentally greatly decreasing the number of unknowns.

In comparison, the t-LLR in Equation (3.22) operates on "parts" of the dynamic image series and

not globally, leveraging an "informed" and local spatiotemporal correlation. This method suggests that if we have information on some similarity among the temporal frames (i.e., similar frames repeated within the dynamic series), we can do a low-rank decomposition of these similar frames (denoted as neighborhood) $\mathbf{X}_c = \mathbf{\Theta}_c \mathbf{\Phi}_c^{\mathrm{T}} \in \mathbb{C}^{M \times P_c}$ and have less intraframe variation than with the global data. Doing a low-rank decomposition for all the neighborhoods separately and putting the images back in their original time-frame position therefore recover the dynamic image series $\mathbf{X} = \sum_{c=1}^{C} \mathbf{\Theta}_c \mathbf{\Phi}_c^{\mathrm{T}} \mathbf{\Pi}_c$. However, one of the main drawbacks is that doing local low-rank decompositions increases the number of unknowns by a factor C (typically 10 in the results shown later in this section), justifying the need of an alignment of the local coordinates in a global manifold, as in the proposed method.

Finally, the SToRM method operates on manifolds, similarly to the proposed method, but in a different way: it proposes a regularization on the smoothness of the manifold through graph Laplacian embedding, as indicated by the second term in Equation (3.23). In order to avoid explicit characterization of the manifold and therefore avoid the pre-image problem (see Section 3.1) the SToRM method operates in the input domain instead of the feature domain to indirectly penalize the underlying manifold.

3.4.2 Numerical simulation study

Numerical phantom

A numerical phantom (Figure 3.2a) was built using the XCAT software [138]. Image intensities of different tissues, including muscle, liver, myocardium, and blood, were assigned to mimic typical contrasts of a balanced Steady-State Free Precession (bSSFP) sequence using proton density, T_1 and T_2 values from the literature. Cardiac and respiratory motions were simulated using the XCAT software, producing a series of images, denoted as $\mathbf{x}_{\mathrm{ref}}(pc, pr), pc = 1, \ldots, 40, pr = 1, \ldots, 30$, that consisted of 40 cardiac phases and 30 respiratory phases (i.e., a combination of 1200 different motion phases). EKG and respiratory signals were then simulated to mimic variations of heart beat rate (60 to 75 beats per minute) and irregular respiratory rate (from 20 to 30 breaths per minute). Finally, dynamic images were synthesized based on the simulated EKG and respiratory signals, i.e.,

$$\mathbf{x}_n = \mathbf{x}_{\mathrm{ref}}(pc_n, pr_n), n = 1, \ldots, N, \tag{3.24}$$

where pc_n and pr_n denote the cardiac and respiratory motion phase of the n-th frame, respectively. The matrix size of each frame was 192 $\times$ 192, resulting in a voxel size of 2.6 mm^2.

Figure 3.2b plots the normalized singular values of the "global" Casorati matrix $\mathbf{X}$ as in Equation (3.2) formed by the simulated dynamic images of all frames (denoted by black solid line) and those of the "local" Casorati matrices $\mathbf{X}_c$ as in Equation (3.6) formed by the simulated dynamic images that were grouped into 10 neighborhoods based on respiratory motion (denoted by colored dashed lines). As can be seen, the singular values of the "local" Casorati matrices $\mathbf{X}_c$ decayed much

faster than those of the "global" Casorati matrix $\mathbf{X}$. Figure 3.2c plots the normalized singular values of the matrix $\mathbf{\Theta}$ as in Equation (3.11) formed by the local coordinates, illustrating the low-rank structure of $\mathbf{\Theta}$. The singular value distributions shown in Figures 3.2b and 3.2c indicate that the intrinsic low-dimensional structure of the dynamic images can be more effectively captured by the proposed LTSA model (i.e., with a smaller rank) than the LR model.

Reconstruction results

To compare the performance of different image reconstruction methods, k-space data were simulated along a radial trajectory with Gaussian noise for each frame. Among the 301 k-space lines of each frame, 7 lines were randomly sampled for image reconstruction while 2 lines along the k_x and k_y axis across the center of the k-space were sampled for the estimation of the temporal bases. This translates to a frame rate of roughly 30 frames per second in a practical imaging experiment. In this experiment, the number of neighborhoods was 10 ($C = 10$), corresponding to a maximum liver displacement of 26 mm as the setup in the XCAT simulation.

Representative images reconstructed by different methods are shown in Figure 3.3. We chose the images reconstructed by the LR method with a rank of 30 using noiseless k-space data of 2114

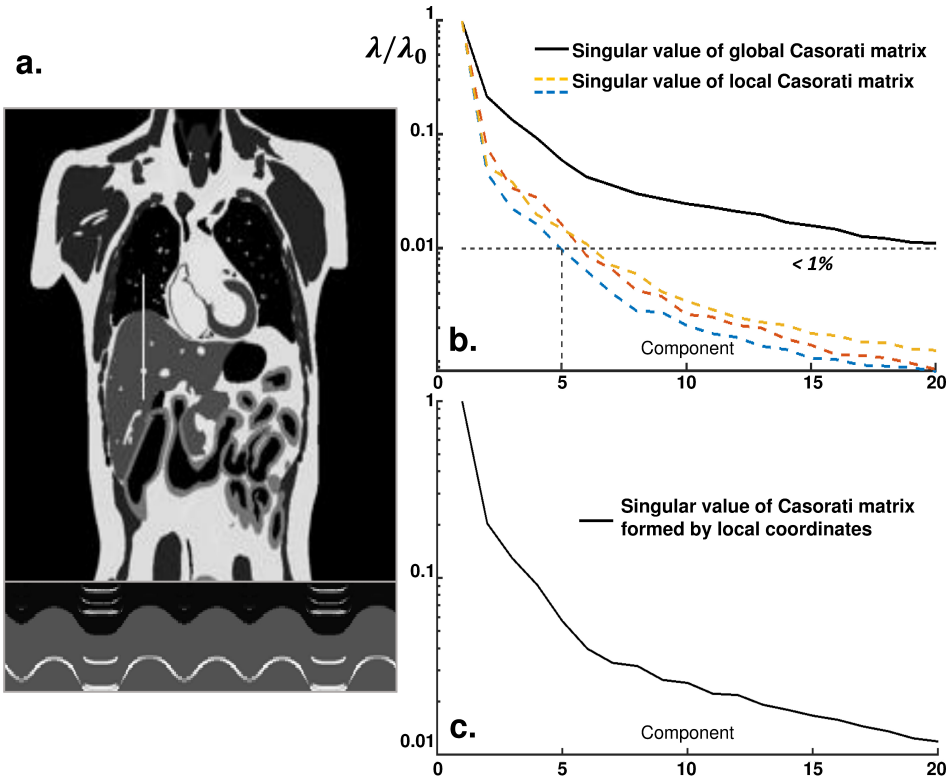


Figure 3.2: (**a**) Representative image of the numerical simulation phantom. Plotted at the bottom is the temporal profile of the simulated dynamic images along the white line at the dome of the liver, showing an irregular respiratory motion pattern. (**b**) Plots of the normalized singular values of the "global" Casorati matrix $\mathbf{X}$ (black solid line) and the "local" Casorati matrices $\mathbf{X}_c$ (colored dashed lines), formed by the simulated dynamic images as in Equation (3.2) and Equation (3.6), respectively. (**c**) Plot of the normalized singular values of the matrix $\mathbf{\Theta}$ formed by the local coordinates as in Equation (3.11).

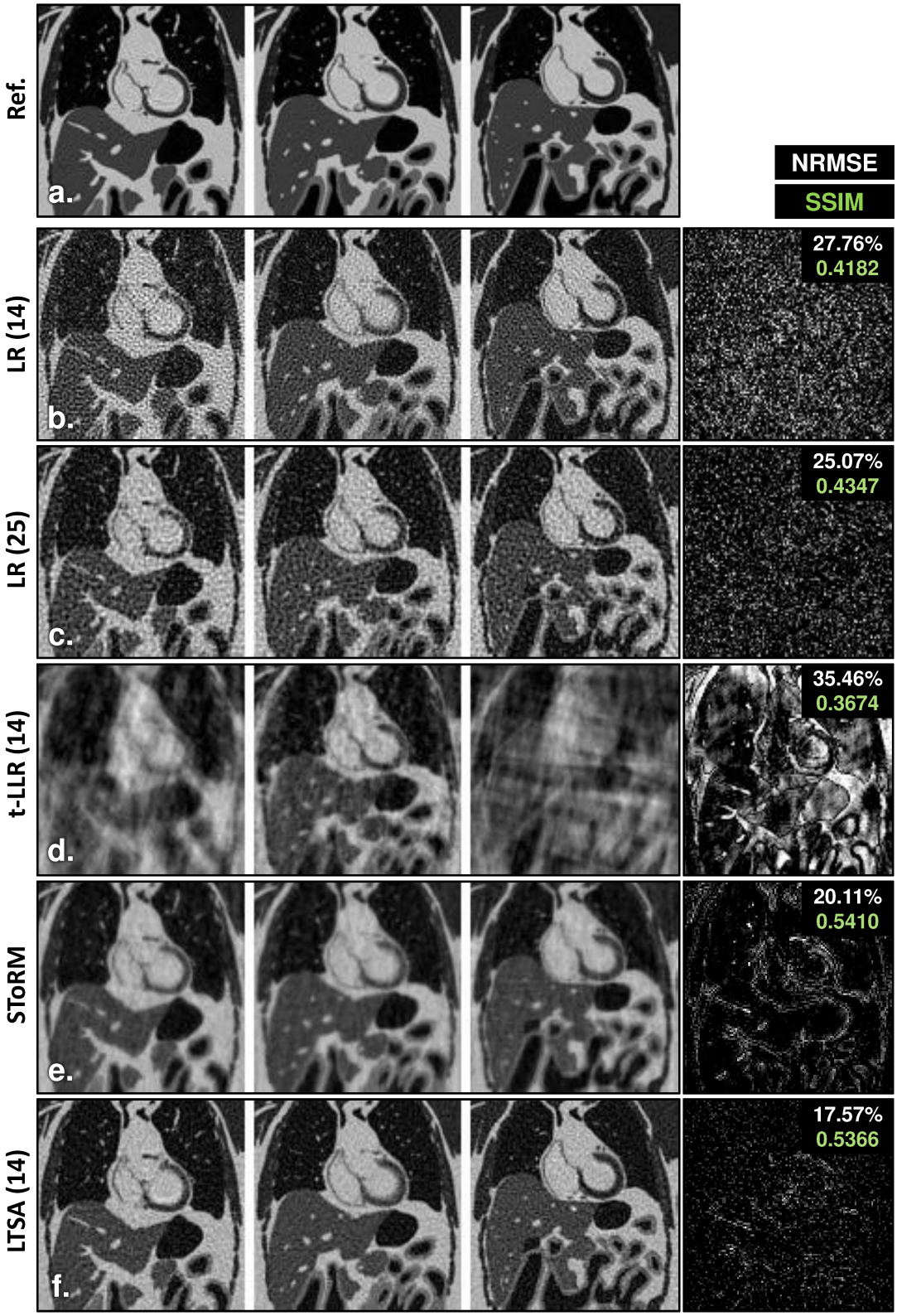


Figure 3.3: Images in the numerical simulation study reconstructed by (**a**) the reference method, (**b**) the LR method with a rank of 14, (**c**) the LR method with a rank of 25, (**d**) the t-LLR method with a rank of 14, (**e**) the SToRM method, and (**f**) the proposed LTSA method with a rank of 14. The first three columns show images in three representative frames. The last column shows representative difference maps between the dynamic images reconstructed by different methods and the reference images. Displayed in the top right corner are the corresponding NRMSEs in white and SSIMs averaged over all the frames in green.

frames (i.e., 49 measurements per spoke) as the reference, shown in Figure 3.3a. The rank was chosen based on the singular value decay of the Casorati matrix formed by the "training" k-space data as in [22]. The normalized singular value of rank 30 is below 0.5%. Based on our previous experience (described in Section 2.2.2 and in [95]), choosing such a rank is sufficient to capture both cardiac and respiratory motions while yielding artifacts-free images. The rank of the reference LR-based image reconstruction was chosen in the same way in the in vivo experiments. Images reconstructed by different methods using noisy k-space data of 600 frames (i.e., 14 measurements per spoke) are shown in Figures 3.3b to 3.3f. The images reconstructed by the LR method with a rank of 14 and 25 (Figures 3.3b and 3.3c) both suffered from noise and aliasing artifacts. The images reconstructed by the t-LLR (Figure 3.3d) method showed significant aliasing artifacts. In comparison, the images reconstructed by the SToRM method (Figure 3.3e) were less noisy but suffered from aliasing artifacts and spatial blurring, which were most noticeable in the difference map. The regularization parameter of the SToRM method controls the trade-off between aliasing artifacts and spatial blurring. Here, we tuned the regularization parameter to produce images with the smallest NRMSE. Compared to the LR and SToRM methods, the proposed LTSA method with a rank of 14 (i.e., the rank of the global coordinate matrix $\mathbf{T}$) reconstructed images (Figure 3.3f) with the best quality and the smallest NRMSE. The SSIMs of the SToRM and the proposed method were similar and significantly higher than the other compared methods.

Figures 3.4 and 3.5 show the NRMSE per frame, representative (y-t) profiles, and the NRMSE of these profiles obtained by the compared methods. Compared to the temporal profile of the reference images (Figure 3.5a), the LR and t-LLR reconstructions with a rank of 14 (Figure 3.5b and 3.5d) failed to capture details of the cardiac and respiratory motion, which led to the two highest NRMSEs among the compared methods (Figure 3.4b). Increasing the rank of the LR reconstruction to 25 better captured the temporal changes of the cardiac and respiratory motion at the cost of noise (Figure 3.5c). The temporal profile using the SToRM method was overly smoothed compared to the reference. In comparison, the proposed LTSA method with a rank of 14 produced the most accurate temporal profiles with the smallest NRMSE (Figure 3.5f) and in terms of NRMSE per frame (Figure 3.4b). Figure 3.4a shows representative images reconstructed by the compared methods during a deep inhalation, which was particularly challenging to recover, except for the proposed method.

3.4.3 2D cardiac imaging experiment

Data acquisition

An *in vivo* 2D cardiac imaging experiment was performed on a healthy volunteer under a study protocol approved by our local Institutional Review Board (IRB). The subject was imaged on a 3T MR scanner (Siemens MAGNETOM Trio) with spine and surface coils for reception and a body coil for transmission. The (k, t)-space data were sampled in the k_x-k_y plane along a radial trajectory with random spoke angles uniformly distributed between 0 and π. The data were acquired continuously

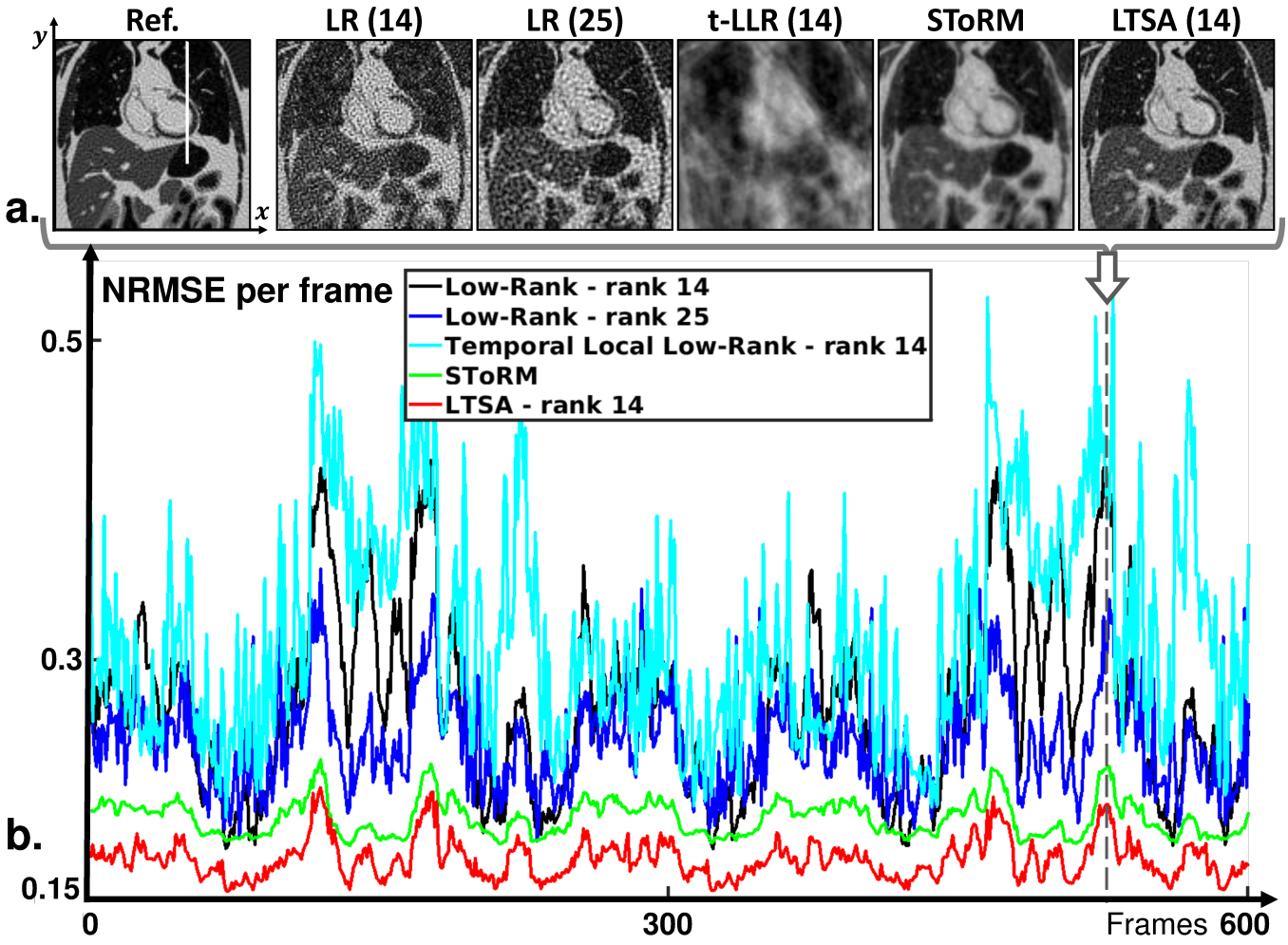


Figure 3.4: NRMSEs of different image reconstruction methods in the numerical simulation study. (**a**) Representative images during a deep inhalation. (**b**) The NRMSEs per frame.

without EKG or respiratory gating using a spoiled gradient recalled echo (SPGR) sequence in the coronal view with the following imaging parameters: TR/TE = 3/1.6 ms, flip angle = 7 degrees, image size = 128 × 128 (201 k-space lines in the k_x-k_y plane), voxel size = 1.9 × 1.9 × 6 mm^3, and acquisition time = 3 minutes. A total of 10 k-space lines were sampled in each frame, including 1 navigator for tracking respiratory motion, 2 lines along the k_x and k_y axis across the center of the k-space for the estimation of temporal bases, and 7 random lines for image reconstruction, resulting in a frame rate of 30 ms per frame. In this experiment, the number of neighborhoods was 8 ($C = 8$), corresponding to a maximum liver displacement of 15.2 mm as expected from a regular breathing pattern (8 voxels at 1.9 mm^2 in-plane resolution).

Reconstruction results

Representative images reconstructed by different methods are shown in Figures 3.6 to 3.8. The reference images were reconstructed by the LR method with a rank of 30 and using the (k, t)-space data acquired in the entire 3-min acquisition (i.e., a total of 6000 frames). Figures 3.6b

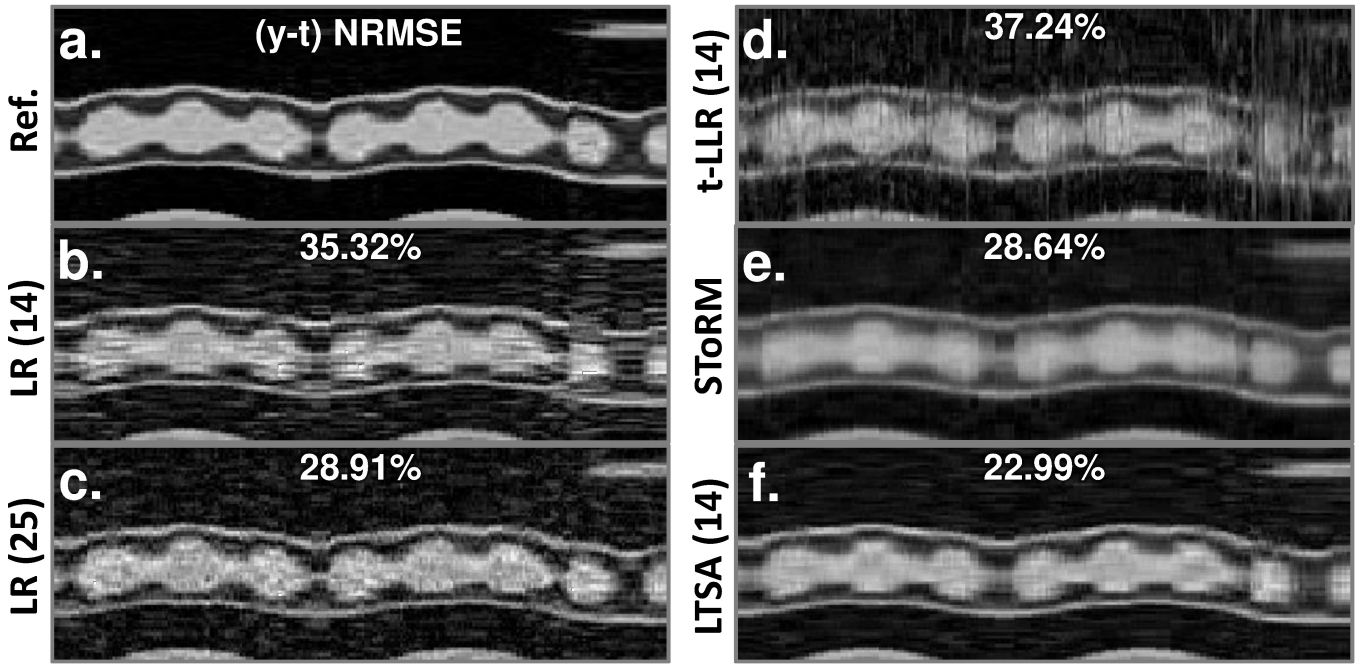


Figure 3.5: Comparison of the (y-t) temporal profiles of different image reconstruction methods in the numerical simulation study. Temporal profiles along the white line across the heart (Figure 3.4a) are shown for (**a**) the reference image, (**b**) the LR method with a rank of 14, (**c**) the LR method with a rank of 25, (**d**) the t-LLR method with a rank of 14, (**e**) the SToRM method, and (**f**) the proposed LTSA method with a rank of 14. The NRMSE of the (y-t) profile is displayed on the top of each image.

to 3.6f show images reconstructed by different methods using (k, t)-space data acquired in a 26-sec acquisition (i.e., a total of 863 frames, resulting in 30 measurements per k-space spoke). The images reconstructed by the LR method with a rank of 14 (Figure 3.6b) suffered from noise and spatial blurring, especially in the heart region (indicated by the yellow arrow in Figure 3.6). The images reconstructed by the LR method with a rank of 20 (Figure 3.6c) had less spatial blurring at the cost of SNR. The images reconstructed by the t-LLR method (Figure 3.6d) showed significant aliasing artifacts presumably due to the large number of unknowns in the t-LLR model. The images reconstructed by the SToRM method (Figure 3.6e) were less noisy but suffered from noticeable spatial blurring as reflected by bright edges in the difference map. The proposed LTSA method produced images (Figure 3.6f) with the best overall quality, the smallest NRMSE, and the highest SSIM among the compared methods.

Figures 3.7 and 3.8 show another set of reconstruction results obtained by the LR and LTSA methods with and without sparsity constraints using even less (k, t)-space data (i.e., a total of 663 frames in a 20-sec acquisition, resulting in 23 measurements per k-space spoke). As expected, incorporating sparsity constraints in the LR method significantly suppressed the noise while preserving the edges of the reconstructed images as shown in Figures 3.7b and 3.7c, resulting in reduced NRMSEs. The benefits of incorporating sparsity into the proposed LTSA method were less significant in the 2D cardiac imaging case. This is not surprising because the LTSA reconstruction without sparsity constraints (Figure 3.7d) had a much better SNR than that of the LR reconstruction (Fig-

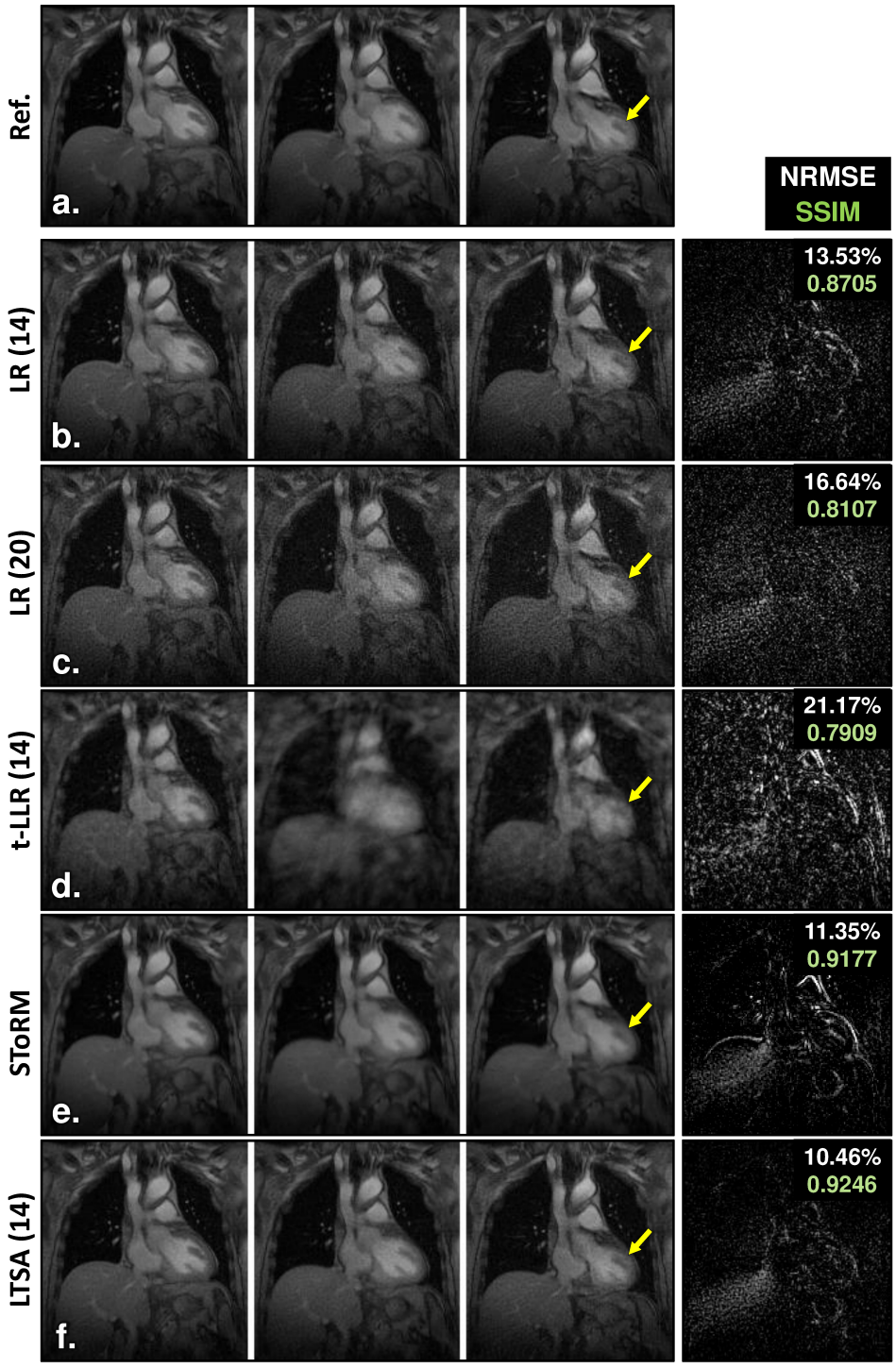


Figure 3.6: Comparison of image reconstruction results of the in vivo 2D cardiac imaging experiment. (**a**) Reference images reconstructed by the LR method with a rank 30 using data collected in a 3-min acquisition. (**b**) to (**f**) Images reconstructed using k-space data collected in a 26-sec acquisition by the LR method with a rank of 14, the LR method with a rank of 25, the t-LLR method with a rank of 14, the SToRM method, and the proposed LTSA method with a rank of 14, respectively. The first three columns show images in three representative frames. The last column shows a representative difference image between the images reconstructed by different methods and the reference images. Displayed in the top right corner are the corresponding NRMSEs in white and SSIMs averaged over all the frames in green.

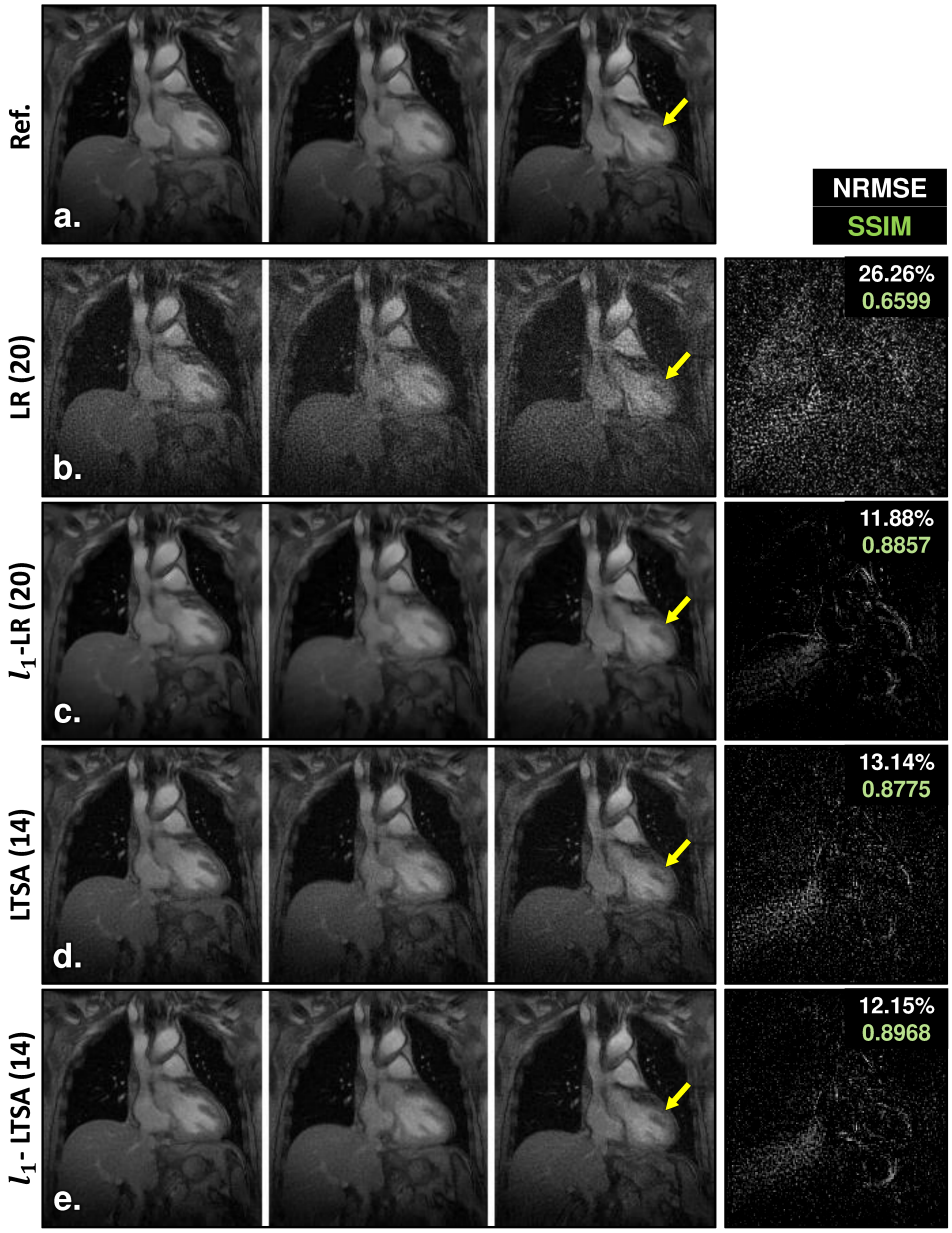


Figure 3.7: Comparison of the in vivo 2D cardiac images reconstructed without and with sparsity constraints. (**a**) Reference images reconstructed by the LR method with a rank 30 using data collected in a 3-min acquisition. (**b**) and (**c**) mages reconstructed by the LR method using k-space data collected in a 20-sec acquisition without and with sparsity constraints, respectively. (**d**) and (**e**) Images reconstructed by the proposed LTSA method using k-space data collected in a 20-sec acquisition without and with sparsity constraints, respectively. The first three columns show images in three representative frames. The last column shows representative difference maps between the dynamic images reconstructed by different methods and the reference images. Displayed in the top right corner are the corresponding NRMSEs in white and SSIMs averaged over all the frames in green.

ure 3.7b). The LTSA reconstruction with sparsity constraints (Figure 3.7e) produced images with similar image quality, NRMSEs, and SSIMs compared to the LR reconstruction with sparsity constraints (Figure 3.7c). However, the proposed method reconstructed sharper temporal profiles in the heart region than the LR method as shown in Figure 3.8, especially in regions indicated by the green arrows.

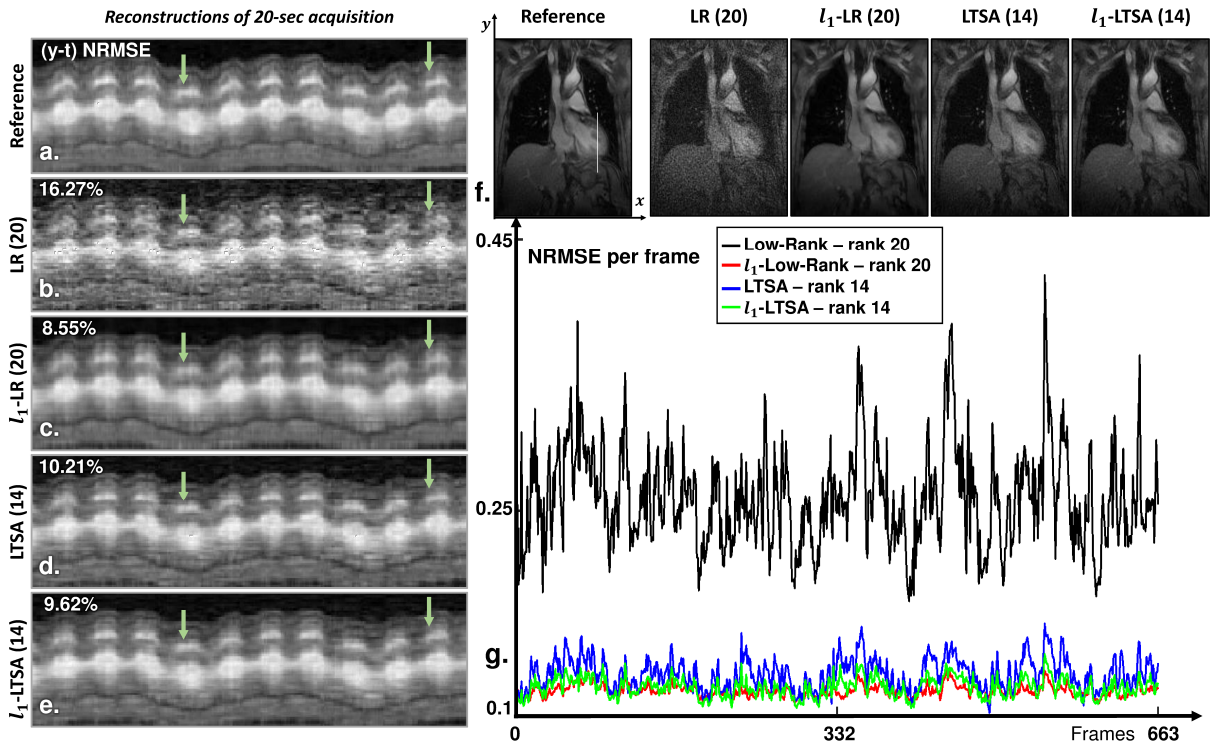


Figure 3.8: Comparison of the (y-t) temporal profiles and NRMSE per frame by different image reconstruction methods in the 2D cardiac imaging experiment. The temporal profiles along the white line across the heart are shown for (**a**) the reference image, (**b**) the LR method without sparsity constraints, (**c**) the LR method with sparsity constraints, (**d**) the LTSA method without sparsity constraints, and (**e**) the LTSA method with sparsity constraints, respectively. Representative images reconstructed by the compared methods are shown in (**f**). The NRMSEs per frame of the reconstructed images are plotted in (**g**) and the NRMSEs of the (y-t) profiles are displayed in the top left corner of (**b**)-(**e**).

3.4.4 3D cardiac imaging experiment

Data acquisition

An *in vivo* 3D cardiac imaging experiment was performed on a healthy volunteer under a study protocol approved by our local IRB. The subject was imaged on a 3T PET/MR scanner (Siemens Biograph, mMR) with spine and surface coils for reception and a body coil for transmission. The (k, t)-space data were sampled along a random stack-of-stars trajectory in the k_x-k_y plane, (i.e., random sampling along a radial trajectory in the k_x-k_y plane and on a Cartesian grid in the k_z axis). The spoke angle and the k_z position were uniformly distributed. The data were acquired continuously without EKG or respiratory gating using a SPGR sequence in the short-axis view. Adiabatic inversion pulses were inserted every 1528 ms for inversion recovery. The rest of the imaging parameters were as follows: TR/TE = 5.5/2.0 ms, flip angle = 5 degrees, image size = 160 × 160 × 32 (251 k-space lines in the k_x-k_y plane), voxel size = 1.9 × 1.9 × 4 mm^3, and acquisition time = 12.4 minutes. A total of 11 k-space lines were sampled in each frame, including 3 k-space lines along the k_x, k_y and k_z axis across the center of the k-space acquired in one TR for the estimation of temporal bases and 8 random k-space lines acquired in 8 TRs for image reconstruction.

The resultant frame rate was 49.5 ms per frame. The number of neighborhoods C was 8, the same as in the 2D in vivo experiment.

Reconstruction results

Representative images reconstructed by the LR and LTSA methods with and without sparsity constraints are shown in Figure 3.9. The entire 12.4-min acquisition was used for reconstruction (i.e., about 15000 frames), resulting in 15 measurements per k-space spoke angle. Note that the sparsity penalty was imposed on the global coordinates in the total variation domain $\mathcal{D}(\mathbf{T})$ in the proposed LTSA-based image reconstruction with sparsity constraints due to the limited memory capacity of the GPU. The images reconstructed by the LR method with a rank of 15 (Figure 3.9a) suffered from noise and aliasing artifacts, especially in the heart region. Incorporating sparsity constraints into the LR method significantly reduced the noise. However, some aliasing artifacts remained, as indicated by the yellow arrow in Figure 3.9b. The images reconstructed by the proposed LTSA method with a rank 15 (Figures 3.9c and 3.9d) had improved visual quality and fewer artifacts. Incorporating the sparsity constraint to the proposed LTSA method (Figure 3.9d) further reduced noise and aliasing artifacts.

Images reconstructed using the LTSA method with the sparsity constraint are shown in Figure 3.10 for representative slices (Figure 3.10a), T_1 contrasts (Figure 3.10b), respiratory phases (Figure 3.10c) and cardiac phases (Figure 3.10d). The T_1 contrast was synthesized via T_1 parametric fitting. As can be seen, the proposed LTSA model successfully captured the high dimensional dynamics of the data, resolving contrast changes, respiratory motion, and cardiac motion in real-time cardiac MR.

3.5 Conclusion and discussions

This chapter presented a novel linear tangent space alignment (LTSA) model-based framework that exploits the intrinsic low-dimensional manifold structure of the underlying spatiotemporal signals for reconstruction of high-resolution, high-frame rate dynamic images from undersampled (k, t)-space data. The proposed LTSA model is a non-linear generalization of the conventional linear subspace models, including the low-rank and low-rank tensor models. The performance of the proposed method was evaluated using numerical simulation studies as well as 2D and 3D *in vivo* cardiac imaging experiments. Compared to state-of-the-art methods, including the low-rank model based methods and SToRM, the proposed LTSA method achieved the best performance in image reconstruction. The proposed LTSA method could prove useful for many other applications beyond dynamic MRI, including multi-parametric MRI and MRSI. For example, the LTSA model has been applied to MRSI spectral quantification in a recent conference abstract [91], where individual temporal bases were estimated to represent the temporal signals of each metabolite, showing superior performance compared to the state-of-the-art spectral quantification methods. An article extending

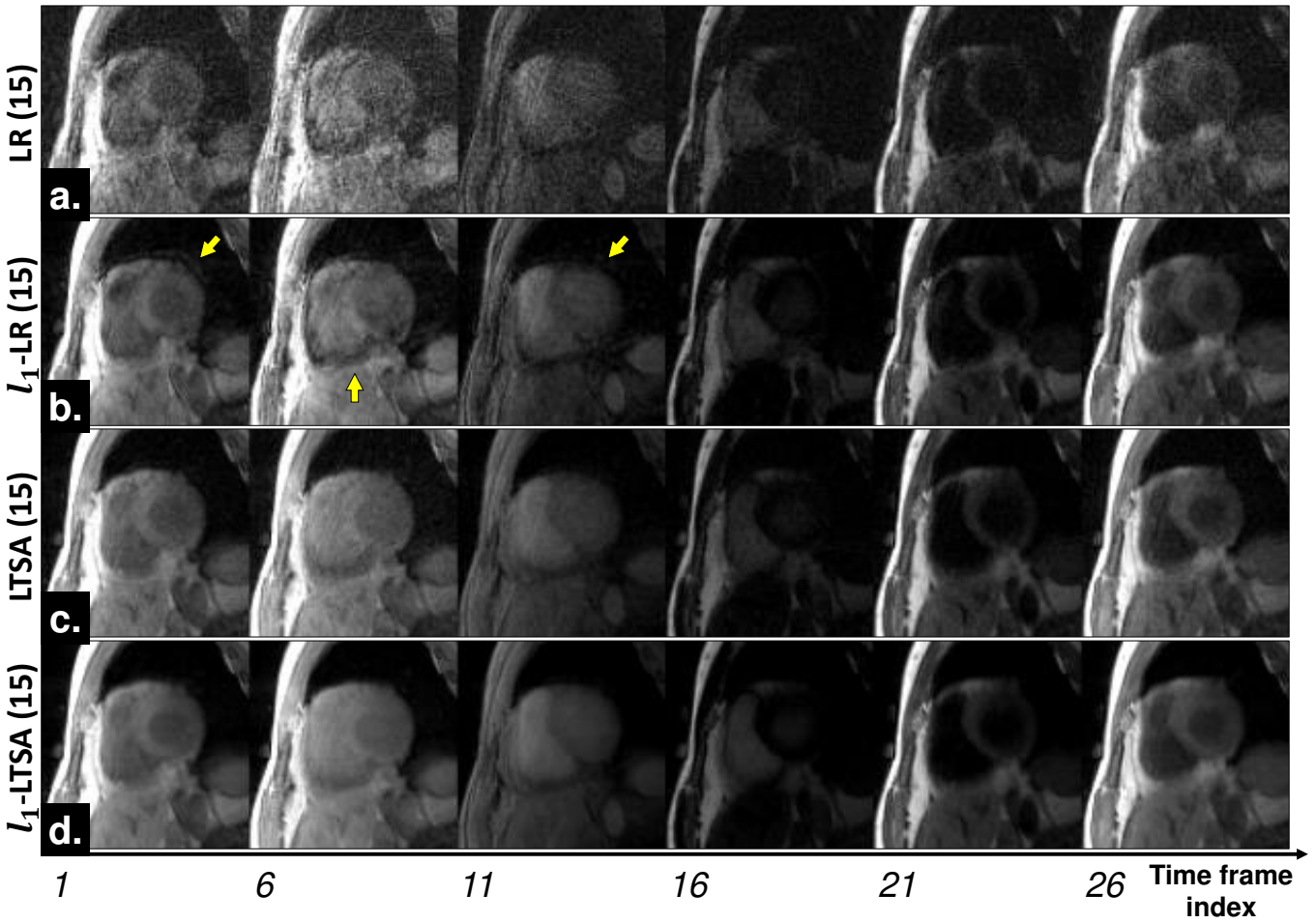


Figure 3.9: Comparison of the in vivo 3D cardiac imaging results. (**a**) and (**b**) Image reconstructed by the LR method without/with sparsity constraints, respectively. (**c**) and (**d**) Image reconstructed by the proposed LTSA method without/with sparsity constraints, respectively.

the work presented in the conference abstract has also been published [93] and can be found in Appendix B.

3.5.1 Comparison with low-rank models

The LTSA model in Equation (3.13) can be reduced to several well-known linear models as special cases. A trivial case is that the LTSA model becomes the LR model (also denoted as subspace model) in Equation (3.2) when there is only one neighborhood, i.e., $C = 1, \mathbf{L}_c = \mathbf{I}, \mathbf{T} = \mathbf{\Theta}$. This basically means that the non-linear manifold learning-based scheme presented in this chapter is a non-linear generalization of the subspace model presented in Chapter 2 with one neighborhood. The results section of this chapter (Section 3.4) therefore presents a critical comparison between the two reconstruction methods. Compared to the global LR model, the LTSA model allows recovering images with a smaller number of unknowns by exploiting the intrinsic nonlinear low-dimensional structure of dynamic images.

One can leverage the low-rank structure of each neighborhood, i.e., Equation (3.9) for image

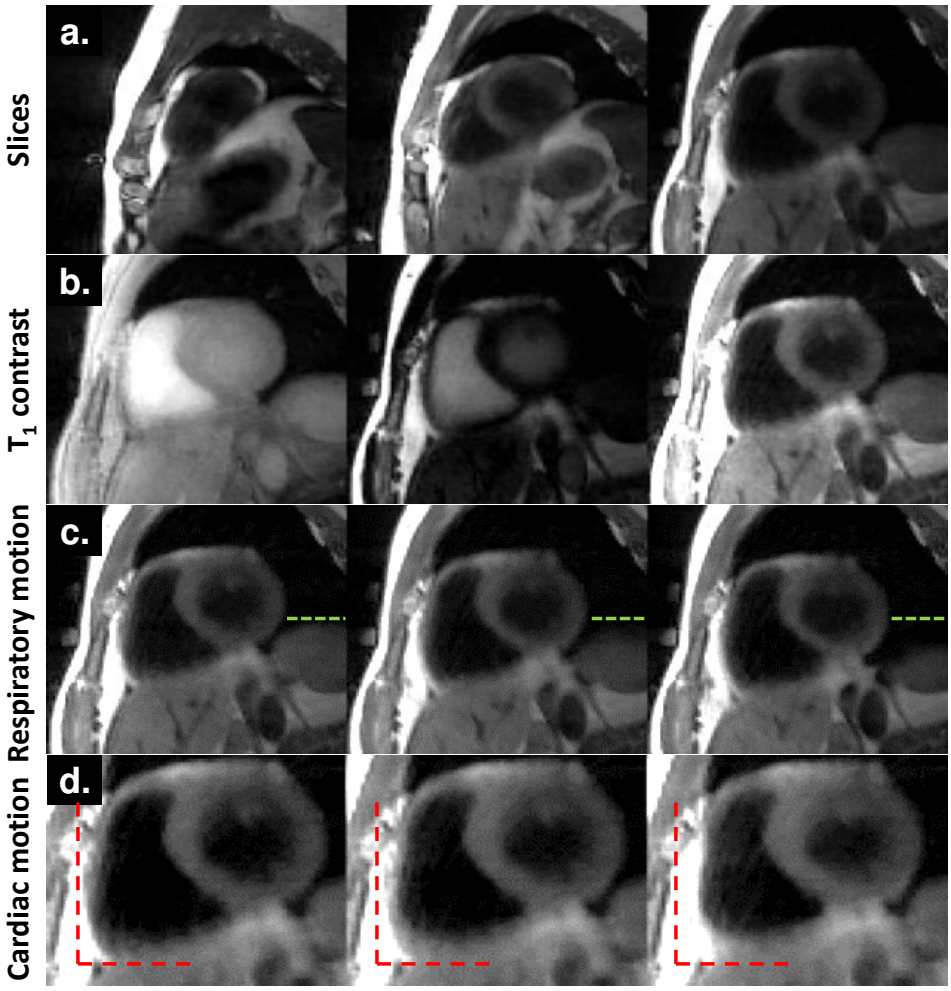


Figure 3.10: Reconstructed 3D cardiac images using the proposed LTSA model with sparsity constraints. Reconstructed images are shown for representative (**a**) slices, (**b**) T_1 contrasts, (**c**) respiratory phases (exhalation to inhalation) and (**d**) cardiac phases (diastole to systole, zoomed on the heart region). The green dashed lines in (**c**) show the position of the tip of the liver at the end-exhalation phase. The red dashed lines indicate the position of the ventricle during diastole to better visualize cardiac contraction.

reconstruction. This leads to a form of local low-rank (LLR) model of dynamic images, which is referred to as the temporal local low-rank (t-LLR) method in Section 3.4.1. Assuming that the temporal bases are predetermined from "training" k-space data, image reconstruction using the t-LLR model requires the estimation of the local coordinates $\mathbf{\Theta}_c$ in each neighborhood, resulting in a total number of unknowns of $M \times D \times C$. In comparison, image reconstruction using the LTSA model requires the estimation of the global coordinates $\mathbf{T}$ and a set of small linear transform matrices $\{\mathbf{L}_c\}_{c=1}^{C}$, resulting in a total number of unknowns of $M \times D + D \times D \times C$. This is significantly smaller than that of the t-LLR model for typical dynamic MR applications, resulting in better performance in image reconstruction as shown in Figures 3.3-3.6.

The more popular "local low-rank" model in the literature is the spatial patch-based local low-rank method [151, 152], which is referred to as the spatial local low-rank (s-LLR) method in this work. The s-LLR method assumes that the dynamic signals in each spatial patch reside in a

low-dimensional subspace. The image reconstruction of the s-LLR method is performed by fitting the measured k-space data with nuclear norm constraints on the dynamic signals in patches to promote low-rankness. Therefore, the proposed method is different from the s-LLR method in both the underlying mathematical model (i.e., *global manifold model* vs. *local low-rank model in patches*) and the image reconstruction process (i.e., *estimation of the manifold global coordinates and linear transforms for alignment* vs. *estimation of the dynamic signals in patches via nuclear norm constraints*). Furthermore, dividing images into spatial patches is a commonly used technique to lift the underlying signals to a higher-dimensional space. For example, applying spatial patches to the low-rank model and the SToRM model of dynamic images leads to the s-LLR method [151, 152] and the iSToRM method [102], respectively. The proposed method is no exception. Extension of the proposed LTSA model to spatial patch-based image reconstruction and comparison with other spatial patch-based methods, e.g., iSToRM, s-LLR, and dictionary learning-based method [128], is an ongoing research.

In summary, the advantage of the proposed LTSA model over the LR and t-LLR models lies in both the freedom to independently define individual temporal basis for each neighborhood and the linear transform aligning local coordinates to low-dimensional global coordinates. The former is a desirable feature for many applications. The latter keeps the number of unknowns of the model roughly independent of the number of neighborhoods. Unlike the LR model that recovers one set of spatial coordinates $\mathbf{\Theta}$ used to span the subspace of whole dynamic image series (with subspace basis $\mathbf{\Phi}$), the LTSA model allows us to recover local coordinates $\mathbf{\Theta}_c$ for each neighborhood, and therefore span the local subspaces formed by each of them (with subspace basis $\mathbf{\Phi}_c$), as shown in Figure 3.11. The relatively small linear transform matrix (see Figure 3.11) aligns these local subspaces, avoiding the unnecessary large amount of unknowns from methods such as t-LLR.

3.5.2 Comparison with low-rank tensor model

Another case is the low-rank tensor (LRT) model of multi-array data [55]. Assume that $\mathbf{X}$ can be arranged into a three-way tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{M\times P\times C}$. Equivalently, $\mathbf{X}$ can be grouped into C neighborhoods, i.e., $\mathbf{X}_c = [\boldsymbol{\mathcal{X}}_{1,c}, \boldsymbol{\mathcal{X}}_{2,c}, \ldots, \boldsymbol{\mathcal{X}}_{P,c}] \in \mathbb{C}^{M\times P}$, where $\boldsymbol{\mathcal{X}}_{p,c}$ denotes a $M \times 1$ vector in $\boldsymbol{\mathcal{X}}$ at the p-th column and c-th slice. Assume that the tensor $\boldsymbol{\mathcal{X}}$ can be written as a low-rank tensor in the Tucker form with a rank of (D, D, C):

$$\boldsymbol{\mathcal{X}} = \sum_{l_1=1}^{D}\sum_{l_2=1}^{D}\sum_{l_3=1}^{C} \boldsymbol{\mathcal{G}}_{l_1,l_2,l_3}\mathbf{u}_{l_1} \circ \mathbf{v}_{l_2} \circ \mathbf{w}_{l_3}, \tag{3.25}$$

where "$\circ$" denotes Kronecker product, $\mathbf{u}_{l_1} \in \mathbb{C}^{M\times 1}$, $\mathbf{v}_{l_2} \in \mathbb{C}^{P\times 1}$, and $\mathbf{w}_{l_3} \in \mathbb{C}^{C\times 1}$ denote the bases of $\boldsymbol{\mathcal{X}}$, $\boldsymbol{\mathcal{G}} \in \mathbb{C}^{D\times D\times C}$ is the core tensor, and $\boldsymbol{\mathcal{G}}_{l_1,l_2,l_3}$ is the entry of $\boldsymbol{\mathcal{G}}$ with index (l_1, l_2, l_3). The c-th

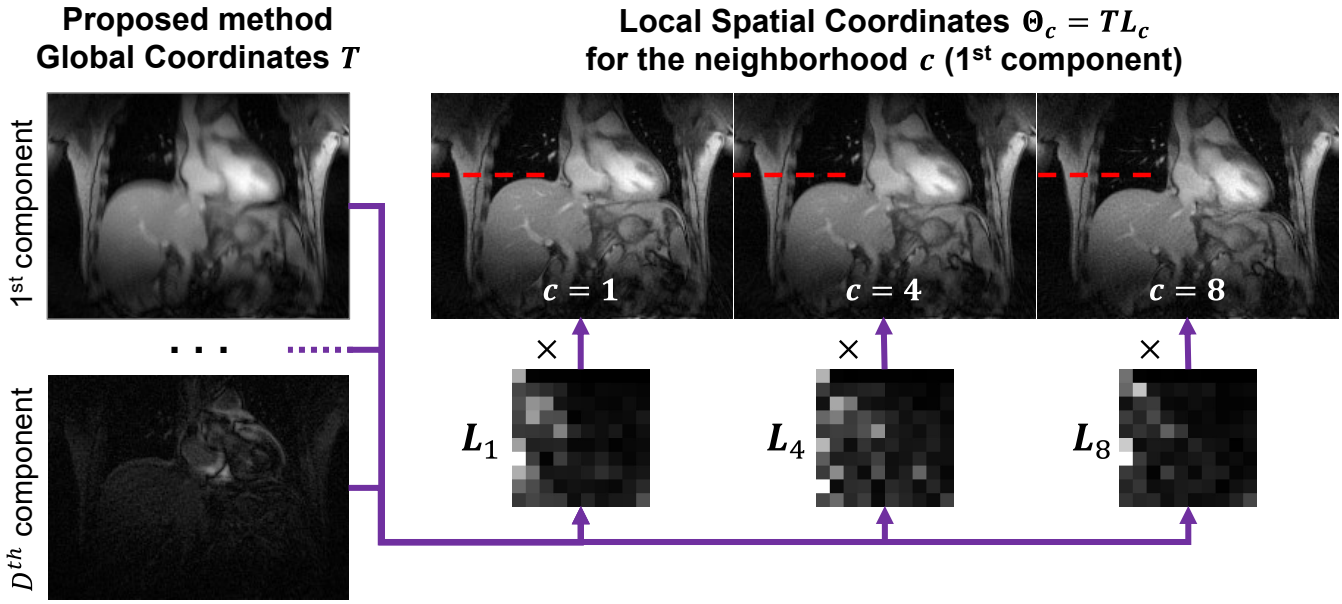


Figure 3.11: Global coordinates (spatial coefficients) and linear transform matrices of the proposed method. The linear transform matrices $\mathbf{L}_c$ are represented using gray levels. In this example, a rank $D = 10$ was used, hence each $\mathbf{L}_c$ matrix having a size 10×10. The 1st component of $\mathbf{T}$ is a temporal average of all the neighborhoods while higher components represent dynamic changes. Our proposed model successfully reconstructs the subspace of each neighborhood by applying $\mathbf{L}_c$ to the global coordinates $\mathbf{T}$. Note that the symbol "$\times$" represents the matrix multiplication operation between the components of $\mathbf{T}$ and $\mathbf{L}_c$. The red dashed lines indicate the position of the liver for the 1st component of the 1st neighborhood.

slice of $\boldsymbol{\mathcal{X}}$ can be then written as:

$$\begin{aligned}\mathbf{X}_c &= \sum_{l_1=1}^{D}\sum_{l_2=1}^{D} \mathbf{u}_{l_1}\left(\sum_{l_3=1}^{C} \boldsymbol{\mathcal{G}}_{l_1,l_2,l_3}\mathbf{w}_{l_3}(c)\right)\mathbf{v}_{l_2}^{\mathrm{T}} \\ &= \mathbf{U}\mathbf{G}_c\mathbf{V}^{\mathrm{T}},\end{aligned} \tag{3.26}$$

where $\mathbf{U} \in \mathbb{C}^{M\times D}$, $\mathbf{G}_c \in \mathbb{C}^{D\times D}$, and $\mathbf{V} \in \mathbb{C}^{P\times D}$. The LTSA model in Equation (3.12) reduces to the LRT model in Equation (3.26) when $\mathbf{T} = \mathbf{U}$, $\mathbf{L}_c^{-1} = \mathbf{G}_c$ and $\boldsymbol{\Phi}_c = \mathbf{V}, \forall c$. In other words, the LTSA model is reduced to a global LRT model of the dynamic images by assuming that all neighborhoods share the same set of temporal bases. In the case of dynamic imaging, the LRT model essentially groups the dynamic images into different motion phases, i.e., $P \times C$ motion phases for the tensor in Equation (3.25), while ignoring interphase variations. On the other hand, the LTSA model allows representing images in each neighborhood using independent temporal bases, enabling real-time imaging.

3.5.3 Comparison with BiLMDM

Both the proposed method and the bi-linear modeling of data manifolds (BiLMDM) method [143] employ a bi-linear model of dynamic MR images for image reconstruction from sparsely sampled

k-space data. However, despite the similarities in mathematical expressions, the two methods are different in model construction and image reconstruction algorithm.

The proposed method is built upon the idea that the local coordinates of a smooth manifold can be aligned with the global coordinates via a linear transform, which leads to the LTSA model in Equation (3.13). With the temporal bases of each neighborhood pre-determined, the bi-linearity of the LTSA model arises in the form of $\mathbf{T}\mathbf{L}_c$, which links the global coordinates of the manifold $\mathbf{T}$ to the local coordinates $\mathbf{\Theta}_c$ of the neighborhood c through a small linear transform matrix $\mathbf{L}_c \in \mathbb{C}^{D\times D}$. In addition, as shown in Section 3.5.1 and Section 3.5.2, such a bi-linear transform builds direct connections to the low-rank, temporal local low-rank, and low-rank tensor models, making the proposed LTSA model a natural non-linear generalization of these linear models.

On the other hand, the BiLMDM method leverages the fact that each data vector on a smooth manifold can be approximated by an affine combination of neighboring *landmark* points, leading to the following model of the Casorati matrix in Equation (3.2):

$$\mathbf{X} = \mathbf{U}\check{\mathbf{\Lambda}}\mathbf{B}, \tag{3.27}$$

where $\check{\mathbf{\Lambda}} \in \mathbb{C}^{D\times M_l}$ represents a low-dimensional latent structure of the dynamic images, which is pre-learned from M_l ($M_l \leq M$) "landmark" points via robust sparse embedding; $\mathbf{B} \in \mathbb{C}^{M_l\times N}$ is a large sparse matrix gathering the coefficients of the affine combinations of the landmark points; and $\mathbf{U} \in \mathbb{C}^{M\times D}$ captures spatial features of the dynamic images. The bi-linearity of the BiLMDM model means that $\mathbf{X}$ is linear with respect to $\mathbf{U}$ or $\mathbf{B}$ when one of them is fixed. In image reconstruction, the BiLMDM method requires the estimation of the spatial feature matrix $\mathbf{U}$ and the sparse affine transformation matrix $\mathbf{B}$.

3.5.4 Future prospects

During this PhD thesis, we developed a novel MR image reconstruction method that can be directly linked to state-of-the art reconstruction models such as the low-rank and low-rank tensor models. Much work has been done by the whole research team to understand better this manifold-learning based method, yet further improvement in some areas are still ongoing efforts. We discuss here some of the ongoing investigations on the LTSA method.

Initialization and effect of the iterative scheme

The LTSA model for MR image reconstruction is a bi-linear model, which makes its optimization not trivial. In the current work, we proposed a simple alternating optimization scheme (with ADMM when using an additional sparsity penalty). However, the optimization problem is bi-convex (i.e., convex in $\mathbf{T}$ when $\mathbf{L}_c$ is fixed and vice versa) and such an algorithm does not guarantee global optimality of the solution, especially when using bad initialization. Therefore, a good first guess of either the global coordinates $\mathbf{T}$ or the linear transform $\mathbf{L}_c$ is of the essence. In our optimization

scheme, we first fix $\mathbf{L}_c$ and recover $\mathbf{T}$ in the first iteration. Therefore, as the first sub-problem is convex for $\mathbf{T}$, the initialization of $\mathbf{T}$ is not relevant and making sure $\mathbf{L}_c$ is correctly initialized is important. The choice to start the optimization with $\mathbf{T}$ and fix $\mathbf{L}_c$ was made due to the relatively small size of the linear transform matrix. There is then less room for error when trying to estimate $\mathbf{L}_c$ than trying to estimate a large matrix such as $\mathbf{T}$. This intuition proved to be successful, and several ways to initialize the linear transform matrices have been explored. To this day we kept in mind two: low-rank initialized and random matrix as described in Section 3.3.3. The random initialization actually yielded good reconstruction results on par with LR initialized reconstructions. We also noticed that in some cases where the LR reconstructed images displayed artifacts, the random initialized LTSA images had better overall quality than the LR initialized.

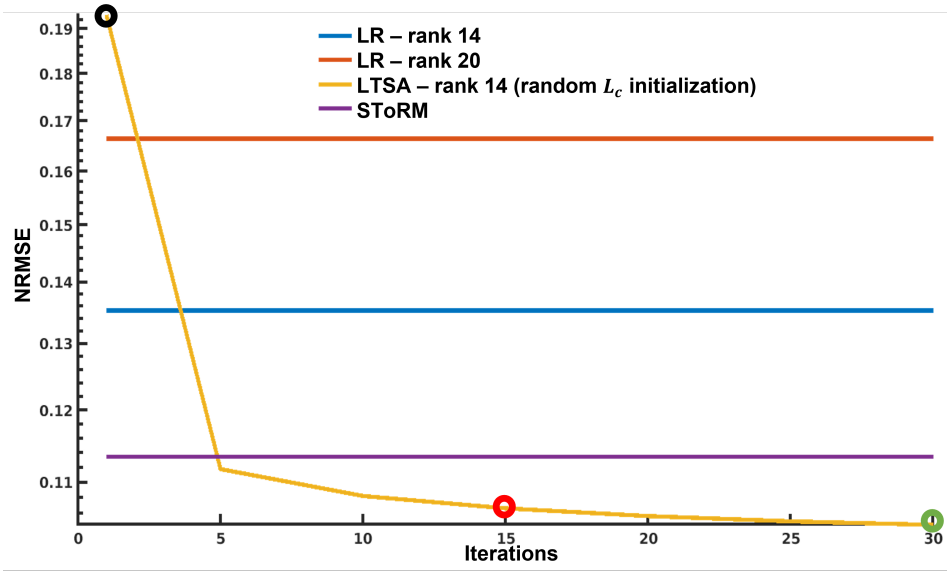


Figure 3.12: NRMSE of the dynamic image series per alternating optimization iteration for the proposed LTSA method (yellow line). The $\mathbf{L}_c$ matrices were initialized with random matrices in this example. The 2D dataset presented in Section 3.4.3 was used with 26 seconds of data for this plot. The NRMSE of the final reconstruction of the LR rank 14 (blue), LR rank 20 (orange) and SToRM (purple) methods are also displayed, demonstrating the advantage of the proposed LTSA method with only a couple of iterations. The images at the iterations indicated by the black, red, and green circles are shown in Figure 3.13. Note that despite the alternating optimization scheme, the proposed method and SToRM had similar reconstruction time.

The fact that the LTSA is a bilinear model makes us pursue two different variables, which leads to this alternating optimization scheme. Having perfect linear transforms $\mathbf{L}_c$ from the beginning would obviously be ideal to alleviate ourselves from an alternating scheme and further increase the speed of the reconstruction. However, the iterative scheme itself is not without merit, as it allows us to get closer to a good solution in a controlled way (i.e., with a chosen number of iterations). Figure 3.12 displays the NRMSE of the dynamic image series $\mathbf{X}$ per alternating optimization iteration for the proposed method and the compared methods. As the LR and the SToRM methods do not have an alternating optimization process, the NRMSE of the resulting reconstructed images is displayed as constant. After only 5 iterations, the proposed LTSA method already yields a lower NRMSE than the NRMSE of the compared methods, and seems to start converging after 20 iterations, with

small gain past 25 iterations. Please also note that despite the alternating iterative scheme, the proposed method had similar reconstruction time than SToRM on CPU workstation and much less memory demand due to the dimensionality reduction it operates. The improvement in terms of image quality is shown in Figure 3.13, which gives better insight on how the iterations affect the reconstruction.

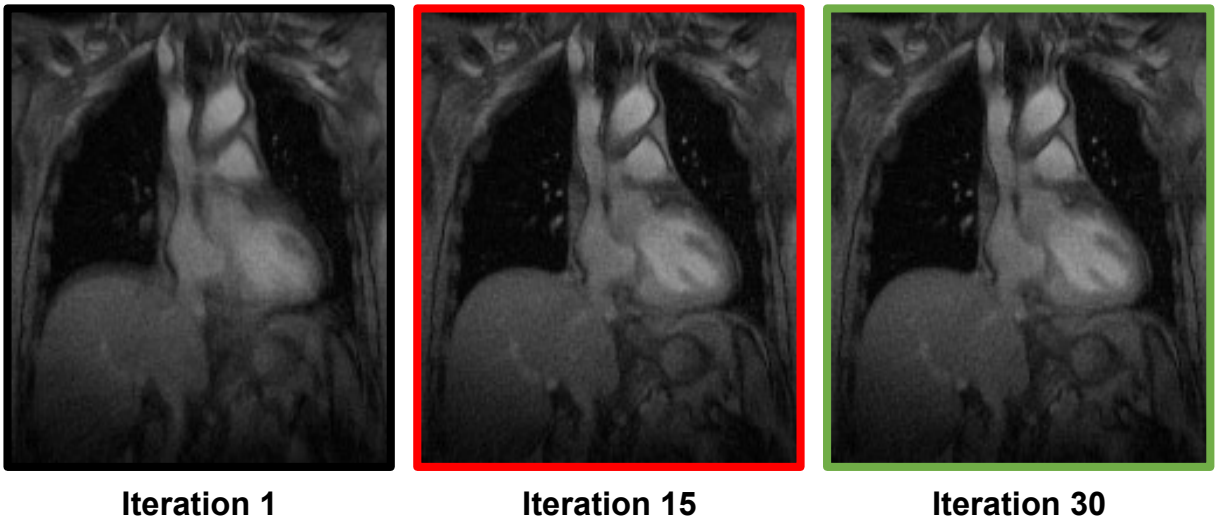


Figure 3.13: Reconstructed images of the proposed method at the iterations 1 (black box), 15 (red box), and 30 (green box). The noise in the image gradually decreases and details are gained through iterations. Corresponding NRMSE can be found in Figure 3.12 for the iterations indicated with the black, red, and green circles.

Spatial patch extension

In the background section where we derive the LTSA model, we first started the derivation with a single pixel in Equation (3.8) and generalized to the whole image by stacking the temporal signals together and forming the matrix $\mathbf{X}_c$. However, we could also think that the spatial dimension (i.e, within pixels) is redundant and assume that a dimensionality reduction on the first dimension of the image (i.e., spatial dimension) is feasible. In such a case, not only would we leverage the spatiotemporal correlation of a rank D matrix for image reconstruction, but we would also utilize the spatial redundancy to reconstruct the whole image from a set of temporal signals $P < M$, which leads to a patch-based version of the LTSA method. Therefore, one of the investigated paths with the LTSA scheme was to determine how well it would perform with patches.

After mathematical derivations, we implemented the patch-based LTSA model and performed image reconstruction. However, the lead was abandoned due mainly to three reasons. First, this new version of LTSA was tested on a CPU workstation and extending to patches was quite expensive in terms of computation time and memory. In addition, integration with the current GPU implementation was not straightforward and could not be done in reasonable timing. Secondly, the preliminary results did not show much improvement compared to the original LTSA-method. We believe this is mostly due to newly introduced parameters (such as the patch size, the number of

patches, the overlap percentage of the patches, etc.) that have not yet been fully optimized. Further optimization of these parameters and understanding what is the optimal way to form patches within the dynamic image series would have been ideal but also time-consuming, which leads us to the third reason. Moving on deep learning-based image reconstruction was a goal of the author during this PhD thesis and further developing the patch-based approach would not have allowed for enough time to actually pursue this route. The choice to move on to deep learning-based image reconstruction was then made, which leads us to the following Chapter 4.

Chapter 4

Non-linear model using neural networks: towards deep learning based reconstruction

Chapters 2 and 3 presented respectively linear and non-linear based approaches for MR image reconstruction. However, even though the LTSA method is based on a non-linear model, both methods actually use conventional linear optimization techniques to solve the optimization problem. This chapter presents an emerging MR image reconstruction field where neural networks (NN) are used for optimization.

4.1 Introduction

Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn [162]. A sub-field of ML is then artificial neural networks (ANNs) [96]. They are computational models inspired by the structure and functioning of the human brain. ANNs are composed of interconnected nodes, called artificial neurons (ANs), which process and transmit information. Conventional ANNs typically use a single (or very few) AN units organized as layers between inputs and outputs. By contrast, Deep learning (DL) is a subset of ANN and ML that uses a large number of layers (i.e., hidden layers) between inputs and outputs, allowing for a more complex representation of the data.

Techniques based on deep neural networks (DNNs) have emerged as powerful tools for optimization over the past two decades, showing promising results in many fields, from natural language processing to image processing [77]. These algorithms can learn features such as patterns, low-dimensional representation, distributions, etc. from large datasets. The network's weights are adjusted during a training process, allowing the network to learn and adapt to the training dataset (notably by solving an optimization problem). The weights learned from the training set are then utilized to perform specific tasks (e.g., classification, segmentation, interpolation, etc.) on new "testing" datasets. The new datasets must however show similarities to the training set in some aspect, e.g., be from the same imaging modalities or have similar noise floor in order to have a relevant

output.

In particular, deep learning techniques have recently been extensively studied and tested in the field of medical imaging to improve the image quality and further accelerate imaging time. In the case of MR image reconstruction, deep neural networks such as recurrent or convolutional neural networks (RNNs or CNNs) [1, 124, 137, 156] or generative adversarial networks (GANs) [25, 90, 94, 166] have been studied for various applications in both static and dynamic MRI. There can be several attack angles such as denoising/deblurring [60], dealiasing [78], low-dimensional features learning [150, 167] or extension of CS using neural networks [50, 67]. These schemes traditionally process data either in the image space or in the k-space, but some methods also do both [30, 157]. Despite promising results, this type of technique can be computationally expensive and may require a large amount of data and resources to train. More precisely, one of the main hurdles in the case of dynamic MR imaging is the lack of available data. Indeed, one would require a large set of high-resolution artifact-free images as labels for network training and acquiring such training labels can be difficult for some dynamic MRI applications such as 3D cardiac imaging. In addition, these networks may be prone to overfitting, meaning that they may perform well on the training data but poorly on new, unseen data. To address these issues, some techniques have been developed such as regularization and model ensembles to improve the performance and generalization of deep learning networks [57, 112, 129, 155].

Overall, deep learning techniques have been shown to be effective at improving the speed and accuracy of static MR imaging, yet more work is still needed for dynamic MR imaging. Despite a potentially long training time, once the network is trained, the reconstruction is usually relatively fast which can be very useful for online reconstruction in clinical settings: the weights of the networks are already determined and only the inference time affects the reconstruction time. To this day, deep learning methods are still being investigated to further improve MR image reconstruction from highly undersampled k-space data. We present here some of the basic concepts of deep neural networks. Then we give a brief overview of a DL-based MR image reconstruction technique. Finally, we show some preliminary results of the current work performed in that field.

4.2 Background

4.2.1 Individual neuron and layer architecture

Neural networks are composed of individual units called neurons that mimic the way the brain processes information and that perform simple tasks [96]. As simple as these individual tasks may be, a sufficient amount of neurons (or nodes) connected together and forming layers can procure a large representation power. A neuron is typically described by four parameters: inputs, weight, bias, and activation function. The computation performed by a neuron is then a simple mathematical operation, such as a dot product of the input data and a set of weights, with an offset (bias), followed by the application of an activation function. While the inputs and activation function are

fixed, weights and biases on the other hand are trainable parameters that are learned during the training process, determining the specific computation that a neuron performs on its input data. This learning process aims at optimizing a loss function that depends on the task.

Weights The weights of a neuron are values that determine the strength of the connection between that neuron and the neurons in the previous layer. The weights are typically represented as an array of values, one for each connection to a neuron in the previous layer

Bias The bias of a neuron is a single value that represents an additional input to the neuron that is independent of the input data. The bias can be thought of as a trainable parameter that determines the overall "shift" or "offset" of the neuron's output.

Activation function The activation function of a neuron is a function that determines the output of the neuron. Activation functions can be used to introduce non-linearity into the network, which allows it to learn more complex patterns in the data. Indeed, without activation functions, a neural network would be limited to learning linear relationships between the input and output data. Some common activation functions include the sigmoid function, the tanh function, and the ReLU (Rectified Linear Unit) function.

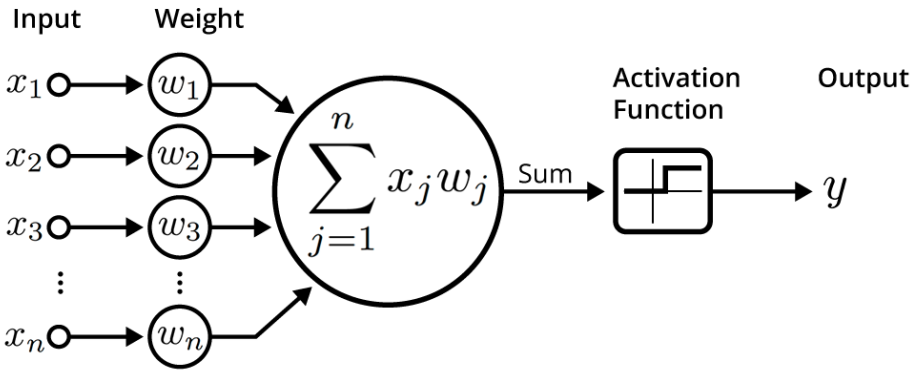


Figure 4.1: Illustration of a neuron unit in an artificial neural network. Source: www.freecodecamp.org

A schematic of an individual neuron unit is shown in Figure 4.1. These neurons can then be stacked and interconnected to form layers and eventually a neural network. Neural networks are traditionally represented with an input layer, one or more hidden layers, and an output layer. The input layer is the first layer of the neural network, it represents the input data. The input data can be raw data such as images, text, or audio, or it can be preprocessed and transformed data. For example, in the case of image processing, each unit in the input layer could represent a single voxel of the input data. The hidden layers are located between the input and output layers and they process the input data features. They transform the input data by combining them according to the neurons' weights and biases. The number of hidden layers and the number of units in each layer are design choices that can significantly impact the performance of the network. The output

layer is then the last layer in the neural network and it produces the final output of the network. This can be a single value, such as a prediction or a probability, or it can be a vector of multiple values, such as a classification label or even an image (e.g., denoised).

4.2.2 Conventional neural network models

As described in Section 4.2.1, neuron units can be connected together to form layers and complex networks. We describe here some of the conventional networks used in deep learning and in image processing tasks.

Perceptron and feed-forward neural network

The simplest network, namely the perceptron [131], consists of a single layer without hidden layers and is mostly used for binary classification tasks. It receives input data and combines them together with weights and biases to produce an output. If the output is above a certain threshold, the perceptron classifies the input as belonging to the first class; otherwise, it classifies the input as belonging to the second class. Similarly, the feed-forward neural network is composed of perceptrons stacked together, forming multiple layers, with the input data flowing through the network in a single direction. This network can be used for a variety of tasks, including classification, regression, and function approximation. Note that the perceptron and feed-forward network are not considered "deep" (because of their limited number of layers) and are the predecessors of current deep neural network architectures.

Multi-layer perceptron (MLP)

A multi-layer perceptron (Figure 4.2) is a type of neural network that consists of multiple hidden layers of interconnected neurons. Each neuron receives inputs from other neurons from the previous layer, which is why they are often denoted as "dense" or "fully-connected" networks. In addition, the hidden layers do not necessarily have the same amount of feature units than the previous layer. Most importantly, MLPs have bidirectional propagation [134, 135] (forward and backward), which is one of the most notable difference with regular feed-forward networks. Such networks can be trained to perform a variety of tasks, such as complex classification, regression, and feature learning.

Convolutional neural network (CNN)

A convolutional neural network (Figure 4.3) is a type of neural network that is specifically designed to process data with a grid-like topology, such as an image [34, 76]. CNNs are composed of multiple layers of interconnected neurons, but the layers are arranged in a way that preserves the spatial relationship between pixels of the input data. Each layer then learns to extract increasingly complex features from the input based on spatial convolutions. This specificity makes CNNs particularly

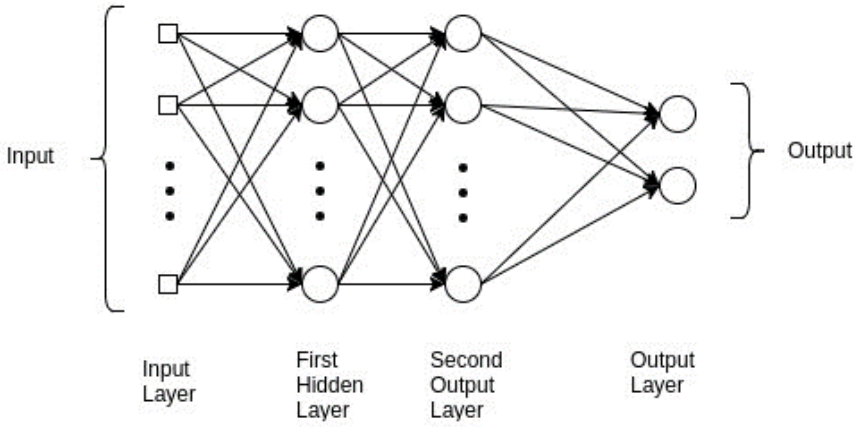


Figure 4.2: Example of an MLP architecture. In this example, there are two hidden layers. Source: analyticsvidhya.com

well-suited for image processing tasks. In the context of image processing, CNNs can be used to denoise images for example.

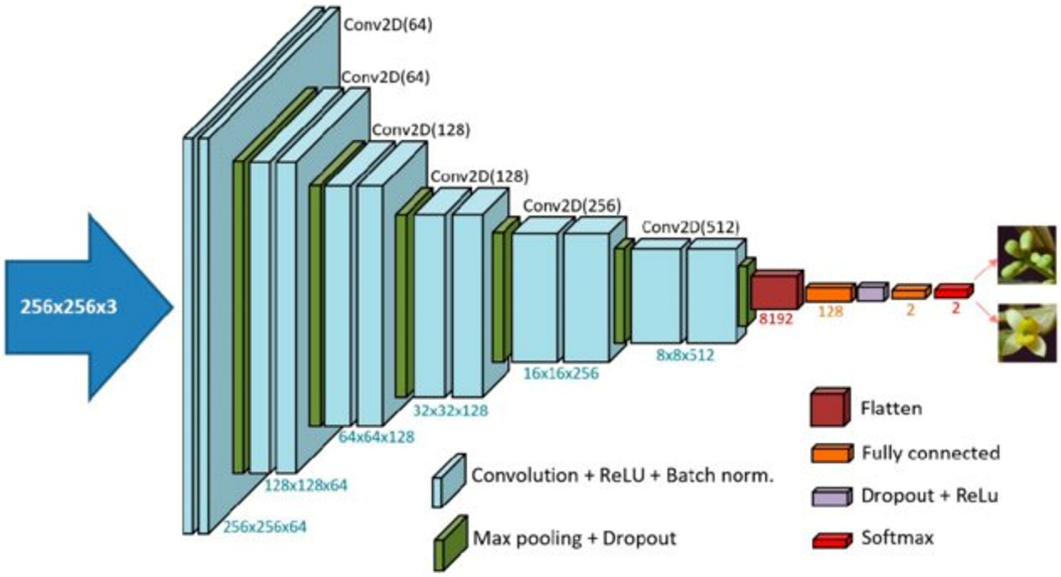


Figure 4.3: Example of a CNN architecture. In this example, an RGB 256×256 image is processed. The RGB color features are stacked in the channel dimension [100].

Recurrent neural network (RNN)

A recurrent neural network (RNN) is a type of neural network that has a "time memory" and is designed to process sequential data, such as time series or natural language [58, 134]. RNNs are composed of multiple layers of neurons, but the connections between neurons are not fully connected like in a standard neural network. Instead, the connections between neurons are designed to form a directed cycle, which allows the RNN to maintain a "memory" of past input and use it to inform the

processing of current input. This makes RNNs well-suited for tasks such as language translation, language modeling, and speech recognition.

Generative Adversarial Network (GAN)

A generative adversarial network (GAN) [40] is a more advanced and recent type of network that is composed of two networks: a generator network and a discriminator network. The generator network is trained to generate new data that are similar to a training dataset, while the discriminator network is trained to distinguish between the generated data and the real data from the training dataset. The two networks are trained in an adversarial manner, with the generator trying to "fool" the discriminator and the discriminator trying to correctly identify the generated data. GANs can be used to generate synthetic images, text, and other types of data. In the context of MR image reconstruction, they can be used to learn a mapping from undersampled or noisy MR data to a high-resolution or noise-free image.

4.2.3 Training types

To operate correctly, the networks presented above first need to "learn" from a training dataset, a step where the weights of the network will be updated depending on the task at hand. There are mainly three training categories for deep-learning based neural networks: supervised, semi-supervised, and unsupervised. We present them here with an MR image reconstruction perspective.

Supervised training Supervised MR image reconstruction techniques involve training a model on a labeled dataset, where the correct output (e.g., a high-resolution or noise-free image) is provided for each input (e.g., undersampled or noisy MR data) in the training set. The goal of the model is to learn a function that can predict the correct output for new, unseen inputs based on the patterns it has learned from the training data. This type of approach is typically the most accurate and is often used when large amounts of labeled data are available.

Semi-supervised training Semi-supervised MR image reconstruction techniques involve training a model on a dataset that contains a mix of labeled and unlabeled examples. The goal is still to learn a function that can predict the correct output for new, unseen inputs, but the model can also use the unlabeled examples to learn additional patterns in the data that may not be present in the labeled examples. This type of approach can be a good compromise when some labeled data are available, but not enough to fully train a supervised learning model.

Unsupervised training Lastly, unsupervised MR image reconstruction techniques do not use any labeled examples and instead rely on the model to learn to discover patterns in the data on its own. Such type of techniques include for example NN-based dimensionality reduction, feature learning (e.g., using auto-encoder networks [134]); or clustering and segmentation (e.g., using

contrastive learning approaches [20]). Unsupervised learning algorithms do not always have a specific output that they are trying to predict, but instead may try to find structures or patterns in the data. This type of approach is often used when labeled data are scarce or not available.

Overall, the choice of which type of dynamic MR image reconstruction technique to use will depend on the specific task and the availability of labeled data. One of the main hurdles when training a neural network for MR image reconstruction is the availability of ground-truth[a], i.e., a large set of high-resolution artifact-free images. Depending on the application, obtaining such a reference may not even be feasible (e.g., 3D cardiac dynamic MRI with high temporal/spatial resolution).

4.2.4 DNN for MR image reconstruction: MoDL

We present in this section an existing unrolling-based MR image reconstruction scheme which lays ground to the foundation of the current work of this PhD thesis, and potential future work. Unrolling-based DNNs involve breaking down an image reconstruction process into a series of steps or "unrolling" the process over a number of iterations, where free-parameters can be learned through training [79]. For example, we can break down an iterative optimization algorithm such as gradient descent. Then, each step of the gradient descent is done by going through some neural network layers. Along that line, we chose to focus on the Model-based Deep Learning (MoDL) [1] framework due to its versatility (designed for generalized inverse problems) and its mathematical foundations. While we focus on this framework here, other schemes of interest that are non-unrolling based and that directly estimate the non-aliased images or the fully-sampled k-spaces do exist [3, 94].

MoDL [1] introduced a systematic approach for architecture design and inverse problem resolution of arbitrary structure. While several contributions can be attributed to MoDL, one of the main contributions of the authors in [1] is that their approach can be used to learn a variety of different model-based optimization algorithms for MR image reconstruction, including algorithms based on different image priors and optimization approaches, with mathematical grounds justifying the model. In addition, MoDL successfully uses and merges gradual improvement from several DNN-based MR image reconstruction such as end-to-end approach, data consistency enforcement, or residual learning.

Forward model

The general optimization problem for image reconstruction with regularization is posed as follows:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \|\mathcal{A}(\mathbf{X}) - \mathbf{s}\|_F^2 + \lambda \|\mathcal{R}(\mathbf{X})\|_F^2 \tag{4.1}$$

[a]Note that in this thesis, the term ground-truth for MR images may refer to reference artifact-free images and not obviously to images obtained from fully-sampled k-spaces (which is impossible to obtain in many cases).

where $\mathcal{A}(\bullet) = \Omega(\mathbf{F}_s \bullet)$ is the forward imaging model. The first term is the data consistency (DC). It enforces fidelity to the measurements $\mathbf{s}$. The second term is a regularization done using the transform $\mathcal{R}$ and weighted by λ. This term commonly involves "prior" knowledge of the data.

In MoDL, a CNN $\mathcal{N}_{\mathbf{w}}$ learning noise and alias pattern in the residual image through the trained weights $\mathbf{w}$ is used for regularization. An important point is that the CNN learns from the residual image and is not estimating the image itself but the "remaining" that is unwanted:

$$\mathcal{N}_{\mathbf{w}}(\mathbf{X}) = \mathbf{X} - \mathcal{D}_{\mathbf{w}}(\mathbf{X}) \tag{4.2}$$

where $\mathcal{D}_{\mathbf{w}}(\mathbf{X})$ is a denoised/dealiased version of $\mathbf{X}$ after going through the denoising/dealiasing CNN. Equation (4.2) can therefore be written as:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \|\mathcal{A}(\mathbf{X}) - \mathbf{s}\|_F^2 + \lambda \|\mathbf{X} - \mathcal{D}_{\mathbf{w}}(\mathbf{X})\|_F^2 \tag{4.3}$$

Note that λ is also a learnable parameter of the network.

The above scheme is then decoupled and solved in an alternating and iterative manner, yielding two sub-problems:

$$\mathbf{Z}_n = \mathcal{D}_{\mathbf{w}}(\mathbf{X}_n) \tag{4.4a}$$

$$\mathbf{X}_{n+1} = \arg\min_{\mathbf{X}} \|\mathcal{A}(\mathbf{X}) - \mathbf{s}\|_F^2 + \lambda \|\mathbf{X} - \mathbf{Z}_n\|_F^2 \tag{4.4b}$$

The sub-problem in Equation (4.4b) enforces data consistency (DC) and can easily be solved using conjugate gradient methods for example. This "layer" updating $\mathbf{X}_n$ is denoted as DC layer. The sub-problem in Equation (4.4a) updates the regularization term used in Equation (4.4b) with the CNN-based denoiser and can be solved with standard backpropagation schemes to update weights with end-to-end training. A schematic of the MoDL scheme is shown in Figure 4.4. In-depth mathematical details can be found in [1].

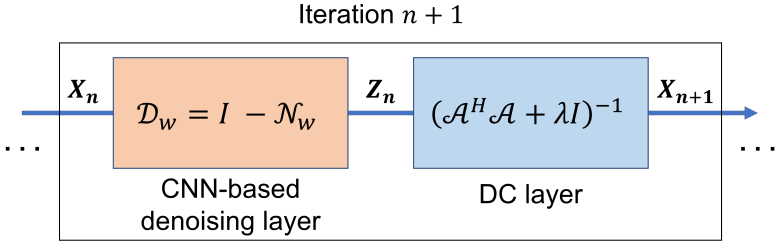


Figure 4.4: Iterative scheme of MoDL at an iteration $n+1$ that alternates between CNN-based denoising ($\mathcal{D}_{\mathbf{w}}$) layer and data consistency (DC) layer.

CNN architecture and implementation details

The CNN-based denoiser in MoDL is composed of N_{layers} layers with 64 convolution (Conv) filters, batch normalization (BN) and rectified linear unit (ReLU) activation function. In [1], the authors show results with different number of layers, and set the optimal number of layers to $N_{layers} = 5$. More specifically, all the layers include (Conv + BN + ReLU), except the last one, that does not have ReLU activation function in order to prevent negative part truncation before the DC layer. A figure of the $\mathcal{D}_{\mathbf{w}}$ layer is shown in Figure 4.5.

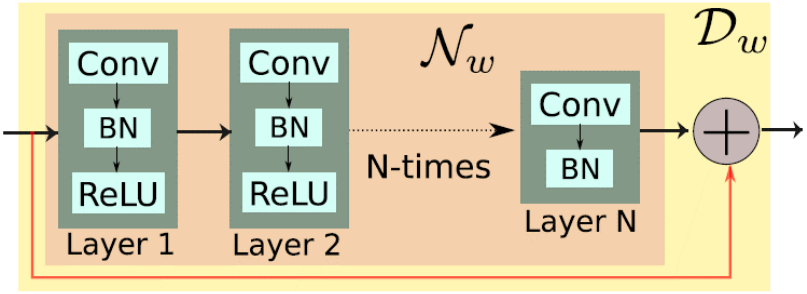


Figure 4.5: CNN denoiser architecture. This layer, shown as the first layer in Figure 4.4, learns aliasing patterns and noise from the difference image (shortcut connection represented by the orange arrow).

The $\mathcal{D}_{\mathbf{w}}$ and DC layers are repeated K times to produce the equivalent of K iterations of the traditional alternating optimization process (gradient descent with decoupling). The optimal K is set to 10 in [1] and the starting image $\mathbf{X}_0$ is a zero-filled Fourier transform of the undersampled k-space data $\mathbf{s}$. Note that the k-space data and the initial value $\mathbf{X}_0$ are complex numbers, which is not always supported by deep learning libraries. A common workaround that the authors of MoDL used is to treat complex data as real vectors with two channels by concatenating the real and imaginary part in the channel dimension (i.e., convert from $\mathbb{C}^{M\times N}$ to $\mathbb{R}^{M\times N\times 2}$).

Another important note is that with 64 convolutional filters times 5 $\mathcal{D}_{\mathbf{w}}$ inner layer, times 10 iterations, the total number of unknowns of the resulting network is very large, hence requiring a large amount of data to train. A main contribution of MoDL regarding this aspect (as ground-truth MR images can sometimes be difficult to obtain) is to share weights between iterations. This means that the weights of the $K = 10$ iterations of successive $\mathcal{D}_{\mathbf{w}}$ layer and DC layers are the same, which considerably reduces the amount of trainable parameters and the complexity of the training. It should be noted that the sharing is not only on the $\mathcal{D}_{\mathbf{w}}$ layer's weight, but also on the DC layer's only weight: λ. More details and results using MoDL can be found in [1].

4.3 Current work on DNNs

We present here preliminary work done during this PhD thesis in the field of deep learning for MR based image reconstruction.

4.3.1 Choice of dataset

The task at hand in this DNN investigation is to reconstruct artifacts-free images from undersampled MR data. To do so, we first need to decide the type of training, which stems from the characteristics of the available data (i.e., amount of data and labels). In this work, we used freely-available MR data with ground-truth to enable DNN-based MR image reconstruction. Many websites or scientific challenges offer MR k-space data with ground-truth for research purpose, as reviewed in [18, 87]. However, almost all the data available online are static images/k-space (i.e., static MRI) and only a handful of websites propose dynamic MR data such as short cardiac cine images (e.g., 10 to 30 frames). Despite our interest in dynamic MRI, the scarceness of dynamic MR data lead us to choose to start our deep learning journey with static MR image reconstruction. This will help gaining valuable experience in a first step, as most of the skills and insight learned will most likely be transferable to dynamic MRI. Note also that most of the k-space data available for deep learning are based on Cartesian sampling due to a lack of deep-learning library compatible NuFFT.

We made the choice of using the fastMRI dataset [169] similarly to MoDL in [1]. The fastMRI[b] database is composed of brain and knee images/k-space data with both single-coil and multi-coil options. The MR data chosen for the reconstruction task were the single-coil knee data. More information on the dataset can be found in Section 4.3.2 – "Training and testing".

4.3.2 Implementation details

A starting point was necessary to explore the power of DNNs in this deep learning journey. Due to the flexibility and mathematical foundations of MoDL, we decided to implement a MoDL-like deep neural network with denoising layer and DC layer in the single-coil case. Note that in the case of single-coiled data with Cartesian trajectory, there is no need of gradient descent in the DC layer as an analytical solution exists [1]. In addition, as we used single-coil data we did not benefit from the acceleration factor from coil combination (i.e., parallel imaging): one of the main goals here was mostly to assess the denoising/dealiasing power of neural networks when given a large database of MR data with ground-truth even without advanced schemes such as parallel imaging.

Network architecture We used $N_{layers} = 5$ similarly to MoDL and several architecture cases were evaluated, such as $K = 1$ (i.e., 1 [$\mathcal{D}_{\mathbf{w}}$ + DC] iteration) and $K = 10$ with $\mathcal{D}_{\mathbf{w}}$ weights shared but independent λ between DC layers. Indeed, we believe that despite the potential usefulness of sharing the large amount of weights from the convolutional layers (when memory and amount of available data are troublesome), sharing the one variable shaping the DC layer and ruling the contribution of the CNN versus the data consistency is not relevant, especially when the gain in memory is negligible. If equal λ's between DC layers yield optimal results, we believe that the neural network training will then assign equal λ weights for each DC layer. Another difference with

[b] More information on the available data can also be found at https://fastmri.org/

the original MoDL is that a non-negative constraint was added on the λ weights to avoid potential negative values that would not make sense from an optimization point-of-view. The algorithm was implemented in Python using the TensorFlow library with Keras API on a NVIDIA GeForce RTX 2080 Ti 11GB. The training batches were loaded on the fly for better memory management using a custom Keras generator.

Undersampling masks As described in the section above, the MR data from the fastMRI database are fully sampled k-space giving ground-truth (artifacts-free) images with a simple Fourier transform (see Section 1.3.1). Therefore, to train the network to denoise/dealias, we first need to produce masks mimicking prospective undersampling in clinical settings. Sampling masks following Cartesian trajectories with random undersampling along the phase-encoding direction were produced randomly for every batch. For every mask, 20 lines were consistently acquired in the center of the k-space to acquire low-frequency components gathering most of the signal energy. The rest of the k-space was randomly undersampled to obtain an undersampling ratio R between 2 and 4. A representative image of the undersampling mask with undersampling ratio $R = 4$ is shown in Figure 4.6.

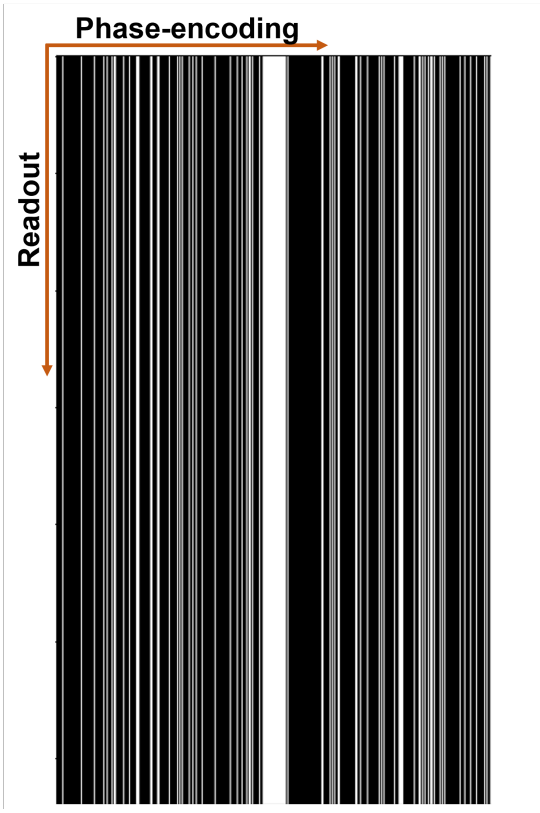


Figure 4.6: Undersampling mask produced for generating retrospectively undersampled MR images. 20 lines in the center of the k-space along the readout direction are constantly acquired, while the rest of the k-space is randomly undersampled. The undersampling ratio of the mask shown in this figure is $R = 4$.

Training and hyper-parameters 973 files were used to train the neural network and 199 files for validation. Note that each file corresponds to a subject and contains between 28 and 42 slices. As some of the first and last slices did not contain relevant anatomical information, the 10 first and 10 last slices of each file were not used for training. In the training step and the validation step, the undersampling masks were randomly produced on the fly. The matrix size was 640×372 and includes k-space data from five sequences for different contrasts and image orientations. A minority of scans showed matrix sizes smaller than 640×372, in which case the k-space was zero-padded. More details can be found in [71]. Undersampled images were produced using undersampling masks as described above.

The $K = 1$ neural network was first pre-trained with undersampled k-space data with undersampling ratio $R = 2$ for a few epochs and then with $R = 4$ after decent initialization. 50 epochs were used for the case where $K = 1$ and $R = 4$, with a batch size of 4, and 8 slices randomly picked for each batch. The full network ($K = 10$) was then trained with masks produced with $R = 4$ for 100 epochs, and later fine-tuned with 50 epochs at a lower learning rate. Due to a higher memory footprint, the batch size was decreased to 2, and 8 slices were randomly picked for each batch. The optimizer was Adam [70] with default Keras parameters, and an exponential decaying learning rate was used in order to avoid overshooting the gradient in late epochs (learning rate decaying from 10^{-2} to 10^{-3}). A gradient clipping method was also used to ensure stability in case of noisy gradients or irregular (e.g., steep) manifold structure (gradient norm was clipped to 1). Results are shown in Section 4.4.

4.4 Preliminary results

4.4.1 Ground-truth and undersampled images

Figure 4.7 displays the ground-truth MR images, i.e., the artifact-free images reconstructed from the fully sampled k-space measurement with an FFT. In order to show the variability among subjects, and to demonstrate the versatility of the denoising/dealiasing, images from 5 subjects from the testing set are shown. As can be seen, images from sequences with various contrasts and from various slice positions are present in the dataset.

The undersampled images are then produced by applying the undersampling mask to the k-space data (see Section 4.3.2). An FFT reconstruction of the zero-filled k-space data gives the images in Figure 4.8. As expected, the images show aliasing artifacts that look similar to blurring artifacts and white noise. Note that, due to the undersampling in the phase-encoding direction, most of the aliasing pattern is along the horizontal (phase-encoding) axis. The aim of the DNN implemented is then to recover the original artifact-free images.

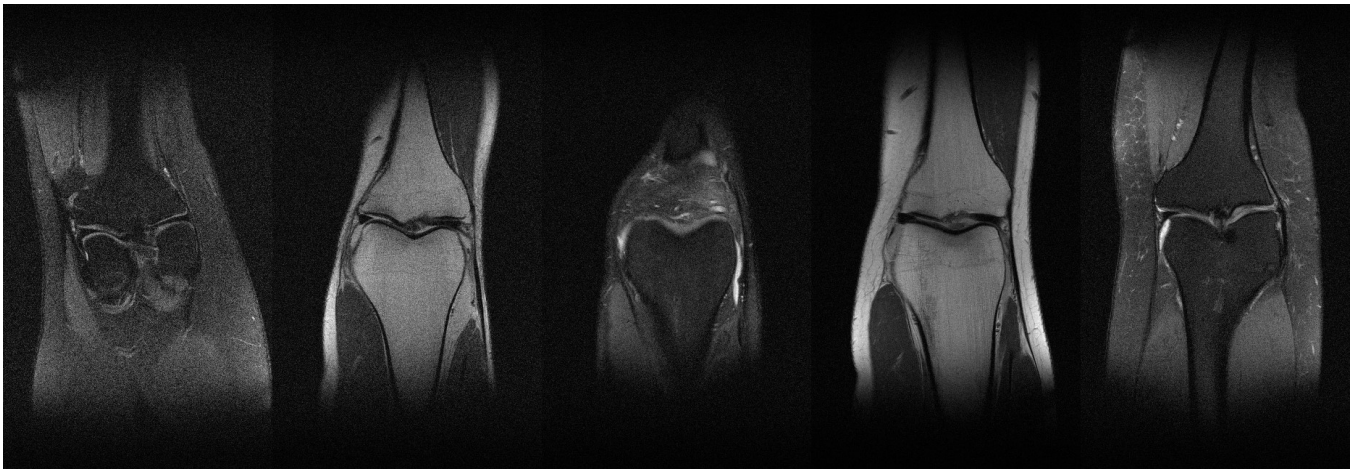


Figure 4.7: Ground-truth MR images of the single-coil knee dataset from the fastMRI database. Images from several sequences with different contrasts and image orientations are present in this dataset.

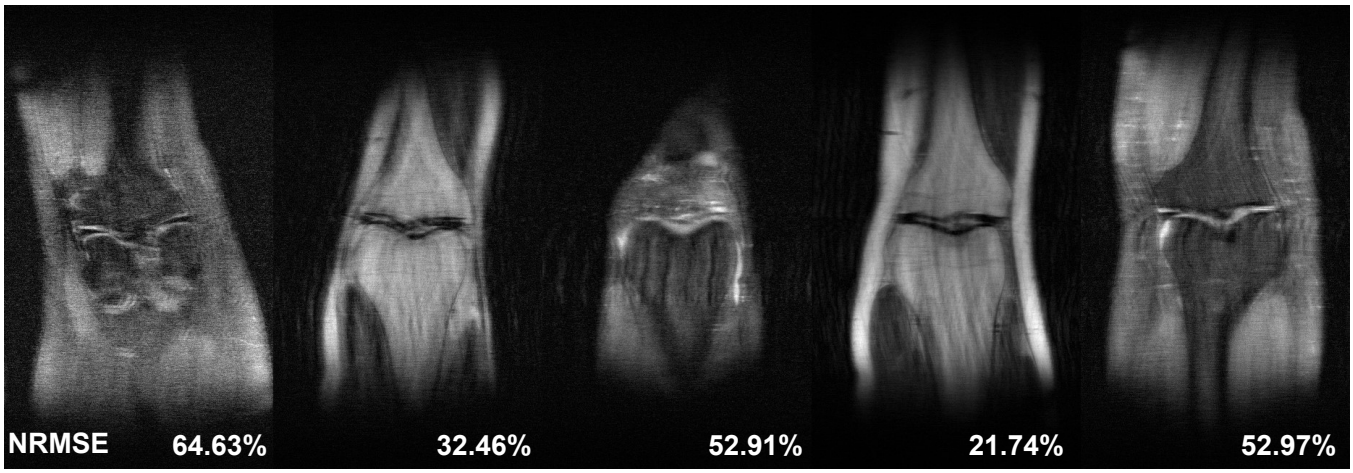


Figure 4.8: Retrospectively undersampled images of the ground-truth MR images shown in Figure 4.7. These images were produced using randomly generated sampling masks with $R = 4$ as described in Section 4.3. Their respective NRMSE is displayed at the bottom of each image.

4.4.2 Pre-training: $K = 1$ unrolled iteration

We show in this section the preliminary results obtained with the MoDL-like DNN detailed in Section 4.3. Training the full neural network (i.e., with 10 unrolled iterations) directly can be challenging as the weights are randomly initialized, and a non-linear optimization problem with so many unknowns could easily be stuck in local minima or have difficulties to converge. Therefore, a first step to prevent this is to pre-train the network with only one unrolled iteration, i.e., $K = 1$. Using only one unrolled iteration considerably reduces the complexity of the problem and can get us closer to the relevant weight distribution. The first row of Figure 4.9 shows the reconstructed images after inference of the trained DNN with $K = 1$. The second row of Figure 4.9 shows the absolute difference images defined by the absolute value of the difference between the ground-truth (or reference) image $\mathbf{X}_{ref}$ and the DNN-reconstructed image $\mathbf{X}_{DNN}$. The reconstructed images still exhibit undersampling artifacts (see second row of Figure 4.9) but appear to have gained edge

information, look less blurred, and have less noise, as indicated by the decreased NRMSE compared to the zero-padded reconstruction in Figure 4.8. The training loss values of the DNN reconstructing the images in Figure 4.9 are shown in Figure 4.10b. Figure 4.10a displays the training and validation loss values in the case where the fixed learning rate is most likely too large for the late epochs, and where there is no clipped gradient, clearly showing the advantage of decaying learning rate along with gradient clipping in the current scenario. Note that these plots represent preliminary results shown here, and could be consolidated with more experience training neural networks.

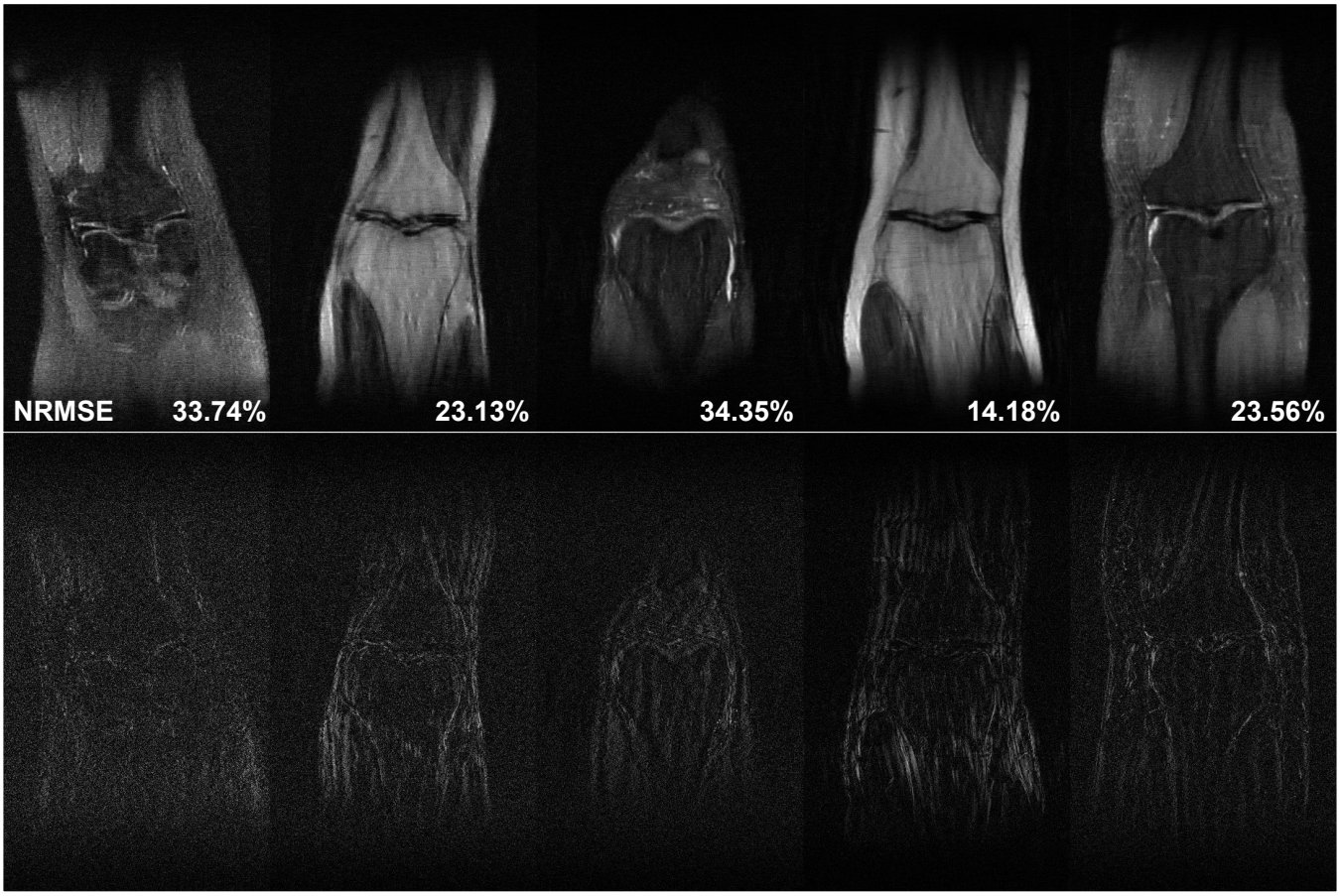


Figure 4.9: First row shows reconstructed images using the DNN architecture presented in Section 4.3 with $K = 1$ unrolled iteration. Their respective NRMSE is displayed at the bottom of each image. Second row shows the absolute difference images defined by $|\mathbf{X}_{ref} - \mathbf{X}_{DNN}|$. The images were reconstructed from undersampled k-space data ($R = 4$). The corresponding zero-filled reconstructed images are presented in Figure 4.8. The reconstructed images shown here still exhibit aliasing artifacts (mostly in the homogeneous regions) yet improvement on the image noise and sharpness is noticeable.

4.4.3 Trained DNN with $K = 10$, shared CNN weights and independent λ weights

After having pre-trained the network with $K = 1$, we used the learned weights to initialize the $K = 10$ network. As in this example we share the CNN-denoiser weights, the $K = 1$ network

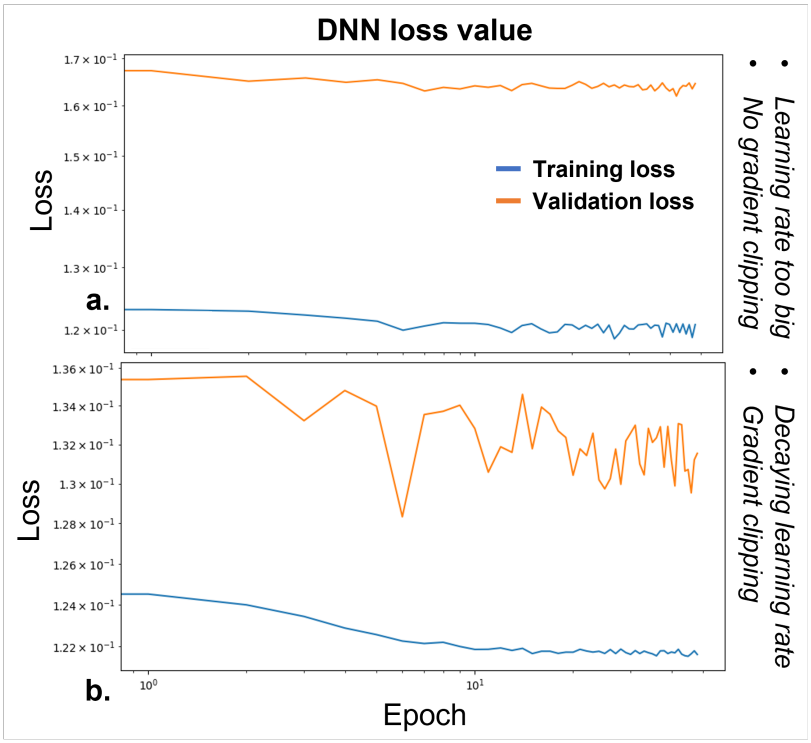


Figure 4.10: Training and validation loss values for two training cases. (**a**) Training and validation loss values in the case of a learning rate too large in the late epochs and where there is no gradient clipping. (**b**) Loss values with the same training (and the same initial learning rate) but with exponentially decaying learning rate and gradient clipping. The training loss values in (**b**) correspond to the DNN weights reconstructing the images in Figure 4.9.

was basically concatenated 10 times, interleaved with DC layers, and the training was performed again. Note that the $[\lambda_0, \cdots, \lambda_9]$ weights were all initialized with the λ from the pre-trained $K = 1$ network but are not bound to stay the same after training. The reconstructed images using the fully trained $K = 10$ DNN can be found in the first row of Figure 4.11, along their respective NRMSE at the bottom. The second row of Figure 4.11 shows the absolute difference images between the groundtruth images in Figure 4.7 and the DNN-reconstructed images with $K = 10$.

The reconstructed images shown have considerably improved overall quality and sharpness, and very little to no aliasing artifact left compared to the original undersampled images in Figure 4.8 or compared to the reconstructed images with $K = 1$ in Figure 4.9. The $[\lambda_0, \cdots, \lambda_9]$ weights after training are given in Table 4.1. As can be seen, the λ values are different between layers, potentially indicating that shared λ weights are not optimal. Note that the value of λ_9 is considerably higher than the other ones. This means that during all the iterations 0 to 8, more emphasis is given to the fidelity of the data to the measurement to recover information from the missing samples, while for the last iteration, more emphasis is given to the estimated result of the CNN denoiser.

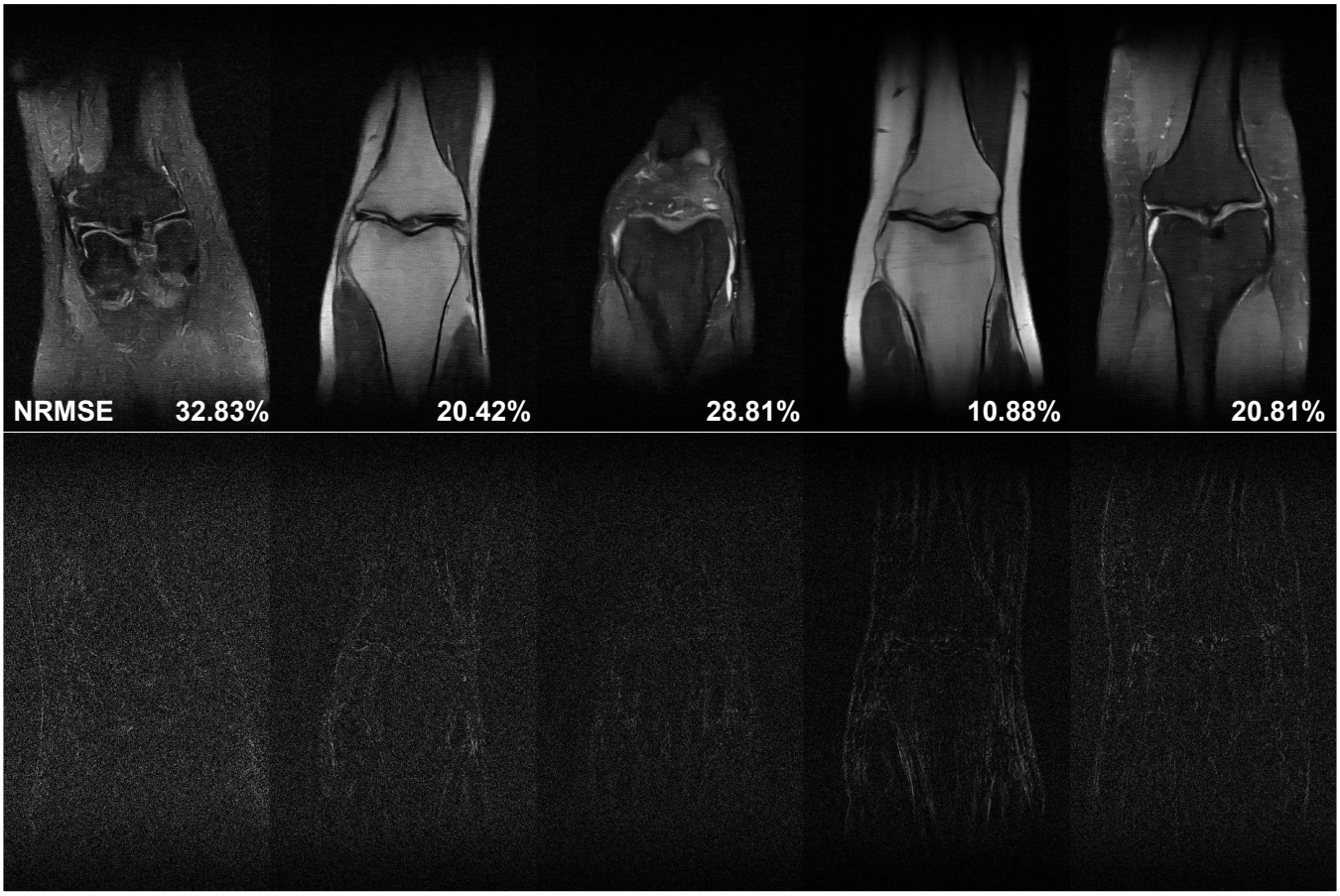


Figure 4.11: First row shows reconstructed images using the DNN architecture presented in Section 4.3 with $K = 10$ unrolled iterations, shared CNN weights, but independent DC weights $[\lambda_0, \cdots, \lambda_9]$. Their respective NRMSE is displayed at the bottom of each image. Second row shows the absolute difference images defined by $|\mathbf{X}_{ref} - \mathbf{X}_{DNN}|$. The images were reconstructed from undersampled k-space data ($R = 4$). The corresponding zero-filled reconstructed images are presented in Figure 4.8.

Table 4.1: Weights λ_0 to λ_9 after training for the DC layers 0 to 9 in the case where $K = 10$ (corresponding to the reconstruction results shown in Figure 4.11). These values weigh the contribution of the CNN regularization versus the fidelity to the original k-space measurements in the DC layer.

λ_0	λ_1	λ_2	λ_3	λ_4
6.4e-07	9.25e-07	8.88e-07	6.75e-07	6.87e-07
λ_5	λ_6	λ_7	λ_8	λ_9
4.47e-07	4.94e-07	2.34e-07	9.48e-07	1.12e-01

4.5 Conclusion and discussions

The work presented in this chapter is a preliminary research in deep learning-based MR image reconstruction and only scratches the tip of the iceberg of a very active field of study. In a first section, we presented the overall notions and network architectures in deep learning. To do so, a literature review of many neural network architectures was performed. This was beneficial to better understand the core concepts behind the more advanced DNN-based schemes such as MoDL. In a

second step, a narrower literature review of DNN-based MR image reconstruction works helped grasp the potential difficulties ahead (e.g., dynamic or static, available dataset, optimization problems, etc.) and a popular DNN-based method (MoDL) was used as a starting point. One of the main outcomes of this work was its invaluable insight on how deep neural networks behave, insight that can only be acquired through trial-and-error, with additional programming skills in TensorFlow and Python. Last but not least, this investigation gave rise to new research ideas on how to use DNNs to further improve MR image reconstruction.

However, there are lacking components in the work presented in this chapter, as this investigation is still ongoing. First, the presented work relies on an existing framework to obtain artifact-free reconstructions. Then, the results shown here could be improved with more training or better hyperparameter tuning. Other common methods such as data augmentation, or simply GPU with larger memory for larger batch size during training would also be helpful. Note that using techniques such as parallel imaging (taking into account coil sensitivity in the forward model) could further help increasing the undersampling ratio.

One of the main critics of the work presented here is the absence of comparison with state-of-the-art reconstruction techniques. This can be explained by the fact that this research route is still fresh in this PhD thesis, and that preliminary investigations are presented in this chapter. Efforts to leverage DNNs for MR image reconstruction are still ongoing, and the author showed interest to pursue this route in future research. Last but not least, the work presented here deals with static MRI, despite a higher interest in dynamic MRI. This can be explained by mainly two reasons. Static MR image reconstruction is an easier task than dynamic MR image reconstruction as dynamic MRI has a fourth (time) dimension that highly decrease the amount of k-space samples per 2D/3D volume. Starting MR research on a new field with a new library and a new programming language was therefore challenging and we made the choice of going step-by-step in that direction: first, an intermediate goal of working on static MRI, then extending to dynamic MRI once more experience and insight is gained. The second reason for working on static MRI as a first step was also the very scarce amount of dynamic MR data and labels available online. Acquiring and using MR data from our own center was also a possibility, but the main hurdle was that creating a database big enough to train a DNN would not have been feasible within the timeline of this PhD thesis.

Several leads can then be pursued to extend the current framework to dynamic imaging. A first lead would be to simply add a prior to the MoDL framework. This was done in [7] with MoDL-SToRM, where there is one CNN regularizer, and one manifold-learning based regularizer (SToRM [118]). However, this "plug-and-use" method just adds the SToRM regularizer and does not take full advantage of the manifold-learning technique: it mostly balances the contribution between the solutions of the SToRM method and the CNN. Another more interesting lead is to extend the LTSA method presented in Chapter 3 to deep learning. One way to do so could be to unroll the LTSA iterative scheme and use neural networks to solve the optimization problem.

Chapter 5

Conclusion and perspectives

MR imaging is a wide-spread modality used in many settings due to its versatility, excellent soft-tissue contrast, and high spatial resolution. In addition, being able to acquire in a dynamic way using dynamic MR imaging techniques is desirable for many applications such as cardiac imaging or dynamic contrast imaging. However, constraints such as limited acquisition times or inherent body motion often lead to measurements sampled below the Nyquist rate in both static and dynamic MRI. Advanced reconstruction techniques are therefore needed to obtain artifact-free images from the undersampled MR data. This PhD thesis presented some advanced schemes for MR image reconstruction and contributed to the development of a new reconstruction method for dynamic MRI, with additional prospective work in deep learning-based approaches.

5.1 Objectives and contributions

The first objective of this PhD thesis was to assess the usefulness, but also the limitations of the linear subspace model (also referred to as low-rank model) in a real-world setting. The main question was: can the subspace model capture irregular types of motion (e.g., bulk motion or irregular respiratory motion) and reconstruct artifact-free images from MR scans presenting such types of motion? An extension of this question was then to assess whether such a reconstruction method could deal with further variability within the data and reconstruct images with dynamic contrast. We answered these questions with two different works presented in Chapter 2.

First, we demonstrated in Section 2.2 the performance of the subspace model for dynamic MR image reconstruction in two challenging cases: bulk motion and irregular respiratory motion. In addition, we showed that the dynamic MR images obtained from a PET/MR scanner could successfully be used for PET data motion correction. The proposed scheme allowed for the reconstruction of high-resolution 3D volumes at a rate of 9.5 volume/s, which was sufficient to capture irregular respiratory motion and bulk motion. Instead of relying on navigator data or other respiratory signal surrogates to estimate respiratory motion (and neglect other types of motion), we relied on real-time dynamic MR images to correct motion in PET data and proved the versatility of the subspace

model. Then, in a separate work presented in Section 2.3, we leveraged the subspace model to successfully reconstruct free-breathing 3D dynamic MR images with dynamic contrast for cardiac applications. Using an efficient MR acquisition scheme developed in [54], the reconstructed 3D images were used for T_1 estimation of the whole heart, enabling high-resolution 3D T_1 mapping on a 3T scanner. This work, combined to the previous one, answers the questions raised in the introduction, demonstrating how the subspace model can be used to efficiently reconstruct high-dimensional data exhibiting complex motion patterns and dynamic contrast.

However, the works presented here are not without limitations. For example, neither manages to resolve cardiac motion (in a free-running way). In the first work on PET data motion correction, the spatial resolution of the images is one of the bottleneck. In addition, the contrast in the heart (i.e., between the myocardium and the blood pool) was poor due to the 3D excitation slab (without inversion pulse). For that study, the spatial resolution was not a problem as the focus was on the kidneys and the resolution was therefore good enough. Decreasing the spatial resolution was thus not wanted, as it would have implied a decreased SNR and the need of a larger FOV (i.e., larger number of unknowns). Both of these consequences would have made the reconstruction more challenging for the subspace model, let alone considering the additional cardiac motion to resolve in a 3D setting. Note that the temporal resolution was also on the lower end to resolve cardiac motion, and decreasing it would mean further undersampling of the data. In the case of cardiac T_1 mapping, respiratory motion and dynamic contrast were resolved for systole and diastole. Using ECG-triggered data was needed here as free-running 3D cardiac imaging with dynamic contrast would have been too undersampled at the given spatial resolution, and too complex to model for this method with such undersampling ratio, as shown in Figure 3.9a and 3.9b. The linear subspace model then arguably shows limitations with higher undersampling ratios and higher data complexity, which clearly demonstrates the need of a more efficient reconstruction scheme with better modeling of the intrinsic data structure.

This led us to our second objective and main contribution of this PhD thesis: overcoming the limitations of the subspace model and developing a novel dynamic MR image reconstruction scheme relying on non-linear manifold learning techniques (Chapter 3). Due to the "naive" linear dimensionality reduction operated in the subspace model, the intrinsic structure of the data may not always be correctly approximated, especially with high undersampling ratios. To overcome these limitations, we relied on a tangent space alignment technique through a linear transform to efficiently model the non-linear structure of the underlying manifold. The improved performances of the proposed LTSA method were demonstrated (e.g., Figures 3.6 and 3.12) and strong mathematical derivations were given, such as links with other subspace methods (Sections 3.5.1 and 3.5.2). Ongoing efforts to further improve the proposed model are being made, as discussed in Section 3.5.4, and we believe many dynamic MR applications could benefit from this model.

Finally, we presented some preliminary work on deep learning-based image reconstruction in Chapter 4 to answer to our third objective and further broaden our horizon. While more work is still needed to claim any contribution on image reconstruction methods based on DNNs, we gave

a brief overview of the current progress and showed how this translated in the development of new skills and insight. We believe that image reconstruction leveraging the representation power of DNNs is fascinating, shows many possibilities for the future, and is a route the author of this PhD thesis would like to pursue.

5.2 Perspectives

This PhD opens several perspective projects that could be interesting to lead in a post-doctorate position and beyond. First, the development of the LTSA method opened the path to many variants and improvements that could be worth taking time to pursue. For example, the LTSA method could be used for other applications such as speech imaging, or for PET data motion correction similarly as Section 2.2 with the subspace model. There is also a wide outlook on further research to improve the LTSA scheme. A path for improvement is for example the choice of neighborhoods: in the presented work, we used respiratory signals to form temporal neighborhoods and reconstruct their subspaces with joint global coordinates. This choice, although intuitive, was studied in a simulation study (Figure 3.2) and proved to be a good candidate for neighborhood formation: the local Casorati matrices formed by the dynamic images showed low-rank properties (i.e., good low-rank approximation), and the Casorati matrix formed by the local coordinates showed low-rank properties as well (i.e., linear alignment of the tangent space feasible). However, other choices to form neighborhoods exist, and some other ways of forming neighborhoods with a better representation of the underlying manifold most likely exist. One of the perspective works along this line is to find an automated / semi-automated way to form neighborhoods using clustering algorithms (e.g., k-means), or even DNN-based methods.

Another perspective project in the LTSA improvement line is to use spatial patches as discussed in Section 3.5.4. Despite some inconclusive results on extending the current framework to spatial patches, we do believe that further research and trials on that end could eventually lead to greater improvement of the current method, leveraging the spatial locality and redundancy. A final potential lead in the LTSA line that would be interesting pursuing is to extend the current LTSA framework to deep learning. Indeed, in the essence, the LTSA method is an iterative scheme. Then, similarly to other unrolling-based methods, we could try to unroll these iterations and use DNNs for each block. Following that direction, many questions arise. What architecture to use? How will the problem be formulated? What type of data is available for training? Do we include automated neighborhood formation in such a scheme? etc. As we can see, following this route, many paths are possible. These many paths will most likely need "trials and errors", but that is also what makes research interesting. A last advantage of this perspective work for the author of this PhD thesis is that it merges two of the objectives presented in the introduction: developing an MR image reconstruction scheme based on non-linear manifold learning and leveraging the representation power of DNNs.

Finally, other perspectives also present themselves with the skills and scientific open-mindedness

acquired during the complex training that is a PhD program. In this PhD thesis, skills such as optimization, programming, and more globally, image reconstruction and inverse problem resolution were developed. A fascinating realization along the way was that most optimization problems are inverse problems that can be solved with similar means. For example, the subspace model can be used in many common applications, such as compression of video displaying spatiotemporal redundancy (e.g., soccer games). This means that the skills learned to this day are transferable to many fields, which further developed interest in fellow imaging modalities. In the case of PET imaging, similar resolution techniques for inverse problems can be used and one of the main difference is then the forward model in the optimization problem. Research on PET image processing is then also one of outlook after this PhD thesis, and working on advanced processing schemes (e.g., using DNNs) of both PET and MR data on PET/MR scanners would be compelling.

Publications

Journal articles

P1. Marin, T.[a], Djebra, Y.[a], Han, P. K., Chemli, Y., Bloch, I., El Fakhri, G., Ouyang, J., Petibon, Y., & Ma, C. (2020). Motion correction for PET data using subspace-based real-time MR imaging in simultaneous PET/MR. Physics in medicine and biology, 65(23), 235022. https://doi.org/10.1088/1361-6560/abb31d

P2. Han, P. K., Marin, T., Djebra, Y., Landes, V., Zhuo, Y., El Fakhri, G., & Ma, C. (2022). Free-breathing 3D cardiac T1 mapping with transmit B1 correction at 3T. Magnetic resonance in medicine, 87(4), 1832–1845. https://doi.org/10.1002/mrm.29097

P3. Djebra, Y.[a], Marin, T.[a], Han, P. K., Bloch, I., El Fakhri, G., & Ma, C. (2022). Manifold Learning via Linear Tangent Space Alignment (LTSA) for Accelerated Dynamic MRI with Sparse Sampling. IEEE Transactions on Medical Imaging, 42(1), 158-169, doi: 10.1109/TMI.2022.3207774

P4. Ma, C., Han, P. K., Zhuo, Y., Djebra, Y., Marin, T., & Fakhri, G. E. (2022). Joint spectral quantification of MR spectroscopic imaging using linear tangent space alignment-based manifold learning. Magnetic resonance in medicine, 10.1002/mrm.29526. Advance online publication. https://doi.org/10.1002/mrm.29526

Conference abstracts

- Han, P., Djebra, Y., Petibon, Y., Marin, T., Ouyang, J., Fakhri, G. El, & Ma, C. (2019). Accelerated MR-Based Motion Field Measurement for PET Motion Correction in PET/MR. Journal of Nuclear Medicine, 60(supplement 1), 176–176. http://jnm.snmjournals.org/content/60/supplement_1/176
- Djebra, Y., Marin, T., Han, P., Chemli, Y., Bloch, I., El Fakhri, G., Ouyang, J., Petibon, Y., & Ma, C. (2020). MR based PET motion correction for irregular respiratory motion. IEEE Nuclear Science Symposium (NSS) and Medical Imaging Conference (MIC).

[a]Equal contribution

- Marin, T., Djebra, Y., Han, P., Bloch, I., El Fakhri, G., Ouyang, J., Petibon, Y., & Ma, C. (2020). PET motion correction using real-time subspace-based MR imaging. Journal of Nuclear Medicine , 61(supplement 1), 373. http://jnm.snmjournals.org/content/61/supplement_1/373.abstract

- Han, P. K., Djebra, Y., Marin, T., El Fakhri, G., Ouyang, J., & Ma, C. (2020). Low-Rank Reconstruction for Multi-Delay Arterial Spin Labeling. Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM). https://archive.ismrm.org/2020/3296.html

- Djebra, Y., Marin, T., Ouyang, J., El Fakhri, G., Ma, C., & Han, P. K. (2020). Compartmental Low-Rank Denoising for Multi-Delay ASL. Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM). https://archive.ismrm.org/2020/3321.html

- Han, P. K., Marin, T., Vanessa, L., Djebra, Y., El Fakhri, G., & Ma, C. (2021). Optimization of Spoiled GRE-based IR Acquisition Scheme for 3D Cardiac T Mapping at 3T. Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM), 3288. https://archive.ismrm.org/2021/3288.html

- Han, P. K., Marin, T., Djebra, Y., El Fakhri, G., Ouyang, J., & Ma, C. (2021). Free-breathing Renal Perfusion Imaging with Multi-Delay Arterial Spin Labeling Using Subspace-Based Fast MR. Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM). https://archive.ismrm.org/2021/1852.html

- Marin, T., Han, P. K., Zhuo, Y., Djebra, Y., Liu, F., El Fakhri, G., & Ma, C. (2021). Three-dimensional cardiac T1 mapping using subspace and sparsity constrained direct estimation. Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM). https://archive.ismrm.org/2021/3607.html

- Marin, T., Djebra, Y., Han, P. K., Landes, V., Zhuo, Y., Su, K.-H., El Fakhri, G., & Ma, C. (2021). Comparison of deformable registration techniques for real-time MR-based motion correction in PET/MR. Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM). https://archive.ismrm.org/2021/2197.html

- Djebra, Y.[a], Bloch, I., El Fakhri, G., & Ma, C. (2021). Manifold learning via tangent space alignment for accelerated dynamic MR imaging with highly undersampled (k, t)-data. Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM). https://archive.ismrm.org/2021/0069.html

- Djebra, Y., Marin, T., Han, P. K., Bloch, I., El Fakhri, G., & Ma, C. (2022). Free-running Simultaneous 3D Cardiac T1 Mapping and Cine Imaging Using a Linear Tangent Space Alignment Model. Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM). https://archive.ismrm.org/2022/4440.html

[a]Selected for oral presentation and awarded Magna Cum Laude Merit Award

Bibliography

[1] Hemant K Aggarwal, Merry P Mani, and Mathews Jacob. "MoDL: Model-Based Deep Learning Architecture for Inverse Problems". In: *IEEE Transactions on Medical Imaging* 38 (2019), pp. 394–405. ISSN: 1558254X. DOI: 10.1109/TMI.2018.2865356. arXiv: 1712.02862.

[2] Abdul Haseeb Ahmed et al. "Free-Breathing and Ungated Dynamic MRI Using Navigator-Less Spiral SToRM". In: *IEEE Transactions on Medical Imaging* 39 (2020), pp. 3933–3943. ISSN: 0278-0062. DOI: 10.1109/TMI.2020.3008329.

[3] Mehmet Akcakaya et al. "Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging". In: *Magnetic Resonance in Medicine* 81.1 (2019), pp. 439–453. DOI: https://doi.org/10.1002/mrm.27420.

[4] American College of Radiology. *Glossary of MRI Terms*. 2005. URL: https://www.acr.org/-/media/ACR/files/radiology-safety/MR-safety/MRGlossary.pdf.

[5] Amir Beck and Marc Teboulle. "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems". In: *SIAM Journal on Imaging Sciences* 2.1 (2009), pp. 183–202. DOI: 10.1137/080716542.

[6] Mikhail Belkin and Partha Niyogi. "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation". In: *Neural Computation* 15 (2003), pp. 1373–1396. ISSN: 0899-7667. DOI: 10.1162/089976603321780317.

[7] Sampurna Biswas, Hemant K. Aggarwal, and Mathews Jacob. "Dynamic MRI using model-based deep learning and SToRM priors: MoDL-SToRM". In: *Magnetic Resonance in Medicine* 82 (2019), pp. 485–494. ISSN: 0740-3194. DOI: 10.1002/mrm.27706.

[8] Martin Blaimer et al. "SMASH, SENSE, PILS, GRAPPA: how to choose the optimal method". In: *Topics in Magnetic Resonance Imaging* 15.4 (2004), pp. 223–236.

[9] F. Bloch. "Nuclear Induction". In: *Phys. Rev.* 70 (7-8 1946), pp. 460–474. DOI: 10.1103/PhysRev.70.460. URL: https://link.aps.org/doi/10.1103/PhysRev.70.460.

[10] Jorge Zavala Bojorquez et al. "What are normal relaxation times of tissues at 3T?" English. In: *Magnetic Resonance Imaging* 35 (2017), pp. 69–80. ISSN: 0730-725X. DOI: 10.1016/j.mri.2016.08.021.

[11] Stephen Boyd. "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers". In: *Foundations and Trends in Machine Learning* 3 (2010), pp. 1–122. ISSN: 1935-8237. DOI: 10.1561/2200000016.

[12] Stephen Boyd et al. "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers". In: *Foundations and Trends in Machine Learning* 3.1 (2011), pp. 1–122. DOI: 10.1561/2200000016.

[13] J.T. Bushberg and J.M. Boone. *The Essential Physics of Medical Imaging*. Online access: Ovid Technologies LWW Doody's Core Collection 2019. Wolters Kluwer Health, 2011. ISBN: 9780781780575.

[14] Ciprian Catana. "Motion correction options in PET/MRI". English. In: *Seminars in Nuclear Medicine* 45.3 (2015), pp. 212–223. ISSN: 0001-2998. DOI: 10.1053/j.semnuclmed.2015.01.001.

[15] Yuchou Chang, Dong Liang, and Leslie Ying. "Nonlinear GRAPPA: A kernel approach to parallel MRI reconstruction". In: *Magnetic Resonance in Medicine* 68.3 (2012), pp. 730–740.

[16] Xin Chen et al. "High-Resolution Self-Gated Dynamic Abdominal MRI Using Manifold Alignment". In: *IEEE Transactions on Medical Imaging* 36.4 (2017), pp. 960–971. DOI: 10.1109/TMI.2016.2636449.

[17] Yong Chen et al. "Single breath-hold 3D cardiac T1 mapping using through-time spiral GRAPPA". In: *NMR Biomed.* 31.6 (2018), pp. 1–13. ISSN: 10991492. DOI: 10.1002/nbm.3923.

[18] Yutong Chen et al. "AI-Based Reconstruction for Fast MRI—A Systematic Review and Meta-Analysis". In: *Proceedings of the IEEE* 110.2 (2022), pp. 224–245. DOI: 10.1109/JPROC.2022.3141367.

[19] Zang-Hee Cho, Joie P Jones, and Manbir Singh. *Foundations of medical imaging*. Vol. 228. 1993. ISBN: 978-0-471-54573-6.

[20] S. Chopra, R. Hadsell, and Y. LeCun. "Learning a similarity metric discriminatively, with application to face verification". In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. Vol. 1. 2005, pp. 539–546. DOI: 10.1109/CVPR.2005.202.

[21] Anthony G Christodoulou et al. "High-resolution cardiovascular MRI by integrating parallel imaging with low-rankand sparse modeling". In: *IEEE Transactions on Biomedical Engineering* 60 (2013), pp. 3083–3092. ISSN: 00189294. DOI: 10.1109/TBME.2013.2266096.

[22] Anthony G Christodoulou et al. "Improved subspace estimation for low-rank model-based accelerated cardiac imaging". In: *IEEE Transactions on Biomedical Engineering* 61 (2014), pp. 2451–2457. ISSN: 15582531. DOI: 10.1109/TBME.2014.2320463.

[23] Anthony G. Christodoulou et al. "Magnetic resonance multitasking for motion-resolved quantitative cardiovascular imaging". In: *Nature Biomedical Engineering* 2 (2018), pp. 215–226. ISSN: 2157846X. DOI: 10.1038/s41551-018-0217-y.

[24] Se Young Chun and Jeffrey A. Fessler. "A simple regularizer for B-spline nonrigid image registration that encourages local invertibility". In: *IEEE Journal of Selected Topics in Signal Processing* 3.1 (2009), pp. 159–169.

[25] Elizabeth K. Cole et al. *Unsupervised MRI Reconstruction with Generative Adversarial Networks*. 2020. DOI: 10.48550/ARXIV.2008.13065. URL: https://arxiv.org/abs/2008.13065.

[26] Gastão Cruz et al. "3D free-breathing cardiac magnetic resonance fingerprinting". In: *NMR Biomed.* 33.10 (2020), pp. 1–16. ISSN: 10991492. DOI: 10.1002/nbm.4370.

[27] Mohammad Dawood et al. "Respiratory motion correction in 3-D PET data with advanced optical flow algorithms". English. In: *IEEE Transactions on Medical Imaging* 27.8 (2008), pp. 1164–1175. ISSN: 0278-0062. DOI: 10.1109/TMI.2008.918321.

[28] Anagha Deshmane et al. "Parallel MR imaging". In: *Journal of Magnetic Resonance Imaging* 36.1 (2012), pp. 55–72. DOI: https://doi.org/10.1002/jmri.23639.

[29] Allen D Elster. *Questions and Answers in MRI*. 2022. URL: https://mriquestions.com/index.html.

[30] Taejoon Eo et al. "KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images". In: *Magnetic Resonance in Medicine* 80.5 (2018), pp. 2188–2201. DOI: https://doi.org/10.1002/mrm.27201.

[31] Li Feng et al. "XD-GRASP: Golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing". English. In: *Magnetic Resonance in Medicine* 75.2 (2016), pp. 775–788. ISSN: 0740-3194. DOI: 10.1002/mrm.25665.

[32] Jeffrey A. Fessler and Bradley P. Sutton. "Nonuniform fast Fourier transforms using min-max interpolation". In: *IEEE Transactions on Signal Processing* 51.2 (2003), pp. 560–574. ISSN: 1053-587x. DOI: 10.1109/TSP.2002.807005.

[33] Allan J. Freeman, Penny A. Gowland, and P. Mansfield. "Optimization of the ultrafast look-locker echo-planar imaging T1 mapping sequence". In: *Magn. Reson. Imaging* 16.7 (1998), pp. 765–772. ISSN: 0730725X. DOI: 10.1016/S0730-725X(98)00011-3.

[34] Kunihiko Fukushima, Sei Miyake, and Takayuki Ito. "Neocognitron: A neural network model for a mechanism of visual pattern recognition". In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC-13.5 (1983), pp. 826–834. DOI: 10.1109/TSMC.1983.6313076.

[35] Roger R. Fulton et al. "Correction for head movements in positron emission tomography using an optical motion-tracking system". In: *IEEE Transactions on Nuclear Science* 49.1 (2002), pp. 116–123. ISSN: 0018-9499. DOI: 10.1109/TNS.2002.998691.

[36] Sairam Geethanath et al. "Compressed Sensing MRI: A Review". In: *Critical Reviews in Biomedical Engineering* 41.3 (2013), pp. 183–204. ISSN: 0278-940X.

[37] Ashley Gillman et al. "PET motion correction in context of integrated PET/MR: Current techniques, limitations, and future projections". English. In: *Medical Physics* 44.12 (2017), e430–e445. ISSN: 0094-2405. DOI: 10.1002/mp.12577.

[38] James F Glockner et al. "Parallel MR imaging: a user's guide". In: *Radiographics* 25.5 (2005), pp. 1279–1297.

[39] F Godenschweger et al. "Motion correction in MRI of the brain". In: *Physics in Medicine & Biology* 61.5 (2016), R32. DOI: 10.1088/0031-9155/61/5/R32. URL: https://dx.doi.org/10.1088/0031-9155/61/5/R32.

[40] Ian Goodfellow et al. "Generative Adversarial Networks". In: *Commun. ACM* 63.11 (2020), 139–144. ISSN: 0001-0782. DOI: 10.1145/3422622.

[41] Leslie Greengard and June-Yub Lee. "Accelerating the Nonuniform Fast Fourier Transform". In: *SIAM Review* 46.3 (2004), pp. 443–454. DOI: 10.1137/S003614450343200X. URL: https://doi.org/10.1137/S003614450343200X.

[42] Robert Grimm et al. "Self-gated MRI motion modeling for respiratory motion compensation in integrated PET/MRI". English. In: *Medical Image Analysis* 19.1 (2015), pp. 110–120. ISSN: 1361-8415. DOI: 10.1016/j.media.2014.08.003.

[43] Mark A. Griswold et al. "Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA)". In: *Magnetic Resonance in Medicine* 47 (2002), pp. 1202–1210. ISSN: 07403194. DOI: 10.1002/mrm.10171.

[44] HáKon Gudbjartsson and Samuel Patz. "The Rician distribution of noisy MRI data". In: *Magnetic Resonance in Medicine* 34.6 (1995), pp. 910–914. DOI: https://doi.org/10.1002/mrm.1910340618. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.1910340618. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1910340618.

[45] Rui Guo et al. "A three-dimensional free-breathing sequence for simultaneous myocardial T1 and T2 mapping". In: *Magn. Reson. Med.* 81.2 (2019), pp. 1031–1043. ISSN: 15222594. DOI: 10.1002/mrm.27466.

[46] Rui Guo et al. "Free-breathing whole-heart multi-slice myocardial T1 mapping in 2 minutes". In: *Magn. Reson. Med.* 85.1 (2021), pp. 89–102. ISSN: 15222594. DOI: 10.1002/mrm.28402.

[47] Philip Haaf et al. "Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: A comprehensive review". In: *J. Cardiovasc. Magn. Reson.* 18.1 (2016), pp. 1–12. ISSN: 1532429X. DOI: 10.1186/s12968-016-0308-4.

[48] Jihun Ham et al. "A kernel view of the dimensionality reduction of manifolds". In: *Proceedings, Twenty-First International Conference on Machine Learning, ICML 2004*. Vol. 12. New York, New York, USA: ACM Press, 2004, pp. 369–376. ISBN: 1581138385. DOI: 10.1145/1015330.1015417.

[49] Jesse I. Hamilton et al. "Cardiac cine magnetic resonance fingerprinting for combined ejection fraction, T1 and T2 quantification". In: *NMR Biomed.* 33.8 (2020), pp. 1–17. ISSN: 10991492. DOI: 10.1002/nbm.4323.

[50] Kerstin Hammernik et al. "Learning a variational network for reconstruction of accelerated MRI data". In: *Magnetic Resonance in Medicine* 79 (2018), pp. 3055–3071. ISSN: 15222594. DOI: 10.1002/mrm.26977. arXiv: 1704.00447.

[51] Paul K. Han et al. "Free-Breathing 3D T1 Mapping of the Whole-Heart Using Low-Rank Tensor Modeling". In: *Proc. ISMRM*. Montreal, 2019, p. 4396.

[52] Paul K. Han et al. "Free-Breathing Three-Dimensional T1 Mapping of the Heart Using Subspace-Based Data Acquisition and Image Reconstruction". In: *41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* Berlin, 2019, p. 2496.

[53] Paul K. Han et al. "Optimization of Spoiled GRE-based IR Acquisition Scheme for 3D Cardiac T Mapping at 3T". In: *Proc. ISMRM*. Virtual, 2021, p. 3288.

[54] Paul Kyu Han et al. "Free-breathing 3D cardiac T1 mapping with transmit B1 correction at 3T". In: *Magnetic Resonance in Medicine* 87.4 (2022), pp. 1832–1845. DOI: https://doi.org/10.1002/mrm.29097. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.29097. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.29097.

[55] Jingfei He et al. "Accelerated High-Dimensional MR Imaging with Sparse Sampling Using Low-Rank Tensors". In: *IEEE Transactions on Medical Imaging* 35 (2016), pp. 2119–2129. ISSN: 1558254X. DOI: 10.1109/TMI.2016.2550204.

[56] Robin M Heidemann et al. "Direct parallel image reconstructions for spiral trajectories using GRAPPA". In: *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 56.2 (2006), pp. 317–326.

[57] Geoffrey E. Hinton et al. "Improving neural networks by preventing co-adaptation of feature detectors". In: *CoRR* abs/1207.0580 (2012). arXiv: 1207.0580. URL: http://arxiv.org/abs/1207.0580.

[58] John J Hopfield. "Neural networks and physical systems with emergent collective computational abilities". In: *Proceedings of the National Academy of Sciences* 79.8 (1982), pp. 2554–2558.

[59] Chuan Huang et al. "Motion compensation for brain PET imaging using wireless MR active markers in simultaneous PET-MR: phantom and non-human primate studies". English. In: *NeuroImage* 91 (2014), pp. 129–137. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2013.12.061.

[60] Qiaoying Huang et al. "Dynamic MRI reconstruction with end-to-end motion-guided network". In: *Medical Image Analysis* 68 (2021), p. 101901. ISSN: 13618423. DOI: 10.1016/j.media.2020.101901.

[61] H. Malcom Hudson and Richard S. Larkin. "Accelerated image reconstruction using ordered subsets of projection data". In: *IEEE Transactions on Medical Imaging* 13.4 (1994), pp. 601–609. DOI: 10.1109/42.363108.

[62] Edward F Jackson et al. "A review of MRI pulse sequences and techniques in neuroimaging". In: *Surgical Neurology* 47.2 (1997), pp. 185–199.

[63] Jihye Jang et al. "Comparison of spoiled gradient echo and steady-state free-precession imaging for native myocardial T1 mapping using the slice-interleaved T1 mapping (STONE) sequence". In: *NMR Biomed.* 29.10 (2016), pp. 1486–1496. ISSN: 10991492. DOI: 10.1002/nbm.3598.

[64] Oren N Jaspan, Roman Fleysher, and Michael L Lipton. "Compressed sensing MRI: a review of the clinical literature". In: *The British journal of radiology* 88.1056 (2015), p. 20150487.

[65] O. Jaubert et al. "Free-running cardiac magnetic resonance fingerprinting: Joint T1/T2 map and Cine imaging". In: *Magn. Reson. Imaging* 68.September 2019 (2020), pp. 173–182. ISSN: 18735894. DOI: 10.1016/j.mri.2020.02.005.

[66] Xiao Jin et al. "Evaluation of motion correction methods in human brain PET imaging–a simulation study based on human motion data". English. In: *Medical Physics* 40.10 (2013), p. 102503. ISSN: 0094-2405. DOI: 10.1118/1.4819820.

[67] Ziwen Ke et al. "Deep Manifold Learning for Dynamic MR Imaging". In: *IEEE Transactions on Computational Imaging* 7 (2021), pp. 1314–1327. ISSN: 23339403. DOI: 10.1109/TCI.2021.3131564. arXiv: 2104.01102.

[68] Peter Kellman and Michael S Hansen. "T1-mapping in the heart: accuracy and precision". In: *J. Cardiovasc. Magn. Reson.* 16.2 (2014), pp. 1–20.

[69] Adam Leon Kesner et al. "Respiratory Gated PET Derived in a Fully Automated Manner From Raw PET Data". In: *IEEE Transactions on Nuclear Science* 56.3 (2009), pp. 677–686. ISSN: 0018-9499. DOI: 10.1109/TNS.2009.2016341.

[70] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2014. DOI: 10.48550/ARXIV.1412.6980.

[71] Florian Knoll et al. "fastMRI: A Publicly Available Raw k-Space and DICOM Dataset of Knee Images for Accelerated MR Image Reconstruction Using Machine Learning". In: *Radiology: Artificial Intelligence* 2.1 (2020), e190007. DOI: 10.1148/ryai.2020190007.

[72] Glenn F Knoll. *Radiation detection and measurement*. John Wiley & Sons, 2010.

[73] Thomas Küstner et al. "MR-based respiratory and cardiac motion correction for PET imaging". English. In: *Medical Image Analysis* 42 (2017), pp. 129–144. ISSN: 1361-8415. DOI: 10.1016/j.media.2017.08.002.

[74] Fan Lam and Zhi-Pei Liang. "A subspace approach to high-resolution spectroscopic imaging". In: *Magnetic Resonance in Medicine* 71 (2014), pp. 1349–1357. ISSN: 07403194. DOI: 10.1002/mrm.25168.

[75] Fan Lam et al. "High-resolution 1 H-MRSI of the brain using SPICE: Data acquisition and image reconstruction". In: *Magnetic Resonance in Medicine* 76 (2016), pp. 1059–1070. ISSN: 07403194. DOI: 10.1002/mrm.26019.

[76] Y. LeCun et al. "Backpropagation Applied to Handwritten Zip Code Recognition". In: *Neural Computation* 1.4 (1989), pp. 541–551. DOI: 10.1162/neco.1989.1.4.541.

[77] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". In: *nature* 521.7553 (2015), pp. 436–444.

[78] Dongwook Lee et al. "Deep residual learning for accelerated MRI using magnitude and phase networks". In: *IEEE Transactions on Biomedical Engineering* 65.9 (2018), pp. 1985–1995. ISSN: 15582531. DOI: 10.1109/TBME.2018.2821699.

[79] Dong Liang et al. "Deep Magnetic Resonance Image Reconstruction: Inverse Problems Meet Neural Networks". In: *IEEE Signal Processing Magazine* 37.1 (2020), pp. 141–151. ISSN: 15580792. DOI: 10.1109/MSP.2019.2950557.

[80] Zhi-Pei Liang. "Spatiotemporal Imaging with Partially Separable Functions". In: *IEEE International Symposium on Biomedical Imaging*. 2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro. 2007, pp. 988–991. DOI: 10.1109/Isbi.2007.357020.

[81] Zhi Pei Liang. "Spatiotemporal imaging with partially separable functions". In: *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro - Proceedings*. Vol. 2. IEEE, 2007, pp. 988–991. ISBN: 1424406722. DOI: 10.1109/ISBI.2007.357020.

[82] Z.P. Liang and P.C. Lauterbur. *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. IEEE Press series in biomedical engineering. SPIE Optical Engineering Press, 2000. ISBN: 9780819435163.

[83] Sajan Goud Lingala et al. "Accelerated Dynamic MRI Exploiting Sparsity and Low-Rank Structure: k-t SLR". In: *IEEE Transactions on Medical Imaging* 30 (2011), pp. 1042–1054. ISSN: 0278-0062. DOI: 10.1109/TMI.2010.2100850.

[84] Chi Liu, Adam M. Alessio, and Paul E. Kinahan. "Respiratory motion correction for quantitative PET/CT using all detected events with internal-external motion correlation". English. In: *Medical Physics* 38.5 (2011), pp. 2715–2723. ISSN: 0094-2405. DOI: 10.1118/1.3582692.

[85] Chi Liu et al. "The impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging". English. In: *Physics in Medicine and Biology* 54.24 (2009), pp. 7345–7362. ISSN: 0031-9155. DOI: 10.1088/0031-9155/54/24/007.

[86] Yihuan Lu et al. "Data-driven voluntary body motion detection and non-rigid event-by-event correction for static and dynamic PET". English. In: *Physics in Medicine and Biology* 64.6 (2019), p. 065002. ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab02c2.

[87] Alexander Selvikvag Lundervold and Arvid Lundervold. "An overview of deep learning in medical imaging focusing on MRI". In: *Zeitschrift für Medizinische Physik* 29.2 (2019). Special Issue: Deep Learning in Medical Physics, pp. 102–127. ISSN: 0939-3889. DOI: https://doi.org/10.1016/j.zemedi.2018.11.002.

[88] Michael Lustig, David Donoho, and John M. Pauly. "Sparse MRI: The application of compressed sensing for rapid MR imaging". In: *Magnetic Resonance in Medicine* 58 (2007), pp. 1182–1195. ISSN: 07403194. DOI: 10.1002/mrm.21391.

[89] Michael Lustig et al. "Compressed Sensing MRI". In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 72–82. DOI: 10.1109/MSP.2007.914728.

[90] Jun Lv, Jin Zhu, and Guang Yang. "Which GAN? A comparative study of generative adversarial network-based fast MRI reconstruction". In: *Philosophical Transactions of the Royal Society A* 379.2200 (2021), p. 20200203.

[91] Chao Ma and Georges El Fakhri. "MRSI spectral quantification using linear tangent space alignment (LTSA)-based manifold learning". In: *Proc Intl Soc Mag Reson Med.* 2022, p. 243.

[92] Chao Ma et al. "High-resolution dynamic 31 P-MRSI using a low-rank tensor model". In: *Magnetic Resonance in Medicine* 78 (2017), pp. 419–428. ISSN: 07403194. DOI: 10.1002/mrm.26762.

[93] Chao Ma et al. "Joint spectral quantification of MR spectroscopic imaging using linear tangent space alignment-based manifold learning". In: *Magnetic resonance in medicine* (2022). ISSN: 0740-3194. DOI: 10.1002/mrm.29526. URL: https://doi.org/10.1002/mrm.29526.

[94] Morteza Mardani et al. "Deep Generative Adversarial Neural Networks for Compressive Sensing MRI". In: *IEEE Transactions on Medical Imaging* 38.1 (2019), pp. 167–179. DOI: 10.1109/TMI.2018.2858752.

[95] Thibault Marin et al. "Motion correction for PET data using subspace-based real-time MR imaging in simultaneous PET/MR". In: *Physics in Medicine & Biology* 65.23 (2020), p. 235022.

[96] Warren S McCulloch and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity". In: *The bulletin of mathematical biophysics* 5.4 (1943), pp. 115–133.

[97] Daniel R. Messroghli et al. "Modified look-locker inversion recovery (MOLLI) for high-resolution T 1 mapping of the heart". In: *Magn. Reson. Med.* 52.1 (2004), pp. 141–146. ISSN: 07403194. DOI: 10.1002/mrm.20110.

[98] Daniel R. Messroghli et al. "Clinical recommendations for cardiovascular magnetic resonance mapping of T1, T2, T2 and extracellular volume: A consensus statement by the Society for Cardiovascular Magnetic Resonance (SCMR) endorsed by the European Association for Cardiovascular Imagin". In: *J. Cardiovasc. Magn. Reson.* 19.1 (2017), pp. 1–24. ISSN: 1532429X. DOI: 10.1186/s12968-017-0389-8.

[99] Reuben Mezrich. "A perspective on k-space". In: *Radiology* 195.2 (1995), pp. 297–316.

[100] Mario Milicevic et al. "Application of Deep Learning Architectures for Accurate Detection of Olive Tree Flowering Phenophase". In: *Remote Sensing* 12.13 (2020). ISSN: 2072-4292. DOI: 10.3390/rs12132120.

[101] Giorgia Milotta et al. "3D whole-heart isotropic-resolution motion-compensated joint T1/T2 mapping and water/fat imaging". In: *Magn. Reson. Med.* 84.6 (2020), pp. 3009–3026. ISSN: 15222594. DOI: 10.1002/mrm.28330.

[102] Yasir Q Mohsin, Sunrita Poddar, and Mathews Jacob. "Free-Breathing Ungated Cardiac MRI Using Iterative SToRM (i-SToRM)". In: *IEEE Transactions on Medical Imaging* 38 (2019), pp. 2303–2313. ISSN: 1558254X. DOI: 10.1109/TMI.2019.2908140.

[103] Andrew J. Montgomery et al. "Correction of head movement on PET studies: comparison of methods". English. In: *Journal of Nuclear Medicine* 47.12 (2006), pp. 1936–1944. ISSN: 0161-5505.

[104] Camila Munoz et al. "Motion-corrected simultaneous cardiac positron emission tomography and coronary MR angiography with high acquisition efficiency". English. In: *Magnetic Resonance in Medicine* 79.1 (2018), pp. 339–350. ISSN: 0740-3194. DOI: 10.1002/mrm.26690.

[105] Ukash Nakarmi et al. "A Kernel-Based Low-Rank (KLR) Model for Low-Dimensional Manifold Recovery in Highly Accelerated Dynamic MRI". In: *IEEE Transactions on Medical Imaging* 36 (2017), pp. 2297–2307. ISSN: 1558254X. DOI: 10.1109/TMI.2017.2723871.

[106] Giovanna Nordio et al. "3D myocardial T1 mapping using saturation recovery". In: *J. Magn. Reson. Imaging* 46.1 (2017), pp. 218–227. ISSN: 15222586. DOI: 10.1002/jmri.25575.

[107] Ryosuke Okuta et al. "CuPy: A NumPy-compatible library for NVIDIA GPU calculations". In: *Workshop on Machine Learning Systems (LearningSys) at Neural Information Processing Systems (NIPS)*. Vol. 6. 2017, pp. 1–7.

[108] Frank Ong and Michael Lustig. "SigPy: A Python Package for High Performance Iterative Reconstruction". In: *Proc Intl Soc Mag Reson Med, Montréal, QC*. 2019.

[109] Jinsong Ouyang, Quanzheng Li, and Georges El Fakhri. "Magnetic resonance-based motion correction for positron emission tomography imaging". eng. In: *Seminars in Nuclear Medicine* 43.1 (2013), pp. 60–67. ISSN: 0001-2998. DOI: 10.1053/j.semnuclmed.2012.08.007.

[110] J.A. Parker. *Image Reconstruction in Radiology*. Taylor & Francis, 2018. ISBN: 9781351081818.

[111] Cynthia B. Paschal and H. Douglas Morris. "K-space in the clinic". In: *Journal of Magnetic Resonance Imaging* 19.2 (2004), pp. 145–159. DOI: https://doi.org/10.1002/jmri.10451. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jmri.10451. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.10451.

[112] Josh Patterson and Adam Gibson. *Deep learning: A practitioner's approach*. " O'Reilly Media, Inc.", 2017.

[113] John Pauly. *Non-Cartesian Reconstruction*. 2007. URL: https://ece-classes.usc.edu/ee591/library/Pauly-Non-Cartesian%20recon.pdf.

[114] Henrik Pedersen et al. "k-t PCA: Temporally constrained k-t BLAST reconstruction using principal component analysis". In: *Magnetic Resonance in Medicine* 62 (2009), pp. 706–716. ISSN: 07403194. DOI: 10.1002/mrm.22052.

[115] Yoann Petibon et al. "Cardiac motion compensation and resolution modeling in simultaneous PET-MR: a cardiac lesion detection study". English. In: *Physics in Medicine and Biology* 58.7 (2013), pp. 2085–2102. ISSN: 0031-9155. DOI: 10.1088/0031-9155/58/7/2085.

[116] Yoann Petibon et al. "Impact of motion and partial volume effects correction on PET myocardial perfusion imaging using simultaneous PET-MR". English. In: *Physics in Medicine and Biology* 62.2 (2016), pp. 326–343. ISSN: 0031-9155. DOI: 10.1088/1361-6560/aa5087.

[117] Yoann Petibon et al. "MR-based cardiac and respiratory motion correction of PET: application to static and dynamic cardiac (18)F-FDG imaging". English. In: *Physics in Medicine and Biology* 64.19 (2019), p. 195009. ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab39c2.

[118] Sunrita Poddar and Mathews Jacob. "Dynamic MRI Using SmooThness Regularization on Manifolds (SToRM)". In: *IEEE Transactions on Medical Imaging* 35 (2016), pp. 1106–1115. ISSN: 1558254X. DOI: 10.1109/TMI.2015.2509245.

[119] Sunrita Poddar et al. "Manifold Recovery Using Kernel Low-Rank Regularization: Application to Dynamic Imaging". In: *IEEE Transactions on Computational Imaging* 5 (2019), pp. 478–491. ISSN: 2573-0436. DOI: 10.1109/tci.2019.2893598.

[120] Robert A Pooley. "Fundamental physics of MR imaging". In: *Radiographics* 25.4 (2005), pp. 1087–1099.

[121] Klaas P. Pruessmann et al. "SENSE: Sensitivity encoding for fast MRI". In: *Magnetic Resonance in Medicine* 42 (1999), pp. 952–962. ISSN: 07403194. DOI: 10.1002/(sici)1522-2594(199911)42:5<952::aid-mrm16>3.3.co;2-j.

[122] Klaas P. Pruessmann et al. "Advances in sensitivity encoding with arbitrary k-space trajectories". In: *Magnetic Resonance in Medicine* 46.4 (2001), pp. 638–651. DOI: https://doi.org/10.1002/mrm.1241. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1241.

[123] Haikun Qi et al. "Free-running 3D whole heart myocardial T1 mapping with isotropic spatial resolution". In: *Magn. Reson. Med.* 82.4 (2019), pp. 1331–1342. ISSN: 15222594. DOI: 10.1002/mrm.27811.

[124] Chen Qin et al. "Convolutional recurrent neural networks for dynamic MR image reconstruction". In: *IEEE Transactions on Medical Imaging* 38 (2019), pp. 280–290. ISSN: 1558254X. DOI: 10.1109/TMI.2018.2863670. arXiv: 1712.01751.

[125] Peng Qu et al. "An improved iterative SENSE reconstruction method". In: *Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering: An Educational Journal* 31.1 (2007), pp. 44–50.

[126] Arman Rahmim, Jing Tang, and Habib Zaidi. "Four-Dimensional Image Reconstruction Strategies in Cardiac-Gated and Respiratory-Gated PET Imaging". English. In: *PET Clinics* 8.1 (2013), pp. 51–67. ISSN: 1556-8598. DOI: 10.1016/j.cpet.2012.10.005.

[127] Christopher M. Rank et al. "Respiratory motion compensation for simultaneous PET/MR based on highly undersampled MR data". English. In: *Medical Physics* 43.12 (2016), p. 6234. ISSN: 0094-2405. DOI: 10.1118/1.4966128.

[128] Saiprasad Ravishankar and Yoram Bresler. "MR image reconstruction from highly undersampled k-space data by dictionary learning". In: *IEEE Transactions on Medical Imaging* 30.5 (2011), pp. 1028–1041. ISSN: 02780062. DOI: 10.1109/TMI.2010.2090538.

[129] Mengye Ren et al. "Learning to Reweight Examples for Robust Deep Learning". In: *CoRR* abs/1803.09050 (2018). arXiv: 1803.09050. URL: http://arxiv.org/abs/1803.09050.

[130] Philip M. Robson et al. "Correction of respiratory and cardiac motion in cardiac PET/MR using MR-based motion modeling". English. In: *Physics in Medicine and Biology* 63.22 (2018), p. 225011. ISSN: 0031-9155. DOI: 10.1088/1361-6560/aaea97.

[131] Frank Rosenblatt. "The perceptron: a probabilistic model for information storage and organization in the brain". In: *Psychological review* 65.6 (1958), p. 386.

[132] Sam T. Roweis and Lawrence K. Saul. "Nonlinear Dimensionality Reduction by Locally Linear Embedding". In: *Science* 290 (2000), pp. 2323–2326. ISSN: 0036-8075. DOI: 10.1126/science.290.5500.2323.

[133] Mathieu Rubeaux et al. "Enhancing Cardiac PET by Motion Correction Techniques". English. In: *Current Cardiology Reports* 19.2 (2017), p. 14. ISSN: 1523-3782. DOI: 10.1007/s11886-017-0825-2.

[134] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. "Learning representations by back-propagating errors". In: *Nature* 323.6088 (1986), pp. 533–536.

[135] David E Rumelhart, James L McClelland, PDP Research Group, et al. *Parallel distributed processing*. Vol. 1. IEEE New York, 1988.

[136] Erik B. Schelbert and Daniel R. Messroghli. "State of the art: Clinical applications of cardiac T1 mapping". In: *Radiology* 278.3 (2016), pp. 658–676. ISSN: 15271315. DOI: 10.1148/radiol.2016141802.

[137] Jo Schlemper et al. "A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction". In: *IEEE Transactions on Medical Imaging* 37 (2018), pp. 491–503. ISSN: 1558254X. DOI: 10.1109/TMI.2017.2760978. arXiv: 1704.02422.

[138] W. P. Segars et al. "4D XCAT phantom for multimodality imaging research". In: *Medical Physics* 37 (2010), pp. 4902–4915. ISSN: 00942405. DOI: 10.1118/1.3480985.

[139] William Paul Segars et al. "4D XCAT phantom for multimodality imaging research". English. In: *Medical Physics* 37.9 (2010), pp. 4902–4915. ISSN: 0094-2405. DOI: 10.1118/1.3480985.

[140] Nicole Seiberlich et al. "Improved radial GRAPPA calibration for real-time free-breathing cardiac imaging". In: *Magnetic resonance in medicine* 65.2 (2011), pp. 492–505.

[141] Fardad Michael Serry et al. "Dual flip-angle IR-FLASH with spin history mapping for B1+ corrected T1 mapping: Application to T1 cardiovascular magnetic resonance multitasking". In: *Magnetic Resonance in Medicine* 86.6 (2021), pp. 3182–3191. DOI: https://doi.org/10.1002/mrm.28935.

[142] Jaime L. Shaw et al. "Free-breathing, non-ECG, continuous myocardial T1 mapping with cardiovascular magnetic resonance multitasking". In: *Magn. Reson. Med.* 81.4 (2019), pp. 2450–2463. ISSN: 15222594. DOI: 10.1002/mrm.27574.

[143] Gaurav Nagesh Shetty et al. "Bi-Linear Modeling of Data Manifolds for Dynamic-MRI Recovery". In: *IEEE Transactions on Medical Imaging* 39 (2020), pp. 688–702. ISSN: 0278-0062. DOI: 10.1109/TMI.2019.2934125. arXiv: 1812.10617.

[144] K.K. Shung, M.B. Smith, and B.M.W. Tsui. *Principles of Medical Imaging*. Elsevier Science, 2012. ISBN: 9780323139939.

[145] Robert L. Siddon. "Fast calculation of the exact radiological path for a three-dimensional CT array". In: *Medical Physics* 12.2 (1985), pp. 252–255.

[146] Daniel K. Sodickson and Warren J. Manning. "Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays". In: *Magnetic Resonance in Medicine* 38 (1997), pp. 591–603. ISSN: 07403194. DOI: 10.1002/mrm.1910380414.

[147] Tao Sun et al. "Body motion detection and correction in cardiac PET: Phantom and human studies". English. In: *Medical Physics* 46.11 (2019), pp. 4898–4906. ISSN: 0094-2405. DOI: 10.1002/mp.13815.

[148] Kyunghyun Sung and Krishna S. Nayak. "Measurement and characterization of RF nonuniformity over the heart at 3T using body coil transmission". In: *J. Magn. Reson. Imaging* 27.3 (2008), pp. 643–648. ISSN: 10531807. DOI: 10.1002/jmri.21253.

[149] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. "A Global Geometric Framework for Nonlinear Dimensionality Reduction". In: *Science* 290 (2000), pp. 2319–2323. ISSN: 0036-8075. DOI: 10.1126/science.290.5500.2319.

[150] Kerem C. Tezcan et al. "MR Image Reconstruction Using Deep Density Priors". In: *IEEE Transactions on Medical Imaging* 38.7 (2019), pp. 1633–1642. ISSN: 1558254X. DOI: 10.1109/TMI.2018.2887072. arXiv: 1711.11386.

[151] J. Trzasko and A. Manduca. "Local versus Global Low-Rank Promotion in Dynamic MRI Series Reconstruction". In: *Proc Intl Soc Mag Reson Med*. Vol. 24. 2011, p. 4371.

[152] Joshua D. Trzasko. "Exploiting local low-rank structure in higher-dimensional MRI applications". In: *Wavelets and Sparsity XV*. Ed. by Dimitri Van De Ville, Vivek K. Goyal, and Manos Papadakis. Vol. 8858. International Society for Optics and Photonics. SPIE, 2013, pp. 551 –558. DOI: 10.1117/12.2027059.

[153] Muhammad Usman et al. "Motion corrected compressed sensing for free-breathing dynamic cardiac MRI". In: *Magnetic Resonance in Medicine* 70.2 (2013), pp. 504–516. DOI: https://doi.org/10.1002/mrm.24463.

[154] Muhammad Usman et al. "Manifold learning based ECG-free free-breathing cardiac CINE MRI". In: *Journal of Magnetic Resonance Imaging* 41.6 (2015), pp. 1521–1527. DOI: https://doi.org/10.1002/jmri.24731.

[155] Li Wan et al. "Regularization of Neural Networks using DropConnect". In: *Proceedings of the 30th International Conference on Machine Learning*. Ed. by Sanjoy Dasgupta and David McAllester. Vol. 28. Proceedings of Machine Learning Research 3. Atlanta, Georgia, USA: PMLR, 2013, pp. 1058–1066. URL: https://proceedings.mlr.press/v28/wan13.html.

[156] Shanshan Wang et al. "Accelerating magnetic resonance imaging via deep learning". In: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. Vol. 36. IEEE, 2016, pp. 514–517. ISBN: 978-1-4799-2349-6. DOI: 10.1109/ISBI.2016.7493320.

[157] Shanshan Wang et al. "DIMENSION: Dynamic MR imaging with both k-space and spatial prior knowledge obtained via multi-supervised network training". In: *NMR in Biomedicine* (2019), pp. 1–16. ISSN: 10991492. DOI: 10.1002/nbm.4131. arXiv: 1810.00302.

[158] Sebastian Weingärtner et al. "Free-breathing combined three-dimensional phase sensitive late gadolinium enhancement and T1 mapping for myocardial tissue characterization". In: *Magn. Reson. Med.* 74.4 (2015), pp. 1032–1041. ISSN: 15222594. DOI: 10.1002/mrm.25495.

[159] Sebastian Weingärtner et al. "Free-breathing multislice native myocardial T1 mapping using the slice-interleaved T1 (STONE) sequence". In: *Magn. Reson. Med.* 74.1 (2015), pp. 115–124. ISSN: 15222594. DOI: 10.1002/mrm.25387.

[160] Sebastian Weingärtner et al. "Simultaneous multislice imaging for native myocardial T1 mapping: Improved spatial coverage in a single breath-hold". In: *Magn. Reson. Med.* 78.2 (2017), pp. 462–471. ISSN: 15222594. DOI: 10.1002/mrm.26770.

[161] Alexander Werling et al. "Fast implementation of the single scatter simulation algorithm and its use in iterative image reconstruction of PET data". In: *Physics in Medicine and Biology* 47.16 (2002), pp. 2947–2960. ISSN: 0031-9155. DOI: 10.1088/0031-9155/47/16/310.

[162] *What is machine learning?* URL: https://www.ibm.com/topics/machine-learning.

[163] Nathan White et al. "PROMO: Real-time prospective motion correction in MRI using image-based tracking". In: *Magnetic Resonance in Medicine* 63.1 (2010), pp. 91–105. DOI: https://doi.org/10.1002/mrm.22176.

[164] Katherine L. Wright et al. "Non-Cartesian parallel imaging reconstruction". In: *Journal of Magnetic Resonance Imaging* 40.5 (2014), pp. 1022–1040. DOI: https://doi.org/10.1002/jmri.24521. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.24521.

[165] Xiao-Long Wu et al. "Impatient MRI: Illinois Massively Parallel Acceleration Toolkit for image reconstruction with enhanced throughput in MRI". In: *IEEE International Symposium on Biomedical Imaging*. IEEE, 2011, pp. 69–72. DOI: 10.1109/ISBI.2011.5872356.

[166] Guang Yang et al. "DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction". In: *IEEE Transactions on Medical Imaging* 37 (2018), pp. 1310–1321. ISSN: 1558254X. DOI: 10.1109/TMI.2017.2785879.

[167] Jaejun Yoo et al. "Time-Dependent Deep Image Prior for Dynamic MRI". In: *IEEE Transactions on Medical Imaging* 40.12 (2021), pp. 3337–3348. ISSN: 1558254X. DOI: 10.1109/TMI.2021.3084288. arXiv: 1910.01684.

[168] Yunhan Yu et al. "Event-by-Event Continuous Respiratory Motion Correction for Dynamic PET Imaging". English. In: *Journal of Nuclear Medicine* 57.7 (2016), pp. 1084–1090. ISSN: 0161-5505. DOI: 10.2967/jnumed.115.167676.

[169] Jure Zbontar et al. "fastMRI: An Open Dataset and Benchmarks for Accelerated MRI". In: *CoRR* abs/1811.08839 (2018). arXiv: 1811.08839. URL: http://arxiv.org/abs/1811.08839.

[170] Zhenyue Zhang and Hongyuan Zha. "Principal Manifolds and Nonlinear Dimensionality Reduction via Tangent Space Alignment". In: *SIAM Journal on Scientific Computing* 26 (2004), pp. 313–338. ISSN: 1064-8275. DOI: 10.1137/S1064827502419154.

[171] Bo Zhao et al. "Image reconstruction from highly undersampled (k ,t)-space data with joint partial separability and sparsity constraints". In: *IEEE Transactions on Medical Imaging* 31 (2012), pp. 1809–1820. ISSN: 0278-0062. DOI: 10.1109/TMI.2012.2203921.

[172] Ruixi Zhou et al. "Dual-excitation flip-angle simultaneous cine and T1 mapping using spiral acquisition with respiratory and cardiac self-gating". In: *Magn. Reson. Med.* 86.1 (2021), pp. 82–96. ISSN: 15222594. DOI: 10.1002/mrm.28675.

[173] Qing Zou et al. "Deep Generative Storm Model for Dynamic Imaging". In: *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*. Vol. 2021. IEEE, 2021, pp. 114–117. ISBN: 978-1-6654-1246-9. DOI: 10.1109/ISBI48211.2021.9433839. arXiv: 2101.12366.

Appendix A

Received: 13 July 2021 | Revised: 12 October 2021 | Accepted: 5 November 2021

DOI: 10.1002/mrm.29097

RESEARCH ARTICLE

Magnetic Resonance in Medicine

Free-breathing 3D cardiac T_1 mapping with transmit B_1 correction at 3T

Paul Kyu Han[1,2] | **Thibault Marin**[1,2] | **Yanis Djebra**[1,2,3] | **Vanessa Landes**[4] | **Yue Zhuo**[1,2] | **Georges El Fakhri**[1,2] | **Chao Ma**[1,2]

[1]Gordon Center for Medical Imaging, Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts, USA

[2]Department of Radiology, Harvard Medical School, Boston, Massachusetts, USA

[3]LTCI, Télécom Paris, Institut Polytechnique de Paris, France

[4]GE Healthcare, Boston, Massachusetts, USA

Correspondence
Chao Ma, Gordon Center for Medical Imaging, Massachusetts General Hospital, 125 Nashua Street, Suite 660, Boston, MA 02114, USA.
Email: cma5@mgh.harvard.edu

Funding information
The National Institutes of Health; Grant/Award Nos. P41EB022544, R01CA165221, R01HL137230, R01HL118261, T32EB013180, and K01EB030045

Purpose: To develop a cardiac T_1 mapping method for free-breathing 3D T_1 mapping of the whole heart at 3 T with transmit B_1 (B_1^+) correction.

Methods: A free-breathing, electrocardiogram-gated inversion-recovery sequence with spoiled gradient-echo readout was developed and optimized for cardiac T_1 mapping at 3 T. High-frame-rate dynamic images were reconstructed from sparse (k,t)-space data acquired along a stack-of-stars trajectory using a subspace-based method for accelerated imaging. Joint T_1 and flip-angle estimation was performed in T_1 mapping to improve its robustness to B_1^+ inhomogeneity. Subject-specific timing of data acquisition was used in the estimation to account for natural heart-rate variations during the imaging experiment.

Results: Simulations showed that accuracy and precision of T_1 mapping can be improved with joint T_1 and flip-angle estimation and optimized electrocardiogram-gated spoiled gradient echo–based inversion-recovery acquisition scheme. The phantom study showed good agreement between the T_1 maps from the proposed method and the reference method. Three-dimensional cardiac T_1 maps (40 slices) were obtained at a 1.9-mm in-plane and 4.5-mm through-plane spatial resolution from healthy subjects (n = 6) with an average imaging time of 14.2 ± 1.6 minutes (heartbeat rate: 64.2 ± 7.1 bpm), showing myocardial T_1 values comparable to those obtained from modified Look-Locker inversion recovery. The proposed method generated B_1^+ maps with spatially smooth variation showing 21%–32% and 11%–15% variations across the septal–lateral and inferior–anterior regions of the myocardium in the left ventricle.

Conclusion: The proposed method allows free-breathing 3D T_1 mapping of the whole heart with transmit B_1 correction in a practical imaging time.

KEYWORDS
cardiac T1 mapping, free-breathing, low-rank, myocardial T1 mapping, spoiled gradient-echo, transmit B1 inhomogeneity

Paul Kyu Han and Thibault Marin contributed equally to this work.

1 | INTRODUCTION

Cardiac T_1 mapping is a powerful cardiovascular MRI technique that allows quantitative assessment of tissue characteristics and underlying pathology of the myocardium. Native (i.e., without the use of exogenous contrast agent) myocardial T_1 characterizes alterations in the structure and intracellular/extracellular components of the myocardium. Native myocardial T_1 is a well-recognized biomarker for quantitative assessment of diseases that alter tissue composition, such as iron deposition, amyloid disease, Anderson-Fabry disease, and myocarditis.[1–3] Extracellular volume fraction, measured from pre-contrast and post-contrast T_1 values, provides quantitative measurement of interstitial expansion and associated diseases, such as amyloidosis, fibrosis, or myocardial edema.[2,3] Extracellular volume fraction is an emerging biomarker for diffuse fibrosis (e.g., heart failure, dilated cardiomyopathy, amyloidosis),[3] which is challenging to detect using qualitative late-gadolinium enhancement methods.

Modified Look-Locker inversion recovery (MOLLI)[4] is a widely used method for 2D cardiac T_1 mapping, which uses adiabatic inversion pulses for magnetization preparation and performs electrocardiogram (ECG)–gated balanced SSFP acquisitions through multiple cardiac cycles with breath-holding. Although MOLLI produces myocardial T_1 maps with high precision,[5] the method is limited to a single-slice imaging per breath-hold. Methods have been developed to extend the conventional 2D MOLLI method to multislice 2D or 3D acquisitions with breath-holding by leveraging the state-of-the-art parallel imaging, simultaneous multislice acquisition, compressed sensing, and non-Cartesian sampling techniques.[6,7] However, these methods suffer from limited through-plane resolution and coverage, often involving long or repetitive breath-holds to obtain volumetric T_1 maps of the heart, which imposes significant burden on patients.

Various methods have been developed to overcome the limitations of breath-holding and allow 3D cardiac T_1 mapping with free-breathing acquisitions. Respiratory and cardiac gating–based T_1 mapping methods have been developed to acquire interleaved multislice 2D[8–10] or segmented 3D k-space data[11–13] at end-diastole with free-breathing, in which effects from respiratory motion were mitigated by prospectively tracking respiratory motion using navigators or self-navigation techniques. Magnetic resonance fingerprinting approaches have been combined with free-breathing ECG-gated acquisitions for multi-parametric cardiac MRI.[14] However, most of these methods are limited by spatial coverage, resolution in slice direction, or imaging time due to the low data-acquisition efficiency of gating. Recently, free-running (i.e., no cardiac or respiratory gating) continuous acquisition methods have been proposed for 2D or 3D cardiac T_1 mapping.[15–20] Of note, Qi et al reported a free-running 3D whole-heart T_1 mapping method[15] that uses translational respiratory motion correction and a patch-based low-rank tensor model to reconstruct 3D T_1 maps with isotropic resolution. The T_1 maps obtained by this method, however, were from 1.5 T, and the method may not translate well to 3 T for reasons discussed subsequently.

Although 3D cardiac T_1 mapping methods developed up until now have been applied mostly at 1.5 T, unique technical challenges arising from more severe B_0 and transmit B_1 (B_1^+) inhomogeneities need to be addressed at 3 T. For instance, spoiled gradient-echo (SPGR) readout with small flip angle (FA) is often used in 3D cardiac imaging at 3 T to avoid B_0 inhomogeneity–caused banding artifacts associated with balanced SSFP readout.[7,16,17,19] However, T_1 mapping with SPGR readout is known to be sensitive to errors in FA caused by imperfect RF pulses and B_1^+ inhomogeneities.[21] The latter is particularly problematic at 3 T, in which B_1^+ variation over the left ventricle with body-coil transmission is reported on the order of 30%–60%,[22] leading to bias in T_1 estimation. Robustness of cardiac T_1 mapping methods with SPGR acquisitions needs to be thoroughly investigated in the presence of B_1^+ inhomogeneity.[23,24]

In this work, we present a new cardiac T_1 mapping method for rapid 3D T_1 mapping of the heart at 3 T. A free-breathing, ECG-gated IR sequence with SPRG readout was developed and optimized in terms of acquisition protocol and excitation FA for accurate and precise cardiac T_1 mapping at 3 T. The optimized scheme was combined with sparse (k,t)-space sampling along a stack-of-stars trajectory to accelerate imaging. A subspace-based image-reconstruction method was used to recover high-frame-rate dynamic images from highly undersampled (k,t)-space data. The effects of FA errors on T_1 mapping were mitigated by joint estimation of T_1 and FA, in which the reconstructed dynamic images were first binned to different respiratory motion phases and then fitted voxel-by-voxel to a signal dictionary generated using Bloch equation simulations. The effects of heart-rate variations on T_1 mapping were reduced by generating signal dictionary with subject-specific timing of data acquisition recorded during imaging experiment. The performance of the proposed method was characterized and validated through numerical simulations, phantom studies, and in vivo experiments on healthy human subjects ($n = 6$). Preliminary accounts of this work have been presented previously in the form of conference abstracts.[25–27]

2 | METHODS

2.1 | Data acquisition

The proposed ECG-gated cardiac T_1 mapping sequence is shown in Figure 1. A nonselective inversion pulse was applied every N+M heartbeats with two different TIs. This scheme is referred to as the N-(M)-N-(M) protocol for simplicity, where N denotes the number of cardiac cycles for acquisition and M denotes the number of cardiac cycles for signal recovery. Data were acquired at end-diastole period using SPGR readout. A special data-acquisition scheme was used for sparse sampling (k,t)-space data along a stack-of-stars trajectory. A limited number of "training" data (e.g., along the k_x, k_y, and k_z directions across the center of the k-space) were acquired with high sampling rate to determine the temporal changes of the underlying signal. Data at all other k-space locations were sparsely sampled over the entire (k,t)-space to ensure that sufficient number of measurements were acquired at each k-space location for subspace-based image reconstruction.

To track respiratory motion, one-dimensional (1D) respiratory navigator signals were acquired in the sagittal plane at the dome of the right hemi-diaphragm at the beginning and end of data acquisition of each cardiac cycle. A spatially selective inversion pulse was applied right after the nonselective inversion pulse to invert the magnetization signals in the same sagittal plane back to the equilibrium state, and therefore to mitigate the contrast changes caused by the nonselective inversion pulses in the navigator signals.

2.2 | Image reconstruction

Image reconstruction of sparsely sampled data was performed by solving the following constrained optimization problem:

$$\hat{\rho}(x,t) = \arg\min_{\rho(x,t)} \|d(k,t) - \Omega\mathcal{F}_s\{\rho(x,t)\}\|_2^2 + \lambda\|\mathcal{F}_t\{\rho(x,t)\}\|_1,$$
$$s.t.\ \rho(x,t) = \sum_{l=1}^{L} u_l(x)\, v_l(t) \tag{1}$$

where $d(k,t)$ denotes the measured data; $\mathcal{F}_s$ denotes the Fourier transform in the spatial domain; Ω denotes the sampling mask in the (k,t)-space; and λ denotes the regularization parameter, which was chosen based on the discrepancy principle.[28] The first term of the cost function in Equation 1 penalizes data inconsistency, while the second term promotes sparsity of the reconstructed dynamic images $\rho(x,t)$ in the spatio-spectral domain.[29] The constraint in Equation 1 represents $\rho(x,t)$ as a partially separable function,[30,31] where $v_l(t)$ denotes the temporal basis function, $u_l(x)$ denotes the corresponding spatial coefficients, and L is the model order. In this work, $v_l(t)$ was estimated separately from training data using singular value decomposition for simplified computation.[29] Image reconstruction problem was then reduced

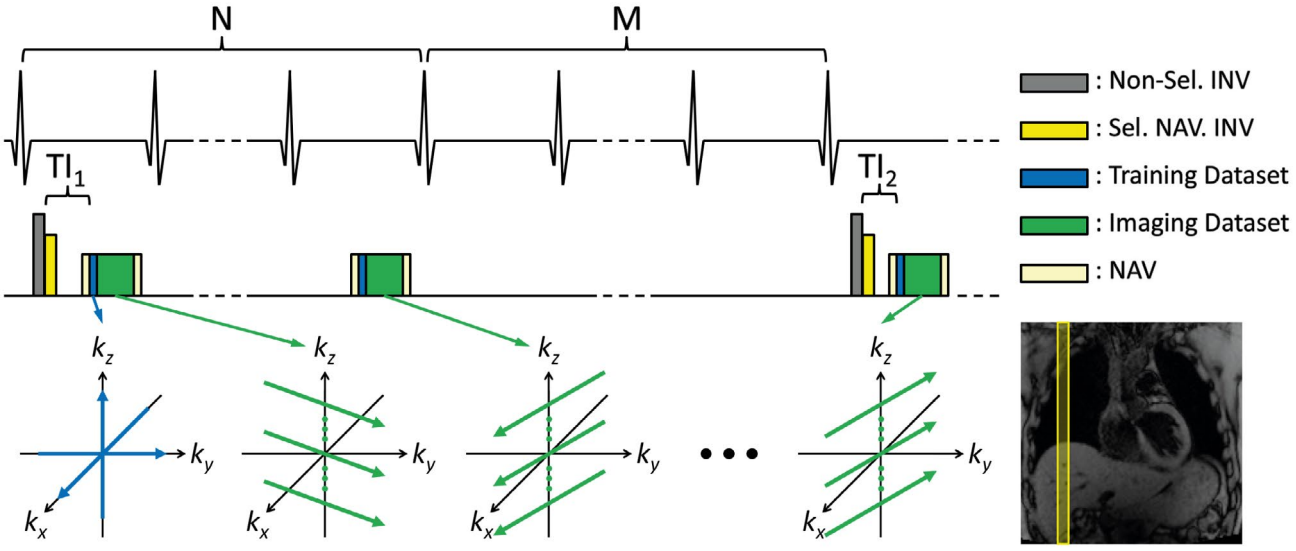


FIGURE 1 Schematic diagram of the proposed data-acquisition scheme. The N-(M)-N-(M) protocol is shown with nonselective inversion pulse applied every N+M heartbeats (where N denotes the number of cardiac cycles for acquisition and M denotes the number of cardiac cycles for signal recovery) with two different TIs. Data acquisition consists of the "training" data set, acquiring a limited number of k-space lines with high sampling rate, and the "imaging" data set, sparsely sampling all other k-space locations for subspace-based image reconstruction. To track respiratory motion, one-dimensional respiratory navigator signals were acquired in the sagittal plane at the dome of the right hemi-diaphragm after a spatially selective inversion pulse was applied in the same sagittal plane to invert the magnetization signals back to the equilibrium state

to determining the spatial coefficients $u_l(x)$ from measured data. The optimization problem in Equation 1 was solved using an alternating direction method of multipliers[32]–based algorithm.[33] For fast computation of $\mathscr{F}_s$, 1D Fourier transform was applied along the k_z direction first, and non-uniform fast Fourier transform[34] was applied in the k_x-k_y plane for slice-by-slice reconstruction. The dynamic images were reconstructed in a coil-by-coil fashion and then combined using the sum-of-squares method to form the final reconstruction. We implemented the image reconstruction algorithm in Python and used the SigPy package[35] to accelerate the computation using GPUs. Reconstructions were performed on four NVIDIA Tesla V100 SXM2 GPUs (parallelized over slice and coil dimensions) with reconstruction time around 1 minute for each slice and coil.

2.3 | Estimation of T_1 and FA

Before parameter estimation, the reconstructed dynamic images were binned to different respiratory-motion phases based on diaphragm position information extracted from the navigator signals. Let $\hat{\boldsymbol{\rho}}_{n,m}$denote the reconstructed dynamic signals at a voxel $x_n$in a selected respiratory phase $m$and $\boldsymbol{\eta}_{n,m} = \left[T_{1,m}\left(x_n\right), \theta_m\left(x_n\right)\right]$ denote the nonlinear parameters (i.e., T_1 and FA). We estimated the $\boldsymbol{\eta}_{n,m}$that best fit the dynamic signals as follows, using the variable projection method[36]:

$$\hat{\boldsymbol{\eta}}_{n,m} = \arg\min_{\boldsymbol{\eta}_{n,m}} \frac{\left|\hat{\boldsymbol{\rho}}_{n,m}\boldsymbol{a}^{\mathrm{H}}\left(\boldsymbol{\eta}_{n,m}\right)\right|^2}{\|\boldsymbol{a}\left(\boldsymbol{\eta}_{n,m}\right)\|_2^2} \tag{2}$$

where $\boldsymbol{a}(\bullet)$ denotes an atom of a dictionary of signals calculated using Bloch equation simulation with varying T_1 and FA values defined on a discrete 2D grid and the actual timing of acquisition recorded during the imaging experiment. The evolution of the bulk magnetization vector was calculated excitation-by-excitation over the course of the entire scan by solving the Bloch equation numerically. The signal bases were formed using the simulated signals associated with the excitations where the spoke at the center of the k-space (e.g., $k_z = 0$) was acquired at each frame. The bases were then binned to a motion phase to form the dictionary for joint T_1 and FA fitting. For the simulation and phantom studies, a dictionary was generated for a range of T_1 from 1 to 3000 ms in increments of 1 ms and for a range of FA from 0 to $2\times FA_o$ in increments of $0.01\times FA_o$, where FA_o is the nominal FA, in the case of joint estimation. For the in vivo study, the dictionary was generated for a range of T_1 from 500 to 2500 ms in increments of 10 ms and for a range of FA from $0.2\times FA_o$ to $1.5\times FA_o$ in increments of $0.01\times FA_o$ in the case of joint estimation, in consideration of expected smaller range of T_1 and FA values and for the sake of reducing computation time in dictionary generation and fitting. The time to generate the dictionary and to fit the data was 6.7 ± 2.1 minutes and 51.1 ± 2.8 seconds, respectively, using 8 Intel Xeon 2.4 GHz CPUs (4-core per CPU) on a workstation.

2.4 | Simulation study

We performed simulation studies to optimize the proposed data acquisition scheme in Figure 1. Bloch equation simulations were performed for various ECG-gated IR schemes with the following parameters in common unless otherwise mentioned: heart-rate = 80 bpm, acquisition window = 180 ms, inversion delay times = 100/180 ms (i.e., delays from the inversion pulse to the beginning of the first acquisition), FA = 6°, and SPGR readout. The effect of B_1^+ inhomogeneity on the accuracy of T_1 estimation was investigated for 8-8, 5-(3)-5-(3), and 10-(3)-10-(3) protocols without noise at different B_1^+ scenarios (i.e., $B_1^+ =$ 0.8/1/1.2). Relative difference (in relation to the ground-truth T_1 value) was used to assess any bias in T_1 estimation. The precision of T_1 estimation was investigated for N-(M)-N-(M) protocols with noise and perfect B_1^+ (i.e., $B_1^+ =$ 1) using Monte Carlo simulations (i.e., 10 000 noise realizations). Normalized SD (nSD) (i.e., SD of estimated T_1 values normalized by SD of noise and inverse of square root of acquisition time) was used to assess the precision of the T_1 estimation. The SD of noise was assumed to be constant for all considered scenarios. Additionally, the effect of nominal FA on T_1 mapping was investigated for 8-8, 5-(3)-5-(3), and 10-(3)-10-(3) protocols with FAs ranging from 1º to 15° in 1° increments. Finally, the effect of heart-rate variation on T_1 mapping was investigated for 10-(3)-10-(3) protocol with heart-rate ranging from 50 to 120 bpm in 5 bpm increments.

Note that the (k,t)-space is highly undersampled and the dynamic images are reconstructed with an explicit low-rank constraint in the proposed method. Therefore, different from MOLLI, in which the k-space of each frame can be fully sampled, the imaging time of the proposed method is given by $T_{Proposed} = \frac{P_s P_z L}{Q_c} \times t_{RR} \times \frac{M+N}{N}$, where P_s and P_z denote the number of spokes in the k_x-k_y plane and phase encoding steps in the k_z axis, respectively, which are determined by the required spatial resolution; L is the rank of the dynamic images; t_{RR} denotes a fixed cardiac cycle duration; and Q_c denotes the number of spokes acquired per cardiac cycle. Here, $P_s P_z L$ is the total number of unknowns of the low-rank model when the temporal basis is predetermined. Note that the inherent rank of the dynamic images reflects the spatial-temporal correlations of the temporal signal variations of all the voxels. It

depends on the respiratory-motion pattern and the T_1 value distributions rather than protocol parameters N and M. Because t_{RR}, P_s, P_z, L, and Q_c were the same for different imaging protocols, the imaging time of the proposed method was calculated as $T_{Proposed} = \frac{M+N}{N}$ in our simulation study for simplicity.

2.5 | Phantom study

A structured phantom consisting of 21 vials of deionized water doped with concentrations of gadolinium (Dotarem) varying from 0 to 0.5 mmol/L was built to validate the performance of the proposed method. Imaging experiment was performed using a 3T MR scanner (MAGNETOM Trio; Siemens Healthcare, Erlangen, Germany) with a body coil for transmission and a 12-channel head coil for reception. Acquisitions were performed using 8-8, 5-(3)-5-(3), and 10-(3)-10-(3) protocols for a simulated heart rate of 80 bpm. Common imaging parameters were FOV = 360 × 304 mm^2, matrix-size = 192 × 162, slice-thickness = 6 mm, FA = 9°, TR/TE = 3.0/1.5 ms, inversion delay times = 100/180 ms, and SPGR readout. Fully sampled (*k,t*)-space data were acquired on a Cartesian grid with a temporal resolution of 30 ms (i.e., 10 phase-encoding lines per frame) to evaluate the performance of different acquisition protocols. The 8-8, 5-(3)-5-(3), and 10-(3)-10-(3) protocols were repeated over 400, 400, and 640 cardiac cycles, respectively, to ensure full sampling of the *(k,t)*-space data. An IR sequence with fast spin echo (FSE) readout was performed to obtain reference T_1 maps with the following imaging parameters: FOV = 360 × 304 mm^2, matrix-size = 192 × 162, slice-thickness = 6 mm, TR = 10,000 ms, echo train length = 7, and TI = 50/100/2 50/500/750/1000/1500/2000/2500/3000 ms. An additional scan with MOLLI[37] was performed for comparison with the following parameters: FOV = 360 × 304 mm^2, matrix-size = 192 × 162, and slice thickness = 6 mm. Regions of interest were drawn within each vial, and the average T_1 value within each vial was used for analysis. Scatter plots were generated to show the correlation between the T_1 values from the different methods and those from the reference IR-FSE method. Bland-Altman analysis was performed to analyze the agreement between the T_1 values from the different methods and those from the reference IR-FSE method.

2.6 | In vivo study

Six healthy volunteers (4 males and 2 females; 32 ± 3 years) were recruited under a study protocol approved by our local institutional review board. Written informed consent was obtained from all subjects before study participation. Imaging experiments were performed using a 3T MR scanner (MAGNETOM Trio) with a body coil for transmission and spine and surface coils for reception. Imaging was performed using 10-(3)-10-(3) protocol with data sampling following a stack-of-stars trajectory. Two frames were acquired per cardiac cycle, each consisting of k-space spokes along the same angle in the k_x-k_y plane over all k_z encodings and three additional training lines at the center of the k-space along the k_x, k_y, and k_z direction, respectively (Figure 1). The spoke angle varied randomly from frame to frame following uniform random distribution. The other imaging parameters were FOV = 308 × 308 × 180 mm^3, matrix size = 160 × 160 × 40, image orientation = short-axis view, FA = 9°, TR/TE = 3.4/1.7 ms, and inversion delay times = 100/180 ms. A relatively large through-slice coverage was chosen to mitigate errors in FA due to imperfect slab excitation profile in the presence of both respiratory and cardiac motions. Data acquired over the first 800 cardiac cycles (corresponding to anticipated scan time of 10 minutes considering average adult heart rate of 80 bpm) were used for reconstruction and analysis. The dynamic images were reconstructed using temporal basis functions $v_l(t)$ estimated from training data with model order L = 15. The model order was chosen based on the singular value decay of the Casorati matrix formed by the training data as in the previous work on using low-rank constraints for image reconstruction.[29,33] The reconstructed dynamic images were then binned into eight respiratory motion bins using the acquired 1D respiratory navigator signals. The number of respiratory bins was chosen to be eight based on our previous experience on using MR for respiratory motion correction in PET.[33] The diaphragm position was first estimated for each frame by fitting a logistic function to the 1D spatial profile near the interface between the liver and lung. Frames were then grouped into bins according to the estimated diaphragm position while ensuring a similar number of frames within each bin. Results obtained from respiratory motion phase at or near end-exhalation were used for analysis. The short-axis view slices were divided into ROIs of 16 segments according to the American Heart Association recommendations[38] for analysis. For comparison, 2D T_1 maps were acquired using MOLLI[37] for five slices in the short-axis view over the apical, midcavity, and basal regions of the heart and for one slice in the long-axis four-chamber view with the following parameters: FOV = 360 × 304 mm^2, matrix size = 192 × 162, and slice thickness = 4.5 mm. All of the five short-axis slices were categorized into apical, midcavity, and the basal regions based on location and were used for analysis. The mean and SD of the T_1 and B_1^+ (defined as the ratio between the measured and nominal FAs) were calculated for each ROI

and were visualized through bull's-eye plot and bar plot. Statistical analysis was performed using Wilcoxon signed-rank test to compare T_1 values obtained by MOLLI and the proposed method.

3 | RESULTS

Results from the simulation study are shown in Figures 2 and 3. When estimating T_1 only with the assumption of perfect B_1^+ (i.e., $B_1^+ = 1$), noticeable bias in T_1 estimation was found in all the investigated imaging protocols in the presence of typical B_1^+ inhomogeneities at 3 T (the blue dashed lines in Figure 2). The 8-8 protocol, which has the highest data-acquisition efficiency (i.e., number of k-space lines acquired per unit time) among all the schemes, showed the highest sensitivity to B_1^+ inhomogeneity (Figure 2A). Insertion of cardiac cycles for signal recovery reduced this bias in T_1 estimation at the cost of imaging time (Figure 2B,C). Joint T_1 and FA estimation led to unbiased T_1 estimation in the simulation study as expected (the red solid lines in Figure 2). Figure 3A shows the precision of T_1 estimation for different imaging protocols. The nSD of the proposed method with joint T_1 and FA estimation was minimized and plateaued around N = 13, 14, 14, and 13 for N-N, N-(1)-N-(1), N-(2)-N-(2), and N-(3)-N-(3) protocols, respectively. The nSD of the proposed method with T_1 estimation only was minimized and plateaued around N = 12, 8, 10, and 12 for N-N, N-(1)-N-(1), N-(2)-N-(2), and N-(3)-N-(3) protocols, respectively. Although resulting in unbiased estimation of T_1, joint T_1 and FA estimation led to larger nSD than the case of estimating T_1 only. Increasing the number of cardiac cycles for signal recovery reduced the nSD observed in joint T_1 and FA estimation, eventually to a level similar to those observed in the case of estimating T_1 only with 10-(3)-10-(3) protocol. Note that joint T_1 and FA estimation was unstable for the 8-8 protocol despite the desired data-acquisition efficiency. Figure 3B shows the effects of FAs on the precision of T_1 estimation for different imaging protocols. The nSD was minimized and plateaued around FA of 9° in the case of joint T_1 and FA estimation for both 5-(3)-5-(3) and 10-(3)-10-(3) protocols. Based on the results, FA of 9° was used for the ECG-gated IR sequence with SPGR readout in the phantom and in vivo experiments. Figure 3C further shows that heartbeat rate has only marginal effects on the precision of T_1 estimation for 10-(3)-10-(3) protocol, which is the protocol selected for the in vivo experiments.

The results from phantom studies are shown in Figure 4 and Supporting Information Figures S1 and S2. Banding artifacts were observed in the estimated T_1 maps

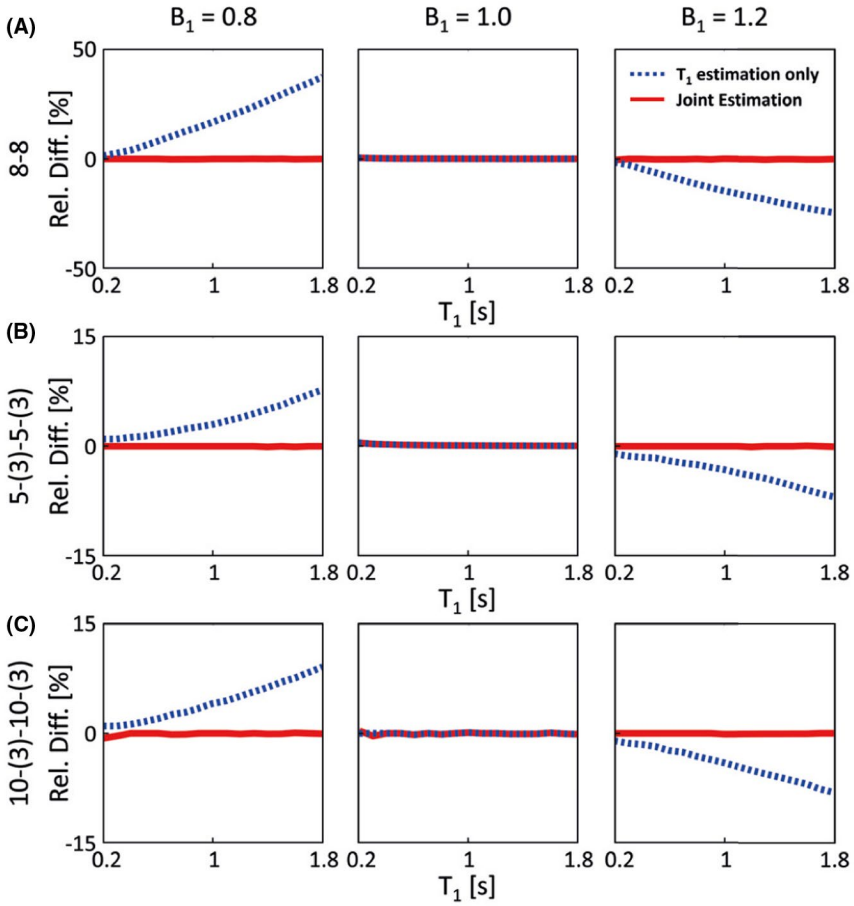


FIGURE 2 Simulation results showing the effect of B_1^+ inhomogeneity on the accuracy of T_1 estimation. Relative difference plots from the 8-8 (A), 5-(3)-5-(3) (B), and 10-(3)-10-(3) (C) protocols are shown for typical B_1^+ inhomogeneities at 3 T (i.e., $B_1^+ = 0.8$, 1, and 1.2). T_1 values obtained by joint T_1 and flip-angle (FA) estimation (red solid line) and T_1 estimation only with the assumption of perfect B_1^+ (blue dashed line) are shown. Small fluctuations are presumed to be due to numerical errors

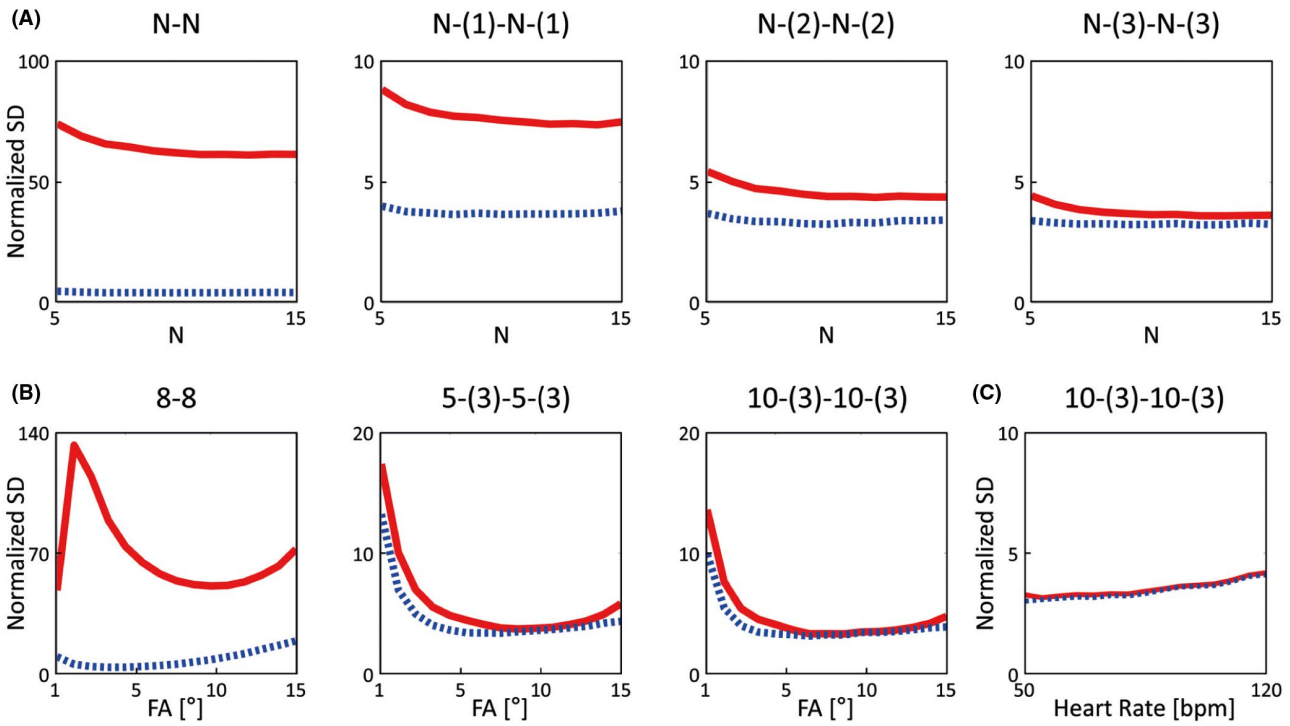


FIGURE 3 Simulation results showing the precision of T_1 estimation for different N-(M)-N-(M) protocols. (A) Effect of N-(M)-N-(M) protocols on normalized SD (nSD). (B) Effect of FA on nSD for the 8-8, 5-(3)-5-(3), and 10-(3)-10-(3) protocols. C, Effect of heart-rate variation on nSD for the 10-(3)-10-(3) protocol. T_1 values obtained by joint T_1 and FA estimation (red solid line) and T_1 estimation only with the assumption of perfect B_1^+ (blue dashed line) are shown

from MOLLI, whereas no noticeable artifacts were shown in the estimated T_1 maps from the proposed method (Figure 4A). The 5-(3)-5-(3) and 10-(3)-10-(3) protocols both showed accurate T_1 mapping in relation to the reference IR-FSE method when T_1 and FA were estimated jointly, as shown in the correlation plots in Figure 4B and Bland-Altman plots in Figure 4C. The 8-8 protocol produced T_1 maps with large variations when T_1 and FA were estimated jointly (Figure 4A), which matched with the simulation results in Figure 3. Supporting Information Figure S1 shows the results obtained from the same experiment but with estimation of T_1 only assuming prefect B_1^+ (i.e., $B_1^+ = 1$). B_1^+ inhomogeneity–caused bias was found in the estimated T_1 maps from IR sequence with SPGR readout as expected. Compared with the case of estimation of T_1 only, joint T_1 and FA estimation reduced the limits of agreement by 77.7, 44.7, and 49.4 ms for the T_1 mapping experiment with 8-8, 5-(3)-5-(3), and 10-(3)-10-(3) protocol, respectively (Figure 4C and Supporting Information Figure S1C). Supporting Information Figure S2 shows results from another phantom study with variation in B_1^+ field strength via control of transmitter voltage. The 5-(3)-5-(3) protocol achieved robust T_1 mapping when T_1 and FA were estimated jointly, despite the variation in B_1^+ field strength. The estimated T_1 from each vial were in good agreement with those estimated from the reference IR-FSE method. Ratios between different nominal FAs and estimated average FAs from each vial were also in good agreement.

Results from the in vivo study are shown in Figures 5–10 and Supporting Information Figures S3 and S4. The average heart-rate of the 6 volunteers was 64.2 ± 7.1 bpm (min: 53.7 bpm, max: 73.5 bpm). The acquisition time for the 6 volunteers was 14.2 ± 1.6 minutes (min: 12.2 minutes, max: 16.4 minutes). Figure 5 shows representative reconstructed images at various slice positions and TIs from subject 1 using the proposed method. No significant artifacts were seen in the reconstructed images across different slices and TI times. After respiratory-motion binning, an average of 153.8 ± 23.4 TIs were observed per respiratory motion bin over the course of 800 heartbeats across all subjects. Figure 6 shows representative short-axis-view T_1 maps from subjects 1 and 2 obtained by the proposed method. Figure 6 also shows four-chamber-view T_1 maps that were generated by reslicing the 3D T_1 map from the proposed method. For comparison, T_1 maps from MOLLI at the same slice position and orientation are shown at the bottom of each subfigure. Overall, the T_1 maps from the proposed method were comparable to those from MOLLI. Note that the nominal in-plane and through-plane resolution from the proposed method was 1.9 mm and 4.5 mm, respectively. Figure 7 further shows 3D T_1 maps from subject 2, covering the whole left ventricle from base to

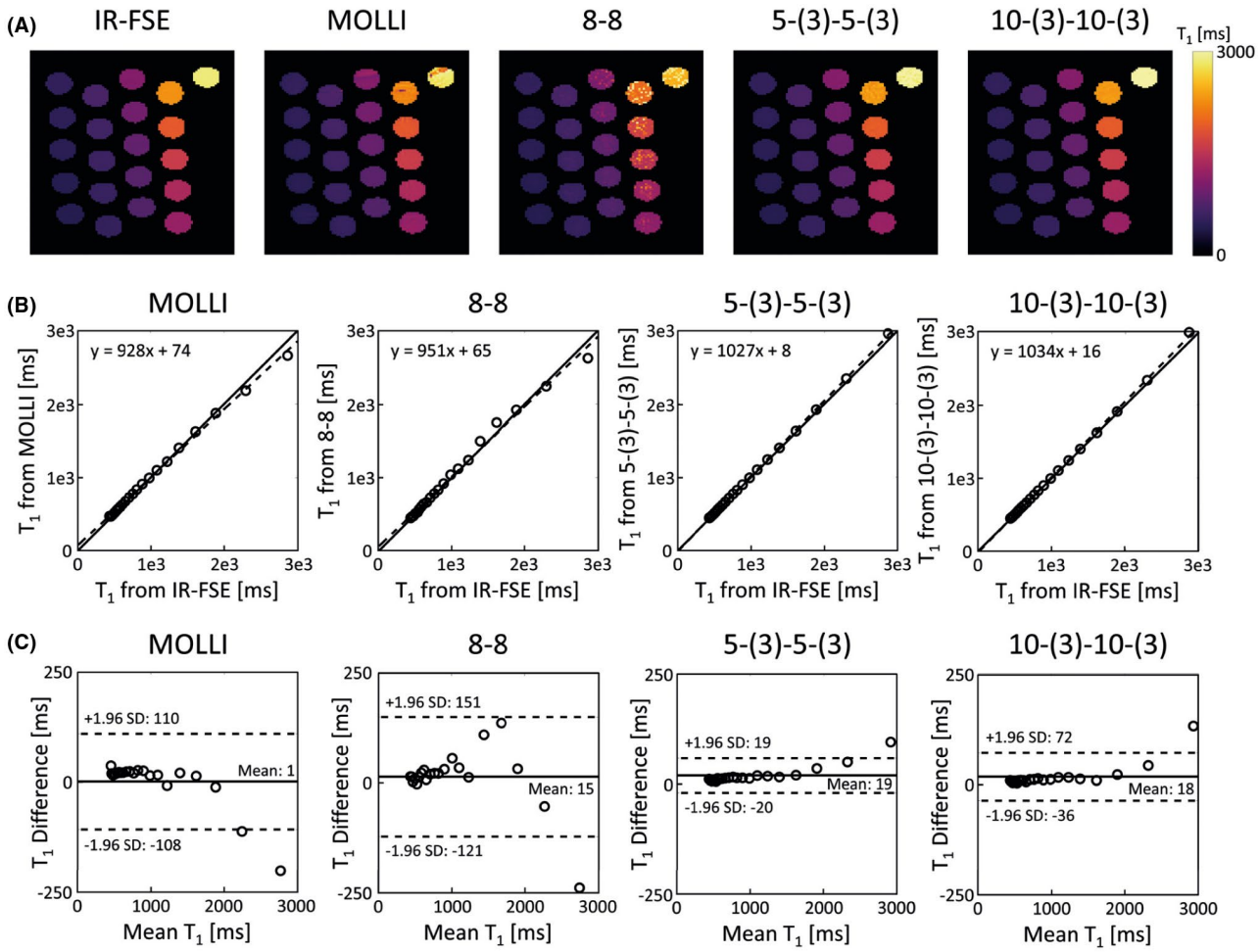


FIGURE 4 Phantom results of electrocardiogram (ECG)–gated inversion-recovery (IR) schemes with spoiled gradient-echo (SPGR) readout and joint T_1 and FA estimation. (A) Estimated T_1 maps from different methods. (B) Scatter plots showing the comparison of estimated T_1 from different methods with those from IR–fast spin echo (FSE). Solid line represents the line of identity, and dashed line represents the line of regression. (C) Bland-Altman plots showing the comparison of estimated T_1 from different methods with those from IR-FSE. Solid line represents the mean difference, and dashed line represents the 95% confidence interval for limits of agreement. Abbreviation: MOLLI, modified Look-Locker inversion recovery

apex. As can be seen, high-quality 3D T_1 maps of the heart were produced using the proposed method. Figure 8 shows a quantitative comparison of T_1 maps from all subjects using the proposed method (with and without joint T_1 and FA estimation) and MOLLI, respectively. The bull's-eye plot and bar plot of the mean and SD of T_1 values from each ROI show a very good agreement between the two methods. This observation is also supported by the Bland-Altman plots shown in Supporting Information Figure S3. Compared with the case of estimation of T_1 only, joint T_1 and FA estimation reduced the SD of T_1 values across 16 myocardial segments by 30.5 ms. Statistical test showed that the T_1 values of the 16 myocardial segments from the proposed method with joint T_1 and FA estimation were not statistically different from MOLLI at 5% significance level ($p = .08$). The mean and SD of septal T_1 values between MOLLI and the proposed method across subjects is compared in Supporting Information Table S1.

Figure 9 shows representative T_1 and B_1^+ maps from subject 2 obtained by the proposed method. Notice the similarity in estimated T_1 values for each tissue type (e.g., myocardium, liver, muscle) even with significant variations in estimated B_1^+ across different regions (Figure 9A). The 3D B_1^+ maps of the heart (Figure 9B) show larger B_1^+ values in lateral/anterior regions than septal/inferior regions, which is consistent with literature.[39] Group analysis of B_1^+ maps acquired from all subjects is shown in Figure 10. B_1^+ variation ranged from 21%–32% and 11%–15% across the septal–lateral and inferior–anterior regions of the myocardium in the left ventricle, respectively. When such B_1^+ inhomogeneities were ignored in T_1 estimation, T_1 values in the septal and inferior regions were overestimated (Supporting Information Figure S4). This was consistent with the simulation results in Figure 2.

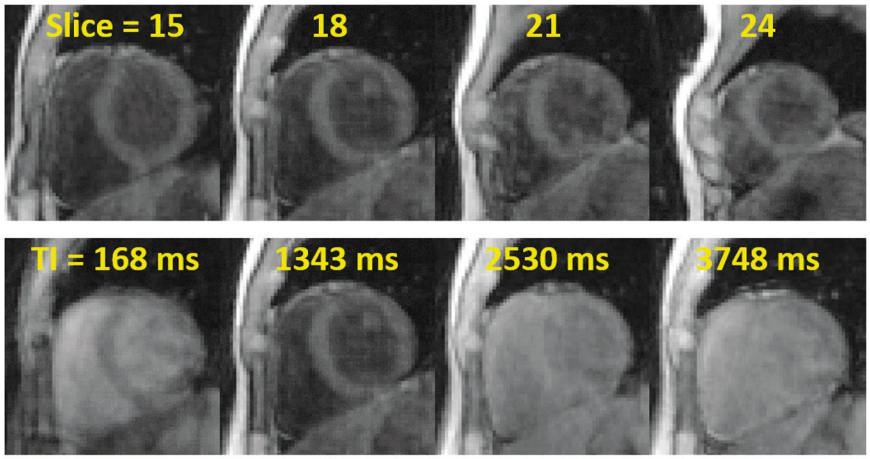


FIGURE 5 In vivo study results of reconstructed images from the proposed method. Representative reconstructed images from subject 1 are shown at various slice positions for a fixed TI of 1343 ms (top row) and various TI times for slice position index of 18 (bottom row). Note that the images were selected from reconstructed images in clock time and may be at different respiratory motion phase

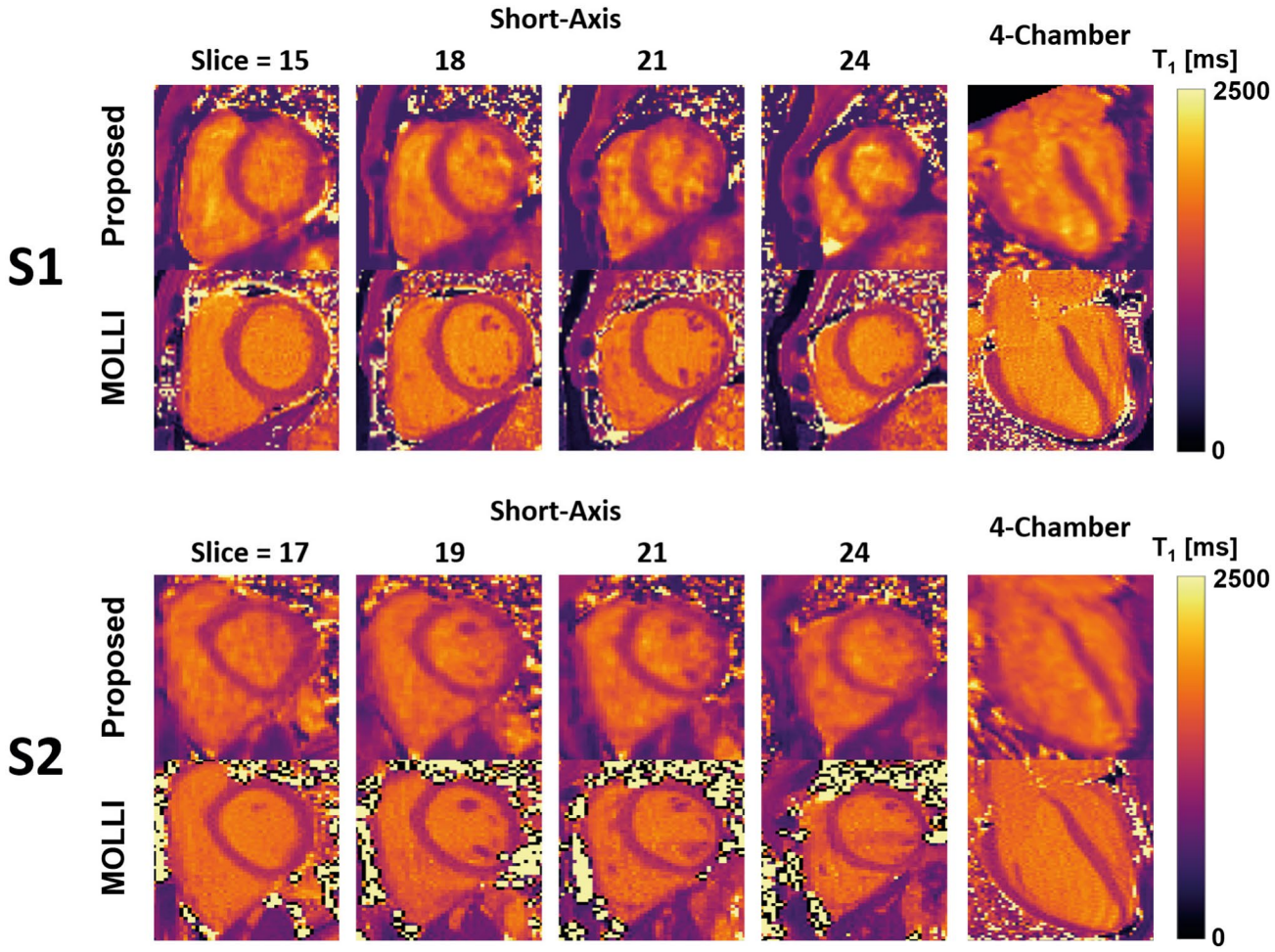


FIGURE 6 T_1 maps from subjects 1 and 2 obtained by the proposed method. Representative T_1 maps are shown for various slice positions in the short-axis view and four-chamber view. T_1 maps from MOLLI at the same slice position and orientation are shown at the bottom of each subfigure for comparison. Note that for the four-chamber view, the T_1 map from MOLLI were acquired with in-plane resolution of 1.5 mm, whereas the T_1 map from the proposed method was generated by reslicing the 3D T_1 map acquired with through-plane resolution of 4.5 mm

4 | DISCUSSION

In this work, a new free-breathing cardiac T_1 mapping method is proposed for robust T_1 mapping of the heart at 3 T. The T_1 maps obtained using the proposed method have a strong correlation and good agreement compared with reference and comparison methods in both the phantom experiments with various conditions (Figure 4 and

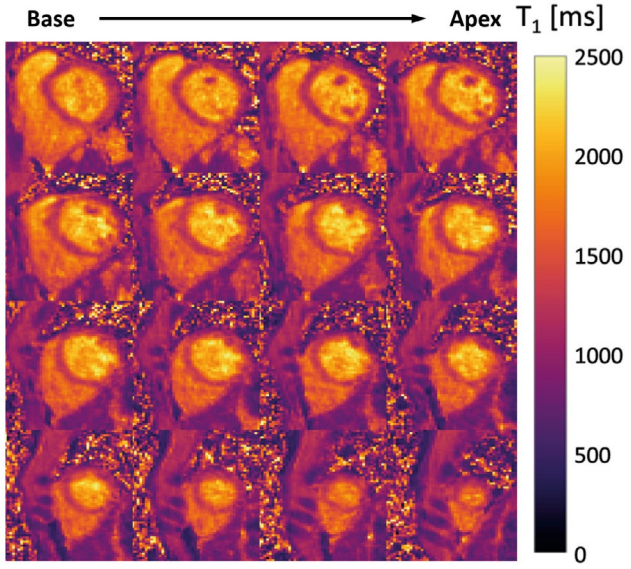


FIGURE 7 Representative 3D T_1 map in the short-axis view covering the whole left ventricle from base to apex

Supporting Information Figure S2) and the in vivo experiments across all subjects (Figures 6 and 8 and Supporting Information Figure S3). Robust T_1 mapping was achieved despite significant B_1^+ variations at 3 T. This is most noticeable by the fact that uniform T_1 distributions for each tissue type (e.g., myocardium, liver, muscle) were achieved across the entire FOV (Figure 9A) and is further supported by the close agreement of T_1 maps between the proposed method and MOLLI (Figure 8), despite the observed B_1^+ variation across the septal–lateral and inferior–anterior regions of the myocardium in the left ventricle (Figure 10). Noticeable bias in T_1 estimation was otherwise observed when B_1^+ inhomogeneities were ignored in T_1 estimation (Figures 2, 4, 8, and 10; Supporting Information Figures S1 and S4). These observations were consistent throughout simulation, phantom, and in vivo results. The estimated B_1^+ values of the myocardium show spatial variations that are consistent with those reported in literature[39] (i.e., the B_1^+ values in the lateral/anterior regions were 10%–30% larger than the septal/inferior regions). However, the B_1^+ distributions within the blood pool should be carefully interpreted, as flow effects were not considered in joint T_1 and FA estimation. As a result, the estimated T_1 and FA values of blood may be biased.

This method has several novel features. First, the proposed method mitigates bias in T_1 estimation caused by errors in FA through joint estimation of T_1 and FA. We carried out systematic numerical simulation studies to optimize the ECG-gated IR sequence with SPGR readout in terms of acquisition protocols and nominal FAs, with the goal of minimizing the SD of the estimated T_1. Second, the proposed method uses the special (k,t)-space sampling scheme and subspace-based image reconstruction to recover dynamic images from undersampled data (i.e., two 3D volumes for every cardiac cycle with data acquisition). This allows mitigating the effects of natural heart-rate variations on T_1 mapping by fitting the reconstructed dynamic signals to a signal dictionary generated with subject-specific timing of data acquisition recorded during imaging experiment. Third, the proposed method is robust to B_0 inhomogeneities, as it uses adiabatic nonselective pulse for inversion and SPGR acquisitions. Altogether, the proposed method achieves free-breathing T_1 mapping in the presence of B_1^+ and B_0 inhomogeneity at 3 T in a practical imaging time.

The proposed method may be potentially useful for quantification of post-contrast T_1 and extracellular volume fraction mapping, in which accurate and precise estimation of T_1 is important. Although results from simulation and phantom studies show that the proposed method can estimate short T_1 values with accuracy and precision, further investigation is necessary to evaluate the performance of proposed method for post-contrast T_1 estimation. Because T_1 relaxation in the tissue changes over time in vivo after contrast agent is injected, the performance of the proposed method needs to be carefully examined for these applications, including investigations in the context of subspace-based reconstruction. An interesting next step would be to investigate the feasibility as well as the performance of the proposed method for these applications in vivo.

In this work, a subspace-based image reconstruction method was used to recover dynamic images in clock time for ECG-gated acquisitions, and respiratory motion was resolved by subsequently binning reconstructed images to different respiratory motion phases. A potentially interesting future work would be to investigate the possibility of treating respiratory motion as an additional temporal dimension using the low-rank tensor model[16] for ECG-gated acquisitions. In the ideal case with constant heart-rate, this would be feasible, as T_1-weighted contrast changes can then be modeled by IR at fixed number of inversion delay times. In reality, however, natural variations of heartbeat rate will require modeling the T_1-weighted contrast changes in clock time for ECG-gated acquisitions over the entire imaging experiment. This can impose technical challenges when attempting to represent the underlying dynamic images using a low-rank tensor model.

Results from the in vivo study showed that T_1 values from the proposed method were not statistically different from MOLLI at 5% significance level ($p = .08$). This indicates that the proposed method has similar bias as MOLLI in the in vivo study. Several factors contribute to the apparent underestimation of T_1 when the proposed method is

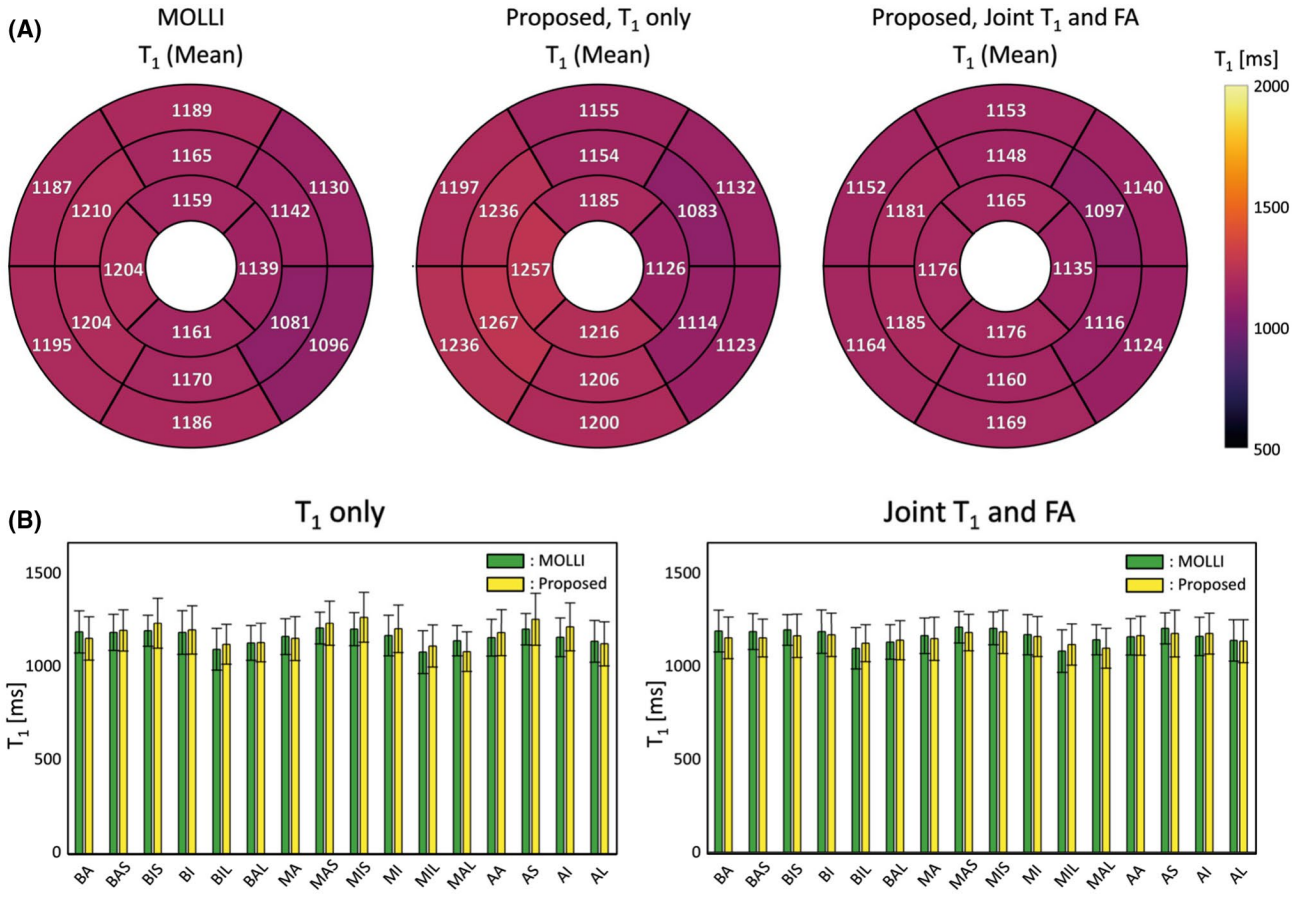


FIGURE 8 In vivo study results showing quantitative comparison of estimated T_1 from MOLLI and those from the proposed method. (A) Sixteen-segment American Heart Association (AHA) bull's eye plots of mean T_1 from MOLLI, proposed method with T_1 estimation only assuming perfect B_1^+ (i.e., $B_1^+ = 1$), and proposed method with joint T_1 and FA estimation. (B) Bar plots showing mean and SD of estimated T_1 from MOLLI and proposed method. Cases of proposed method with T_1 estimation only assuming perfect B_1^+ (i.e., $B_1^+ = 1$) and joint T_1 and FA estimation are shown. BA, BAS, BIS, BI, BIL, BAL, MA, MAS, MIS, MA, MIL, MAL, AA, AS, AI, AL each denotes basal anterior, basal anteroseptal, basal inferoseptal, basal inferior, basal inferolateral, basal anterolateral, midcavity anterior, midcavity anteroseptal, midcavity inferoseptal, midcavity inferior, midcavity inferolateral, midcavity anterolateral, apical anterior, apical septal, apical inferior, and apical lateral regions of the myocardium in the left ventricle

used for in vivo cardiac T_1 mapping. First, magnetization-transfer effects can lead to underestimated T_1 values in the proposed method as in MOLLI, because both methods use IR-based acquisition schemes for T_1 mapping. Second, the proposed method currently assumes perfect inversion pulse. Adding inversion efficiency to parametric fitting may potentially improve the accuracy of T_1 estimation at the cost of computation time (i.e., a larger dictionary of bases needed for parametric fitting). Further investigation is necessary to study the potential source of bias in the proposed method in the in vivo settings.

The current work has several limitations that warrant further investigation. First, the proposed method involves ECG-gated acquisition, which is susceptible to ECG mistriggering and may suffer from image blurring due to cardiac motion. The former could be addressed by adaptive heartbeat-rate prediction as in the double-gating technique.[40] The latter can be mitigated by retrospectively discarding k-space data acquired outside the end-diastole window based on recorded ECG signals. Free-running (i.e., no cardiac or respiratory gating) based continuous acquisition scheme may be preferable over the ECG-gated acquisition schemes for maximizing data-acquisition efficiency. However, in the free-running continuous acquisition scheme, apparent inversion recovery rate is strongly coupled with FA. Therefore, accurate estimation of FA is expected to be critical for accurate T_1 estimation with free-running acquisition in the presence of B_1^+ inhomogeneity at 3 T. This research direction is currently under investigation. Second, FA estimation by the proposed method was validated using a phantom experiment with variation in B_1^+ field strength by changing the transmitter voltage (Supporting Information Figure S2). Validation with a reference cardiac FA or B_1^+ mapping method (e.g., actual flip-angle imaging

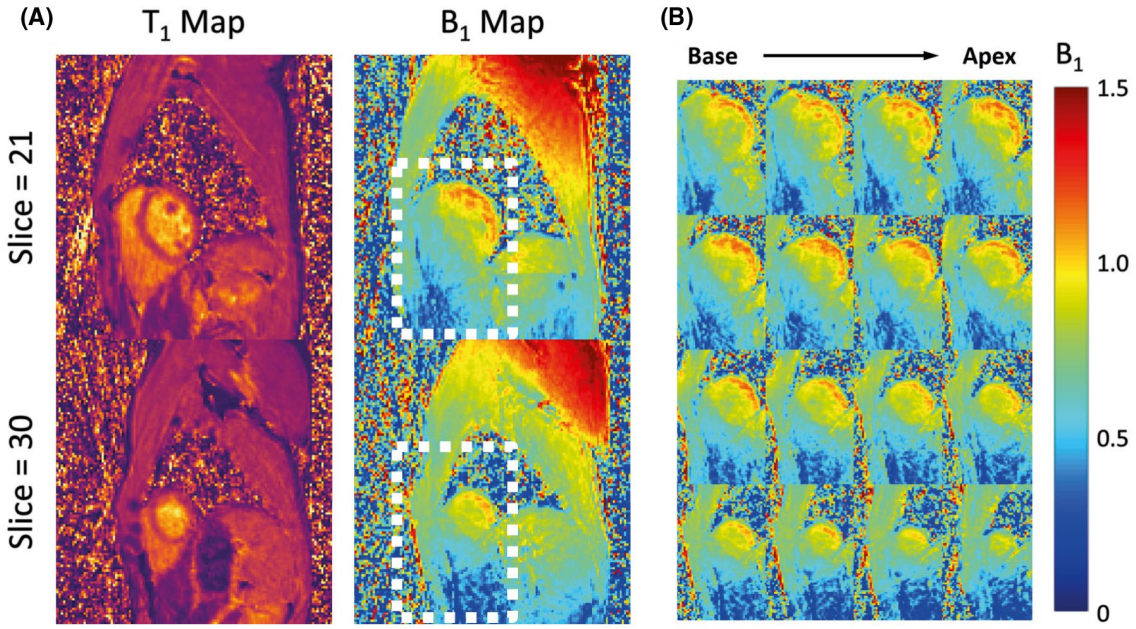


FIGURE 9 Representative 3D B_1^+ map (defined as the ratio between the measured and nominal FAs) obtained by the proposed method. (A) Short-axis view T_1 and B_1^+ maps at two slice positions. Notice the smooth variation of B_1^+ across regions. Also notice the similarity in estimated T_1 for each tissue type, even with variation in the estimated B_1^+ across different regions. (B) Short-axis view B_1^+ maps of the heart at various slice positions from base to apex

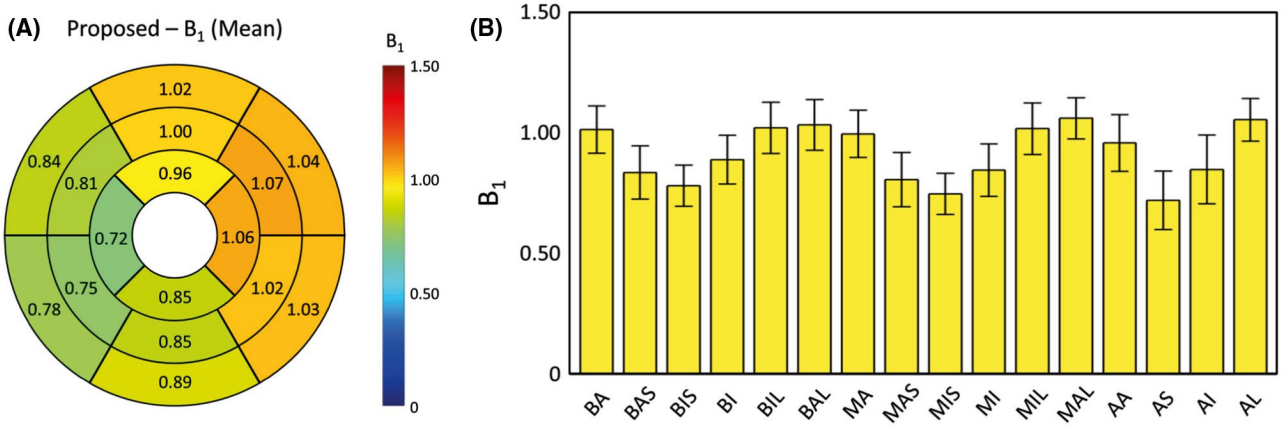


FIGURE 10 In vivo study results of estimated B_1^+ (defined as the ratio between the measured and nominal FAs) from all of the subjects. (A) 16-segment AHA bull's eye plot of normalized mean B_1^+ (i.e., B_1^+ normalized by the mean B_1^+ estimated from the midcavity anterior region of the myocardium in the left ventricle for each subject) estimated from the proposed method. Notice the difference in estimated B_1^+ between the septal–lateral and inferior–anterior regions of the myocardium. (B) Bar plot showing mean and SD of normalized B_1^+ estimated from the proposed method

method,)[41] in human subject studies is necessary to further verify the performance of FA mapping results from the proposed method. Third, performance of the proposed method was validated in vivo with a small number of healthy subjects (n = 6). Studies with a larger number of healthy subjects and patients are necessary to assess the accuracy and reproducibility of the proposed method and to evaluate its value in clinical applications. Future work can also include investigation with different sparsity constraints (e.g., finite difference or wavelet transform), investigation with different sampling schemes, and multisite/multivendor validations of the proposed method and findings from this work.

5 | CONCLUSIONS

A new free-breathing cardiac T_1 mapping method was proposed and optimized for fast 3D T_1 mapping of the

whole heart at 3 T with transmit B_1 correction in practical imaging time.

CONFLICT OF INTEREST

Vanessa Landes is an employee at GE Healthcare.

ORCID

Paul Kyu Han https://orcid.org/0000-0003-1650-1347
Chao Ma https://orcid.org/0000-0002-2016-8084

REFERENCES

1. Schelbert EB, Messroghli DR. State of the art: clinical applications of cardiac T1 mapping. *Radiology*. 2016;278:658-676.
2. Haaf P, Garg P, Messroghli DR, Broadbent DA, Greenwood JP, Plein S. Cardiac T1 mapping and extracellular volume (ECV) in clinical practice: a comprehensive review. *J Cardiovasc Magn Reson*. 2016;18:1-12.
3. Messroghli DR, Moon JC, Ferreira VM, et al. Clinical recommendations for cardiovascular magnetic resonance mapping of T1, T2, T2 and extracellular volume: a consensus statement by the society for cardiovascular magnetic resonance (SCMR) endorsed by the European Association for Cardiovascular Imaging. *J Cardiovasc Magn Reson*. 2017;19:1-24.
4. Messroghli DR, Radjenovic A, Kozerke S, Higgins DM, Sivananthan MU, Ridgway JP. Modified Look-Locker inversion recovery (MOLLI) for high-resolution T1 mapping of the heart. *Magn Reson Med*. 2004;52:141-146.
5. Kellman P, Hansen MS. T1-mapping in the heart: accuracy and precision. *J Cardiovasc Magn Reson*. 2014;16:1-20.
6. Weingärtner S, Moeller S, Schmitter S, et al. Simultaneous multislice imaging for native myocardial T1 mapping: improved spatial coverage in a single breath-hold. *Magn Reson Med*. 2017;78:462-471.
7. Chen Y, Lo WC, Hamilton JI, et al. Single breath-hold 3D cardiac T1 mapping using through-time spiral GRAPPA. *NMR Biomed*. 2018;31:1-13.
8. Weingärtner S, Roujol S, Akçakaya M, Basha TA, Nezafat R. Free-breathing multislice native myocardial T1 mapping using the slice-interleaved T1 (STONE) sequence. *Magn Reson Med*. 2015;74:115-124.
9. Guo R, Cai X, Kucukseymen S, et al. Free-breathing whole-heart multi-slice myocardial T1 mapping in 2 minutes. *Magn Reson Med*. 2021;85:89-102.
10. Jang J, Bellm S, Roujol S, et al. Comparison of spoiled gradient echo and steady-state free-precession imaging for native myocardial T1 mapping using the slice-interleaved T1 mapping (STONE) sequence. *NMR Biomed*. 2016;29:1486-1496.
11. Weingärtner S, Akçakaya M, Roujol S, et al. Free-breathing combined three-dimensional phase sensitive late gadolinium enhancement and T1 mapping for myocardial tissue characterization. *Magn Reson Med*. 2015;74:1032-1041.
12. Nordio G, Henningsson M, Chiribiri A, Villa ADM, Schneider T, Botnar RM. 3D myocardial T1 mapping using saturation recovery. *J Magn Reson Imaging*. 2017;46:218-227.
13. Guo R, Chen Z, Herzka DA, Luo J, Ding H. A three-dimensional free-breathing sequence for simultaneous myocardial T1 and T2 mapping. *Magn Reson Med*. 2019;81:1031-1043.
14. Cruz G, Jaubert O, Qi H, et al. 3D free-breathing cardiac magnetic resonance fingerprinting. *NMR Biomed*. 2020;33:1-16.
15. Qi H, Jaubert O, Bustin A, et al. Free-running 3D whole heart myocardial T1 mapping with isotropic spatial resolution. *Magn Reson Med*. 2019;82:1331-1342.
16. Christodoulou AG, Shaw JL, Nguyen C, et al. Magnetic resonance multitasking for motion-resolved quantitative cardiovascular imaging. *Nat Biomed Eng*. 2018;2:215-226.
17. Shaw JL, Yang Q, Zhou Z, et al. Free-breathing, non-ECG, continuous myocardial T1 mapping with cardiovascular magnetic resonance multitasking. *Magn Reson Med*. 2019;81:2450-2463.
18. Milotta G, Bustin A, Jaubert O, Neji R, Prieto C, Botnar RM. 3D whole-heart isotropic-resolution motion-compensated joint T1/T2 mapping and water/fat imaging. *Magn Reson Med*. 2020;84:3009-3026.
19. Hamilton JI, Jiang Y, Eck B, Griswold M, Seiberlich N. Cardiac cine magnetic resonance fingerprinting for combined ejection fraction, T1 and T2 quantification. *NMR Biomed*. 2020;33:1-17.
20. Jaubert O, Cruz G, Bustin A, et al. Free-running cardiac magnetic resonance fingerprinting: joint T1/T2 map and cine imaging. *Magn Reson Imaging*. 2020;68:173-182.
21. Freeman AJ, Gowland PA, Mansfield P. Optimization of the ultrafast look-locker echo-planar imaging T1 mapping sequence. *Magn Reson Imaging*. 1998;16:765-772.
22. Sung K, Nayak KS. Measurement and characterization of RF nonuniformity over the heart at 3T using body coil transmission. *J Magn Reson Imaging*. 2008;27:643-648.
23. Zhou R, Weller DS, Yang Y, et al. Dual-excitation flip-angle simultaneous cine and T1 mapping using spiral acquisition with respiratory and cardiac self-gating. *Magn Reson Med*. 2021;86:82-96.
24. Serry FM, Ma S, Mao X, et al. Dual flip-angle IR-FLASH with spin history mapping for B1+ corrected T1 mapping: application to T1 cardiovascular magnetic resonance multitasking. *Magn Reson Med*. 2021;86:3182-3191.
25. Han PK, Horng DE, Petibon Y, et al. Free-breathing 3D T1 mapping of the whole-heart using low-rank tensor modeling. In: Proceedings of the 27th Annual Meeting of ISMRM, Montréal, Canada, 2019. p 4396.
26. Han PK, Horng DE, Marin T, et al. Free-breathing three-dimensional T1 mapping of the heart using subspace-based data acquisition and image reconstruction. In: Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Berlin, Germany, 2019. p 2496.
27. Han PK, Marin T, Vanessa L, Djebra Y, El Fakhri G, Ma C. Optimization of spoiled GRE-based IR acquisition scheme for 3D cardiac T mapping at 3T. In: Proceedings of the 29th Annual Meeting of ISMRM [Virtual], 2021. p 3288.
28. Vogel CR. *Computational Methods for Inverse Problems*. Society for Industrial and Applied Mathematics; 2002.
29. Zhao B, Haldar JP, Christodoulou AG, Liang ZP. Image reconstruction from highly under sampled (k, t)-space data with joint partial separability and sparsity constraints. *IEEE Trans Med Imaging*. 2012;31:1809-1820.
30. Liang Z-P. Spatiotemporal imaging with partially separable functions. *IEEE Int Symp Biomed Imaging*. 2007;2:988-991.
31. Haldar JP, Liang Z-P. Spatiotemporal imaging with partially separable functions: a matrix recovery approach. In: Proceedings of the IEEE International Symposium on Biomedical Imaging:

From Nano to Macro. IEEE, Rotterdam, Netherlands, 2010. pp 716-719.
32. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found Trends Mach Learn*. 2010;3:1-122.
33. Marin T, Djebra Y, Han PK, et al. Motion correction for PET data using subspace-based real-time MR imaging in simultaneous PET/MR. *Phys Med Biol*. 2020;65:235022.
34. Fessler JA, Sutton BP. Nonuniform fast fourier transforms using min-max interpolation. *IEEE Trans Signal Process*. 2003;51:560-574.
35. Ong F, Lustig M. SigPy: a python package for high performance iterative reconstruction. In Proceedings of the ISMRM 27th Annual Meeting, Montreal, Quebec, Canada; 2019. Abstract 4819.
36. Golub G, Pereyra V. Separable nonlinear least squares: the variable projection method and its applications. *Inverse Probl*. 2003;19:R1.
37. Kellman P, Arai AE, Xue H. T1 and extracellular volume mapping in the heart: estimation of error maps and the influence of noise on precision. *J Cardiovasc Magn Reson*. 2013;15:1-12.
38. Cerqueira MD, Weissman NJ, Dilsizian V, et al. Standardized myocardial sementation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the cardiac imaging committee of the council on clinical cardiology of the American Heart Association. *Circulation*. 2002;105:539-542.
39. Rincón-Domínguez T, Menini A, Solana AB, et al. Accelerated multi-snapshot free-breathing B1+ mapping based on the dual refocusing echo acquisition mode technique (DREAM): an alternative to measure RF nonuniformity for cardiac MRI. *J Magn Reson Imaging*. 2019;49:499-507.
40. Do HP, Yoon AJ, Fong MW, Saremi F, Barr ML, Nayak KS. Double-gated myocardial ASL perfusion imaging is robust to heart rate variation. *Magn Reson Med*. 2017;77:1975-1980.
41. Yarnykh VL. Actual flip-angle imaging in the pulsed steady state: a method for rapid three-dimensional mapping of the transmitted radiofrequency field. *Magn Reson Med*. 2007;57:192-200.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

FIGURE S1 Phantom results of electrocardiogram (ECG)–gated inversion-recovery (IR) schemes with spoiled gradient-echo (SPGR) readout and T_1 estimation only assuming perfect B_1^+ (i.e., $B_1^+ = 1$). (A) Estimated T_1 maps from the different methods. (B) Scatter plots showing comparison of estimated T_1 from the different methods with those from IR-fast spin echo (FSE). Solid line represents line of identity, and dashed line represents line of regression. (C) Bland-Altman plots showing comparison of estimated T_1 from different methods with those from IR-FSE. Solid line represents the mean difference and dashed line represents the 95% confidence interval for limits of agreement

FIGURE S2 Phantom study results with variation in B_1^+ field strength via control of transmitter voltage. (A) Estimated T_1 maps from IR-FSE and 5-(3)-5-(3) protocol with 80%, 100%, and 120% of reference transmitter voltage. (B) Vial number positions. (C) Estimated B_1^+ maps (defined as the ratio between the measured and nominal flip angles [FAs]) from the 5-(3)-5-(3) protocol with 80%, 100%, and 120% of the reference transmitter voltage. (D) Estimated average FA within each vial from the 5-(3)-5-(3) protocol with 80% (green line and circle), 100% (blue line and circle), and 120% (red line and circle) of reference transmitter voltage. Note that the ratio between the different B_1^+ field strength and the estimated average FAs from each vial are in good agreement

FIGURE S3 In vivo study results of Bland-Altman plots comparing estimated T_1 from the proposed method with those from modified Look-Locker inversion recovery (MOLLI). Results from the anterior, septal, inferior, and lateral regions of the myocardium in the left ventricle are shown. Solid line represents the mean difference and dashed line represents the 95% confidence interval for limits of agreement

TABLE S1 Comparison of mean and SD of septal T_1 values between MOLLI and the proposed method across subjects

How to cite this article: Han PK, Marin T, Djebra Y, et al. Free-breathing 3D cardiac T_1 mapping with transmit B_1 correction at 3T. *Magn Reson Med*. 2021;00:1–14. doi:10.1002/mrm.29097

Appendix B

Received: 21 July 2022 | Revised: 7 October 2022 | Accepted: 24 October 2022

DOI: 10.1002/mrm.29526

RESEARCH ARTICLE

Magnetic Resonance in Medicine

Joint spectral quantification of MR spectroscopic imaging using linear tangent space alignment-based manifold learning

Chao Ma[1,2] | **Paul Kyu Han**[1,2] | **Yue Zhuo**[1,2] | **Yanis Djebra**[1,2,3] | **Thibault Marin**[1,2] | **Georges El Fakhri**[1,2]

[1]Gordon Center for Medical Imaging, Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts, USA

[2]Department of Radiology, Harvard Medical School, Boston, Massachusetts, USA

[3]LTCI, Telecom Paris, Institut Polytechnique de Paris, Paris, France

Correspondence
Chao Ma, Gordon Center for Medical Imaging, Massachusetts General Hospital, 125 Nashua Street, Boston, MA, 02114, USA.
Email: cma5@mgh.harvard.edu

Funding information
National Cancer Institute, Grant/Award Number: R01CA165221; National Institute of Biomedical Imaging and Bioengineering, Grant/Award Numbers: K01EB030045, P41EB022544, R01EB033582

Purpose: To develop a manifold learning-based method that leverages the intrinsic low-dimensional structure of MR Spectroscopic Imaging (MRSI) signals for joint spectral quantification.

Methods: A linear tangent space alignment (LTSA) model was proposed to represent MRSI signals. In the proposed model, the signals of each metabolite were represented using a subspace model and the local coordinates of the subspaces were aligned to the global coordinates of the underlying low-dimensional manifold via linear transform. With the basis functions of the subspaces predetermined via quantum mechanics simulations, the global coordinates and the matrices for the local-to-global coordinate alignment were estimated by fitting the proposed LTSA model to noisy MRSI data with a spatial smoothness constraint on the global coordinates and a sparsity constraint on the matrices.

Results: The performance of the proposed method was validated using numerical simulation data and in vivo proton-MRSI experimental data acquired on healthy volunteers at 3T. The results of the proposed method were compared with the QUEST method and the subspace-based method. In all the compared cases, the proposed method achieved superior performance over the QUEST and the subspace-based methods both qualitatively in terms of noise and artifacts in the estimated metabolite concentration maps, and quantitatively in terms of spectral quantification accuracy measured by normalized root mean square errors.

Conclusion: Joint spectral quantification using linear tangent space alignment-based manifold learning improves the accuracy of MRSI spectral quantification.

KEYWORDS
linear tangent space alignment, manifold learning, MRSI, spectral quantification

1 | INTRODUCTION

MR Spectroscopic Imaging (MRSI) is a label-free metabolic imaging tool that measures the spatial distribution of metabolites in the human body using the NMR phenomena. MRSI spectral quantification aims to estimate metabolite concentrations from noise and nuisance signal corrupted spectra and is a critical step in quantitative MRSI.[1-3] However, it is a challenging problem because of the intrinsically low signal-to-noise ratio (SNR) of MRSI, the nonlinear nature of the underlying parameter estimation problem, and the residual nuisance signals. MRSI spectral quantification has been extensively investigated in the literature. Early efforts include nonparametric methods, such as the Hankel Singular Value Decomposition- (HSVD) based methods[4-7] that exploit the linear predictability of MRSI signals for spectral quantification. The power of spectral prior knowledge has been later recognized, leading to several widely used spectral quantification methods, such as the variable projection (VARPRO)-based method,[8] LCModel,[9] AMARES,[10] QUEST,[11] and AQSES.[12] These methods differ in terms of the domain of the signal model (i.e., frequency vs. time), the optimization algorithm, and the additional constraints imposed on spectral parameters. Recently, deep learning-based methods[13-18] have been proposed for MRSI spectral quantification. However, the above methods estimate the parameters of each voxel's spectrum independently, resulting in large variations in the estimated metabolite concentration maps when the SNR of the MRSI data is low.

Over the past decades, technique developments in data acquisition and image reconstruction have significantly improved the spatial resolution and imaging speed of MRSI, making high-resolution MRSI practical in clinical settings.[19-21] This opens new opportunities to take advantage of spatial and spectral prior knowledge of MRSI signals for spectral quantification. Spatial correlations among spectra have been exploited indirectly by initializing or constraining spectral quantification at each voxel with the parameters estimated from the neighboring voxels.[9,22] Alternatively, joint spectral quantification methods have been proposed to directly impose spatial constraints on the estimated parameter maps while simultaneously quantifying the spectra of all MRSI voxels.[23-26] However, the underlying parameter estimation problem remains nonlinear in these methods, resulting in large-scale, nonlinear, nonconvex optimization problems that are difficult to solve efficiently.

Recently, Li et al. have proposed a subspace-based method for MRSI spectral quantification,[27] which is referred to as the subspace method in the rest of the paper. In this method, the spatiotemporal distribution of each metabolite is represented by a subspace model, which translates the MRSI spectral quantification problem into a subspace estimation problem and enables efficient and effective use of spatiotemporal priors for accurate MRSI spectral quantification. However, modeling the spatiotemporal distributions of metabolites as independent linear subspaces leads to a large number of unknowns, making the resultant subspace estimation problem challenging when the SNR of the MRSI data is low.

Manifold learning methods assume the data of interest reside in a low-dimensional manifold embedded in a high-dimensional space. Manifold learning methods have been applied to pattern recognition, classification, denoising, and image reconstruction in supervised, semi-supervised, and unsupervised settings, showing superior performance over the conventional linear model-based methods.[28-34] In MRSI, deep learning-based methods have been proposed to leverage the low-dimensional manifold structure of MRSI data using neural networks for image reconstruction[35] and signal separation.[36] The deep neural network-based methods require a large labeled dataset for network training, which is often obtained from population experimental MRSI data augmented with simulation data.

This work presents an unsupervised learning method that takes advantage of the intrinsic low-dimensional manifold structure of MRSI signals for joint spectral quantification. More specifically, we propose using a linear tangent space alignment (LTSA) model[37] to represent the spatiotemporal distribution of MRSI signals. In the proposed model, the spatiotemporal distribution of each metabolite is represented by a subspace model as in the subspace method. The key innovation of the proposed model is that the local coordinates of these subspaces can be aligned to the global coordinates of a low-dimensional manifold via linear transform when the underlying manifold is smooth and regular.[37] The proposed model significantly reduces the number of unknowns in spectral quantification compared to the subspace model. In the proposed method, the temporal bases of the LTSA model are estimated via quantum mechanics simulation. The global coordinates and the associated linear transform matrices are determined by fitting the LTSA model to noisy MRSI data while imposing spatial smoothness constraints on the global coordinates. We validated the proposed method using simulation data and in vivo MRSI data acquired from healthy volunteers on a 3T MR scanner. We compared the performance of the proposed method with the QUEST method[11] and the subspace method.[27]

2 | THEORY

2.1 | Spectral quantification problem

Without loss of generality, MRSI signals can be written in the $x-t$ domain as:

$$\mathbf{s}(\mathbf{x}_n, t_m) = \sum_{q=1}^{Q} \rho_q(\mathbf{x}_n, t_m) + \epsilon(\mathbf{x}_n, t_m), \tag{1}$$

where $\rho_q(\mathbf{x}_n, t_m)$ denotes the time domain signal of the qth metabolite at spatial location $\mathbf{x}_n$, Q is the total number of metabolites of interest, and $\epsilon(\mathbf{x}_n, t_m)$ is the measurement noise that is assumed to be independent and identically distributed (i.i.d.) Gaussian. When prior knowledge of the spectrum is available, $\rho_q(\mathbf{x}_n, t_m)$ can be expressed as

$$\rho_q(\mathbf{x}_n, t_m) = a_{q,n} e^{-i2\pi f_{q,n} t_m + t_m / T^*_{2,q,n}} \psi_q(t_m), \tag{2}$$

where $a_{q,n}, f_{q,n}$, and $T^*_{2,q,n}$ denote the amplitude, frequency shift, and T^*_2 relaxation time of the qth metabolite, respectively, and $\psi_q(t_m)$ is the time-domain basis function of the qth metabolite that can be obtained experimentally[9] or via quantum mechanics simulation.[11] Note that the signal model in Equation (2) can be generalized to a mixture of Lorentzian and Gaussian model to account for lineshape distortions in practical settings.[38]

Joint spectral quantification of the spectra from all MRSI voxels can be formulated as the following optimization problem:[26]

$$\begin{aligned} \underset{\left\{a_{q,n} f_{q,n}, T^*_{2,q,n}\right\}}{\arg\min} & \sum_{n=1}^{N}\sum_{m=1}^{M} \Big| \mathbf{s}(\mathbf{x}_n, t_m) \\ & - \sum_{q=1}^{Q} a_{q,n} e^{-i2\pi f_{q,n} t_m + t_m / T^*_{2,q,n}} \psi_q(t_m) \Big|_2^2 \\ & + R(a_{q,n}, f_{q,n}, T^*_{2,q,n}), \end{aligned} \tag{3}$$

where N and M are the total number of voxels and time-domain samples, respectively. The first term in Equation (3) is a data fidelity term. The second is a regularization term introduced to impose spatial prior knowledge (e.g., piece-wise smoothness) on the estimated parameter maps. Note that the optimization problem in Equation (3) can be solved voxel by voxel, for example, as in the QUEST method,[11] when the regularization term is dropped. The joint spectral quantification problem in Equation (3) is a large-scale, nonlinear, nonconvex optimization problem, which is difficult to solve efficiently.

This work adopts an alternative strategy first proposed by Li et al.[27] to tackle the spectral quantification problem in a two-step process:

1. Separation of each metabolite's signal from the noise corrupted measurement:

$$\begin{aligned} \{\hat{\rho}_q(\mathbf{x}_n, t_m)\}_{q,n,m} = \underset{\{\rho_q(\mathbf{x}_n, t_m)\}}{\arg\min} & \sum_{n=1}^{N}\sum_{m=1}^{M} \Big\| \mathbf{s}(\mathbf{x}_n, t_m) \\ & - \sum_{q=1}^{Q} \rho_q(\mathbf{x}_n, t_m) \Big\|_2^2. \end{aligned} \tag{4}$$

2. Voxel-wise estimation of spectral parameters (i.e., $a_{q,n}$, $f_{q,n}$, and $T^*_{2,q,n}$) from the separated spectrum:

$$\begin{aligned} \left\{\hat{a}_{q,n}, \hat{f}_{q,n}, \hat{T}^*_{2,q,n}\right\} = \underset{a_{q,n} f_{q,n}, T^*_{2,q,n}}{\arg\min} & \sum_{m=1}^{M} \Big\| \hat{\rho}_q(\mathbf{x}_n, t_m) \\ & - a_{q,n} e^{-i2\pi f_{q,n} t_m + t_m / T^*_{2,q,n}} \psi_q(t_m) \Big\|_2^2. \end{aligned} \tag{5}$$

Note that Equation (4) alone is an ill-posed optimization problem but can be solved with additional spatiotemporal constraints, that is, the subspace model in Reference 27, on the metabolite signals. The advantage of this strategy is that once each metabolite's signal is separated from the measured data, the following spectral quantification problem in Equation (5) is a voxel-wise optimization problem with only two nonlinear unknowns (i.e., $f_{q,n}$ and $T^*_{2,q,n}$). Such optimization problem can be efficiently solved using the VARPRO method.[39] The rest of the paper focuses on solving the optimization problem in Equation (4) by imposing a novel LTSA model of the metabolite signals.

2.2 | LTSA model of metabolite signal

We assume the metabolite signals $\{\rho_q(\mathbf{x}_n, t_m)\}_{q,n,m}$ reside in a D-dimensional ($D \ll M$) manifold $\mathcal{G}$:

$$\boldsymbol{\rho}_n = g(\boldsymbol{\tau}_n), \tag{6}$$

where $\boldsymbol{\rho}_n = [\rho_1(\mathbf{x}_n, t_1), \cdots, \rho_q(\mathbf{x}_n, t_m), \cdots, \rho_Q(\mathbf{x}_n, t_M)]^{\mathrm{T}} \in \mathbb{C}^{MQ\times 1}$ is a vector concatenating the time-domain signal of each metabolite at $\mathbf{x}_n$, $\boldsymbol{\tau}_n \in \mathbb{C}^{D\times 1}$ is a vector in the feature space, $^{\mathrm{T}}$ denotes the transpose operation, and g is a nonlinear function mapping the global coordinate of the manifold $\boldsymbol{\tau}_n$ to the time-domain signal $\boldsymbol{\rho}_n$. We further assume the manifold $\mathcal{G}$ is regular, that is, its Jacobian matrix is always of full rank D.[37]

The signals of the qth metabolite (i.e., $\boldsymbol{\rho}_{q,n} = [\rho_q(\mathbf{x}_n, t_1), \cdots, \rho_q(\mathbf{x}_n, t_M)]^{\mathrm{T}} \in \mathbb{C}^{M\times 1}, n = 1, \ldots, N$) form a neighborhood (i.e., neighborhood $q, q = 1, \ldots, Q$) in the frequency domain as illustrated in Figure 1A. This is most intuitive when there is no overlap between the spectra of different metabolites while the mathematical model

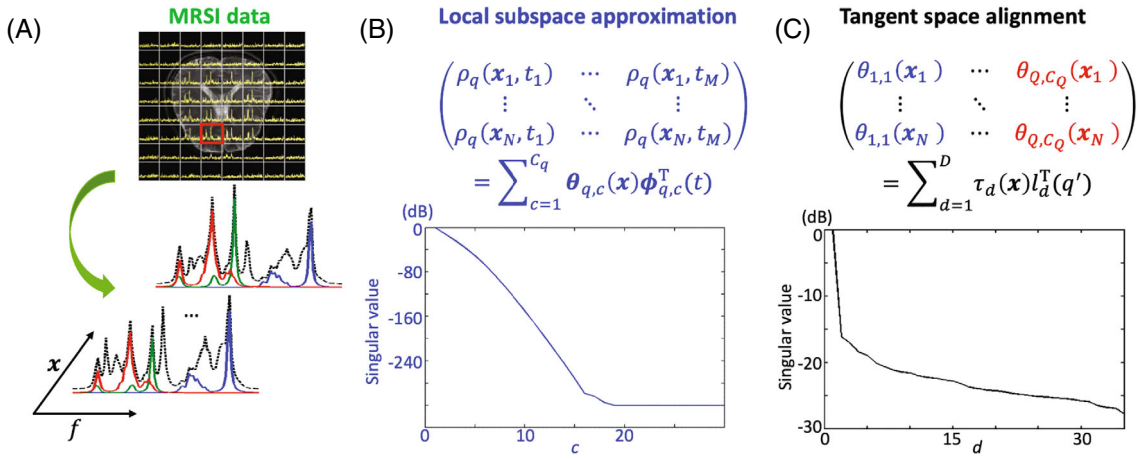


FIGURE 1 Schematic diagram of the proposed MR spectroscopic imaging (MRSI) spectral quantification method. We assume the spectra of metabolites from an MRSI experiment live in a low-dimensional smooth manifold. The intrinsic low-dimensional structure of the manifold can be learned conceptually as follows: (1) determination of the local coordinates of the manifold via local subspace approximation and (2) determination of the global coordinates of the manifold via tangent space alignment

derived below allows overlap between neighborhoods. Note that representing the signals in the time or frequency domain does not affect the underlying geometry of the signals since the Fourier transform is a unitary transformation. We choose to represent the signals in the time domain for the convenience of data processing.

If the underlying manifold $\mathcal{G}$ is smooth, the signals in qth the neighborhood can be approximated locally using the first-order Taylor expansion:

$$\boldsymbol{\rho}_{q,n} \approx \overline{\boldsymbol{\rho}}_q + \boldsymbol{\Pi}_q \mathcal{J}(\overline{\boldsymbol{\tau}}_q)(\boldsymbol{\tau}_n - \overline{\boldsymbol{\tau}}_q), \tag{7}$$

where $\overline{\boldsymbol{\rho}}_q$ denotes the centroid of the signals in the qth neighborhood, $\overline{\boldsymbol{\tau}}_q$ denotes the corresponding global coordinate of $\overline{\boldsymbol{\rho}}_q$, that is, $\overline{\boldsymbol{\rho}}_q = \boldsymbol{\Pi}_q g(\overline{\boldsymbol{\tau}}_q)$, $\boldsymbol{\Pi}_q \in \mathbb{R}^{M\times MQ}$ is a selection operator, and $\mathcal{J}(\overline{\boldsymbol{\tau}}_q) \in \mathbb{C}^{MQ\times D}$ denotes the Jacobian matrix of g evaluated at $\overline{\boldsymbol{\tau}}_q$.

Equation (7) indicates that the signals in the qth neighborhood reside in a local subspace, the dimension of which is upper bounded by the dimension of the underlying manifold, that is, less than $D+1$ accounting for the nonzero centroids of the metabolite signals. More specifically, the following Casorati matrix $\mathbf{P}_q \in \mathbb{C}^{N\times M}$ formed by the signals in the qth neighborhood:

$$\mathbf{P}_q = \begin{bmatrix} \rho_q(\mathbf{x}_1,t_1) & \cdots & \rho_q(\mathbf{x}_1,t_M) \\ \vdots & \ddots & \vdots \\ \rho_q(\mathbf{x}_N,t_1) & \cdots & \rho_q(\mathbf{x}_N,t_M) \end{bmatrix}, \tag{8}$$

can be approximated by a low-rank matrix via SVD:

$$\mathbf{P}_q = \sum_{c=1}^{C_q} \boldsymbol{\theta}_{q,c}(\mathbf{x})\boldsymbol{\phi}_{q,c}^{T}(t) = \boldsymbol{\Theta}_q \boldsymbol{\Phi}_q^{T}, \tag{9}$$

where C_q is the rank of the matrix ($C_q \le D+1$), $\boldsymbol{\phi}_{q,c}(t) = [\phi_{q,c}(t_1), \cdots, \phi_{q,c}(t_M)]^{T} \in \mathbb{C}^{M\times 1}$ denotes the temporal basis of the local subspace, and $\boldsymbol{\theta}_{q,c}(\mathbf{x}) = [\theta_{q,c}(\mathbf{x}_1), \cdots, \theta_{q,c}(\mathbf{x}_N)]^{T} \in \mathbb{C}^{N\times 1}$ denotes the corresponding local coordinates.

In theory, the temporal basis $\boldsymbol{\phi}_{q,c}(t)$ can be determined by explicitly calculating the Jacobian matrix of g, which is impractical, however, because g is defined implicitly. Alternatively, the temporal basis $\boldsymbol{\phi}_{q,c}(t)$ can be estimated using quantum mechanics simulation as in Reference 27. Briefly, instead of using in vivo MRSI signals, the Casorati matrix in Equation (8) is synthesized using simulation signals expressed as in Equation (2), where $\psi_q(t_m)$ is obtained by quantum mechanics simulation, $a_{q,n} = 1, f_{q,n}$, and $T^*_{2,q,n}$ are random variables uniformly distributed in a typical range matched to in vivo experiment setup. Figure 1B shows a representative plot of the singular values of such Casorati matrix, demonstrating the low-rank property of the matrix.

Equation (9) uses a local subspace model to represent signals in each neighborhood, resulting in a large number of unknowns when quantifying in vivo MRSI data. The key innovation of the proposed method is leveraging the fact that the local coordinates $\{\boldsymbol{\theta}_{q,c}(\mathbf{x})\}_{c=1}^{C_q}$ of the qth neighborhood can be aligned with the global coordinates of the manifold by a linear transform.

Substituting Equation (7) to Equation (8), yields

$$\begin{aligned} \mathbf{P}_q &= \mathbf{1}\left(\boldsymbol{\Pi}_q g(\overline{\boldsymbol{\tau}}_q) - \boldsymbol{\Pi}_q \mathcal{J}(\overline{\boldsymbol{\tau}}_q)\overline{\boldsymbol{\tau}}_q\right)^{T} + \tilde{\mathbf{T}}\left(\boldsymbol{\Pi}_q \mathcal{J}(\overline{\boldsymbol{\tau}}_q)\right)^{T} \\ &= \left[\tilde{\mathbf{T}}, \mathbf{1}\right]\left[\mathcal{J}(\overline{\boldsymbol{\tau}}_q), g(\overline{\boldsymbol{\tau}}_q) - \mathcal{J}(\overline{\boldsymbol{\tau}}_q)\overline{\boldsymbol{\tau}}_q\right]^{T}\boldsymbol{\Pi}_q^{T}, \\ &= \mathbf{T}\mathbf{J}_q^{T}\boldsymbol{\Pi}_q^{T}, \end{aligned} \tag{10}$$

where $\mathbf{1}$ is an N-dimensional column vector of all 1's, $\tilde{\mathbf{T}} = [\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_N]^{\mathrm{T}} \in \mathbb{C}^{N\times D}$ consists of the global coordinates of the manifold, and $\mathbf{T} \in \mathbb{C}^{N\times(D+1)}$ and $\mathbf{J}_q \in \mathbb{C}^{MQ\times(D+1)}$ are the augmented global coordinate and Jacobian matrix, respectively, as defined in Equation (10). Comparing Equation (10) with Equation (9) and leveraging the fact that $\boldsymbol{\Phi}_q$ consists of orthogonal bases, yields

$$\begin{aligned}\boldsymbol{\Theta}_q &= \mathbf{T}\mathbf{J}_q^{\mathrm{T}}\boldsymbol{\Pi}_q^{\mathrm{T}}\boldsymbol{\Phi}_q \\ &= \mathbf{T}\mathbf{L}_q,\end{aligned} \tag{11}$$

where and $\mathbf{L}_q \in \mathbb{C}^{(D+1)\times C_q}$ denotes a linear transform matrix for the alignment of local coordinates. Intuitively, Equation (11) leads to a low-rank representation of the Casorati matrix formed by the local coordinates, as illustrated in Figure 1C.

Substituting Equation (11) to Equation (9) yields the following LTSA model of the metabolite signal:

$$\mathbf{P}_q = \mathbf{T}\mathbf{L}_q\boldsymbol{\Phi}_q^{\mathrm{T}}, \tag{12}$$

where $q = 1, \dots, Q$. Assuming $C_q = D + 1$ for $\forall q$ and the temporal bases predetermined, the subspace model in Equation (9) requires the estimation of the local coordinates $\{\boldsymbol{\Theta}_q\}_{q=1}^{Q}$, resulting in a total number of unknowns of $N \times (D + 1) \times Q$. In comparison, the LTSA model in Equation (11) requires the estimation of the global coordinates T and a set of small matrices $\{L_q\}_{q=1}^{Q}$, resulting in a total number of unknowns of $N \times (D + 1) + (D + 1) \times (D + 1) \times Q$. Therefore, the proposed LTSA model significantly reduces the number of unknowns in spectral quantification compared to the subspace model.

2.3 | LTSA-based spectral quantification

Assuming the temporal basis matrices $\{\boldsymbol{\Phi}_q\}_{q=1}^{Q}$ are predetermined, the proposed spectral quantification is performed by solving the following constrained least-square optimization problem:

$$\begin{aligned}\underset{\{\mathbf{L}_q\},\mathbf{T}}{\arg\min} &\left\|\mathbf{S} - \sum_{q=1}^{Q}\mathbf{T}\mathbf{L}_q\boldsymbol{\Phi}_q^{\mathrm{T}}\right\|_{\mathrm{F}}^2 + \mu\sum_{q=1}^{Q}\|\mathrm{vec}\{\mathbf{L}_q\}\|_1 \\ &+ \lambda\|\mathbf{W}\mathbf{D}\mathbf{T}\|_{\mathrm{F}}^2,\end{aligned} \tag{13}$$

where $\mathbf{S}$ denotes the Casorati matrix formed by the measured data, $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm of a matrix, μ and λ are regularization parameters. The first term of Equation (13) is a data fidelity term; the second penalizes the sparsity of the linear transform matrices, leading to a sparse representation of the local coordinates that can further improve the performance of the proposed method; and the last penalizes the spatial smoothness of the global coordinates $\mathbf{T}$, where $\mathbf{D}$ is a finite difference operator in the space domain and $\mathbf{W}$ denotes edge weights derived from an anatomical reference image.[40,41]

We propose an Alternating Direction Methods of Multipliers- (ADMM)[42] based algorithm to solve the optimization problem in Equation (13), which solves the following sub-problems in an alternative fashion:

$$\mathbf{T}^{(k+1)} = \underset{\mathbf{T}}{\arg\min}\left\|\mathbf{S} - \sum_{q=1}^{Q}\mathbf{T}\mathbf{L}_q^{(k)}\boldsymbol{\Phi}_q^{\mathrm{T}}\right\|_{\mathrm{F}}^2 + \lambda\|\mathbf{W}\mathbf{D}\mathbf{T}\|_{\mathrm{F}}^2, \tag{14}$$

$$\begin{aligned}\left\{\mathbf{L}_q^{(k+1)}\right\}_{q=1}^{Q} = \underset{\{\mathbf{L}_q\}}{\arg\min}&\left\|\mathbf{S} - \sum_{q=1}^{Q}\mathbf{T}^{(k+1)}\mathbf{L}_q\boldsymbol{\Phi}_q^{\mathrm{T}}\right\|_{\mathrm{F}}^2 \\ &+ \frac{\beta}{2}\sum_{q=1}^{Q}\|\mathbf{L}_q - \mathbf{G}_q^{(k)} + \boldsymbol{\eta}_q^{(k)}\|_{\mathrm{F}}^2,\end{aligned} \tag{15}$$

$$\begin{aligned}\mathbf{G}_q^{(k+1)} = \underset{\mathbf{G}_q}{\arg\min}\,&\mu\|\mathrm{vec}\{\mathbf{G}_q\}\|_1 \\ &+ \frac{\beta}{2}\left\|\mathbf{L}_q^{(k+1)} - \mathbf{G}_q^{(k)} + \boldsymbol{\eta}_q^{(k)}\right\|_{\mathrm{F}}^2,\end{aligned} \tag{16}$$

$$\boldsymbol{\eta}_q^{(k+1)} = \boldsymbol{\eta}_q^{(k)} + \mathbf{L}_q^{(k+1)} - \mathbf{G}_q^{(k+1)}, \tag{17}$$

where $\mathbf{G}_q$ is the augmented Lagrangian split variable, $\boldsymbol{\eta}_q$ is the dual variable, and β is a relaxation parameter. Equations (14) and (15) are convex optimization problems, which are solved using the conjugate gradient method. Equation (16) is solved using soft thresholding. The convergence analysis of the algorithm can be found in Reference 43.

The optimization problem in Equation (13) is nonconvex because of the bilinearity of the LTSA model. Proper initialization is required to solve Equation (13) using the ADMM algorithm. In our implementation, we initialize the algorithm based on the subspace-based method.[27] More specifically, the local coordinates are estimated by solving the following optimization problem:

$$\left\{\hat{\boldsymbol{\Theta}}_q\right\}_{q=1}^{Q} = \underset{\{\boldsymbol{\Theta}_q\}}{\arg\min}\left\|\mathbf{S} - \sum_{q=1}^{Q}\boldsymbol{\Theta}_q\boldsymbol{\Phi}_q^{\mathrm{T}}\right\|_{\mathrm{F}}^2 + \lambda\sum_{q=1}^{Q}\|\mathbf{W}\mathbf{D}\boldsymbol{\Theta}_q\|_{\mathrm{F}}^2, \tag{18}$$

and then grouped into a Casorati matrix $\hat{\boldsymbol{\Theta}} = [\hat{\boldsymbol{\Theta}}_1, \cdots, \hat{\boldsymbol{\Theta}}_Q]$. Performing SVD on this Casorati matrix yields the initial guess of $\mathbf{T}$ and $\{\mathbf{L}_q\}_{q=1}^{Q}$. $\mathbf{G}_q$ and $\boldsymbol{\eta}_q$ are initialized as zero matrices.

3 | METHODS

We performed numerical simulation studies and in vivo proton MRSI (^{1}H-MRSI) experiments to validate the performance of the proposed method. Four healthy

volunteers (one for Section 3.2 and three for Section 3.3, two males and two females, 46 ± 12 years) were recruited under a study protocol approved by our local institutional review board. Written informed consent was obtained from all subjects. Imaging experiments were performed using a 3T MR scanner (Prisma, Siemens Healthcare) equipped with a 20-channel headcoil for reception.

For comparison, we implemented the QUEST method[11] by solving the optimization problem in Equation (3) without the regularization term and the subspace method[27] by solving the optimization problem in Equation (17). The time-domain basis functions of the three compared methods were obtained by quantum mechanical simulations using GAVA.[44]

3.1 | Performance evaluation using a numerical simulation phantom

We performed numerical simulation studies to evaluate the accuracy of the proposed spectral quantification method and compared it with the QUEST and the subspace methods. A two-dimensional (2D) numerical metabolite phantom was built for this purpose. MPRAGE images acquired from a healthy subject (Subject 2 in Section 3.3) were used to create the anatomical reference image and generate the probability maps of gray matter, white matter, and CSF using the SPM software.[45] GAVA[44] was used to simulate metabolite spectra of an FID-based ^{1}H-MRSI acquisition at 3T, including NAA, Cho, Cr, mIns, GABA, Glu, and Gln. The simulated spectra were assigned to gray matter and white matter based on the metabolite concentrations and T_2 relaxation times of the two tissues reported in the literature.[46] Random frequency shift uniformly distributed spatially between $[-5, 5]$ Hz was introduced in the numerical phantom to mimic the effect of inaccurate knowledge of the spectral position of each metabolite, which is typically caused by the residual B_0 inhomogeneity after B0 inhomogeneity correction and field drift compensation in practical MRSI experiments. T_2' time variations uniformly distributed between [60,100] ms were also added to the phantom.

The numerical phantom was used to simulate 2D MRSI data with 72×72 spatial encodings, 512 spectral encodings, and a sampling bandwidth of 2200 Hz. Gaussian noise was added to produce spectra with different SNRs. Metabolite concentration was calculated by integrating the magnitude of the separated metabolite spectrum in all the compared methods. The accuracy of the compared spectral quantification methods was measured by calculating the normalized root mean square error (NRMSE) of the estimated metabolite concentration maps.

In the subspace method, the dimension of the subspace of each metabolite signal (i.e., the rank of $\mathbf{\Phi}_q, q = 1, \ldots, 7$ in Equation (17) was 5, which was chosen based on the singular value decay of the Casorati matrix formed by the simulated metabolite signals.[27] The regularization parameter λ in Equation (17) that produced the smallest NRMSE was selected. In the proposed method, the dimension of the subspace of each metabolite signal was also 5. The dimension of the global coordinates (i.e., the rank of $\mathbf{T}$ in Equation (12) was 7. This was chosen based on the singular value decay of the Casorati matrix $\hat{\mathbf{\Theta}} = [\hat{\mathbf{\Theta}}_1, \cdots, \hat{\mathbf{\Theta}}_Q]$, which was formed using the local coordinates obtained by the subspace method. The regularization parameter λ in Equation (13) that produced the smallest NRMSE was selected. The regularization parameter μ was selected empirically, noting that the proposed method was not sensitive to the choice of μ in our experience. The typical computation time for spectral quantification using the QUEST, subspace, and LTSA methods was 400, 20, and 70 seconds, respectively, using 10 cores of an Intel Xeon Gold 6148 2.4 GHz CPU on a workstation.

3.2 | Performance evaluation using in vivo MRSI-based simulation

We further evaluated the performance of the proposed spectral quantification method using in vivo MRSI data acquired from a healthy volunteer. Because of the limited SNR at 3T, it is difficult to obtain high-resolution, high-SNR, ^{1}H-MRSI data using the conventional MRSI methods even with a large number of averages. To address this issue, we used the SPICE (SPectroscopic Imaging by exploiting spatiospectral CorrElation) method[41,47] to create reference MRSI data for comparison purposes.

More specifically, in the data acquisition of SPICE, high-resolution ^{1}H-MRSI (k, t)-space data were acquired using an FID, echo-planar-spectroscopic-imaging (EPSI) sequence with the following imaging parameters: field of view = 240×240 mm^2, slice thickness = 10 mm, pulse repetition time = 360 ms, echo time = 3.6 ms, flip angle = 45°, spatial encoding matrix = 64×64, spectral encoding size = 300, echo time = 0.8 ms (bipolar acquisition), sampling bandwidth = 100 kHz, and number of averages = 60. The EPSI acquisition was performed with WET pulses for water signal suppression but without lipid suppression. Low-resolution ^{1}H-MRSI (k, t)-space data were also acquired using an FID, phase encoding-based MRSI sequence to estimate the subspace basis functions in the SPICE image reconstruction.[41] The imaging parameters of the low-resolution MRSI data were the same as those of the high-resolution data except that spatial encoding matrix = 32×32 (elliptical sampling), spectral encoding

size = 460, sampling bandwidth = 2000 Hz, and eight outer volume suppression pulses for lipid suppression. Additional MPRAGE images were acquired as anatomical reference images for the SPICE image reconstruction.

In the data processing of SPICE, Fourier-transformed image reconstruction of the (k, t)-space data was first performed to measure and correct B_0 field inhomogeneities for each measurement using the residual water signals. Coil combination was performed using the SVD-based method in Reference 48. The B_0 inhomogeneity corrected, coil combined, MRSI data were then averaged and transformed back to the (k, t)-space. We followed the union-of-subspaces method in References 49,50 to remove the nuisance water and lipid signals as well as the baseline macromolecular signals from the averaged (k, t)-space data. The reference MRSI data were reconstructed using the SPICE method from the nuisance signal removed data.

The reference concentration maps of NAA, Cho, Cr, mIns, and Glx (Gln+Glu), were generated by performing spectral quantification on the reference MRSI data using the proposed method. This was because the metabolite concentration maps obtained by the proposed method showed fewer visible artifacts compared to the QUEST and the subspace methods (see Sections 4 and 5 for more details). Gaussian noise was then added to the reference MRSI data to create MRSI data with lower SNRs for evaluating the performance of the proposed method and the comparisons with the QUEST and the subspace methods. The metabolite concentration was calculated by integrating the magnitude of the separated metabolite spectrum as in the simulation study. In the subspace and the proposed methods, the dimension of the subspace of each metabolite signal was 5. The dimension of the global coordinates was 10. These parameters and the other regularization parameters were selected the same way as described in Section 3.1.

3.3 | Application to spectral quantification of SPICE reconstruction

Sections 3.1 and 3.2 aim at evaluating the performance of the proposed method in quantifying MRSI data reconstructed by the conventional Fourier transform-based methods, where the noise of the MRSI data is considered as Gaussian noise uniformly distributed in the $x - t$ domain. SPICE is an emerging technique for high-resolution MRSI but has noise characteristics different from the conventional Fourier transform-based reconstruction methods due to the low-rank based image reconstruction.[50] More specifically, the spectrum reconstructed by SPICE has high SNR because the spectral basis of the low-rank model that determines the SNR of the reconstructed spectrum is estimated from high-SNR training data. The noise of the measured (k, t)-space data mainly contribute to the spatial variations of the MRSI data reconstructed by SPICE. In this Section, in vivo ^{1}H-MRSI experimental studies were performed to demonstrate the usefulness of the proposed method in quantifying high-resolution MRSI data reconstructed by the SPICE method.

To this end, three healthy volunteers were recruited. The imaging protocol included five consecutive EPSI scans, each acquiring high-resolution ^{1}H-MRSI (k, t)-space data with the same imaging parameters as described in Section 3.2 except that the spectral encoding size was 320 and the number of averages was 24. Low-resolution ^{1}H-MRSI (k, t)-space data and anatomical reference images were also acquired as described in Section 3.2.

The reference MRSI data were created from the B_0 inhomogeneity corrected, coil combined, nuisance signal removed (k, t)-space data averaged out of all the 120 measurements following the procedure in Section 3.2. The reference concentration maps of NAA, Cho, Cr, mIns, and Glx (Gln+Glu) were generated by performing spectral quantification on the reference MRSI data using the proposed method. Four independent MRSI (k,t)-space datasets were created by selecting and averaging out 6 of the 24 measurements from each of the five EPSI scans (i.e., $6 \times 5 = 30$ averages in total) and were then used to reconstruct MRSI data at lower SNRs using SPICE. Lastly, spectral quantification was performed on the SPICE reconstructed MRSI data using the three compared spectral quantification methods. In the subspace and the proposed methods, the dimension of the subspace of each metabolite signal was 5. The dimension of the global coordinates was 10. These parameters and the other regularization parameters were selected the same way as described in Section 3.1.

4 | RESULTS

4.1 | Numerical simulation results

MRSI spectral quantification results obtained using the numerical phantom described in Section 3.1 are shown in Figures 2 and 3, where the average SNR of the NAA peak was set to be 10. At the individual voxel level, the three compared methods achieved similar quantification accuracy, as shown in the residual signal plots in Figure 3. However, the metabolite concentration maps estimated by the voxel-wise QUEST method showed larger spatial variations than the joint spectral quantification methods, that is, the subspace and the proposed methods, as shown in

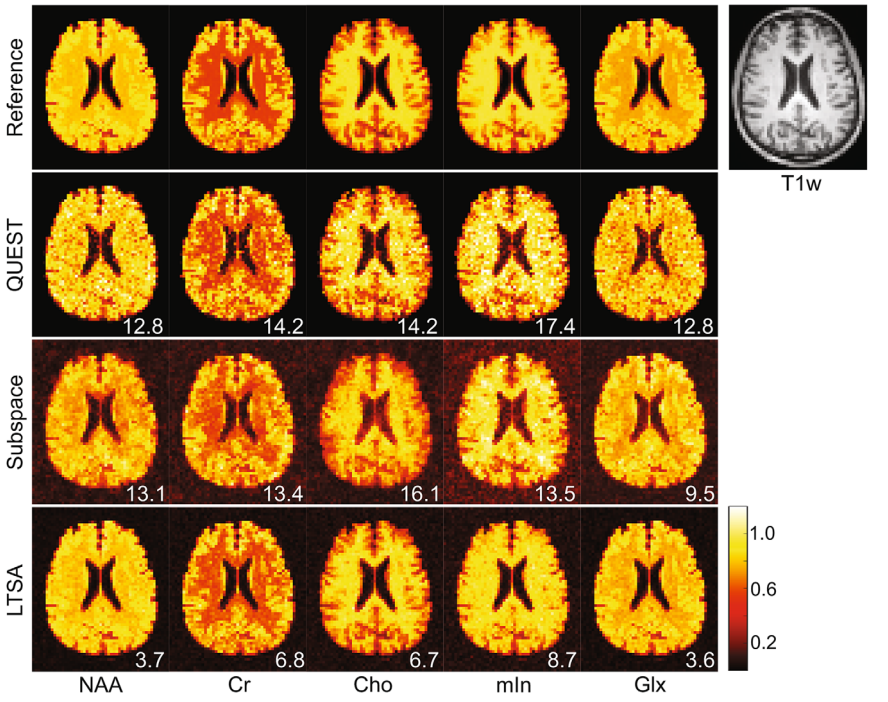


FIGURE 2 Spectral quantification results obtained using the numerical phantom: Comparison of the metabolite concentration maps estimated by the QUEST, the subspace, and the proposed methods. The corresponding normalized root mean square errors (in percentage) are listed in the bottom right corner of each subfigure.

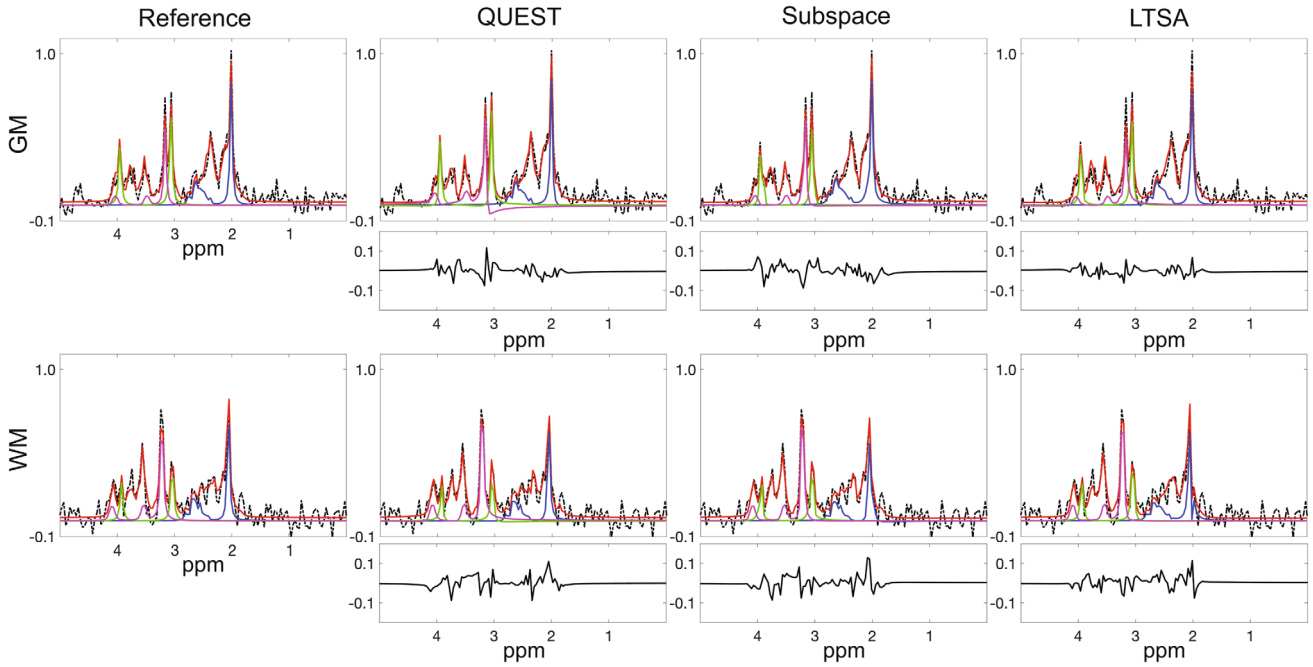


FIGURE 3 Spectral quantification results obtained using the numerical phantom: Spectral fitting results for representative spectra from voxels in the gray matter (GM) and white matter (WM), respectively. The color coding is as follows. Black-dashed line: the real part of the noisy spectrum; red-solid line: the sum of the fitted spectra; blue-, magenta-, and green-solid line: the fitted spectrum of the NAA, Cho, and Cr, respectively. A plot of the difference between the sum of the fitted spectra and the ground truth is shown under each spectrum plot in black color.

Figure 2. This comparison highlighted the advantages of incorporating spatial constraints into MRSI spectral quantification. Compared to the subspace method, the proposed method further improved quantification accuracy qualitatively as shown in Figure 2 and quantitatively as measured by the NRMSEs of the estimated metabolite concentration maps. Supporting Information Figures S1 and S2 show the quantification results and the corresponding linear transform matrices with and without a sparsity constraint on the linear transform matrices, demonstrating the benefits of promoting the sparsity of L_q in Equation (13).

4.2 | In vivo MRSI-based simulation results

MRSI spectral quantification results obtained using the in vivo MRSI data described in Section 3.2 are shown in Figures 4 and 5, where the average SNR of the NAA peak was set to be 10. Consistent with the simulation study results, the three compared methods achieved similar quantification accuracy at the individual voxel level, as shown in Figure 5. Compared to the reference metabolite maps, the metabolite maps estimated by the QUEST method (the second row of Figure 4) showed large spatial variations, most noticeably in the estimated mIns and Glx maps. There were also noticeable biases in the estimated Cho and Cr maps. In comparison, the subspace method (the third row of Figure 4) reduced the spatial variations in the estimated metabolite concentration maps. However, the concentration maps obtained

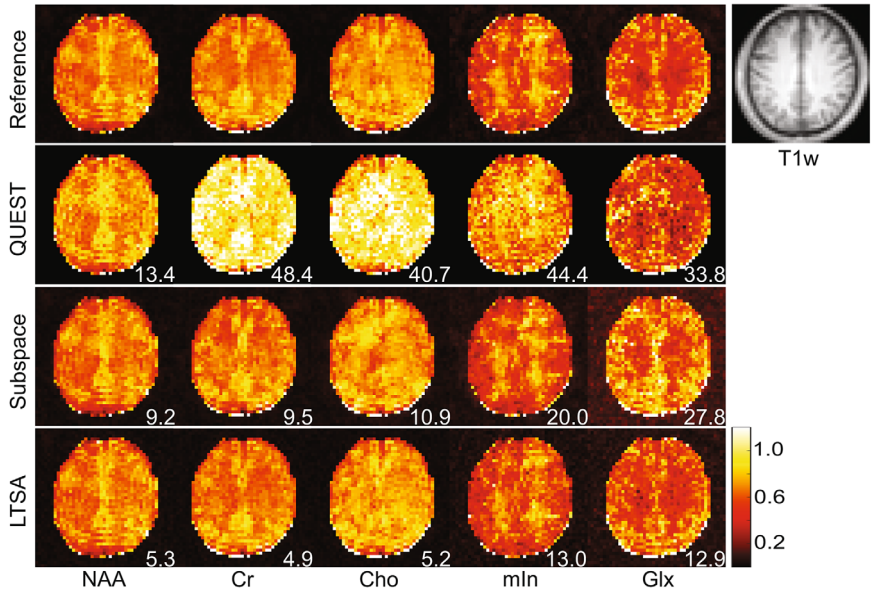


FIGURE 4 Spectral quantification results obtained using the in vivo MR spectroscopic imaging data: Comparison of the metabolite concentration maps estimated by the QUEST, the subspace, and the proposed methods. The corresponding normalized root mean square errors (in percentage) are listed in the bottom right corner of each subfigure.

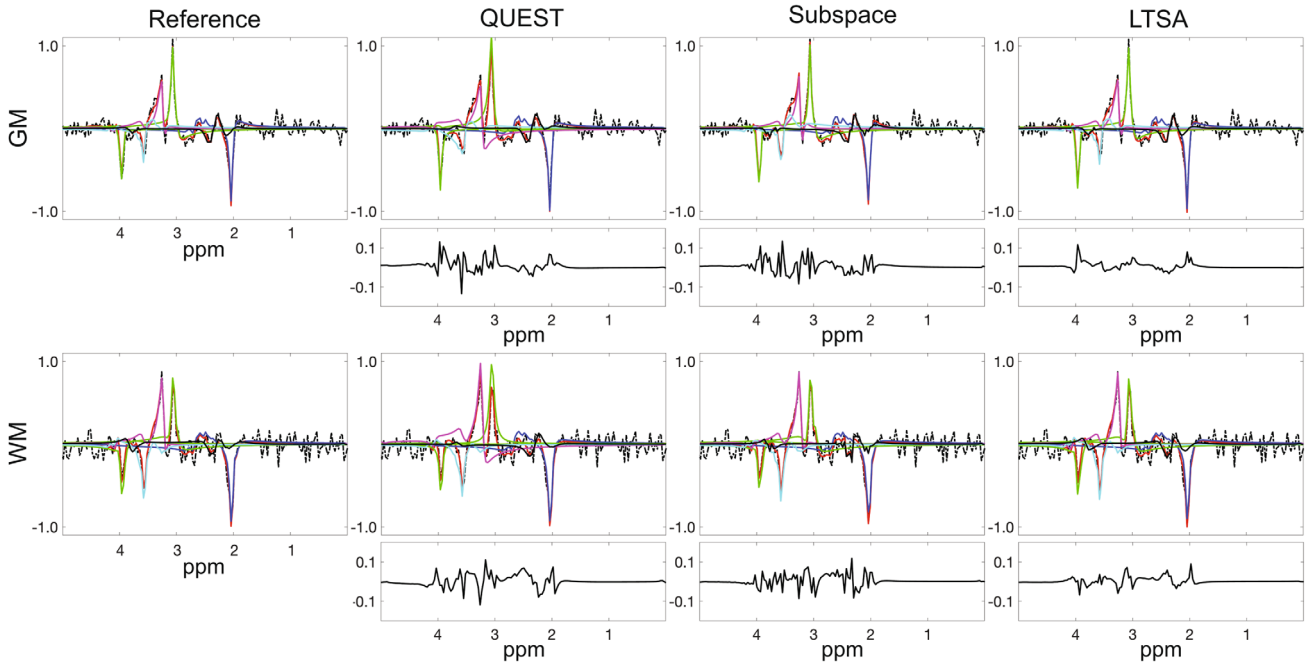


FIGURE 5 Spectral quantification results obtained using the in vivo MR spectroscopic imaging data: Spectral fitting results for representative spectra from voxels in the gray matter (GM) and white matter (WM), respectively. The color coding is as follows. Black-dashed line: the real part of the noisy spectrum; red-solid line: the sum of the fitted spectra; blue-, magenta-, green-, cyan-, and black-solid line: the fitted spectrum of the NAA, Cho, Cr, mIn, and Glx, respectively. A plot of the difference between the sum of the fitted spectra and the ground truth is shown under each spectrum plot in black color.

by the subspace method showed spatial blurring artifacts most noticeably in the estimated Cho map. Among the three compared methods, the proposed method achieved the best quantification accuracy as measured by the NRMSEs of the estimated metabolite concentration maps.

Monte Carlo simulations were performed using the in vivo MRSI data described in Section 3.2 to investigate the bias and standard deviation (SD) of the metabolite concentration maps estimated by the compared methods. Spectral quantification by different methods was repeated with 100 noise realizations in the simulation. The results are shown in Figure 6, where the results of Cr were similar to those of Cho and were thus not shown for better readability of the figure. As can be seen, the proposed method achieved the smallest SDs in the estimated metabolite concentration maps among the compared methods, which was consistent with the observations from Figure 4. Since the reference maps were created using the proposed method, comparing the bias maps favored the proposed method as expected. However, the bias by the subspace and the proposed methods were overall at a similar level. The QUEST method led to significantly larger bias and SDs than the subspace and the proposed methods (note the difference in the color-bar scales). The bias of the QUEST method might be caused by the mismatch between the metabolite basis functions and the in vivo spectra. We performed the same Monte Carlo simulation study using the numerical phantom described in Section 3.1. The results are shown in Supporting Information Figure S3. As expected, the bias of the QUEST method was much smaller than that of the in vivo data-based study because there was no mismatch between the metabolite basis functions and the simulated spectra. The proposed method achieved superior performance over the compared methods as shown in the maps of bias and SDs in Supporting Information Figure S3, which was consistent with the observations from Figure 6.

Figure 7 and Suppporting Information Figure S4 show the averaged bias and SD of the compared methods obtained by Monte Carlo simulations at three SNR levels (i.e., SNR = 5, 10, and 15) using

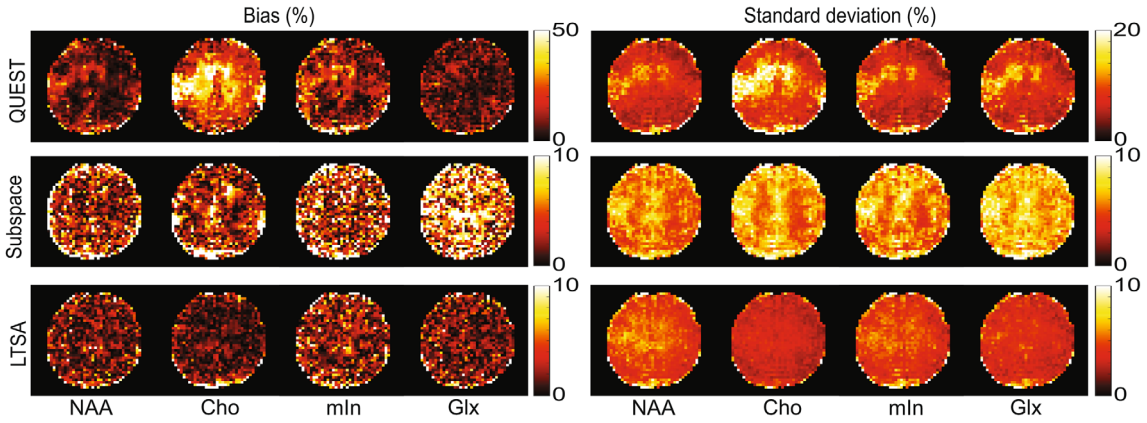


FIGURE 6 Monte Carlo simulation results obtained using the in vivo MR spectroscopic imaging data: Bias and SD maps of the metabolite concentration maps obtained by the QUEST, the subspace, and the proposed method, respectively. The average signal-to-noise ratio of the NAA peak was fixed to 10 for this simulation.

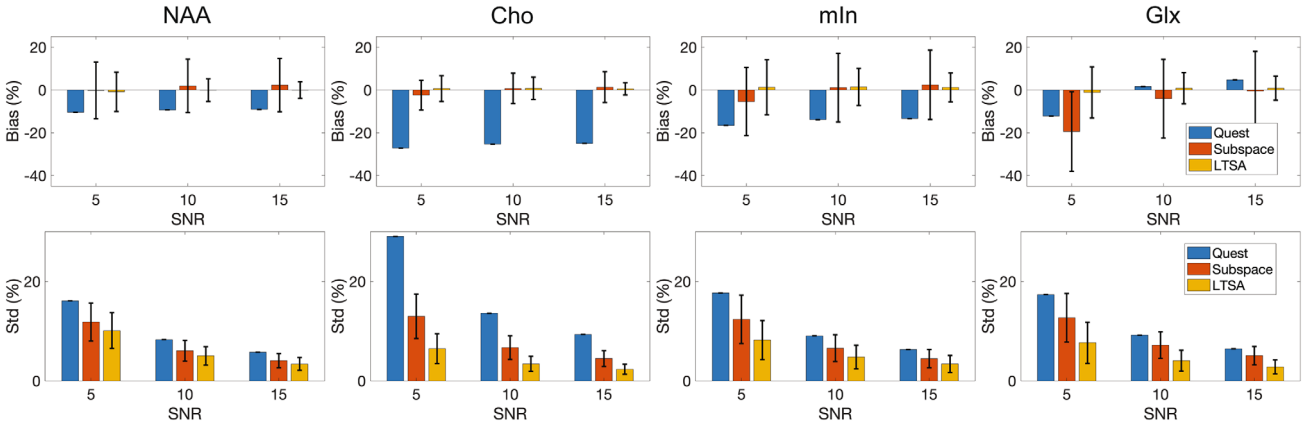


FIGURE 7 Monte Carlo simulation results obtained using the in vivo MR spectroscopic imaging data: Averaged bias and standard deviation achieved by the QUEST, the subspace, and the proposed method, respectively, at three signal-to-noise ratio levels. Note that the error-bars of the bias and SD of the QUEST method were off the charts and thus omitted.

the in vivo MRSI data and the numerical phantom, respectively. The proposed method reduced the bias and SDs in all the investigated cases compared to the QUEST and the subspace methods.

4.3 | Spectral quantification of SPICE reconstruction

Figures 8 and 9 show the in vivo results of spectral quantification for SPICE from one healthy subject (Subject 2). Figure 8 shows the metabolite concentration maps estimated by the compared methods. The subspace and the proposed methods significantly improved the accuracy of spectral quantification for the SPICE reconstructed MRSI data compared to the QUEST method. The metabolite concentration maps estimated by the subspace method showed blurring artifacts, most noticeably in the Cho and mIn maps. The proposed method achieved the best performance in spectral quantification for SPICE. As a side note, the loss of metabolite signal near the skull at the bottom of the image was likely due

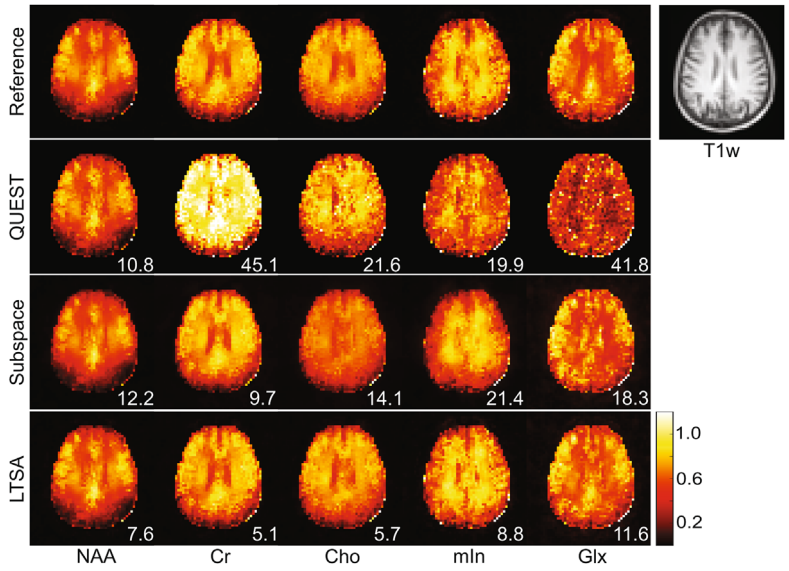


FIGURE 8 Spectral quantification for SPICE (Subject 2): Comparison of the metabolite concentration maps estimated by the QUEST, the subspace, and the proposed methods. The corresponding normalized root mean square errors (in percentage) are listed in the bottom right corner of each subfigure.

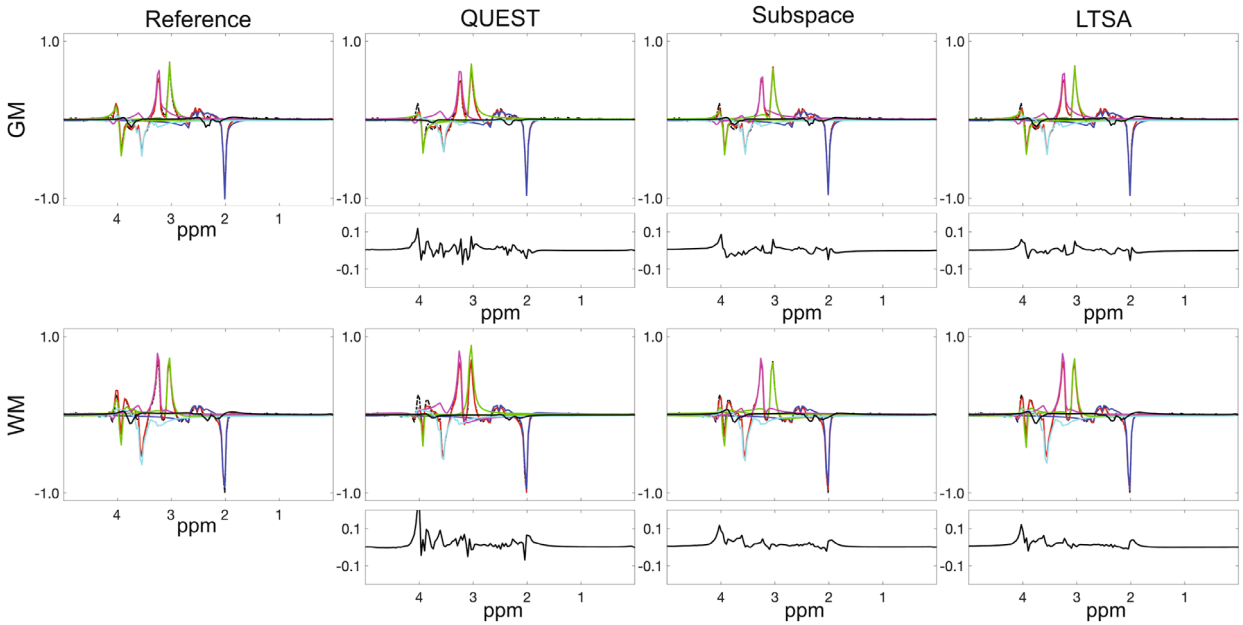


FIGURE 9 Spectral quantification for SPICE (Subject 2): Spectral fitting results for representative spectra from voxels in the gray matter (GM) and white matter (WM), respectively. The color coding is as follows. Black-dashed line: the real part of the SPICE reconstructed spectrum; red-solid line: the sum of the fitted spectra; blue-, magenta-, green-, cyan-, and black-solid line: the fitted spectrum of the NAA, Cho, Cr, mIn, and Glx, respectively. A plot of the difference between the sum of the fitted spectra and the ground truth is shown under each spectrum plot in black color.

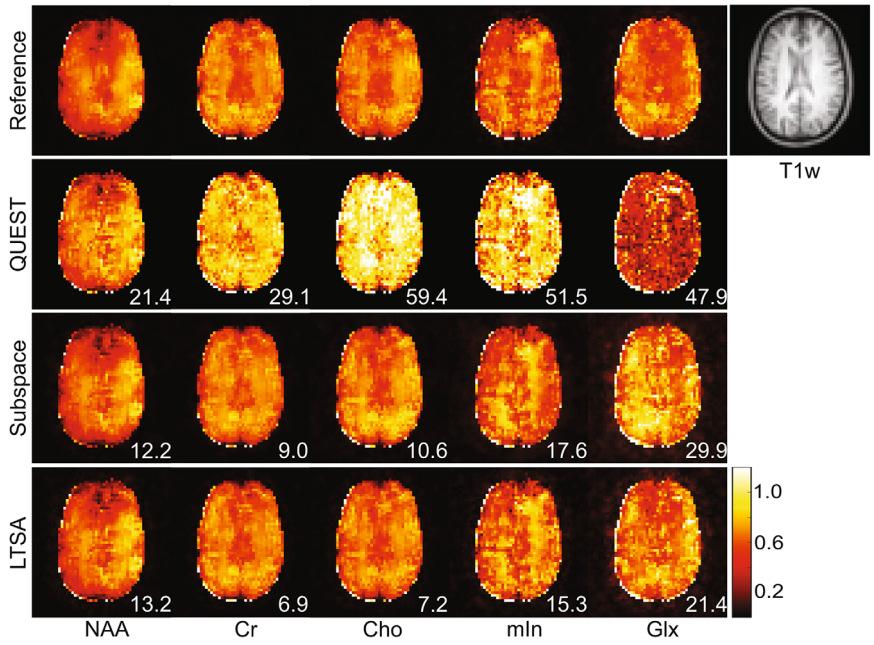


FIGURE 10 Spectral quantification for SPICE (Subject 3): Comparison of the metabolite concentration maps estimated by the QUEST, the subspace, and the proposed methods. The corresponding normalized root mean square errors (in percentage) are listed in the bottom right corner of each subfigure.

to overly aggressive nuisance signal removal. Figure 9 shows spectral fitting results for representative spectra from voxels in the gray matter and white matter, respectively. Figure 10 shows the metabolite concentration maps estimated by the compared methods from another healthy subject (Subject 3). Similar to the results in Figure 8, the proposed method achieved the best performance in spectral quantification among the compared methods. Note that besides the spatial blurring artifacts, the subspace method also showed noticeable artifacts in the estimated Glx map. Supporting Information Figure S5 shows the NRMSEs of the estimated metabolite concentration maps achieved by the compared methods among the three recruited healthy volunteers. The proposed method achieved the most accurate spectral quantification for all the subjects. The improvement offered by the proposed method over the QUEST and the subspace methods were most significant in the estimated metabolite concentration maps of Cr, Cho, and Glx.

5 | DISCUSSIONS

5.1 | Comparison with the QUEST and subspace methods

This work presents a novel manifold learning-based method for MRSI spectral quantification. The MRSI signals reside in an intrinsically low dimensional manifold, which can be justified by the parametric model in Equation (2) that employs a small amount of nonlinear parameters to represent the metabolite signals over the entire FID. However, using this nonlinear parametric model for joint spectral quantification leads to a large-scale, nonlinear, nonconvex optimization problem that is difficult to solve efficiently.[26] The subspace method[27] addresses this issue by representing signals of each metabolite as a subspace, resulting in a convex optimization problem that can be solved efficiently with a global minimum. However, since the subspace of each metabolite is modeled independently, the number of unknowns of the subspace method increases quickly as the number of metabolites of interest increases, which is often the case in in vivo proton MRSI. Compared to the subspace method, the proposed method introduces a bi-linear LTSA model of metabolite signals that aligns the local coordinates of each subspace to the global coordinates of the underlying manifold via linear transforms represented by small-size matrices. As a result, the number of unknowns of the proposed method remains roughly independent of the number of metabolites. The significantly reduced amount of unknowns in the proposed method leads to more accurate spectral quantification than the subspace method, as can be seen in the results of the numerical simulation studies and the in vivo ^{1}H-MRSI experiments in Section 4.

5.2 | Nonideal spectral lineshapes

Although a less severe problem in high-resolution MRSI, intra-voxel field inhomogeneities introduce lineshape distortions to spectral signals. The proposed method is not limited to the Lorentzian spectral model. More specifically, in the first step of the proposed method, each metabolite's

signal is separated from the noise corrupted measurement using the proposed LTSA model by solving the optimization problem in Equation (13). Although the temporal basis of the LTSA model is estimated from a Casorati matrix formed by spectral signals with Lorentzian shapes, the subspace spanned by the estimated basis functions can represent spectral signals with non-Lorentzian shapes. In the second step, spectral parameters are estimated voxel-by-voxel from the separated spectrum by solving the optimization problem in Equation (5). When needed, Equation (5) can be generalized to a mixture of Lorentzian and Gaussian models to account for lineshape distortions. In the current work, we focused on solving the spectral signal separation problem in the first step using the proposed LTSA model. Exploring how to properly handle the non-ideal spectral lineshapes in the voxel-wise spectral parameter estimation problem of the second step is beyond the scope of the current work.

To illustrate that the subspace model of spectral signals can represent Lorentzian and non-Lorentzian spectral signals, a set of six temporal basis functions were estimated from a Casorati matrix formed by Cr signals with Lorentzian lineshapes as described in Section 2.2. The estimated temporal basis functions were then used to approximate a signal with a Lorentzian lineshape in the form of $s_L(t) = e^{-t/T_2 - t/T_{2,p}} \sum_{q=1}^{2} e^{-i2\pi f_q t}$, where $T_2 = 200$ ms and $T_{2,p} = 50$ ms, and a signal with a mixture of Lorentzian and Gaussian lineshape (i.e., the Voigt model[38]) in the form of $s_V(t) = e^{-t/T_2 - t^2/g} \sum_{q=1}^{2} e^{-i2\pi f_q t}$, where $g = 5000$ ms^2. The time- and frequency-domain signals of the Lorentzian and non-Lorentzian signals are shown in Supporting Information Figure S6A,B and Figure S6D,E, respectively. The corresponding approximation errors are shown as the insert at the bottom of each subfigure. Finally, Supporting Information Figure S6C,F plots the NRMSEs in percentage when the temporal basis functions were used to approximate the Lorentzian signals with $T_{2,p}$ in the range of [50,150] ms and the non-Lorentzian signals with g in the range of $[5000, 15000]$ ms^2, respectively. As can be seen, the subspace model can accurately (i.e., NRMSEs < 1 %) represent spectral signals with Lorentzian or non-Lorentzian lineshapes and with a large range of linewidths.

5.3 | Residual nuisance signals

Nuisance signal removal is often done in the MRSI data preprocessing step and/or image reconstruction step.[41,49,51] Residual nuisance signals in the MRSI data can affect the accuracy of the proposed LTSA-based spectral quantification method because the spectra of residual nuisance signals are not orthogonal to the subspaces of the metabolites. In this case, the proposed method can be extended with additional subspaces to account for the residual nuisance signals in the MRSI data.

More specifically, we write the Casorati matrix formed by the MRSI data in the presence of nuisance signals as following:

$$\begin{aligned}\mathbf{P} &= \mathbf{P}_{\text{meta}} + \mathbf{P}_{\text{nuisance}} \\ &= \sum_{q=1}^{Q} \mathbf{T}\mathbf{L}_q \mathbf{\Phi}_q^{\mathrm{T}} + \mathbf{\Theta}_{\text{ns}} \mathbf{\Phi}_{\text{ns}}^{\mathrm{T}},\end{aligned} \tag{19}$$

where the nuisance signals are represented as a subspace model (i.e., $\mathbf{P}_{\text{nuisance}} = \mathbf{\Theta}_{\text{ns}}\mathbf{\Phi}_{\text{ns}}^{\mathrm{T}}$) as in Reference 49, $\mathbf{\Phi}_{\text{ns}}$ denotes the temporal basis of the nuisance signals, and $\mathbf{\Theta}_{\text{ns}}$ denotes the corresponding spatial coefficients. In Equation (19), the nuisance signals are modeled by a separate subspace because the spatiotemporal distribution of the nuisance signals can be affected by the suppression techniques used the MRSI data acquisition and thus may not belong to the underlying manifold of the metabolite signals.

Spectral quantification using the extended model in Equation (19) is performed by solving the following optimization problem:

$$\begin{aligned}\underset{\{\mathbf{L}_q\},\mathbf{T},\mathbf{\Theta}_{\text{ns}}}{\arg\min} &\left\| \mathbf{S} - \left(\sum_{q=1}^{Q} \mathbf{T}\mathbf{L}_q \mathbf{\Phi}_q^{\mathrm{T}} + \mathbf{\Theta}_{\text{ns}} \mathbf{\Phi}_{\text{ns}}^{\mathrm{T}} \right) \right\|_{\mathrm{F}}^2 \\ &+ \mu \sum_{q=1}^{Q} \|\text{vec}\{\mathbf{L}_q\}\|_1 + \lambda_1 \|\mathbf{W}\mathbf{D}\mathbf{T}\|_{\mathrm{F}}^2 + \lambda_2 \|\mathbf{W}\mathbf{D}\mathbf{\Theta}_{\text{ns}}\|_{\mathrm{F}}^2.\end{aligned} \tag{20}$$

The optimization problem in Equation (20) can be solved using a strategy that alternates the following two steps: (1) minimizing the cost function over $\mathbf{L}_q$ and $\mathbf{T}$ with $\mathbf{\Theta}_{\text{ns}}$ fixed, which can be done through the algorithm described in Equations (14)–(17); and (2) minimizing the cost function over $\mathbf{\Theta}_{\text{ns}}$ with $\mathbf{L}_q$ and $\mathbf{T}$ fixed, which leads to the same least-square optimization problem as in the subspace method.

A numerical simulation study was performed to illustrate the effects of residual water signals on the accuracy of spectral quantification using the proposed LTSA method (i.e., solving the optimization problem in Equation (13)) and the extended LTSA method (i.e., solving the optimization problem in Equation (20)). The water signals were extracted from EPSI data using the Hankel Singular Value Decomposition method,[4] which were acquired without water suppression in the same imaging session where the MPRAGE images were collected to create the numerical phantom in Section 3.1. The extracted water signals were added to the numerical phantom at different levels to evaluate the effects of the water signals on spectral quantification using the

proposed method. The temporal basis of the water signals in the extended LTSA model was obtained in the same way as the metabolite signals. The rank of the water signals was 6.

The results are shown in Supporting Information Figures S7–S11. As expected, the quantification error of the LTSA method increased as the level of residual water signals increased. The quantification accuracy of mIn was most significantly affected by the residual water signals because mIn's spectrum is closest to that of water. The robustness to residual water signals was significantly improved with the extended LTSA method that used an additional subspace to account for the residual water signals. However, a full-fledged evaluation of the robustness of the extended LTSA method to residual water and lipid signals in in vivo experiments is beyond the scope of this current work.

5.4 | Limitation and other future work

One limitation of our current work is that the reference metabolite concentration maps in Sections 3.2 and 3.3 were generated using the proposed method, which made the comparisons among the three spectral quantification methods favor the proposed method in terms of NRMSE and bias. Ideally, the QUEST method should be used as the reference method because it is unbiased when the basis functions of metabolite signals match the experiment setup. However, because of the limited SNR of ^{1}H-MRSI at 3T, applying QUEST to high-resolution MRSI data resulted in very noisy metabolite concentration maps as illustrated in Figures 4, 8, and 10. Reducing resolution will not fundamentally address this issue and mitigate the merits of joint spectral quantification, which has clear advantage over the voxel-wise spectral quantification methods supported by the results in this work and the literature.[23-26] On the other hand, the metabolite concentration maps estimated by the subspace method showed noticeable artifacts as illustrated in Figures 4, 8, and 10. Nevertheless, in the in vivo data-based Monte Carlo simulation study as described in Section 3.2, the proposed method achieved significantly reduced SD compared to the subspace method as shown in Figures 6 and 7. The bias of the proposed method was smaller than the subspace method as expected but was overall at a similar level. In addition, when applying the proposed method to spectral quantification for SPICE as described in Section 3.3, the resultant metabolite concentration maps showed the least noticeable artifacts.

There are two other future directions that warrant further investigations. First, the macromolecular signals which were removed along with the residual nuisance signals were not considered in our current work. However, the macromolecular signals contain important physiological and pathological information.[52,53] The proposed method can be extended to quantification of macromolecular signals by including the temporal basis of the macromolecular signals, which can be obtained using in vivo experiments[54-56] or using simulated macromolecular components.[57,58] Second, the proposed LTSA model can enable direct estimation of metabolite concentration maps from noise-corrupted and undersampled (k, t)-space data by including Fourier transform and (k,t)-space sampling operators to the data fidelity term of Equation (13). This is an ongoing research and will be reported in a future publication.

6 | CONCLUSIONS

Joint spectral quantification using LTSA-based manifold learning improves the accuracy of MRSI spectral quantification.

ACKNOWLEDGMENTS

This work was supported in part by the National Institutes of Health (grant numbers: R01CA165221, P41EB022544, K01EB030045, and R01EB033582). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

DATA AVAILABILITY STATEMENT

The MATLAB code used for the proposed spectral quantification method can be downloaded at https://gordon.mgh.harvard.edu/ltsa_mri.zip. The numerical phantom used in the simulation study of this work can be downloaded at https://gordon.mgh.harvard.edu/ltsa_mri.zip. The MATLAB code and the numerical phantom can also be accessed by directly contacting the corresponding author.

ORCID

Chao Ma https://orcid.org/0000-0002-2016-8084
Paul Kyu Han https://orcid.org/0000-0003-1650-1347
Yue Zhuo https://orcid.org/0000-0002-5739-1896
Yanis Djebra https://orcid.org/0000-0003-0756-571X
Thibault Marin https://orcid.org/0000-0002-8669-1003
Georges El Fakhri https://orcid.org/0000-0002-9005-6993

REFERENCES

1. Near J, Harris AD, Juchem C, et al. Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: experts' consensus recommendations. *NMR Biomed*. 2021;34:e4257.

2. Maudsley AA, Andronesi OC, Barker PB, et al. Advanced magnetic resonance spectroscopic neuroimaging: experts' consensus recommendations. *NMR Biomed*. 2021;34:e4309.
3. Lin A, Andronesi O, Bogner W, et al. Minimum reporting standards for in vivo magnetic resonance spectroscopy (MRSinMRS): experts' consensus recommendations. *NMR Biomed*. 2021;34:e4484.
4. Barkhuijsen H, Beer R, Ormondt D. Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals. *J Magn Reson*. 1987;73:553-557.
5. Pijnappel WW, Boogaart A, Beer R, Ormondt D. SVD-based quantification of magnetic resonance signals. *J Magn Reson*. 1992;97:122-134.
6. Barkhuijsen H, Beer R, Bovée WMM, Ormondt D. Retrieval of frequencies, amplitudes, damping factors, and phases from time-domain signals using a linear least-squares procedure. *J Magn Reson*. 1985;61:465-481.
7. Vanhuffel S, Chen H, Decanniere C, Hecke P. Algorithm for time-domain NMR data fitting based on total least squares. *J Magn Reson*. 1994;110:228-237.
8. Veen JW, Beer R, Luyten PR, Ormondt D. Accurate quantification of in vivo 31P NMR signals using the variable projection method and prior knowledge. *Magn Reson Med*. 1988;6:92-98.
9. Provencher SW. Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magn Reson Med*. 1993;30:672-679.
10. Vanhamme L, Boogaart A, Huffel S. Improved method for accurate and efficient quantification of mrs data with use of prior knowledge. *J Magn Reson*. 1997;129:35-43.
11. Ratiney H, Sdika M, Coenradie Y, Cavassila S, Ormondt D, Graveron-Demilly D. Time-domain semi-parametric estimation based on a metabolite basis set. *NMR Biomed*. 2005;18:1-13.
12. Poullet JB, Sima DM, Simonetti AW, et al. An automated quantification of short echo time MRS spectra in an open source software environment: AQSES. *NMR Biomed*. 2007;20:493-504.
13. Bhat H, Sajja BR, Narayana PA. Fast quantification of proton magnetic resonance spectroscopic imaging with artificial neural networks. *J Magn Reson*. 2006;183:110-122.
14. Hatami N, Sdika M, Ratiney H. Magnetic resonance spectroscopy quantification using deep learning. *Lecture Notes in Computer Science–2018 MICCA*. Springer; 2018:467-475.
15. Courvoisier S, Klauser A, Lichard P, Kocher M, Lazeyras F. High-resolution magnetic resonance spectroscopic imaging quantification by convolutional neural network. Proceedings of the 27th Annual Meeting of the ISMRM; 2019:0519.
16. Gurbani SS, Sheriff S, Maudsley AA, Shim H, Cooper LA. Incorporation of a spectral model in a convolutional neural network for accelerated spectral fitting. *Magn Reson Med*. 2019;81:3346-3357.
17. Turco F, Zubak I, Slotboom J. Deep learning based MRS metabolite quantification: CNN and ResNet versus non linear least square fitting. Proceedings of the 30th Annual Meeting of the ISMRM; 2021:2018.
18. Lee HH, Kim H. Bayesian deep learning-based 1H-MRS of the brain: metabolite quantification with uncertainty estimation using Monte Carlo dropout. *Magn Reson Med*. 2022;88:38-52.
19. Bogner W, Otazo R, Henning A. Accelerated MR spectroscopic imaging-a review of current and emerging techniques. *NMR Biomed*. 2021;34:4314.
20. Li Y, Wang T, Zhang T, et al. Fast high-resolution metabolic imaging of acute stroke with 3D magnetic resonance spectroscopy. *Brain*. 2021;143:3225-3233.
21. Guo R, Ma C, Li Y, et al. High-resolution label-free molecular imaging of brain tumor. Proceedings of the 43rd International Conference of the IEEE EMBC; 2021:3049-3052.
22. Soher BJ, Young K, Govindaraju V, Maudsley AA. Automated spectral analysis III: application to in vivo proton MR spectroscopy and spectroscopic imaging. *Magn Reson Med*. 1998;40:822-831.
23. Croitor Sava AR, Sima DM, Poullet JB, Wright AJ, Heerschap A, Huffel S. Exploiting spatial information to estimate metabolite levels in two-dimensional MRSI of heterogeneous brain lesions. *NMR Biomed*. 2011;24:824-835.
24. Kelm BM, Kaster FO, Henning A, et al. Using spatial prior knowledge in the spectral fitting of MRS images. *NMR Biomed*. 2012;25:1-13.
25. Laruelo A, Chaari L, Tourneret JY, et al. Spatio-spectral regularization to improve magnetic resonance spectroscopic imaging quantification. *NMR Biomed*. 2016;29:918-931.
26. Ning Q, Ma C, Lam F, Liang ZP. Spectral quantification for high-resolution MR spectroscopic imaging with spatiospectral constraints. *IEEE Trans Biomed Eng*. 2017;64:1178-1186.
27. Li Y, Lam F, Clifford B, Liang ZP. A subspace approach to spectral quantification for MR spectroscopic imaging. *IEEE Trans Biomed Eng*. 2017;64:2486-2489.
28. Ham J, Lee D, Mika S, Schölkopf B. A kernel view of the dimensionality reduction of manifolds. Proceedings, the 21st International Conference on Machine Learning, ICML 2004; 2004:369-376; ACM Press, New York.
29. Usman M, Atkinson D, Kolbitsch C, Schaeffter T, Prieto C. Manifold learning based ECG-free free-breathing cardiac CINE MRI. *J Magn Reson Imaging*. 2015;41:1521-1527.
30. Poddar S, Jacob M. Dynamic MRI using smoothness regularization on manifolds (SToRM). *IEEE Trans Med Imaging*. 2016;35:1106-1115.
31. Nakarmi U, Wang Y, Lyu J, Liang D, Ying L. A kernel-based low-rank (KLR) model for low-dimensional manifold recovery in highly accelerated dynamic MRI. *IEEE Trans Med Imaging*. 2017;36:2297-2307.
32. Biswas S, Aggarwal HK, Jacob M. Dynamic MRI using model-based deep learning and SToRM priors: MoDL-SToRM. *Magn Reson Med*. 2019;82:485-494.
33. Shetty GN, Slavakis K, Bose A, Nakarmi U, Scutari G, Ying L. Bi-linear modeling of data manifolds for dynamic-MRI recovery. *IEEE Trans Med Imaging*. 2020;39:688-702.
34. Djebra Y, Marin T, Han PK, El Fakhri G, Bloch I, Ma C. Manifold learning via tangent space alignment (LTSA) for accelerated dynamic MR imaging with sparse sampling. *IEEE Trans Med Imaging*. 2022. doi:10.1109/TMI.2022.3207774
35. Lam F, Li Y, Peng X. Constrained agnetic resonance spectroscopic imaging by learning nonlinear low-dimensional models. *IEEE Trans Med Imaging*. 2020;39:545-555.
36. Li Y, Wang Z, Sun R, Lam F. Separation of metabolites and macromolecules for short-TE 1H-MRSI using learned component-specific representations. *IEEE Trans Med Imaging*. 2021;40:1157-1167.

37. Zhang Z, Zha H. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. *SIAM J Sci Comput*. 2005;26:313-338.
38. Vanhamme L, Sundin T, Hecke P, Huffel S. MR spectroscopy quantitation: a review of time-domain methods. *NMR Biomed*. 2001;14:233-246.
39. Golub G, Pereyra V. Separable nonlinear least squares: the variable projection method and its applications. *Inverse Probl*. 2003;19:R1-R26.
40. Haldar JP, Hernando D, Song SK, Liang ZP. Anatomically constrained reconstruction from noisy data. *Magn Reson Med*. 2008;59:810-818.
41. Lam F, Ma C, Clifford B, Johnson CL, Liang ZP. High-resolution 1H-MRSI of the brain using SPICE: data acquisition and image reconstruction. *Magn Reson Med*. 2016;76:1059-1070.
42. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found Trends Mach Learn*. 2010;3:1-122.
43. He J, Liu Q, Christodoulou AG, Ma C, Lam F, Liang ZP. Accelerated high-dimensional MR imaging with sparse sampling using low-rank tensors. *IEEE Trans Med Imaging*. 2016;35:2119-2129.
44. Soher BJ, Young K, Bernstein A, Aygula Z, Maudsley AA. GAVA: spectral simulation for in vivo MRS applications. *J Magn Reson*. 2007;185:291-299.
45. Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp*. 1994;2:189-210.
46. Graaf RA. *In Vivo NMR Spectroscopy: Principles and Techniques*. 2nd ed. John Wiley and Sons, Ltd; 2007.
47. Lam F, Liang ZP. A subspace approach to high-resolution spectroscopic imaging. *Magn Reson Med*. 2014;71:1349-1357.
48. Bydder M, Hamilton G, Yokoo T, Sirlin CB. Optimal phased-array combination for spectroscopy. *Magn Reson Imaging*. 2008;26:847-850.
49. Ma C, Lam F, Johnson CL, Liang ZP. Removal of nuisance signals from limited and sparse 1H MRSI data using a union-of-subspaces model. *Magn Reson Med*. 2016;75:488-497.
50. Ma C, Lam F, Ning Q, Johnson CL, Liang ZP. High-resolution 1H-MRSI of the brain using short-TE SPICE. *Magn Reson Med*. 2017;77:467-479.
51. Bligic B, Gagioski B, Kok T, Adalsteinsson E. Lipid suppression in CSI with spatial priors and highly undersampled peripheral k-space. *Magn Reson Med*. 2013;69:1501-1511.
52. Behar KL, Rothman DL, Spencer DD, Petroff OA. Analysis of macromolecule resonances in 1H NMR spectra of human brain. *Magn Reson Med*. 1994;32:294-302.
53. Graham GD, Hwang JH, Rothman DL, Prichard JW. Spectroscopic assessment of alterations in macromolecule and small-molecule metabolites in human brain after stroke. *Stroke*. 2001;32:2797-2802.
54. Mader I, Seeger U, Karitzky J, Erb M, Schick F, Klose U. Proton magnetic resonance spectroscopy with metabolite nulling reveals regional differences of macromolecules in normal human brain. *J Magn Reson Imaging*. 2002;16:538-546.
55. Birch R, Peet AC, Dehghani H, Wilson M. Influence of macromolecule baseline on 1H MR spectroscopic imaging reproducibility. *Magn Reson Med*. 2017;77:34-43.
56. Lam F, Li Y, Clifford B, Liang ZP. Macromolecule mapping of the brain using ultrashort-TE acquisition and reference-based metabolite removal. *Magn Reson Med*. 2018;79:2460-2469.
57. Seeger U, Klose U, Mader I, Grodd W, Nägele T. Parameterized evaluation of macromolecules and lipids in proton MR spectroscopy of brain diseases. *Magn Reson Med*. 2003;49:19-28.
58. Cudalbu C, Mlynárik V, Gruetter R. Handling macromolecule signals in the quantification of the neurochemical profile. *J Alzheimers Dis*. 2012;31:S101-S115.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

Figure S1. Spectral quantification results obtained using the numerical phantom: Comparison of the metabolite concentration maps estimated by the proposed methods with and without a sparsity penalty on the linear transform matrices in Equation (13. The corresponding NRMSEs (in percentage) are listed in the bottom right corner of each sub-figure. As can be seen, promoting the sparsity of the linear transform matrices further improved the quantification accuracy of the proposed method. The corresponding linear transform matrices are shown in Figure S2.

Figure S2. The linear transform matrices estimated without (A) and with (B) a sparsity penalty.

Figure S3. Monte Carlo simulation results obtained using the numerical phantom: Bias and standard deviation maps of the metabolite concentration maps obtained by the QUEST, the subspace, and the proposed method, respectively. The average SNR of the NAA peak was fixed to 10 for this simulation.

Figure S4. Monte Carlos simulation results obtained using the numerical phantom: Averaged bias and standard deviation achieved by the QUEST, the subspace and the proposed method, respectively, at three SNR levels.

Figure S5. Spectral quantification for SPICE: NRMSEs of the estimated metabolite concentration maps achieved by the QUEST, the subspace, and the proposed method, respectively, among the recruited subjects. Note that the error-bars of the bias and standard deviation of the QUEST method were off the charts and thus omitted.

Figure S6. Representation of different lineshapes using a subspace model. A set of 6 temporal basis functions were estimated from a Casorati matrix formed by Cr signals with Lorentzian lineshapes as described in Section 2.2. The estimated temporal basis functions were used to approximate a signal with a Lorentzian line shape in the form of $s_L(t) = e^{-t/T_2 - t/T_{2,p}} \sum_{q=1}^{2} e^{-i2\pi f_q t}$, where $T_2 =$ 200 ms and $T_{2,p} = 50$ ms and a signal with a mixture of Lorentzian and Gaussian line shape (i.e., the Voigt model) in the form of $s_V(t) = e^{-t/T_2 - t^2/g} \sum_{q=1}^{2} e^{-i2\pi f_q t}$, where $g = 5000$ ms^2. The time- and frequency-domain signals of

the Lorentzian and non-Lorentzian signals are shown in (A) & (B) and (D) & (E), respectively. The corresponding approximation errors are shown as the insert at the bottom of each subfigure. Finally, (C) and (F) plot the NRMSE in percentage when the temporal basis functions were used to approximate the Lorentzian signals with $T_{2,p}$ in the range of [50,150] ms and non-Lorentzian signals with g in the range of [5000, 15000] ms^2, respectively. As can be seen, the subspace model can accurately (i.e., NRMSEs < 1 %) represent spectral signals with Lorentzian or non-Lorentzian lineshapes and with a large range of linewidths.

Figure S7. Spectral quantification results obtained by the proposed LTSA method using the numerical phantom with different level of residual water signals. The water scale was defined as the ratio of the L_2 norm of the water signals over that of the metabolite signals. The corresponding NRMSEs (in percentage) are listed in the bottom right corner of each sub-figure. As expected the quantification error increased as the level of residual water signals increased. The quantification accuracy of mIn was most significantly affected by the residual water signal because mIn's spectrum is closest to that of water. Representative spectral fitting results are shown in Figure S8.

Figure S8. Representative spectral fitting results for the simulation experiment in Figure S7. The color coding is as follows. Black-dashed line: the real part of the noisy spectrum; red-solid line: the sum of the fitted spectra; blue-, magenta-, and green-solid line: the fitted spectrum of the NAA, mIn, and Cr, respectively. A plot of the difference between the sum of the fitted spectra and the ground truth is shown under each spectrum plot in black color. Note the distortion of the mIn spectra due to the residual water signal.

Figure S9. Spectral quantification results obtained by the extended LTSA method using the numerical phantom with different level of residual water signals. The simulation experiment setup was the same as that in Figure S7. The corresponding NRMSEs (in percentage) are listed in the bottom right corner of each sub-figure. Compared to the results in Figure S7, the robustness to residual water signals was significantly improved with the extended LTSA method that used an additional subspace to account for the residual water signals. Representative spectral fitting results are shown in Figure S10.

Figure S10. Representative spectral fitting results for the simulation experiment in Figure S10. The color coding is as follows. Black-dashed line: the real part of the noisy spectrum; red-solid line: the sum of the fitted spectra; blue-, magenta-, green-, and cyan-solid line: the fitted spectrum of the NAA, mIn, Cr, and water respectively.

Figure S11. Quantification errors of NAA and mIn by the proposed LTSA and extended LTSA method, respectively, with different level of residual water signals.

How to cite this article: Ma C, Han PK, Zhuo Y, Djebra Y, Marin T, Fakhri GE. Joint spectral quantification of MR spectroscopic imaging using linear tangent space alignment-based manifold learning. *Magn Reson Med*. 2022;1-17. doi: 10.1002/mrm.29526

ECOLE
DOCTORALE

Titre : Imagerie IRM Dynamique Accélérée Utilisant Des Modèles d'Apprentissage Automatique Linéaire Et Non-Linéaire Pour La Reconstruction d'Images

Mots clés : IRM Dynamique, Reconstruction d'Image, Sous-espace, Variétés, Apprentissage Automatique, Respiration Libre

Résumé : L'imagerie par Résonance Magnétique (IRM) dynamique est d'une grande valeur pour le diagnostic médical grâce à sa polyvalence en termes de contraste, sa haute résolution spatiale, son rapport signal/bruit élevé et permet l'obtention non invasives d'images multi-planaires. Elle peut être utile pour l'imagerie du cerveau et du cœur entre autres, ainsi que pour la détection d'anomalies. De plus, la disponibilité croissante de machines de Tomographie par Émission de Positrons (TEP) / IRM permet l'acquisition simultanée de données de TEP et IRM pour des informations complémentaires. Cependant, un défi majeur en IRM dynamique est la reconstruction d'images à partir de données d'espace-k échantillonnées en dessous de la fréquence de Nyquist. De nombreuses méthodes ont été proposées pour l'imagerie IRM sous-échantillonnée, notamment l'imagerie parallèle et le compressed sensing. Le premier objectif de cette thèse est de montrer le potentiel et l'utilité du modèle de sous-espace linéaire pour l'imagerie IRM sous respiration libre. Ce modèle peut théoriquement capturer des mouvements respiratoires et cardiaques réguliers. Cependant, des mouvements irréguliers peuvent survenir, tels qu'une respiration erratique ou un mouvement global causé par l'inconfort du patient. Une première question se pose donc naturellement : un tel modèle peut-il capturer ces types de mouvement et, si oui, peut-il reconstruire les images IRM sans artefacts ? Nous démontrons dans cette thèse comment le modèle de sous-espace peut efficacement reconstruire des images à partir de données d'espace-k fortement sous-échantillonnées. Une première application est présentée où nous reconstruisons des images IRM dynamiques avec haute résolution spatiale et temporelle et les utilisons pour corriger le mouvement des données TEP. Une deuxième application sur la cartographie T1 cardiaque est présentée. Des données sous-échantillonnées ont été acquises à l'aide d'une séquence inversion-récupération sous respiration libre, et des images IRM 3D dynamiques du cœur entier ont été reconstruites. Le deuxième objectif de cette thèse est de comprendre les limites du modèle de sous-espace linéaire et de développer un nouveau modèle qui pallie ces limitations. Le modèle de sous-espace suppose que les données de haute dimension résident dans un sous-espace linéaire qui capture les corrélations spatiotemporelles des images dynamiques. Ceci repose sur un modèle de réduction de dimension linéaire et ne prend pas en compte les caractéristiques intrinsèquement non linéaires du signal. Des modèles basés sur l'apprentissage de variétés ont donc été explorés et visent à apprendre la structure intrinsèque du signal en résolvant des problèmes de réduction de dimensionnalité non linéaires. Nous présentons dans cette thèse une stratégie alternative pour la reconstruction d'images IRM basée sur l'apprentissage de variétés. La méthode proposée apprend la structure des variétés via un alignement linéaire des espaces tangents (LTSA) et peut être interprétée comme une généralisation non linéaire du modèle de sous-espace. Des validations ont été effectuées sur des études de simulation numérique ainsi que sur des expériences d'imagerie cardiaque 2D et 3D in vivo, démontrant des performances améliorées par rapport à l'état-de-l'art. Les deux premiers objectifs présentent respectivement des modèles linéaires et non linéaires, mais ces méthodes utilisent des techniques d'optimisation linéaire conventionnelles pour résoudre le problème de reconstruction. L'utilisation de réseaux de neurones profonds pour l'optimisation peut procurer une meilleure puissance de représentation non linéaire. Des premiers résultats sur les approches basées sur l'apprentissage profond sont présentés dans cette thèse et l'état-de-l'art est discuté. Le dernier chapitre présente les conclusions, discute des contributions de l'auteur et détaille les perspectives de recherche potentielles ouvertes par le travail effectué dans cette thèse.

Title : Accelerated Dynamic MR Imaging Using Linear And Non-Linear Machine Learning-Based Image Reconstruction Models

Keywords : Dynamic MRI, Image Reconstruction, Subspace, Manifold, Machine Learning, Free-breathing

Abstract : Dynamic Magnetic Resonance (MR) imaging is of high value in medical diagnosis thanks to its contrast versatility, high spatial resolution, high Signal-to-Noise Ratio (SNR), and allows for non-invasive multi-planar images of the body. It can be particularly useful for imaging the brain, heart, spine, and joints, as well as for detecting abnormalities. In addition, the increasing availability of Positron Emission Tomography (PET)/MR machines enables simultaneous acquisition of PET and MR data for better reconstruction and complementary information. However, a key challenge in dynamic MRI is reconstructing high-dimensional images from sparse k-space data sampled below the Nyquist sampling rate. Many methods have been proposed for accelerated imaging with sparse sampling, including parallel imaging and compressed sensing. The first objective of this thesis is to show the potential and usefulness of the linear subspace model for free-breathing MR imaging. Such a model can in principle capture regular respiratory and cardiac motion. However, when dealing with lengthy scans, irregular motion patterns can occur, such as erratic breathing or bulk motion caused by patient discomfort. A first question thus naturally arises: can such a model capture irregular types of motion and, if so, can it reconstruct images from a dynamic MR scan presenting bulk motion and irregular respiratory motion? We demonstrate in this thesis how the subspace model can efficiently reconstruct artifact-free images from highly undersampled k-space data with various motion patterns. A first application is presented where we reconstruct high-resolution, high frame-rate dynamic MR images from a PET/MR scanner and use them to correct motion in PET data, capturing complex motion patterns such as irregular respiratory patterns and bulk motion. A second application on cardiac T1 mapping is presented. Undersampled k-space data were acquired using a free-breathing, ECG-gated inversion recovery sequence, and dynamic 3D MR images of the whole heart were reconstructed leveraging the linear subspace model. The second objective of this thesis is to understand the limits of the linear subspace model and develop a novel dynamic MR reconstruction scheme that palliates these limitations. More specifically, the subspace model assumes that high-dimensional data reside in a low-dimensional linear subspace that captures the spatiotemporal correlations of dynamic MR images. This model relies on a linear dimensionality reduction model and does not account for intrinsic non-linear features of the signal, which may show its limits with higher undersampling rates. Manifold learning-based models have therefore been explored for image reconstruction in dynamic MRI and aim at learning the intrinsic structure of the input data that are embedded in a high-dimensional signal space by solving non-linear dimensionality reduction problems. We present in this thesis an alternative strategy for manifold learning-based MR image reconstruction. The proposed method learns the manifold structure via linear tangent space alignment (LTSA) and can be interpreted as a non-linear generalization of the subspace model. Validation on numerical simulation studies as well as in vivo 2D and 3D cardiac imaging experiments were performed, demonstrating improved performances compared to state-of-the-art techniques. The two first objectives present respectively linear and non-linear models yet both methods use conventional linear optimization techniques to solve the reconstruction problem. In contrast, using deep neural networks for optimization may procure non-linear and better representation power. Early results on deep learning-based approaches are presented in this thesis and state-of-the-art techniques are discussed. The last chapter then presents conclusions, discusses the author's contributions, and considers the potential research perspectives that have been opened up by the work presented in this thesis.

Institut Polytechnique de Paris
91120 Palaiseau, France