PARALLEL TRANSMISSION (PTX) TECHNIQUES AND APPLICATIONS

ON A TRANSCEIVER COIL ARRAY IN HIGH-FIELD MRI

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Xianglun Mao

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. David J. Love, Co-Chair

    School of Electrical and Computer Engineering

Dr. Joseph V. Rispoli, Co-Chair

    Weldon School of Biomedical Engineering

    School of Electrical and Computer Engineering

Dr. Thomas M. Talavage

    Weldon School of Biomedical Engineering

    School of Electrical and Computer Engineering

Dr. Eric A. Nauman

    School of Mechanical Engineering

Approved by:

    Dr. Dimitrios Peroulis

        Head of the School Graduate Program

ACKNOWLEDGMENTS

About six years ago, I couldn't imagine that one day I will start writing my PhD thesis and defend in front of an academic committee. I joined Purdue University in 2013, but I never know that eventually I can graduate with a PhD degree. This six years program, a.k.a., Masters and PhD program, is one significant component of my entire life.

When I joined Purdue as an innocent college graduate, I am eager to learn more knowledge and to broaden my skill set in the graduate school. I would very much like to thank Dr. Thomas M. Talavage, who has guided me during my Masters program. He introduced me to many of his colleagues, then I had a chance to meet the field of MRI and start my research and career path in this research topic.

I had a chance to meet my committee chairs, Dr. David J. Love and Dr. Joseph V. Rispoli in the beginning of my PhD life. It was a pleasure that I can work with them not only to conduct research in MRI and communication field, but also to study their wisdom to better understand the world of science. During the six years graduate school, I made multiple mistakes and they never gave up on me. I deeply appreciated what they had done to guide and support me. It is certainly impossible that I could make this work without their support.

Dr. David J. Love, Dr. Joseph V. Rispoli, Dr. Thomas M. Talavage, and Dr. Eric Nauman are my committee members, and I received a lot of help and suggestions from them, on this dissertation and other projects went on. I am grateful for all project experiences and research discussions, they certainly made me into a well-trained personnel who has the ability to conduct independent and collaborative research.

Further, I would like to thank all the people I worked with during my PhD life. The current colleagues, Ikbeom Jang, Nicole Vike, Sumra Bari, Pratik Kashyap, Andres

Finally, I would like to thank my parents. Their financial and spiritual supports helped me during the past six years. This PhD life is very stressful and bitter, but yet the result is joyful.

Vincent Xianglun Mao
West Lafayette, IN
08/31/2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AF	Acceleration Factor
CG	Conjugate Gradient
CT	Computed Tomography
DTI	Diffusion Tensor Imaging
DOF	Degrees-Of-Freedom
fMRI	functional Magnetic Resonance Imaging
EM	Electromagnetic
EMF	Electromagnetic Force
EPI	Echo-Planar Imaging
FDTD	Finite-Difference Time-Domain
FOV	Field of View
FFT	Fast Fourier Transform
FNN	Feedforward Neural Network
GRAPPA	Generalized Auto-calibrating Partially Parallel Acquisition
IEC	International Electrotechnical Commission
KKT	Karush-Kuhn-Tucker
kNN	k-Nearest Neighbors
KRR	Kernelized Ridge Regression
MIMO	Multiple-Input Multiple-Output
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
NMR	Nuclear Magnetic Resonance
NN	Nearest Neighbor

NRMSE	Normalized Root-Mean-Square Error
pTx	Parallel Transmission
RF	Radio Frequency
ROI	Region of Interest
SAR	Specific Absorption Rate
SNR	Signal-to-Noise Ratio
SENSE	SENSitivity Encoding
SMASH	SiMultaneous Acquisition of Spatial Harmonics
SNR	Signal-to-Noise Ratio
SoS	Sum-of-Squares
SWI	Susceptibility-Weighted Imaging
TE	Echo Time
TR	Repetition Time
UHF	Ultra-High Field
VOP	Virtual Observation Point

ABSTRACT

Mao, Xianglun P.h.D., Purdue University, December 2019. Parallel Transmission (pTx) Techniques and Applications On a Transceiver Coil Array in High-field MRI . Major Professor: David J. Love & Joseph V. Rispoli.

Magnetic resonance imaging (MRI) has become an invaluable tool in health care. Despite its popularity, there is still an ever-increasing need for faster scans and better image quality. Multi-coil MRI, which uses multiple transmit and/or receive coils, holds the potential to address many of these MRI challenges. Multi-coil MRI systems can utilize parallel transmission (pTx) technology using multi-dimensional radio-frequency (RF) pulses for parallel excitation. The pTx platform is shown to be superior in high field MRI. Therefore, this dissertation is focused on the RF pulse design and optimization on an MRI system with multiple transceiver coils.

This dissertation addresses three major research topics. First, we investigate the optimization of pTx RF pulses when considering both transmitters and receivers of the MRI system. We term this framework multiple-input multiple-output (MIMO) MRI. The RF pulse design method is modeled by minimizing the excitation error while simultaneously maximizing the signal-to-noise ratio (SNR) of the reconstructed MR image. It further allows a key trade-off between the SNR and the excitation accuracy. Additionally, multiple acceleration factors, different numbers of used receive coils, maximum excitation error tolerance, and different excitation patterns are simulated and analyzed within this model. For a given excitation pattern, our method is shown to improve the SNR by 18-130% under certain acceleration schemes, as compared to conventional parallel transmission methods, while simultaneously controlling the excitation error in a desired scope.

Second, we propose a pTx RF pulse design method that controls the peak local specific absorption rates (SARs) using a compressed set of SAR matrices. RF power, peak local SARs, excitation accuracy, and SNR are simultaneously controlled in the designed pTx RF pulses. An alternative compression method using $k-$means clustering algorithm is proposed for an upper-bounded estimation of peak local SARs. The performance of the pTx design method is simulated using a human head model and an eight-channel transceiver coil array. The proposed method reduces the 10-g peak local SAR by 44.6-54.2%, as compared to the unconstrained pTx approach, when it has a pre-defined lower bound of SNR and an upper bound of excitation error tolerance. The $k-$means clustering-based SAR compression model shows its efficiency as it generates a narrower and more accurate overestimation bound than the conventional SAR compression model.

Finally, we propose two machine learning based pTx RF pulse design methods and test them for the ultra-fast pTx RF pulse prediction. The two methods proposed are the kernelized ridge regression (KRR) based pTx RF pulse design and the feedforward neural network (FNN) based pTx RF pulse design. These two methods learn the training pTx RF pulses from the extracted key features of their corresponding B_1^+ fields. These methods are compared with other supervised learning methods (nearest-neighbor methods, etc.). All learned pTx RF pulses should be reasonably SAR-efficient because training pTx RF pulses are SAR-efficient. Longer computation time and pre-scan time are the drawbacks of the current pTx approach, and we address this issue by instantaneously predicting pTx RF pulses using well-trained machine learning models.

1. INTRODUCTION

Magnetic resonance imaging (MRI) is a radio-frequency (RF) transmission and reception technology that is widely used for clinical diagnosis and medical research. The images provided by MRI systems show various tissue contrasts and therefore can help clinicians to visually inspect for abnormalities inside the human body. Some body imaging techniques, such as computed tomography (CT), are often involves with ionizing radiation and thus are harmful to the human-being. MRI, on the other hand, is harmless to the scanned humans or animals because it is a non-ionizing radiation technique. MRI typically requires a longer scan time than CT imaging, but images produced by MRI are often superior to those produced by CT. Researchers in the MRI field are continuously improving image acquisition methods, making MRI more valuable in daily clinical and research tasks. There are various types of advanced MRI techniques, including functional MRI (fMRI) [1], diffusion tensor imaging (DTI) [2], and susceptibility-weighted imaging (SWI) [3, 4]. All of them reveal the internal physiological characteristics of the human body through nuclear magnetic resonance (NMR) signals [5]. Generally, MRI works by imaging signals from hydrogen nuclei inside a human or animal body. The body is placed in a homogeneous static magnetic field (known as the B_0 field). The MRI scanner transmits a weak high frequency RF field (known as B_1 field) to selectively excite certain regions of a human/animal body, through one or more transmit coils. After applying the linear gradient field, MRI signals are detected from the excited region and are encoded in the $k-$space domain. The final anatomical MR image can be obtained by processing the raw $k-$space data.

The field strength of MRI systems continues to evolve, and it is important for us to characterize different field strengths in the MRI field. Most commercial scanners are 1.5 Tesla (1.5 T) or 3 T. Human brain research often utilizes fields up to 7 T [6, 7]. There are also several machines operating up to 11.7 T [8]. Although it

may induce arguments, the following characterization of field strength is reasonable to most researchers. 1-3 T is referred to as the high field, 3-7 T is referred to as the very high field, 7 T field or above is referred to as the ultra-high field (UHF). High field or UHF MRI systems can provide extra benefits, such as higher image resolution, higher signal-to-noise ratio (SNR), and better image contrast. Nevertheless, high field MRI always induces higher RF power deposition which may cause serious heating problems in the human body. Some high field techniques can counteract the heating problem [9, 10], and other techniques can reduce the image acquisition time [11–13] or improve the image quality [14–16]. These techniques are the major investigation area of this dissertation.

Multi-transmit coils in MRI systems can be used to reduce the dielectric effect and minimize the deposited RF power. The dielectric effect refers to the interaction between the matter and the concomitant electric field inside the MRI scanner, and this phenomenon will result in abnormal bright and dark artifacts in scanned images. However, the major advantage of multi-transmit coils is that they enable the parallel transmission (pTx), which takes advantage of multiple independent RF sources to design a homogeneous B_1 field or a uniform tip-angle pattern. The pTx technique has shown to be effective in combating magnetic field inhomogeneity in high field MRI [11, 17–26]. RF shimming is another useful technique to reduce the RF field inhomogeneity by designing the RF shim weight for each individual channel in the transmit array [27–29]. Parallel MRI systems use tailored RF pulses to facilitate the parallel excitation within a short RF pulse duration and an adequate spectral range [11, 18–21]. For example, in the widely-used transmit sensitivity encoding (Transmit SENSE) technique [20], a complex RF pulse for each transmit channel is designed to implement the parallel excitation so that the resulting transverse magnetization is almost identical to the desired magnetization (which is usually uniform). The designed RF pulse in the Transmit SENSE is often controlled in a short time duration so that the total image acquisition time is minimized.

Multiple receive coils can also be used in the MRI system to enhance the imaging. Different types of receive coils, such as surface coils, linear coils, and quadrature coils, are implemented to facilitate parallel imaging techniques. Multiple receive coils are often combined into a receive coil array to receive MRI signals [30]. Parallel imaging technology can be applied in the phased array to either speed up the image reconstruction or improve the image quality [12, 13, 31]. These techniques reduce the image acquisition time by collecting less $k-$space data from each receive coil, and reconstructing the undistorted image by combining images from all receive coils.

Despite extensive research on multi-coil transmission and/or reception, there have been only minor advancements in multi-coil processing at *both* sides of the RF chain. Transceiver coil array investigation has been the focus of several papers [16, 32–34]. These studies suggest that proper transceiver coil design can effectively enable transmit and/or receive techniques while maintaining a high SNR in the reconstructed MR image. Parallel transmit/receive technology is another area of research needing investigation in the multi-coil MRI. Prior research has shown that parallel RF transmission and reception can provide mean-squared error (MSE) benefits when the transmitters and receivers are simultaneously optimized [26]. To follow up this research, in **Chapter 3**, we focus on parallel RF transmission and reception techniques in a multi-coil MRI system. We target to improve the MR image quality and acquisition speed by developing a fast and robust pTx RF pulse design strategy by exploiting the well-developed multiple-input multiple-output (MIMO) communication theory. This multi-coil pTx RF pulse design is implemented using multiple transceive coils, and we denote this technique as MIMO MRI.

MIMO MRI is a framework where we simultaneously optimize transmitters and receivers in the MRI system. This framework contributes to the investigation of transmit-side multi-coil RF pulse design by jointly considering transmitters and receivers in the MRI system. Traditional multi-coil RF design approaches allow for region of interest (ROI) specification, and also have adaptability to mitigate the field inhomogeneity. Apart from these advantages, this MIMO MRI model provides a

theoretical method to estimate and maximize the SNR, while maintaining the excitation error in a controllable range. It further studies the potential effects of different acceleration schemes and different receive elements within the MIMO MRI model.

MRI, much like any other commercial RF device, exposes the user to electromagnetic (EM) radiation. RF exposure is one of the biggest concerns in the MRI system since the majority of the transmitted RF power will result in body heat [35–37], and body heating might cause burns in human tissues. Therefore, transmitted RF pulses need to be power constrained to prevent from possible body heating. These pTx RF waveforms in multi-coil MRI have extra degrees-of-freedom (DOFs) compared to RF waveforms in single-coil MRI, and these extra DOFs could be used to control the RF exposure. The most important measure of human RF exposure for MRI is the specific absorption rate (SAR), which is introduced to describe the RF exposure as the dissipated power per unit mass [36–38].

The International Electrotechnical Commission (IEC) [39] provides guidelines to regulate the MR safety by defining different SAR limits. Since the SAR varies throughout the tissue, regulations focus on different measures based on average SAR. Global SAR limits, partial body SAR limits, and 10 gram (i.e., 10-g) local SAR limits on different operating modes are listed in IEC guidelines. The global (whole body averaged) SAR is important since it can affect the core temperature in the scanned human body, and thus it is an important online monitoring parameter during the scanning. The partial body SAR is quantified as the averaged SAR of the exposed body part ($>>$ 10-g), which is also important for online monitoring. The local SAR of each location in the human body is the averaged SAR across 10-g of surrounding tissue. It is a significant factor reflecting the localized heating, and while it is difficult to assess in real-time, it should be carefully controlled during the RF transmission [36, 37]. In 2013, the U.S. Food and Drug Administration (FDA) accepted IEC standards and agreed to use IEC standards in the MRI manufacturing. Moreover, the International Commission on Non-Ionizing Radiation Protection (ICNIRP) [40] and the Institute of Electrical and Electronics Engineers (IEEE) [41] also use IEC standards.

The RF exposure problem becomes even more complicated in high field or UHF MRI systems. There are numerous academic papers published on the topics about pTx platforms with SAR constraints [10,42,43]. However, the development of optimal and practical pTx RF pulses that adapt to SAR constraints is still an important area of research. In **Chapter 4**, we further propose a pTx RF pulse design method that jointly optimizes transmitted RF pulses and the MRI receive voltage, and it is designed to capture peak local SARs using a compressed set of SAR matrices. RF power, peak local SARs, excitation accuracy, and SNR are simultaneously controlled in the designed pTx RF pulses. Further, an alternative compression method using k-means clustering is proposed to determine the compressed set of SAR matrices for an adequate, conservative estimation of peak local SARs.

There exist several useful methods to address the problem of B_1 field inhomogeneity, and one popular research trend is to design pTx RF pulses in the multi-coil MRI environment. The computation for tailored pTx RF pulses is usually intensive and time-consuming, due to the fact that the RF pulse optimization problem is non-convex. In **Chapter 5**, we integrate machine learning techniques into the pTx platform to investigate the relation between the optimized pTx pulses and the target tip-angle pattern in which they are tailored. The designed algorithm can directly predict pTx RF pulses using the extracted features from those given transmit sensitivity profiles. We show it is possible to quickly predict pTx RF pulses and extract useful features since both methods have linear time complexity.

Overall, the work presented in this dissertation proposed multiple pTx RF pulse design methods in a high field or UHF MRI. These methods have the potential to 1) combine the parallel transmit and receive technology together for enhanced imaging, 2) control peak local SARs in a sufficient way, 3) reduce the overall computational complexity by integrating machine learning methods into the pTx framework. This study is achieved by exploring the use of parallel RF transmission and reception technologies on a multi-channel transceiver coil array in MRI, and by investigating the application of machine learning methods in the slice-selective pTx RF pulse design.

2. BACKGROUND

2.1 MRI Theory

MRI uses both static and rotating magnetic fields to generate NMR signals which are detectable by resonant receive coils and thus yield anatomical images. The static magnetic field (i.e., B_0 field) is generated by the main magnet inside the MRI scanner, which is strong and reasonably homogeneous. B_0 field can induce the polarization of spins (i.e., hydrogen nuclei (^1H)) inside the human body. Meanwhile, the rotating magnetic field (i.e., B_1 field) performs the excitation of spins inside the human body and generates MRI signals. The receive side of the MRI system usually contains a phased coil array to detect MRI signals, and then the signals are digitized in the RF receiver system. An additional computer interface is added to perform the image reconstruction, i.e., to convert detected MRI signals into readable medical images.

With the development of MRI techniques, the static field strength is continuously increasing. High field MRI is commonly used for daily diagnosis and medical research. High field MRI scanners are with the field strength of 1-3 T, as they provide high image resolution and short image acquisition time. Furthermore, UHF MRI ($\geq$ 7 T) scanners are also entering into the MRI field as they produce better image quality even than the high field MRI scanners. Typical advantages of high field or UHF scanners include but are not limited to better image quality, faster MRI scan, less patient motion due to shorter times, more advanced imaging techniques available, and more details in smaller ROIs due to stronger signals.

Static B_0 field is one of the fundamental components in MRI. Suppose we have a strong and reasonably uniform B_0 field, $B_0(\boldsymbol{r}) = B_0$, along a certain direction, where $\boldsymbol{r} := (x, y, z)$. Conventionally, this direction should be the $z-$axis, i.e.,

$$B_0(\boldsymbol{r}) = \begin{bmatrix} 0 \\ 0 \\ B_0 \end{bmatrix}, \tag{2.1}$$

where B_0 is the strength of the static field. The application of B_0 field will cause the spins precessing about the the $z-$axis, and the spins are referred to as the nuclei with odd number of protons (with angular momentum). In MRI, hydrogen nuclei are used as spins because they are abundant in water molecules in a human body. There is a magnetic moment vector, $\vec{\mu} := (\mu_x, \mu_y, \mu_z)$, associated with each spin

$$\vec{\mu} = \gamma \vec{J}, \tag{2.2}$$

where γ is the gyromagnetic ratio and $\vec{J}$ is the angular momentum vector of that spin. Recall that the hydrogen nucleus (^1H) has a γ value of 2.68×10^8rad/s/tesla.

The magnetic moment can cause spins precessing about a certain direction, usually the $z-$axis, if under a strong magnetic field. The precession rate is determined by the Larmor frequency, which is defined as the angular frequency of the nuclear precession

$$\omega_0 = \gamma B_0, \tag{2.3}$$

where ω_0 is the Larmor frequency. It is known that the Larmor frequency has a linear relationship with the B_0 field strength. The B_0 field strength of an MRI scanner can range from hundreds of millitesla to tens of Tesla. Commonly used clinical and research scanners are 1.5 T, 3 T, and 7 T. Therefore, the Larmor frequency of those scanners are $64MHz$ at 1.5 T, $128MHz$ at 3 T, and $298MHz$ at 7 T.

If we only apply the static field B_0, then the whole scanned body is resonant at the Larmor frequency w_0. This is called the non-selective excitation. If we need to selectively excite certain regions inside the scanned body, then we need to add a linear gradient field. The linear gradient field consists of three channels that are along x, y, and z axis, respectively. The gradient system in a MRI scanner generates the three-dimensional gradient waveforms to encode the MRI signal to follow the desired

trajectory in the $k-$space, and the $k-$space data can be later converted to readable images.

Those gradient waveforms are applied to allow the Larmor frequency to be spatially dependent over the selected excitation region. To be more specific, the Larmor frequency in the ROI is changing associated with the provided gradient waveforms, and only the spins with the target Larmor frequency can be excited. The Larmor frequency in the ROI is given as $w_0 = \gamma(B_0 + zG_z)$, where z is the slice position and G_z is the $z-$directional gradient. For example, if a target $x - y$ plane $(z = z')$ needs to be excited, then we should apply the RF field with the Larmor frequency of $w' = \gamma(B_0 + z'G_z)$. In this scenario, only the spins in the $z = z'$ slice are excited. This is called the slice-selective excitation.

For the non-selective excitation, only the B_0 field is applied to ensure the whole scanned body is excited. But this excitation is very time-consuming, which is why researchers prefer using the selective excitation. For the slice-selective excitation, it is crucial to excite one single plane to obtain a 2D image. The designed RF pulse should be or a sinc pulse (or something similar). Therefore, the desired excitation pattern can become a rectangular function that excite a thin slice with a small slice thickness Δz, ensuring that enough spins can be excited in that slice. This is because there is a Fourier-like relationship between the excitation pattern and the RF pulses.

Apart from the applied gradient field, another important component in the MRI scanner is the rotating RF field (i.e., B_1 field). In MRI, the B_1 field is used to excite spins inside the human body. This spins can be used to generate spatially and temporally varied magnetization (i.e., $\vec{M}(\boldsymbol{r}, t)$) because they have angular momentum. The B_1 field needs to tip down this magnetization to generate detectable MRI signals. This process is what we called RF excitation. RF excitation needs the input of B_1 field, which is generated by the designed RF pulses. The rotating B_1 field needs to cooperate with the linear gradient field, generated by the gradient system, to spatially encode the MR signal in the $k-$space domain.

Commercial MRI scanners target to obtain uniform B_0 field and B_1 field because they can be used to enhance the imaging. However, the internal physiology of different human tissues generate different magnetic susceptibilities, and this phenomenon always lead to non-uniform B_0 field [44, 45]. B_0 field inhomogeneity is becoming a major problem in UHF MRI systems because of the off-resonance effect, when spins precess at a frequency different than intended. As a result, this off-resonance effect causes distortions in resulting images. This effect is difficult to remove, especially in high field because of the hardware implementation issue. This is no perfectly uniform B_0 field even in the latest MRI setup. B_1 field inhomogeneity is another severe issue since it can cause regional increased or decreased MRI signal, resulting in a non-uniform excitation pattern. B_1 field inhomogeneity is also inevitable since every RF coil has a nonuniform field pattern, and these nonuniform patterns become more evident in high field MRI. This is the reason that multiple advanced techniques (birdcage coil, RF shimming, multi-coil MRI, etc.) are developed to counteract this inhomogeneity. To better understand these advanced imaging techniques, we should firstly investigate the concept of RF excitation.

The Bloch equation describes the relation between the magnetization and the applied B_1 field during the RF excitation (relaxation effects are ignored) [46]

$$\frac{d}{dt}\vec{M}(\boldsymbol{r},t) = \vec{M}(\boldsymbol{r},t) \times \gamma\vec{B}(\boldsymbol{r},t), \tag{2.4}$$

where the magnetization is given by

$$\vec{M}(\boldsymbol{r},t) = \begin{bmatrix} M_x(x,y,z,t) \\ M_y(x,y,z,t) \\ M_z(x,y,z,t) \end{bmatrix}. \tag{2.5}$$

The applied B_1 field is given as

$$\vec{B}(\boldsymbol{r},t) = \begin{bmatrix} \text{Re}(B_1(t)) \\ \text{Im}(B_1(t)) \\ B_0 + xG_x(t) + yG_y(t) + zG_z(t) \end{bmatrix}, \tag{2.6}$$

where $B_1(t)$ is the applied B_1 field, which will generate the $\vec{B}(\boldsymbol{r}, t)$ field. $G_x(t)$, $G_y(t)$ and $G_z(t)$ are gradients in x, y, and z directions.

According to (2.4) and (2.5), the complex transverse magnetization is given by

$$M_{xy}(\boldsymbol{r}, t) = M_x(\boldsymbol{r}, t) + iM_y(\boldsymbol{r}, t). \tag{2.7}$$

The transverse magnetization, $M_{xy}(\boldsymbol{r}, t)$, can be used to obtain the corresponding RF pulses because there is a linear relationship between them.

If B_1 field is circularly polarized in one specific RF coil, then it can also be written as

$$B_1(t) = b_1(t)e^{-i\gamma B_0 t}, \tag{2.8}$$

where $b_1(t)$ is the transmitted RF pulse with the magnitude of $|B_1(t)|$ (i.e., B_1 field strength) and $e^{-i\gamma B_0 t}$ denoted the phase changed along with the B_0 field and the time.

The applied B_1 field and the linear gradient field will induce the transverse component of the magnetization, and thus generate detectable MRI signals. Take 2D imaging for example, we assume that there is no z gradient in the linear gradient field (i.e., $k_z(t) = 0$). The 2D MRI signal model can be written as

$$s(t) = \int \int \left[\int M_{xy}(x, y, z, 0)e^{-t/T_2(x,y,z)}dz \right] e^{-i2\pi[xk_x(t)+yk_y(t)]}dxdy, \tag{2.9}$$

where $s(t)$ is the encoded MRI signal, $T_2(x, y, z)$ is the T_2 decay in different locations, and $(k_x(t), k_y(t), k_z(t))$ are three components in the $k-$space domain. The excitation $k-$space is defined as

$$k_x(t) = \frac{\gamma}{2\pi} \int_t^T G_x(\tau)d\tau, \tag{2.10}$$

$$k_y(t) = \frac{\gamma}{2\pi} \int_t^T G_y(\tau)d\tau, \tag{2.11}$$

$$k_z(t) = \frac{\gamma}{2\pi} \int_t^T G_z(\tau)d\tau. \qquad (2.12)$$

It can be concluded from (2.9) that the 2D MR image can be written as

$$m(x,y) = \int M_{xy}(x,y,z,0)e^{-t/T_2(x,y,z)}dz, \qquad (2.13)$$

where $m(x,y)$ is the 2D MR image. The 2D MR image defined in (2.13) can be obtained through the relationship defined in (2.9). MR images are encoded in the $k-$space. The $k-$space data spacing determines the field-of-view (FOV), and its trajectory determines the quality of the reconstructed images. For example, under-sampled $k-$space data (i.e., big spacing between $k-$space data samples) will result in reduced FOV image and aliasing artifacts, since it violates the *Nyquist* criterion. Additionally, echo-planar imaging (EPI) trajectory will result in geometric distortions in images.

In typical cases, a simple inverse fast Fourier transform (FFT) can be applied on the $k-$space data to obtain MR images. However, this usually requires the $k-$space data to be uniformly distributed in its domain. If the $k-$space data are not uniformly sampled, RF pulses should be computed using other methods, such as FFT with $k-$space density compensation [47], or the iterative FFT method [48], or the nonuniform FFT method [49]. There also exist other image reconstruction methods based on the approximation of the image domain or the $k-$space domain, when ignoring the B_0 field deviation effects [50, 51].

It is known that we can use Fourier analysis to solve for the relation between RF pulses and the excited magnetization. The excitation $k-$space can have different trajectories, such as Cartesian, spiral, EPI (i.e., zig-zag), radial, and even random sampling trajectories. Some trajectories are beneficial for the RF excitation performance (e.g., spiral), and some are beneficial for the parallel imaging performance (e.g., Cartesian, EPI).

Each $k-$space trajectory has unique undersampling technique. In EPI trajectory, undersampling is achieved by acquiring less number of phase-encoding lines. In spi-

ral, undersampling is achieved by reducing the number of spiral turns. No matter what technique the researcher implement, it will always result in signal loss in reconstructed images because only a portion of the imaging data is acquired. However, this undersampling technique can be used to improve the image acquisition. For instance, $k-$space undersampling can be used to reduce the RF pulse duration in the RF transmission [11], and it also can be used to substantially reduce the reconstruction time in the MR signal reception stage [12, 13]. All these techniques tend to reduce the imaging time while maintaining the reconstruction performance.

2.2   RF Transmission in MRI

In a conventional single channel MRI system, the RF transmission can be described as follows. The RF system inside the MRI scanner firstly generates a reasonably homogeneous B_1 field to excite spins inside the human body. The applied B_1 field is perpendicular to the main static field (i.e., B_0 field). The continuous application of B_1 field will tip the net magnetization away from its initial alignment with B_0, and thus generate the measurable transverse magnetization. The amplitude and phase of the MR signal can be measured from the precessing transverse magnetization. When a large number of transmit coils are used, the applied B_1 field is considered to be a linear superposition of transmit field sensitivities from all individual transmit coils. These individual transmit field sensitivities can be used to generate almost homogeneous B_1 field and reduce the possible dielectric effect.

2.2.1   Small-tip-angle RF Pulse Design

Transmitted RF pulses in the MRI system are used to induce the net magnetization on the subject being scanned and generate detectable received signals. The induced magnetization can be used to enhance or suppress the generated MRI signal. Therefore, RF pulse design is a important process to design the RF pulse for each existing channel of the MRI system to achieve the target magnetization profile,

$\vec{M}(\boldsymbol{r},t)$ (i.e., (2.5)). The RF pulse design is usually implemented based on the underlying physics of the Bloch equation (i.e., (2.4)). Both RF field and gradients shall be programmed by the RF pulse designer to achieve the desired pattern, $\vec{M}(\boldsymbol{r},t)$, at the end of the RF pulse application.

However, it is generally difficult to obtain optimal RF pulses and gradient waveforms by directly solving the Bloch equation. Theoretically, the desired transverse magnetization and initial equilibrium state can be arbitrary, which makes it difficult to generate a universal solution of optimal RF pulses and gradients. Directly solving for optimal RF pulses using the Bloch equation is computationally complex, and this is the reason why researchers use other alternative methods to solve for RF pulses.

The small-tip-angle approximation [52] method is proposed to simplify the calculation of the Bloch equation. It assumes the z direction of the magnetization is unperturbed by RF pulses. In other words, $M_z(\boldsymbol{r},t) \approx M_0$, where M_0 is the magnetization at the equilibrium state. Therefore, the Bloch equation can be transferred to a first order differential equation. This small-tip-angle regime is made when the tip angle is small enough so that we can assume the sine of the tip angle equals to the tip angle itself. This angle is usually under $30°$, but the results work well for the angles up to $90°$ [52].

In this work, all the derivations are made using the small-tip-angle regime. For large tip angles ($> 90°$), several methods have been proposed to implement the RF excitation, including the constrained optimization on gradient waveforms [47], and approximations on the Bloch simulation [53, 54]. Additionally, there are some proposed methods design RF pulses and gradient waveforms on specific types of $k-$space trajectories [55–58].

In small-tip-angle regime, the transverse magnetization at the end of the RF pulse is given by

$$M_{xy}(\boldsymbol{r},T) = i\gamma M_0 \int_0^T b_1(t)e^{i\gamma\Delta B_0(\boldsymbol{r})(T-t)}e^{i\boldsymbol{r}\cdot\boldsymbol{k}(t)}dt, \qquad (2.14)$$

where T is the RF pulse duration, $M_{xy}(\boldsymbol{r}, T)$ is given in (2.7), $e^{i\gamma\Delta B_0(\boldsymbol{r})(T-t)}$ represents the phase accumulated due to the B_0 field deviation, and $\boldsymbol{k}(t)$ represents the 3-dimensional k-space trajectory $(k_x(t), k_y(t), k_z(t))$.

Since $M_z(\boldsymbol{r}, t)$ stays unchanged during the RF excitation, we only need to acquire the target transverse magnetization, $M_{xy}(\boldsymbol{r}, t)$, if we need to obtain the final target magnetization, $\vec{M}(\boldsymbol{r}, t)$. Using (2.14), we can solve for optimal RF pulses using methods that solve for convex optimization problems. This optimization problem is what we called the magnetization matching problem. In addition to this problem, there are several other aspects that RF pulse designers need to consider in reality. The designed gradient waveform is one of the key issues. The amplitude and the slew rate of gradients should be limited so that human body can tolerate it. The peak and integrated RF power should also be limited to ensure the scan safety. Additionally, the RF pulse duration should also be carefully controlled during the image acquisition. Longer RF pulse duration will increase the tissue relaxation effect and increase the acquisition time.

2.2.2 Parallel Transmission in MRI

As we mentioned earlier, B_1 field inhomogeneity always exists in practice, especially in high field or UHF MRI scanners. Several methods exist for the B_1 field inhomogeneity correction, e.g., [59–61]. For a single channel MRI system, we can compensate the B_1 field inhomogeneity by extending the pulse duration or increasing the RF power. In this case, the linear gradient field can help to reduce the spatial variation in the B_1 field. In an MRI system with multiple channels, it is desired to design an RF pulse for each transmit channel and eliminate the B_1 field inhomogeneity by taking the weighted sum of all individual transmit fields. Each transmit coil's sensitivity profile, i.e., B_1^+ field, is non-uniform, and its strength is higher in the regions near the physical coil and lower in the regions away from the coil. By

combining all transmit fields together, it is able to generate a homogeneous B_1 field and a desired tip-angle pattern (i.e., a desired magnetization pattern).

The pTx technology can improve the accuracy of the RF excitation and allows for efficiently using the RF pulses. Multiple applications can be developed upon the pTx framework, for example, it can be used to accelerate the scanning by reducing the FOV in each transmit coil [11]. However, the resulting FOV is not reduced after we combine all coil images together. Further, pTx can also be used to mitigate signal loss susceptibility artifacts by incorporating phase correction techniques [58, 62, 63]. Fig. 2.1 illustrates the schematic of the 8-channel pTx framework in an MRI system, where the spiral $k-$space trajectory is undersampled by one half. To be more specific, the pTx technology stores the Fourier transform of the target magnetization in a desired $k-$space trajectory and uses appropriate coil sensitivity profiles to generate the target excited magnetization. Ideally, the excited magnetization should be very close to the target magnetization.

Fig. 2.2 illustrates the general RF pulse design procedure in an MRI system (single-channel or multi-channel). The computing system take RF pulse design parameters (B_0 field map, individual transmit field sensitivities, $k-$space trajectory, ROI mask, etc.) and scanner parameters (gradient amplitudes, slew rates, sampling rate, etc.) as the input, and it also applies multiple software constraints (RF power, SAR, excitation accuracy, etc.) during the RF pulse design procedure. Eventually, the computing system will output the designed RF pulses and gradient waveforms to users.

The RF pulse design problem in pTx MRI platform is similar to the problem described in (2.14). When N_t transmit coils with different sensitivity profiles are used, the resulting transverse magnetization can be expressed as a linear superposition of the individual transverse magnetization from each transmit channel

$$M_{xy}(\boldsymbol{r}, T) = i\gamma M_0 \sum_{p=1}^{N_t} S_p(\boldsymbol{r}) \int_0^T b_{1,p}(t) e^{i\gamma \Delta B_0(\boldsymbol{r})(T-t)} e^{i\boldsymbol{r}\cdot\boldsymbol{k}(t)} dt, \qquad (2.15)$$
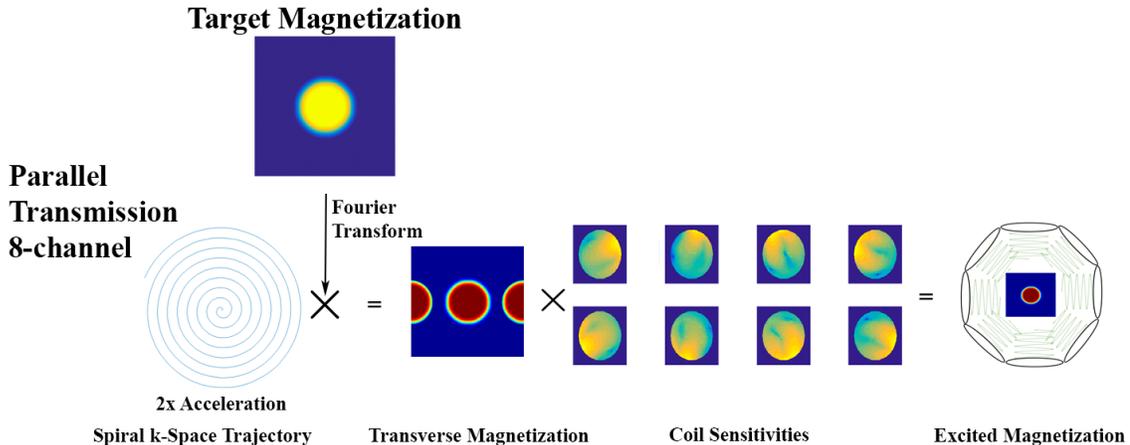
Fig. 2.1. Parallel Transmission (pTx) framework with 8-channel transmit coils, a spiral $k-$space trajectory, and an acceleration factor of 2. Multi-transmit coils are applied to remove the possible aliasing effects that caused by the $k-$space undersampling. The purpose of the pTx technology is to excite the transverse magnetization to be very close to the target magnetization pattern, and it accelerates the excitation by applying the $k-$space undersampling in each of the coil. The resulting excited transverse magnetization is obtained as the product of the coil sensitivities and the Fourier transform of the target magnetization in the $k-$space.

where $S_p(\boldsymbol{r})$ is the transmit sensitivity of the pth coil in which an RF pulse $b_{1,p}(t)$ is placed. Graphically, (2.15) is equivalent to the framework shown in Fig. 2.1. Conjugate gradient (CG) iterations can be performed to design the appropriate RF pulse, $b_{1,p}(t)$, to achieve the desired excitation pattern. Like the input-output model described in Fig. 2.2, we should further apply multiple constraints (excitation accuracy, global/local SAR, etc.) in the optimization problem in (2.15).

To solve for pTx RF pulses, we can design the magnetization matching problem with Tikhonov regularization terms. These regularization terms are used to control the transmitted RF power and other constraints such as global and peak local SARs. The details of the pTx RF pulse design optimization will be discussed in **Chapter 3 and 4**.
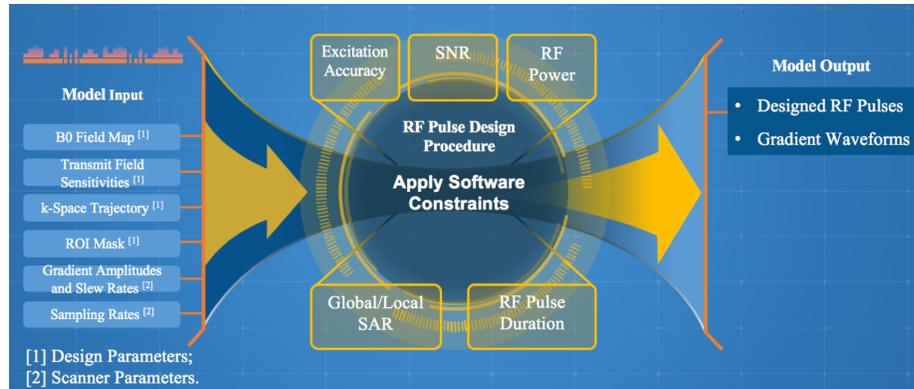
Fig. 2.2. The input-output model of the pTx RF pulse design procedure in the RF system of the MRI. The RF pulse design requires the scanner parameters (amplitudes and slew rates of the gradient, sampling rates, etc.) as the model input so the designed RF pulses have to obey these regulations. B_0 field map, transmit sensitivities, $k-$space trajectory, and ROI masks are also considered as the model input so that the pTx technique (2.15) can be implemented. Different software constraints (global/local SAR, RF power, excitation accuracy, SNR, etc.) shall be considered and simultaneously controlled in the design procedure. The designed RF pulses should be energy efficient so that they will not induce serious heating in the human body, therefore global/local SAR and RF power needed to be carefully controlled. The major difficulties of the RF pulse design are the proper RF excitation and the control of the applied constraints.

2.3 Reception in MRI

This section will mainly discuss the MRI received signal model. It is known that measurable MRI signals are proportional to the transverse component of the magnetization. Therefore, it is significant to distinguish the transverse and longitudinal magnetization after the excitation. During the MRI, after spins are tipped down, those spins keep precessing about the $z-$axis, but spins will not stay precessing forever. The longitudinal magnetization, $M_z(\boldsymbol{r}, t)$, will return to its initial state, meanwhile the transverse magnetization, $M_{xy}(\boldsymbol{r}, t)$, will decay but still precessing. The longitudinal relaxation, also known as the T_1 relaxation, is the process of the longitudinal magnetization recovering to its initial equilibrium, M_0. The constant

value of T_1 is the time that the longitudinal magnetization reaches to about 63% of the equilibrium value. The transverse relaxation, also known as the T_2 relaxation, is the process of the transverse magnetization going into an exponential decay. The constant value of T_2 is the time that the transverse magnetization decreases to about 37% of its initial value. Here, T_1 is typically about hundreds of milliseconds, and T_2 is much shorter and is about tens of milliseconds. These two relaxation effects always exist in the resulting MR images.

T_1 and T_2 are important parameters in the MR imaging, as their values can provide information for the clinical diagnosis. For example, longer myocardial T_1 values might be the signal of some forms of pathology, including chronic cardiac disorders and cardiogenic pulmonary edema [64]. Further, T_2 measures can also be used for the diagnosis of myocardial edema [65]. Short myocardial T_2 regions might indicate myocardial hemorrhages [66]. However, these measurements are usually estimated after we reconstruct the regular MR images.

In general, an iterative FFT reconstruction method [48] or a nonuniform FFT reconstruction method [49] can be used to reconstruct the MR image. However, these methods may be susceptible to noises and the amount of collected $k-$space data. This is the reason we develop the state-of-art parallel imaging technique. Parallel imaging is a widely-used technique that majorly used to reduce the image acquisition time, using less $k-$space data and a receive coil array to maintain the image quality in a short reconstruction time. Details about the MRI signal model, reconstruction methods, and different parallel imaging techniques will be discussed in this section.

2.3.1   MRI Signal Model

The RF excitation will induce measurable MR signals in resonant coils. We denote the peak of the measurable MR signal after the RF excitation as *echo*. The echo time (TE) is the time between the center of the RF pulse to the peak of the signal (i.e., center of the *echo*) detected in resonant coils. *Echo* does not need to be unique in

the pulse sequences, and multiple *echoes* can be applied in pulse sequences to acquire images with T_2 or T_2^* decay. Typical pulse sequences do not acquire all the data in one readout session but repeat the readout for several times. The pulse sequence repeats the same RF excitation but with different phase-encoding gradients, and thus multiple phase-encoding lines can be generated in the $k-$space domain and the full $k-$space data can be obtained for image reconstruction. The time length for each repetition is referred to as the repetition time (TR).

According to (2.9) and (2.13), a more general form of the ideal MRI signal model can be written as follow

$$s(t) = \int m(\boldsymbol{r})e^{-i2\pi[\boldsymbol{r}\cdot\boldsymbol{k}(t)]}d\boldsymbol{r}. \tag{2.16}$$

The desired image $m(\boldsymbol{r})$ should contain the relaxation decays, and it is related to receive coil sensitivity profiles, TE, and TR. In the typical spin echo imaging, different values of TE and TR are used to control the image contrast and determine the weighting of the MR image, such as T_1-weighting, T_2-weighting, T_2^*-weighting, and proton density weighting.

Given the MRI signal model is like (2.16), then the reconstructed MR image, $\tilde{m}(\boldsymbol{r})$, can be computed via the following equation

$$\tilde{m}(\boldsymbol{r}) = \int_{k-space} (s(\boldsymbol{k}(t)) + \tilde{n}(t))e^{i2\pi[\boldsymbol{r}\cdot\boldsymbol{k}(t)]}d\boldsymbol{k}, \tag{2.17}$$

where $\tilde{n}(t)$ is the channel noise. (2.17) provides useful insights into the MR image reconstruction. It shows that MR images can be reconstructed even if we only collect a finite number of samples in $s(\boldsymbol{k}(t))$. Different $k-$space trajectories can be applied to improve the reconstruction performance, even with a limited amount of $k-$space data.

2.3.2 Parallel Imaging in MRI

If multiple receive coils are used in the MRI system, then parallel imaging techniques should be implemented for image reconstruction. During the readout session

of the multi-coil MRI, each receive coil measures the voltage as a result of the magnetic flux through that coil. The transverse magnetization at each location in the excited object contributes to the magnetic flux through each receive coil, and hence contributes to the measured voltage of that location. The measured NMR voltage at every location is the weighted sum of measured voltages from all receive coils. The MR image intensity at each location depends on the measured voltage, the coil geometry and sensitivity, and the transverse magnetization in that location.

Parallel imaging techniques have the ability to change the way MRI is used in both clinic and research, and they can be used to assist multiple new applications of MRI, such as fMRI and DTI. The three most widely-applied parallel imaging techniques are SiMultaneous Acquisition of Spatial Harmonic (SMASH) [31], SENSitivity-Encoding (SENSE) [12], and Generalized Auto-calibrating Partially Parallel Acquisition (GRAPPA) [13].

In parallel imaging, MRI data are usually obtained using a receive coil array instead of a single volume coil. Each one of the receive coil elements is sensitive to the tissue that is close to itself. By combining all sub-images from all receive coils, we can reconstruct the final undistorted image. The performance of the parallel imaging is usually quantified by the acceleration factor. The acceleration factor in the parallel imaging is defined as the ratio of the undersampled $k-$space data to the fully sampled $k-$space data. However, researchers generally considered the acceleration factor to be the amount of reconstruction time reduced in the reception stage (i.e., acceleration factor equals to two if the reconstruction time is reduced by one half).

SENSE reconstruction is one widely used parallel imaging technique that was first proposed in [12]. It is required to know the coil placement and coil receive sensitivities ahead of the scanning, and usually a gradient-recalled echo pre-scan can acquire those information. After this pre-scan, SENSE begins to acquire partial $k-$space data for each individual receive coil, and it will generate aliased images in all receive channels. This means fold-over artifacts can be visually seen in each coil image. If we take N_r receive coils, for example, N_r voxels alias each other because

each coil acquires a reduced FOV image. Nevertheless, we multiply each of these voxels by the corresponding coil sensitivity and then add them together to produce the unaliased image

$$U = \sum_{i=1}^{N_r} C_i I_i, \tag{2.18}$$

where U is the unaliased voxel, C_i is the coil sensitivity for coil i, and I_i is the intensity values of voxels collected by coil i.

GRAPPA is another widely used technique to accelerate the MRI scan by requiring fewer numbers of phase-encoding lines in the $k-$space domain because it can regenerate those missing phase-encoding lines [13]. GRAPPA found that the information of missing phase-encoding lines can be found in their neighboring $k-$space points, so it reconstructs the image using the knowledge of the existing $k-$space data and receive sensitivity profiles. The process of GRAPPA reconstruction can be described as follows. First, an auto-calibration signal line is collected for each receive coil. Second, each coil's reconstruction kernel is determined in order to estimate those missing phase-encoding lines. The reconstruction kernel in each receive coil is normally the local known data region next to the missing data region. Finally, the $k-$space data in reconstruction kernels from all receive coils are used to fit the auto-calibration signal line in each receive coil [13]. The resulting fit will generate the weighting factor to obtain those missing lines in each specific coil, using those reconstruction kernels.

Therefore, GRAPPA basically uses neighboring $k-$space data points to regenerate those missing $k-$space data and thus obtain the full $k-$space map. An inverse Fourier transform is applied when GRAPPA obtained the full $k-$space map, so it can generate undistorted images. Unlike SENSE, which is working on aliased images, GRAPPA is directly applied on the raw $k-$space data.

2.4 SAR Management in MRI

2.4.1 SAR Definition and Calculation

There is always a E field component generated associated with the applied B_0 and B_1 field, and E field will emerge RF energy deposition in human tissues. As a result, transmitted RF power in MRI will likely dissipated as heat in human body [35], and the heating will probably result in tissue damage, cardiovascular stress, etc. Therefore, in applications involving RF exposure, it is critical to minimize the human's internal temperature rise to avoid tissue damage and any other possible health side-effects.

Since temperature rise is hard to measure in the scanned object, SAR is introduced and used as a substitute figure (i.e., dissipated power per unit mass) to quantify the safety limit in a local 10-g tissue volume V

$$SAR(\boldsymbol{r}) = \frac{1}{V} \int_V \frac{\sigma(\boldsymbol{r})}{2\rho(\boldsymbol{r})} |\boldsymbol{E}(\boldsymbol{r})|^2 dV, \tag{2.19}$$

where σ is the tissue electrical conductivity, ρ is the tissue density, and $\boldsymbol{E}(\boldsymbol{r})$ is the magnitude of the electric field at location $\boldsymbol{r}$.

This equation is often considered as the local SAR definition. Additionally, the global (body-averaged) SAR is defined as:

$$SAR_{global} = \frac{1}{V_{body}} \int SAR(\boldsymbol{r}) d\boldsymbol{r}, \tag{2.20}$$

where V_{body} refers to the volume of the scanned human body.

2.4.2 SAR Modeling in Parallel Transmission (pTx)

The 10-g local SAR and global SAR are two conventional definitions for SAR modelling in MRI systems, and they are depend on E field distributions. However, E field cannot be directly measured in the MRI system, so we need to find an alternative way to estimate E field. EM simulation is therefore implemented to estimate E field

prior to the RF pulse design procedure. To be more specific, we obtain the electric field sensitivities from the EM simulation and then estimate global/local SAR values using aforementioned equations.

In the multi-coil MRI system, the electrical field, $\boldsymbol{E}(\boldsymbol{r})$, can be written as a sum of weighted electric field sensitivities. This is like $\boldsymbol{B}_1^+(\boldsymbol{r})$ field, which can be written as a sum of weighted individual transmit field sensitivities of all transmit channels. The electric field sensitivity from each channel is obtained from the EM simulation but need to be normalized into a desirable scope (i.e., $[0,1]$). The transmit field, $\boldsymbol{B}_1^+(\boldsymbol{r})$, and the electric field, $\boldsymbol{E}(\boldsymbol{r})$, can be written as

$$
\boldsymbol{B}_1^+(\boldsymbol{r}) = \boldsymbol{S}(\boldsymbol{r})\boldsymbol{\mathcal{B}} = [S_1(\boldsymbol{r})...S_{N_t}(\boldsymbol{r})] \begin{bmatrix} \boldsymbol{b}_{1,1} \\ \vdots \\ \boldsymbol{b}_{1,N_t} \end{bmatrix}, \tag{2.21}
$$

$$
\begin{aligned}
\boldsymbol{E}(\boldsymbol{r}) &= \boldsymbol{S}_E(\boldsymbol{r})\boldsymbol{\mathcal{B}} \\
&= \begin{bmatrix} S_{Ex,1}(\boldsymbol{r})...S_{Ex,N_t}(\boldsymbol{r}) \\ S_{Ey,1}(\boldsymbol{r})...S_{Ey,N_t}(\boldsymbol{r}) \\ S_{Ez,1}(\boldsymbol{r})...S_{Ez,N_t}(\boldsymbol{r}) \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_{1,1} \\ \vdots \\ \boldsymbol{b}_{1,N_t} \end{bmatrix},
\end{aligned} \tag{2.22}
$$

where $\boldsymbol{\mathcal{B}}$ is the transmitted RF waveform across N_t channels.

To predict the SAR distribution in the scanned body, we need to combine (2.19) and (2.22) to obtain the local SAR expression

$$
SAR(\boldsymbol{r}) = \frac{1}{T} \int \boldsymbol{\mathcal{B}}^H \underbrace{\left[\frac{1}{V} \int_V \frac{\sigma(\boldsymbol{r})}{2\rho(\boldsymbol{r})} \boldsymbol{S}_E^H(\boldsymbol{r}) \boldsymbol{S}_E(\boldsymbol{r}) dV \right]}_{\boldsymbol{Q}(\boldsymbol{r})} \boldsymbol{\mathcal{B}} dt, \tag{2.23}
$$

where H is the conjugate transpose operator, T is the pulse length. This "Q-matrix formalization" is initially introduced to build the SAR model for the hyperthermia treatment [67], but it is then found to be very useful in the SAR modeling in RF pulse design [68, 69]. This formalization has the potential to be extended to the real-time SAR monitoring in the MRI data acquisition.

It is believed that there is a huge breakthrough from the single-channel transmission to the multi-channel transmission, and it is not only about the improved performance but also the reduced imaging time. Nevertheless, pTx also brings us more unsolved issues, computational complexity, B_1 field mapping, high RF power deposition, etc.

One of the biggest concerns about the pTx platform is the SAR management. Compared to a single-channel RF transmission, SAR might be much more problematic in a multi-channel RF transmission because more RF power is involved. If a human scan is conducted, then the SAR measure will substantially increase if we switch from a single-channel system to a pTx system, and a high SAR value usually has potential risks to the human body. Two types of SAR are generally considered for MRI scans, global SAR and local SAR. As defined earlier, the former one indicates the body-averaged SAR, and the latter one indicates the SAR of each local region of the body. Global SAR is generally easier to compute than the local SAR, while the local SAR can reveal more information than the global SAR. The major difficulty in the pTx framework is to constrain global/local SARs without affecting the excitation performance or the image quality.

Prior research [18] proposed the SAR management technique in the pTx MRI system, and it pointed out that the DOFs in the pTx platform can be used to achieve the desired tip-angle pattern and meanwhile minimize the local/global SAR. The local SAR can be written as a quadratic function containing RF pulses and the electric field. Therefore, in the general framework of parallel excitation, such as the *spatial domain* method [11], the *Transmit SENSE* method [20], and the MIMO MRI method [70,71], the local SAR expression can be written as a quadratic form containing RF waveforms, $\boldsymbol{\mathcal{B}}$, from all transmit channels

$$SAR(\boldsymbol{r}_m) = \frac{1}{T} \int \boldsymbol{\mathcal{B}}^H(t) \boldsymbol{Q}(\boldsymbol{r}_m) \boldsymbol{\mathcal{B}}(t) dt. \tag{2.24}$$

If we discretize the expression in (2.24), this expression can be rewritten into the following form

$$SAR(\boldsymbol{r}_m) = \frac{1}{\mathcal{K}} \sum_{k=1}^{\mathcal{K}} \boldsymbol{\mathcal{B}}^H(t_k)\boldsymbol{Q}(\boldsymbol{r}_m)\boldsymbol{\mathcal{B}}(t_k) \tag{2.25}$$

$$= \sum_{p=1}^{N_t} \sum_{q=1}^{N_t} Q_{p,q}(\boldsymbol{r}_m) \cdot \underbrace{\frac{1}{\mathcal{K}} \sum_{k=1}^{\mathcal{K}} b_{1,p}^*(t_k)b_{1,q}(t_k)} \tag{2.26}$$

$$= \sum_{p=1}^{N_t} \sum_{q=1}^{N_t} Q_{p,q}(\boldsymbol{r}_m) \cdot E\left[b_{1,p}^*(t_k)b_{1,q}(t_k)\right]. \tag{2.27}$$

Here, we need to compute the averaged cross power of each two RF pulses to obtain the local SAR value (RF pulse is complex-valued and thus the conjugate operator $*$ is used in the expression). However, this expression can be further simplified to avoid calculating the averaged cross power

$$SAR(\boldsymbol{r}_m) = \frac{1}{\mathcal{K}}\boldsymbol{\mathcal{B}}^H\boldsymbol{Q_d}(\boldsymbol{r}_m)\boldsymbol{\mathcal{B}}, \tag{2.28}$$

where $\boldsymbol{Q_d}(\boldsymbol{r}_m) = diag(\boldsymbol{Q}(\boldsymbol{r}_m), ..., \boldsymbol{Q}(\boldsymbol{r}_m)) \in \mathbb{C}^{(N_t \cdot \mathcal{K}) \times (N_t \cdot \mathcal{K})}$.

2.4.3   Relations in SAR and the Electric Fields

SAR values can be directly estimated from the electric field (i.e., E field). In an MRI system, the E field is arose from two independent sources, the *conservative E field* and the *magnetically induced E field*. We are interested in *magnetically induced E field* in modern MRI systems, especially in high-field or UHF MRI systems. This is because the *magnetically induced E field* is dominant in these machines as its strength increases linearly with the strength of the Larmor frequency.

The *magnetically induced E field* is caused by the changing magnetic flux. This E field highly depends on the amplitude and phase of transmitted RF pulses of all channels, and it occurs in every location of the subject body exposed to the B_1 field. Meanwhile, the *conservative E field* is caused by free electric charges in the coil winding [72]. In latest RF coil setups, high *conservative E field* is always observed at the capacitors in the coils. Modern RF coil design methods have the ability to

minimize the effect of the *conservative E field* by introducing more capacitors and evenly distributing them in coils [36, 73, 74].

For the majority of cases, the *magnetically induced E field* is considered to be the dominant E field induced in the MRI system. However, the *conservative E field* can be greater than the *magnetically induced E field* in some MRI setups, such as MRI with small coils, or MRI with very low field. This suggests that we should keep an eye on the *conservative E field* in those cases, but the major investigation focus in this dissertation is still the *magnetically induced E field*.

The SAR increases substantially along with the strength of *magnetically induced E field*, and thus it increases with the frequency of the B_0 and B_1 fields [36, 45, 74]. For lower B_0 fields, the associated electric fields are of small strength because of the long wavelength of transmitted RF waveforms. However, for higher B_0 fields, both SAR values and electric fields' strengths are increased dramatically. As for B_1 fields, the causes of the increased SAR values and E fields' strengths can be summarized into the following four items, Farady's law [45], dielectric dispersion [75], inhomogeneous B_1 field, and decreased quadrature efficiency [76]. Faraday's law and dielectric dispersion phenomenons are difficult to solve in reality because they are caused by the physiological properties of human tissues. The latter two phenomenons, i.e., inhomogeneous B_1 field and decreased quadrature efficiency, are determined by the B_1 field design and can be controlled in pTx MRI systems. It is worth-mentioning that quadrature excitation can generate an ideal B_1^+ field and an almost zero B_1^- field in 1.5 T MRI [76]. However, at 3 T MRI, the average amplitude of the B_1^- field increased to about 40% of the B_1^+ field [36]. This means that high field MRI often requires to have high RF powers to generate the desired B_1^+ field because of the low quadrature efficiency.

2.4.4   SAR Regulations

The IEC [39] provides guidelines to regulate the MR safety by defining the standard global SAR, partial body SAR, and 10-g local SAR limits on different operating modes. The global (whole body averaged) SAR and partial body SAR (exposed body averaged) can reflect the core temperature in the scanned/exposed human body, and thus they are important parameters for online monitoring. Local SAR is a significant factor reflecting the localized heating, while it is difficult to assess in real-time, it should be carefully controlled during the RF transmission [36, 37].

There are three modes in the IEC standards, *Normal Mode*, *First Level Controlled Mode*, and *Second Level Controlled Mode*. *Normal Mode* is generally for clinical purpose and it requires the B_0 field strength to be $\leq$ 3 T. *First Level Controlled Mode* is usually for research purpose and it requires the B_0 field strength to be $\leq$ 4 T. *Second Level Controlled Mode* is rarely used in practice since it has a crucial risk in the human body and it always requires an explicit medical supervision.

RF exposure becomes sophisticated in high field or UHF MRI systems with the pTx platform, because amplitude and phase of the multi-dimensional RF pulses can induce large variations in both global and local SAR. This is reason we need to distinguish the use of different IEC operating modes. Different IEC modes and SAR regulations are summarized in Table 2.1, and these SAR values are averaged in a six-minute time window [39]. Note that IEC standards are required to have an ambient temperature less than or equal to 25°, and the SAR limits should be lower if it has a higher ambient temperature.

To better understand the physical meaning of the SAR unit, it is given that the resting metabolic SAR value is 1.3 W/kg, but the metabolic rate for highly trained athletes may go up to 18 W/kg if they are under vigorous exercise [77]. The maximum allowable energy deposited during the MRI exam is 14,400 Ws/kg, which equals to expose the body at the maximum global SAR environment in the *First Level Controlled Mode* for one hour [36].

Table 2.1.

IEC SAR Regulations with Different Operating Modes

	Global Body SAR	Global Head SAR	Partial Body SAR (Adult)	Partial Body SAR (Baby Head)	Local Head and Torso SAR	Local Extremity SAR
IEC Normal Mode	2.0 W/kg	3.2 W/kg	6.6 W/kg	3.4 W/kg	10.0 W/kg	20.0 W/kg
IEC First Level Controlled Mode	4.0 W/kg	3.2 W/kg	7.4 W/kg	5.0 W/kg	20.0 W/kg	40.0 W/kg
IEC Second Level Controlled Mode	>4.0 W/kg	>3.2 W/kg	>7.4 W/kg	>5.0 W/kg	>20.0 W/kg	>40.0 W/kg

2.4.5 Global and Local SAR

SAR values play a pivot role in the human MRI scans. Higher global SAR values often indicate higher core temperatures in the human body, but the actual heat diffusion is different from person to person. Higher temperature inside the human body will cause an increased respiratory rate, increased heart rate, and faster blood flow. Severe body heat will even cause the irreversible tissue damage. Therefore, SAR values need to be explicitly monitored during the MRI scan. RF exposure and the possible body heating are especially sensitive for the people with limited thermoregulatory capability, such as neonates, elderly patients, pregnant patients, anesthetized patients, and patients with diabetes. For these patients, MRI operators need to choose the correct global/local SAR limits in order to achieve the real-time supervision.

As discussed in previous sections, global and local SARs are commonly used in the RF exposure measurement in MRI. Global SAR values represent the overall RF exposure level in the whole body, and it reflects the core temperature level to some degree. There are two commonly used methods to measure the global SAR values in real-time [78], the calorimetric method and pulse energy method. The calorimetric method is used to validate the global SAR values, which measures the temperature, and it measures in the calorimetric phantom model (heat isolated model). The pulse

energy method measures the forward and reflected RF power, which can be used either in a phantom or in a human body.

Local SAR, unlike global, is difficult to measure in real-time scanning. Researchers tend to threshold local SAR values offline and construct a RF pulse designer that will never exceed local SAR limits. To implement this, local SAR values and E fields should be firstly estimated from the EM field simulation using numerical models. Second, by using these numerical models, researchers can threshold and control local SAR values in the RF pulse design and thus avoid the possible localized heating in the actual scanning. Localized heating is a really important issue in MRI scanners. In some cases, although body-averaged SAR values can be maintained in a lower level, the peak local SAR values can go beyond local SAR limits. Locations with high local SAR values are often found in places with high E field strength. For example, those locations are generally observed in the proximity to the physical coils. This is because this location has a high strength of *conservative E field*. High local SAR values are also located in the eddy current pathway in the human body, where the strength of *magnetically induced E field* is higher than other locations.

The universal solution of the local SAR control is difficult to generate, and thus researchers and clinicians tend to use different numerical models as references to design suitable RF pulses to prevent the possible localized heating under any circumstance. Their proposed RF pulse designers should guarantee to satisfy any local SAR constraint for any arbitrary human model and/or coil setup. There certainly are some regions in which researchers need to pay more attention in the RF pulse design procedure. For example, body regions closely contact to physical coils or foreign materials should be carefully monitored because they are often found to have high local SAR values, and even thermal burns in some academic findings [79, 80].

3. MIMO MRI MODELING

3.1 Introduction

Over the last 30 years, the clinical applications of MRI systems have become widespread, and research on multi-coil MRI has already yielded commercial deployment in clinical settings. The most popular multi-coil techniques used for MR image reception are forms of parallel imaging [81]. Parallel imaging is a robust method used to accelerate the acquisition of MRI data and has resulted in many new MRI applications [12, 13, 31]. Other research has addressed MRI systems using multiple independent transmitter coils under the theory of pTx [17, 20, 82]. Transmission of RF pulses with transmit coil arrays has been shown to reduce RF energy deposition, facilitating compliance with SAR safety limits [18, 83].

Multi-coil RF pulse design enables the acceleration of multidimensional excitation pulses by making use of multiple independently transmit coils [84]. To date, the majority of pulse designs assume operation in the small-tip-angle regime [52]. Pioneering pulse design methods in small-tip-angle regime were introduced by Katscher et al. [20], Zhu [18], and Grissom et al. [11]. Katscher et al. [20] proposed *Transmit SENSE*, which introduced parallel transmission with arbitrarily shaped transmit coils. *Transmit SENSE* can be applied to general cases of three dimensional imaging and arbitrary k-space trajectories. In contrast, the method used by Zhu [18] uses spatially selective excitation, allowing for both multidimensional excitation acceleration and SAR control during parallel excitation.

Despite extensive research on multi-coil transmission and/or reception, there have been only minor advancements in multi-coil processing at *both* sides of the scanner. Transceive multi-coil design for both transmission and reception has been the focus of several papers [16, 32–34]. These studies suggest that proper transceiver coil design

can effectively enable transmit and/or receive techniques while maintaining a high SNR. Other research has proposed a MSE-based excitation pattern design for parallel transmit and receive MR image reconstruction [26].

In this paper, we address multi-coil RF pulse designs using multiple transceivers. We refer to this technique as MIMO MRI. This research contributes to the investigation of transmit-side multi-coil RF pulse design by jointly considering the operation of multi-coil transmitters and receivers.

Specifically, we propose an alternative RF pulse design method for MIMO MRI, similar to the spatial domain method introduced by Grissom, et al. [11, 54]. Our method is specifically designed for the scenario of multiple transceive coils using an SNR cost function. Traditional multi-coil RF design approaches allow for spatially-varying excitation error weighting, such as ROI specification, and also have adaptability to the main magnetic field inhomogeneity. This MIMO MRI model provides a theoretical method to estimate and maximize the SNR at the receive side, while maintaining the excitation error in a controllable/manageable range. This paper outlines the theoretical framework of the MIMO MRI model, using numerical simulations for verification. Preliminary research on this method using an echo planar imaging trajectory was presented in an IEEE conference proceeding [70].

This paper is organized as follows: Section II reviews the fundamental MRI theory and introduces the MIMO MRI model setup; Section III outlines the novel RF pulse design method; several computer simulations are employed to verify this theoretical study in Section IV; discussions and conclusions are presented in Sections V and VI.

3.2 Theory

We consider an MRI system with a transceive coil array, where a number of coil elements are used for signal transmission, N_t, and another number of elements are used for signal reception, N_r. We are explicitly interested in systems optimized for

the case where $\min(N_t, N_r) > 1$, but results still apply when $\min(N_t, N_r) = 1$. All derivations in this section are made using the small-tip-angle approximation.

3.2.1 Parallel Excitation with Multi-Coils

Parallel excitation uses multiple coils with each transmitting a different RF pulse to excite spins over a certain bandwidth, yielding visible MR images at the receive side. During the spin excitation, the multi-dimensional RF pulse follows a certain k-space trajectory to enable imaging.

MR signal intensities are proportional to the transverse magnetization and dependent on the user-defined excitation parameters, e.g., repetition time (T_R), echo time (T_E), and slice selection. Each coil transmits a complex RF waveform with the pth coil sending $b_{1,p}(t)$. These RF pulses generate the transverse magnetization $M(\boldsymbol{r})$ (e.g., $\boldsymbol{r} = [x, y]^\top$ in 2D imaging), as determined from the Bloch equation. The transverse magnetization can be expressed as

$$M(\boldsymbol{r}) = \sum_{p=1}^{N_t} S_p(\boldsymbol{r}) M_p(\boldsymbol{r}), \tag{3.1}$$

where $S_p(\boldsymbol{r})$ is the transmit sensitivity map of the pth transmit coil at position $\boldsymbol{r}$, and $M_p(\boldsymbol{r})$ is the transverse plane magnetization produced by the pth transmit waveform.

Similar to methods outlined in [52] and [11],

$$M_p(\boldsymbol{r}) = i\gamma M_0 \int_0^T b_{1,p}(t) e^{i\gamma \Delta B_0(\boldsymbol{r})(T-t)} e^{i\boldsymbol{r} \cdot \boldsymbol{k}(t)} dt, \tag{3.2}$$

where γ is the gyromagnetic ratio, M_0 is the magnetization at the equilibrium state, T is the pulse length, $\boldsymbol{k}(t)$ represents the two-dimensional k-space trajectory ($[k_x(t), k_y(t)]^\top$), and $e^{i\gamma \Delta B_0(\boldsymbol{r})(T-t)}$ represents the phase accrued due to the static field deviation $\Delta B_0(\boldsymbol{r})$. We normally assume the static field is uniform over the magnetization domain.

Similar to approaches in [11] and [85], we discretize the problem. We assume that all position vectors, $\boldsymbol{r}$, are restricted to a finite set, $\mathcal{R} = \{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_\mathcal{R}\}$, and time values

are represented by $t \in \{t_1, t_2, \ldots, t_K\}$, where $t_k = t_1 + (k-1)\Delta t$ for a sample time, Δt. Using these assumptions, we define

$$\boldsymbol{m}_p = [M_p(\boldsymbol{r}_1)\ M_p(\boldsymbol{r}_2)\ \cdots\ M_p(\boldsymbol{r}_{\mathcal{R}})]^\top \tag{3.3}$$

$$= \boldsymbol{A}\boldsymbol{b}_{1,p}, \tag{3.4}$$

where $\boldsymbol{A} \in \mathbb{C}^{\mathcal{R} \times K}$ has entry (j,k) defined as

$$a_{jk} = i\gamma M_0 \Delta t e^{i\gamma \Delta B_0(\boldsymbol{r}_j)(T-t_k)} e^{i\boldsymbol{r}_j \cdot \boldsymbol{k}(t_k)}, \tag{3.5}$$

and $\boldsymbol{b}_{1,p} = [b_{1,p}(t_1)\ \cdots\ b_{1,p}(t_K)]^\top$. The transverse magnetization ($\boldsymbol{m} \in \mathbb{C}^{\mathcal{R} \times 1}$) can then be expressed as

$$\boldsymbol{m} = \sum_{p=1}^{N_t} \tilde{\boldsymbol{S}}_p \boldsymbol{m}_p = \sum_{p=1}^{N_t} \tilde{\boldsymbol{S}}_p \boldsymbol{A}\boldsymbol{b}_{1,p} = \tilde{\boldsymbol{S}}_A \boldsymbol{\mathcal{B}}, \tag{3.6}$$

where $\tilde{\boldsymbol{S}}_p = diag(\boldsymbol{S}_p(\boldsymbol{r}_1), \boldsymbol{S}_p(\boldsymbol{r}_2), \ldots \boldsymbol{S}_p(\boldsymbol{r}_{|\mathcal{R}|}))$, $\tilde{\boldsymbol{S}}_A = \left[\tilde{\boldsymbol{S}}_1 \boldsymbol{A}\ \cdots\ \tilde{\boldsymbol{S}}_{N_t} \boldsymbol{A}\right]$.

Using this discrete version, we can analyze the transverse magnetization by optimizing the matrix RF pulse

$$\boldsymbol{\mathcal{B}} = \begin{bmatrix} \boldsymbol{b}_{1,1} \\ \vdots \\ \boldsymbol{b}_{1,N_t} \end{bmatrix}. \tag{3.7}$$

Thus, the problem lies in the construction of the RF pulses in *both space and time*.

Previously, researchers have approached the design of $\boldsymbol{\mathcal{B}}$ to excite a given pattern $\boldsymbol{m}_{ref}$ [86], or to mitigate undersampling artifacts [20]. To meet a reference magnetization, the problem becomes one of solving the optimization ('H' is the conjugate transpose operator)

$$\boldsymbol{\mathcal{B}}_{ref} = \underset{\boldsymbol{\mathcal{B}}}{\operatorname{argmin}} \left(\left\| \tilde{\boldsymbol{S}}_A \boldsymbol{\mathcal{B}} - \boldsymbol{m}_{ref} \right\|^2 + \alpha \boldsymbol{\mathcal{B}}^{\mathrm{H}} \boldsymbol{\mathcal{B}} \right). \tag{3.8}$$

Integrated RF power is controlled via a Tikhonov regularization term $\alpha \boldsymbol{\mathcal{B}}^{\mathrm{H}} \boldsymbol{\mathcal{B}}$, where α is the tuning parameter. This magnetization pattern matching problem can be solved either by the pseudoinverse or the conjugate gradient (CG) method. Once $\boldsymbol{\mathcal{B}}_{ref}$ is obtained, all designed RF pulses can be obtained.

3.2.2 Parallel Reception with Multi-Coils

Parallel reception in an MRI system allows data to be separately and simultaneously acquired from multiple coils. Fast parallel reception methods are referenced as parallel imaging techniques, such as SiMultaneous Acquisition of Spatial Harmonics (SMASH) [31], SENSitivity-Encoding (SENSE) [12], and GeneRalized Autocalibrating Partially Parallel Acquisition (GRAPPA) [13].

The ideal image-combining procedure in the receive coil array is shown in Fig. 3.1. This image combination method is similar to the sum-of-squares (SoS) method [30]. However, the SoS method usually causes the reconstructed image to appear dark at locations further away from the receive coils [87]. Hence, receiver weights are applied to avoid non-uniformity of the image intensity. In addition, each individual channel is digitized so channel combinations can be performed digitally. Various image reconstruction algorithms can be applied to complete the image reconstruction; here, we used SENSE to obtain the electromotive force (emf) output, $\boldsymbol{V}$.

Often in the literature, B_1^- fields are used as the receive sensitivity maps for SENSE reconstruction, but the framework presented herein distinguishes the difference between the B_1^- fields and the receive sensitivity maps. The B_1^- fields should not only contain the coil patterns, but also the complex-valued RF waveform. The amplitude of the applied RF pulses contribute to the final strength of the B_1^- field. For q^{th} receive coil, we assume the corresponding receive coil sensitivity map is $S_q'(\boldsymbol{r})$. We define the B_1^- field, detected by the qth transceive coil, as

$$\widetilde{B}_{1,q}^-(\boldsymbol{r}) = S_q'(\boldsymbol{r}) \int_0^T b_{1,q}(t) e^{i\gamma \Delta B_0(\boldsymbol{r})(T-t)} dt. \tag{3.9}$$

The body/coil geometry and the electromagnetic parameters (e.g., permittivity and conductivity), and circular polarization components are considered in the receive sensitivity maps. B_1^+ and B_1^- fields are often observed to be different at higher field strengths [88–90], so we use sensitivity maps to account for the distinction. The time
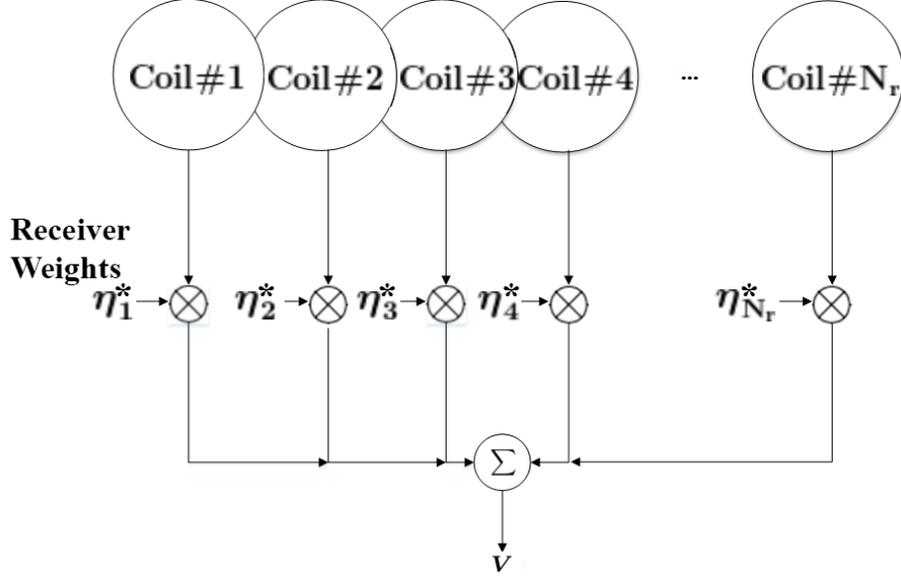
Fig. 3.1. Block diagram of a phased receive coil array. N_r receive coils with outputs phase shifted and summed through a set of transformers. η_q is the digitized receiver weight of qth receive coil. $\boldsymbol{V}$ is the complex electromotive force (emf) of the NMR signals received after applying the receiver weights.

integral of $b_1(t)$ is equivalent to a homogeneous scaling factor to the B_1^- field (assume uniform B_0 field), and it has no interference with image reconstruction [12].

Consider an axial plane 2D MRI, for which the $z-$component of the gradient waveforms is zero during the readout. If $C_q(\boldsymbol{r})$ is defined as the 2D MR signal received from the qth coil in the image domain, it can be represented by [44,91]

$$C_q(\boldsymbol{r}) \approx \delta_z M(\boldsymbol{r}) e^{-T_E/T_2(\boldsymbol{r})} \widetilde{B}_{1,q}^-(\boldsymbol{r}) + \mu_q(\boldsymbol{r}), \tag{3.10}$$

where M is the transverse plane magnetization generated by multiple transmit coils (see (3.1)), T_2 is the transverse relaxation time constant, δ_z is the slice thickness, $\mu_q(\boldsymbol{r})$ is the noise in q^{th} channel during the MR signal reception, and we assume $\mu_q(\boldsymbol{r})$ follows a Gaussian distribution.

After applying receiver weights $\{\eta_q, q \in [1, N_r]\}$, the effective MR output signal received from each receive channel is (slice thickness and the T_2 weighting coefficients are ignored for simplicity; '*' is the conjugate operator)

$$C_{eff,q}(\boldsymbol{r}) = \eta_q^*(\boldsymbol{r})M(\boldsymbol{r})\widetilde{B}_{1,q}^-(\boldsymbol{r}) + \eta_q^*(\boldsymbol{r})\mu_q(\boldsymbol{r}). \tag{3.11}$$

If the MR signals obtained from the receive coils are $\{C_{eff,q}(\boldsymbol{r}), q \in [1, N_r]\}$, then the complex receive voltage after the image combination, $V(\boldsymbol{r})$, is given as

$$V(\boldsymbol{r}) = M(\boldsymbol{r})\sum_{q=1}^{N_r} \eta_q^*(\boldsymbol{r})\widetilde{B}_{1,q}^-(\boldsymbol{r}) + \sum_{q=1}^{N_r} \eta_q^*(\boldsymbol{r})\mu_q(\boldsymbol{r}), \tag{3.12}$$

$$= V_{desired}(\boldsymbol{r}) + \sum_{q=1}^{N_r} \eta_q^*(\boldsymbol{r})\mu_q(\boldsymbol{r}). \tag{3.13}$$

Specifically, the combined noise is defined as the spatially-variant additive Gaussian noise. If for each location $\boldsymbol{r}$, the thermal noise across all coils is $\boldsymbol{\mu}(\boldsymbol{r}) \in \mathbb{R}^{N_r \times 1}$, then it should satisfy the following two conditions:

(a) $E[\boldsymbol{\mu}(\boldsymbol{r})] = 0$,

(b) $E[\boldsymbol{\mu}(\boldsymbol{r})\boldsymbol{\mu}^{\mathrm{H}}(\boldsymbol{r})] = \boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ is the covariance matrix.

The noiseless MRI receive voltage can be described in the matrix form:

$$\boldsymbol{V}_{desired}(\boldsymbol{r}) = \boldsymbol{m}(\boldsymbol{r})\boldsymbol{\eta}^{\mathrm{H}}(\boldsymbol{r})\widetilde{\boldsymbol{B}}_1^-(\boldsymbol{r}). \tag{3.14}$$

3.3 Methods

To date, researchers have implemented parallel excitation and parallel reception in the same MRI clinical machine. However, the optimal use of multi-coil MRI systems remains unknown. This novel MIMO MRI model aims to address this gap.

3.3.1   SNR Optimization Method

When using the SENSE method in the MIMO MRI model, two types of noise remain problematic during MRI acquisition: channel noise and thermal fluctuation noise.

There are two types of channel noises that commonly affect the reconstructed images: noise in k-space sample values and noise in the sensitivity data. Sensitivity-related noise can be reduced to a negligible level by smoothing, and the sample data related noise can be minimized by utilizing different optimization methods, e.g. Parallel magnetic resonance imaging with Adaptive Radius (PARS) method [15]. The MIMO MRI model works to minimize channel and thermal noise, as well as maximize the SNR after image combination.

In the MIMO MRI model, we aim to match a desired magnetization pattern while obtaining the maximized SNR of the MRI receive signal. We use N_t transmit coils for RF pulse transmission, and N_r receive coils for image reception. However, we assume that $N_r = N_t$ in this RF pulse design procedure, and we use the coil selection matrix to select the corresponding receive coils for image reconstruction.

The SNR after the image combination can be defined as the ratio of the absolute value of noiseless MRI signal voltage to the standard deviation of noise voltage,

$$SNR(\boldsymbol{r}) = \frac{|V_{desired}(\boldsymbol{r})|}{\sqrt{\Sigma_{total}(\boldsymbol{r})}}. \tag{3.15}$$

$V_{desired}$ contains the complex receive voltage of the selected ROI. For each voxel, the NMR signal voltage is defined as the emf induced from that specific location, which is defined in (3.14). The covariance matrix Σ_{total} at each location is defined as

$$\Sigma_{total}(\boldsymbol{r}) = \sum_{j=1}^{N_r}\sum_{k=1}^{N_r} \eta_j^*(\boldsymbol{r})\eta_k(\boldsymbol{r})\Sigma_{jk} = \boldsymbol{\eta}^{\mathrm{H}}(\boldsymbol{r})\boldsymbol{\Sigma}\boldsymbol{\eta}(\boldsymbol{r}), \tag{3.16}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix of $\boldsymbol{\mu}(\boldsymbol{r})$, η_j is the receiver weight of jth receive coil. Σ_{jk} represents the correlated ($j \neq k$) and uncorrelated ($j = k$) thermal noise between

coils j and k. We assume Σ_{jk} is the noise covariance matrix induced by additive Gaussian noises in the simulation.

For simplicity, we ignore the constants which remain unchanged from voxel to voxel, and we further define the *coil SNR, SNR_c*, as

$$|SNR_c(\boldsymbol{r})| = \frac{\left|M(\boldsymbol{r}) \sum_{q=1}^{N_r} \eta_q^*(\boldsymbol{r}) \widetilde{B}_{1,q}^-(\boldsymbol{r})\right|}{\sqrt{\sum_{j=1}^{N_r} \sum_{k=1}^{N_r} \eta_j^*(\boldsymbol{r}) \eta_k(\boldsymbol{r}) \Sigma_{jk}}}, \tag{3.17}$$

where $\widetilde{B}_{1,q}^-(\boldsymbol{r})$ is defined in (3.9).

The goal of this MIMO MRI model is to achieve the desired excitation pattern, i.e., match the induced transverse magnetization $\boldsymbol{m}$ (see (3.6)) to the desired excitation pattern $\boldsymbol{m}_{ref}$ (see (3.8)), while simultaneously maximizing the SNR_c in (3.17) to achieve high SNR in the receive signal. SNR_c is a useful metric for a RF pulse design algorithm since it preserves image quality of the reconstructed MR images and further provides the relative SNR of each voxel in the image. However, it is not convenient to apply the SNR_c matrix within the optimization, as it does not provide a single value to optimize. To combat this problem, we define the global SNR ($gSNR$) as the mean value of SNR_c^2 in a specific ROI. The definition of $gSNR$ will be given later in this section.

The efficient RF pulse design algorithm in this paper matches the magnetization pattern $\boldsymbol{m}_{ref}$ and the $gSNR$ term simultaneously. In the optimization, we first take the square of the SNR_c so that (3.17) becomes

$$SNR_c^2(\boldsymbol{r}) = \frac{|M(\boldsymbol{r})|^2 \sum_{j=1}^{N_r} \sum_{k=1}^{N_r} \eta_j^*(\boldsymbol{r}) \eta_k(\boldsymbol{r}) \widetilde{B}_{1,j}^-(\boldsymbol{r}) \widetilde{B}_{1,k}^-(\boldsymbol{r})^{\mathrm{H}}}{\sum_{j=1}^{N_r} \sum_{k=1}^{N_r} \eta_j^*(\boldsymbol{r}) \eta_k(\boldsymbol{r}) \Sigma_{jk}}. \tag{3.18}$$

Second, we discretize the term SNR_c^2 to a $\mathcal{R} \times 1$ vector $\boldsymbol{SNR}_c^2$, which is given as

$$\boldsymbol{SNR}_c = \begin{bmatrix} \boldsymbol{SNR}_c(\boldsymbol{r}_1) & \boldsymbol{SNR}_c(\boldsymbol{r}_2) & \cdots & \boldsymbol{SNR}_c(\boldsymbol{r}_{\mathcal{R}}) \end{bmatrix}^{\top}. \tag{3.19}$$

For each voxel in the selected ROI, its corresponding SNR is given as

$$|\boldsymbol{SNR}_c(\boldsymbol{r})|^2 = \frac{|\boldsymbol{m}(\boldsymbol{r})|^2 \, \boldsymbol{\eta}(\boldsymbol{r})^{\mathrm{H}} \widetilde{\boldsymbol{B}}_1^-(\boldsymbol{r}) \widetilde{\boldsymbol{B}}_1^-(\boldsymbol{r})^{\mathrm{H}} \boldsymbol{\eta}(\boldsymbol{r})}{\boldsymbol{\eta}(\boldsymbol{r})^{\mathrm{H}} \boldsymbol{\Sigma} \boldsymbol{\eta}(\boldsymbol{r})}. \tag{3.20}$$

Additionally, $\widetilde{\boldsymbol{B}}_1^-(\boldsymbol{r})$ is defined as a column vector that contains the receive fields at a specific position $\boldsymbol{r}$ from all receive coils (from (3.9)), given by

$$\widetilde{\boldsymbol{B}}_1^-(\boldsymbol{r}) = \begin{bmatrix} \boldsymbol{A}_1'(\boldsymbol{r})\boldsymbol{b}_{1,1} & \boldsymbol{A}_2'(\boldsymbol{r})\boldsymbol{b}_{1,2} & \cdots & \boldsymbol{A}_{N_r}'(\boldsymbol{r})\boldsymbol{b}_{1,N_r} \end{bmatrix}^{\top}, \tag{3.21}$$

where the system matrix $\boldsymbol{A}_q'(\boldsymbol{r}_j) \in \mathbb{C}^{1 \times K}$ has k entries

$$\boldsymbol{A}_q'(\boldsymbol{r}_j, k) = S_q'(\boldsymbol{r}_j) e^{i\gamma \Delta B_0(\boldsymbol{r}_j)(T - t_k)}. \tag{3.22}$$

We assume $N_r = N_t$ in this scenario, then (3.20) can be written as

$$|\boldsymbol{SNR}_c(\boldsymbol{r})|^2 = \frac{|\boldsymbol{m}(\boldsymbol{r})|^2 \, \boldsymbol{\eta}(\boldsymbol{r})^{\mathrm{H}} \boldsymbol{E} \boldsymbol{D}(\boldsymbol{r}) \boldsymbol{\mathcal{B}} \boldsymbol{\mathcal{B}}^{\mathrm{H}} \boldsymbol{D}(\boldsymbol{r})^{\mathrm{H}} \boldsymbol{E}^{\mathrm{H}} \boldsymbol{\eta}(\boldsymbol{r})}{\boldsymbol{\eta}(\boldsymbol{r})^{\mathrm{H}} \boldsymbol{\Sigma} \boldsymbol{\eta}(\boldsymbol{r})}, \tag{3.23}$$

where $\boldsymbol{E} \in \mathbb{R}^{N_t \times N_t}$ is the receive coil selection matrix, and $\boldsymbol{D}(\boldsymbol{r}) \in \mathbb{C}^{N_r \times (N_r K)}$ is defined as $\boldsymbol{D}(\boldsymbol{r}) = diag(\boldsymbol{A}_q'(\boldsymbol{r}))$.

The expression of SNR_c^2 is a ratio of quadratic Hermitian forms, one that commonly arises in optimization problems. We can determine the optimal SNR by finding the point when the gradient of $|\boldsymbol{SNR}_c|^2$ with respect to $\boldsymbol{\eta}(\boldsymbol{r})$ is zero (N_r equations). Solving for $\boldsymbol{\eta}$ yields optimum receiver weights. Using the methods in [92] and [93], the optimum receive weights for each voxel are given by

$$\boldsymbol{\eta}(\boldsymbol{r}) = \boldsymbol{\Sigma}^{-1} \widetilde{\boldsymbol{B}}_1^-(\boldsymbol{r}) = \boldsymbol{\Sigma}^{-1} \boldsymbol{E} \boldsymbol{D}(\boldsymbol{r}) \boldsymbol{\mathcal{B}}. \tag{3.24}$$

We can substitute the expression for $\boldsymbol{\eta}(\boldsymbol{r})$ from (3.24) into (3.20), and the equation becomes

$$|\boldsymbol{SNR}_c(\boldsymbol{r})|^2 = |\boldsymbol{m}(\boldsymbol{r})|^2 \left| \widetilde{\boldsymbol{B}}_1^-(\boldsymbol{r})^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \widetilde{\boldsymbol{B}}_1^-(\boldsymbol{r}) \right| \tag{3.25}$$

$$= |\boldsymbol{m}(\boldsymbol{r})|^2 \left| \boldsymbol{\mathcal{B}}^{\mathrm{H}} \boldsymbol{D}(\boldsymbol{r})^{\mathrm{H}} \boldsymbol{E}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \boldsymbol{E} \boldsymbol{D}(\boldsymbol{r}) \boldsymbol{\mathcal{B}} \right|. \tag{3.26}$$

The $gSNR$ is defined as the averaged value of the SNR_c^2 matrix in a specific ROI:

$$gSNR(\boldsymbol{\mathcal{B}}) = \frac{1}{\mathcal{R}} \left| \boldsymbol{\mathcal{B}}^{\mathrm{H}} \sum_{\boldsymbol{r}} |\boldsymbol{m}(\boldsymbol{r})|^2 \boldsymbol{D}(\boldsymbol{r})^{\mathrm{H}} \boldsymbol{E}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \boldsymbol{E} \boldsymbol{D}(\boldsymbol{r}) \boldsymbol{\mathcal{B}} \right|. \tag{3.27}$$

Note that $\boldsymbol{m} \approx \boldsymbol{m}_{ref}$ because the simulated pattern should be very close to the target pattern after the excitation. Therefore, (3.27) is in quadratic form if we assume $\boldsymbol{m} = \boldsymbol{m}_{ref}$. In this case, the dominant eigenvector ($\boldsymbol{\mathcal{B}}_{eig}$) of its system matrix can be used as RF pulse samples to maximize $gSNR$.

Recall that RF pulse samples, $\boldsymbol{\mathcal{B}}_{ref}$ (see (3.8)), achieve the desired excitation pattern in MRI parallel excitation, and $\boldsymbol{\mathcal{B}}_{eig}$ obtained from (3.27) maximizes the $gSNR$ of the output signal in MRI parallel reception. Thus, a gradient descent algorithm is constructed to optimize a real β coefficient correlated with $\boldsymbol{\mathcal{B}}_{eig}$ and $\boldsymbol{\mathcal{B}}_{ref}$, to update the resulting RF pulse samples, $\boldsymbol{\mathcal{B}}_{opt}$. The finalized RF pulse samples, $\boldsymbol{\mathcal{B}}_{opt}$, can maximize $gSNR$ while controlling the excitation error and the RF power (ϵ_{max} is the maximum excitation error allowed; P is the maximum power constraint).

Note that we checked $\boldsymbol{\mathcal{B}}_{eig}$ (and scale back if necessary) to see if it satisfies the power constraint in (3.30) before the optimization starts. This is because $\boldsymbol{\mathcal{B}}_{eig}$ with any constant scale factor is also an eigenvector with the same eigenvalue regarding to the same Hermitian matrix. This scaling method will help controlling the RF power constraints.

$$\boldsymbol{\mathcal{B}}_{opt} = \underset{\boldsymbol{\mathcal{B}}, \beta}{\operatorname{argmax}} \, gSNR(\boldsymbol{\mathcal{B}}), \tag{3.28}$$

$$\text{where } \boldsymbol{\mathcal{B}} = \beta \boldsymbol{\mathcal{B}}_{ref} + \sqrt{1 - \beta^2} \boldsymbol{\mathcal{B}}_{eig}, \; |\beta| \leq 1;$$

$$\text{subject to } \frac{\left\| \tilde{\boldsymbol{S}}_A \boldsymbol{\mathcal{B}} - \boldsymbol{m}_{ref} \right\|^2}{\|\boldsymbol{m}_{ref}\|^2} - \epsilon_{max} \leq 0, \tag{3.29}$$

$$\boldsymbol{\mathcal{B}}^{\mathrm{H}} \boldsymbol{\mathcal{B}} - P \leq 0. \tag{3.30}$$

The maximum excitation error is given by the normalized root-mean-square Error (NRMSE) [54] between the desired pattern and the simulated transverse plane magnetization pattern, and is given in the following equation:

$$\mathrm{NRMSE} = \frac{\|\boldsymbol{m}_{ref} - \boldsymbol{m}_{sim}\|^2}{\|\boldsymbol{m}_{ref}\|^2}. \tag{3.31}$$

The Karush-Kuhn-Tucker (KKT) conditions can be used to verify its optimality. We examined the Lagrange dual problem of (3.28) and the duality gap. The resulting $\mathcal{B}_{opt}$ might be sub-optimal, but it can substantially improve the resulting SNR.

3.3.2   Acceleration Factor in MIMO MRI model

Acceleration factor (AF) is defined as the maximum factor for which the imaging time is reduced with acceptable loss in SNR. We define AF_e as the AF in the parallel excitation stage of the MIMO MRI model.

In reality, acceleration in parallel excitation is achieved via undersampling, resulting in a reduction of the excitation FOV for each individual coil's excitation pattern. The practical AF can be defined as the ratio of the desired excitation FOV (FOV_f) to the accelerated excitation FOV (FOV_a) of a given k-space trajectory per transmit coil [11, 18, 54], which can be expressed as

$$AF_e = \frac{FOV_f}{FOV_a}. \tag{3.32}$$

In the MIMO MRI model, there is no acceleration during the MR image reception. Instead, we experimented with different numbers of receive elements (N_r) and investigated the effect of multi-coil reception in MRI.

3.4   Simulations and Results

The Bloch simulation is performed to test the new RF pulse design method. For the excitation, simulations were implemented over a 64×64 grid covering a $24 \times 24 \ cm^2$ region, assuming that all simulations were carried out by an eight-channel transceiver coil array [84]. Transceive sensitivity patterns were obtained using FDTD simulation at 3.0 T environment in [84] and [54].

$N_t = 8$ transmit coils were used to reduce the parallel excitation-associated B_1 field inhomogeneity, and $N_r \leq 8$ coils were used for parallel reception. The corresponding magnitudes of B_1^+ patterns from the eight-channel parallel excitation are shown in

Fig. 3.2. The transmit and receive sensitivity maps are extracted from the given B_1^+ patterns [90]. The receive field of each qth coil $(\widetilde{B}_{1,q}^-)$, defined in (3.9), was calculated based on the designed RF pulses and the receive sensitivity maps.

In the simulation, we designed RF pulses to excite the desired slice in a sphere phantom. The RF pulses were designed not only to reduce erroneous excitation outside the desired region, but also to maximize the $gSNR$ in the receive signal, $\boldsymbol{V}$, when a specific excitation pattern was used. Different ϵ_{max} values (0.05-0.15) were used for each simulation. During the image reconstruction, we assumed the noise across all receive coils to be $\boldsymbol{\mu}(\boldsymbol{r})$, where $\boldsymbol{\mu}(\boldsymbol{r}) \sim \mathcal{N}(0, 0.5)$.
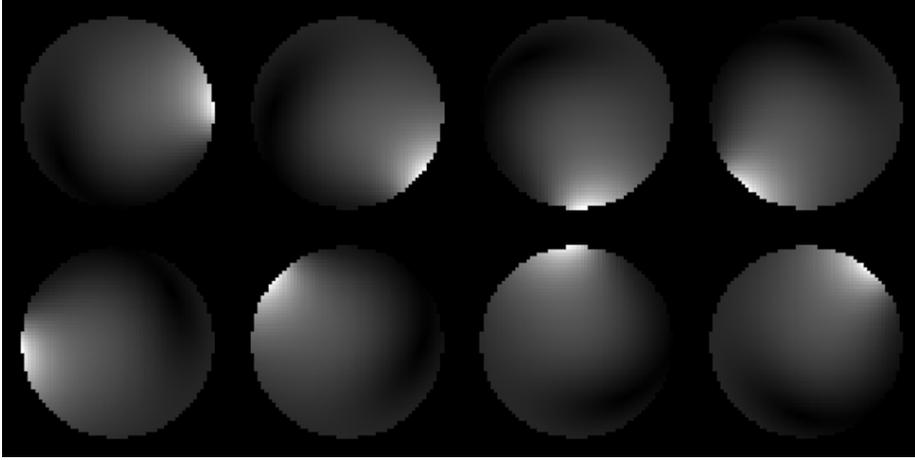
Fig. 3.2. B_1^+ patterns measured on a spherical phantom using an eight-channel transceiver coil array. The circular region indicates the ROI mask $(radius = 10\,cm)$ has been applied.

3.4.1 Acceleration Factor

The desired excitation pattern in this simulation, shown in Fig. 3.3(a), was a smoothed circular slice in a sphere phantom $(radius \approx 4.5\,cm)$. The excitation pattern was convoluted with a Gaussian kernel to reduce the Gibbs ringing artifacts. Fig. 3.3(b) shows the transverse plane magnetization after exciting the plane with

a ($\pi/2$) RF pulse. The applied ROI mask was a circular region shown in Fig. 3.2 ($radius = 10\ cm$).
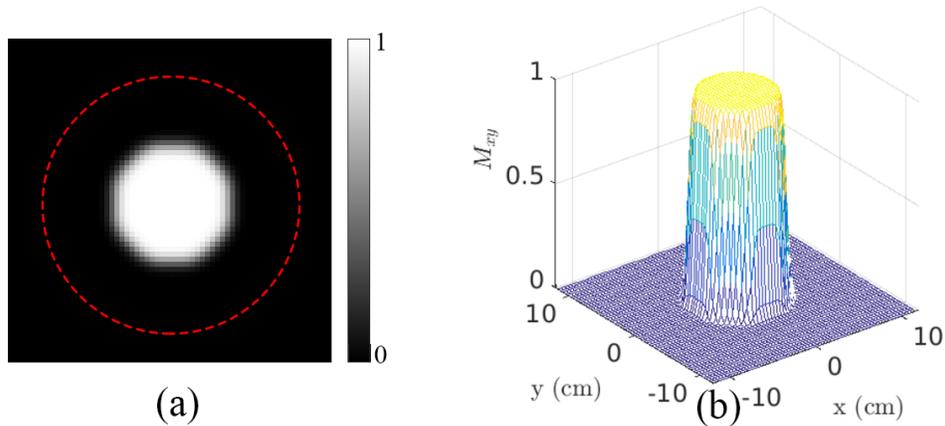
Fig. 3.3. Desired excitation pattern and transverse plane magnetization used in the simulations. (a) the desired excitation pattern within a given ROI (red dashed circle), which is a smoothed slice of a sphere phantom ($radius \approx 4.5\ cm$); (b) the transverse plane magnetization (M_{xy}), which is a smoothed circle block, whose peak was scaled to $\pi/2$, corresponding to 90° tip angle.

We designed RF pulses with a single-shot, spiral-out excitation k-space trajectory. This k-space trajectory was used with the following GE Discovery MR750 3.0 T scanner parameters: $5\ G/cm$ maximum gradient amplitude, $20\ G/cm/ms$ maximum slew rate, and $4\ \mu s$ sampling period [94]. To accelerate the excitation, we used the trajectory with variable sampling density, resulting in reduced FOV (FOV_a) for each individual transmit coil. The acceleration factor was defined as the ratio of the accelerated FOV and the desired FOV per transmit coil (given in (3.32)). The spiral trajectory was chosen for these simulations to obtain a spatial resolution of $0.75\ cm$ for the case of no acceleration. In addition, the spatial resolution increased as sampling density decreased.

In this simulation, we investigated the excitation performance by applying different acceleration schemes during parallel excitation. Although these accelerations

translate to undersampling the excitation k-space, aliasing effects were minimal since the optimization cost function and constraint functions limited the NRMSE. Furthermore, the ROI definition provided additional degrees of freedom for maintaining excitation accuracy. Different AF_e values were applied during the simulation to investigate the changes in $gSNR$. For $AF_e = 6$, the final $gSNR$ improved by 18.36%, and the NRMSE between the desired magnetization pattern and the simulated pattern was 0.12. Overall, the improvement of $gSNR$ decreased as AF_e increased.

As AF_e increased from 1 to 10, the pulse length T decreased from $9.3\,ms$ (no reduction in each receive FOV) to $0.88\,ms$ (10-fold reduction in each receive FOV). The left column in Fig. 3.4 shows the simulated receive voltage after image combination, $\boldsymbol{V}$, produced by the designed RF pulses; the right column is the transverse plane magnetization after excitation. We aimed to maximize the $gSNR$ of $\boldsymbol{V}$, while maintaining a controlled erroneous excitation (NRMSE ≤ 0.12). The results in Fig. 3.4 show that $gSNR$ decreased and that NRMSE was always less than 0.12 when AF_e was incremented from 2 to 6. However, the $gSNR$ generated by the SNR-controlled RF design method improved compared to the multi-coil RF design method without SNR control.

Because the maximum excitation error tolerance was set with ϵ_{max}, some erroneous excitation outside the desired cylinder always remained. Erroneous excitation became apparent when $AF_e \geq 7$ (over 85.7% reduction in individual k-space data), and corresponded to decreased efficiency of the SNR-controlled RF design method. The transverse magnetization in the right column of Fig. 3.4 suggests that there were minimal biases in the center of the excitation pattern.

Two different RF pulse design methods (i.e., with SNR control and without SNR control) were compared in Fig. 3.5 and Fig. 3.6. The RF pulse design method without SNR control was determined by minimizing the NRMSE between the desired magnetization pattern and the simulated magnetization pattern, which was the solution obtained from (3.8). The results illustrate that these two methods generated different SNRs during image reconstruction. When AF_e was increased, the $gSNR$ generated
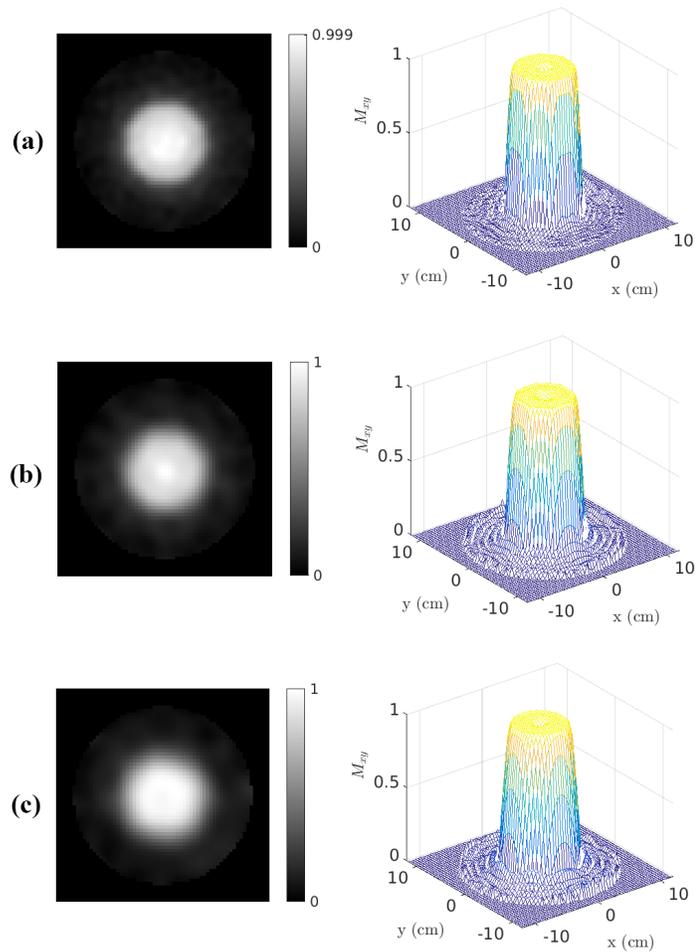
Fig. 3.4. The left column shows the simulated MRI receive voltage, V, produced by the designed RF pulses. The right column shows the transverse plane magnetization (M_{xy}) obtained through the exact Bloch simulation; excitation outside the cylinder is the erroneous excitation. (a) $N_t = 8, N_r = 8, AF_e = 2, T = 4.7\,ms, gSNR = 17.45\,dB$ (improved 130.52%), NRMSE = 0.10; (b) $N_t = 8, N_r = 8, AF_e = 4$, $T = 2.3\,ms, gSNR = 11.88\,dB$ (improved 59.89%), NRMSE = 0.10; (c) $N_t = 8, N_r = 8, AF_e = 6, T = 1.6\,ms, gSNR = 9.09\,dB$ (improved 18.36%), NRMSE = 0.12. The $gSNR$ is shown to be substantially improved in the SNR-controlled method with the NRMSE constraint of 0.12.

by the SNR-controlled method was always higher than the method without SNR control; but the improvement decreased as AF_e increased. The $gSNR$s generated by

these two methods are shown in Fig. 3.5, related to the RF transmission time (i.e., pulse length T). High accelerations led to shorter RF pulse lengths and lower SNRs. We observed that $gSNR$ obtained in the SNR-controlled method decreased, although it was always greater than the $gSNR$ obtained in the method with no SNR control. The multi-coil RF pulse design method without SNR control maintained the $gSNR$ when $T > 2\,ms$, but dropped when $T \leq 2\,ms$.
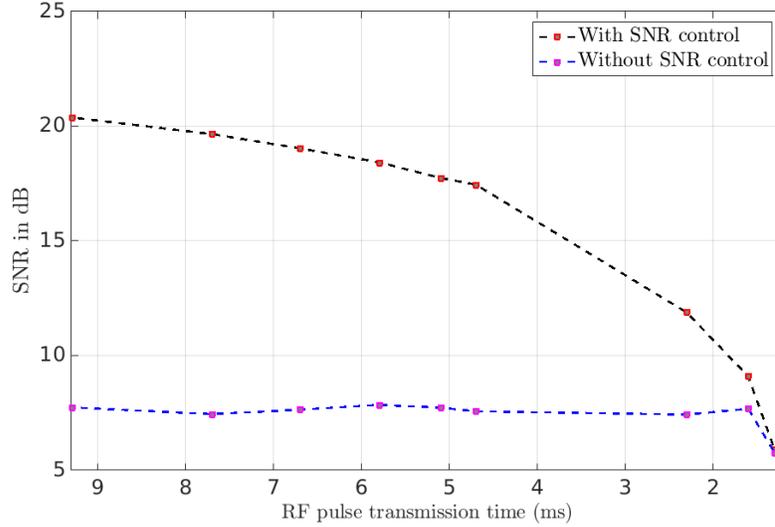
Fig. 3.5. Two different $gSNR$s ($gSNR_{w/\ control}$, $gSNR_{w/o\ control}$) are plotted associated with the pulse length T. Varying AF_e (1-7.5) was applied in these simulations. $N_t = 8, N_r = 8$ elements were used, and the NRMSE $\in [0.10, 0.12]$ in all simulations. The $gSNR$ obtained from the SNR-controlled method was improved when $T > 2\,ms$.

3.4.2   Receive Coil Array

In the second simulation, we investigated the $gSNR$ associated with the receive coils. In Fig. 3.6, we observed that the $gSNR$ increased when the number of receive elements was increased from 1 to 8. The $gSNR$ in the SNR-controlled method increased faster than the method without SNR control, suggesting that the SNR-controlled method could substantially improve SNR when the number of receive coils

was sufficient. All simulations were implemented when NRMSE $\leq \epsilon_{max}$. For the majority of cases, we optimized the RF pulses until NRMSE $= \epsilon_{max}$ to fulfill the KKT conditions.
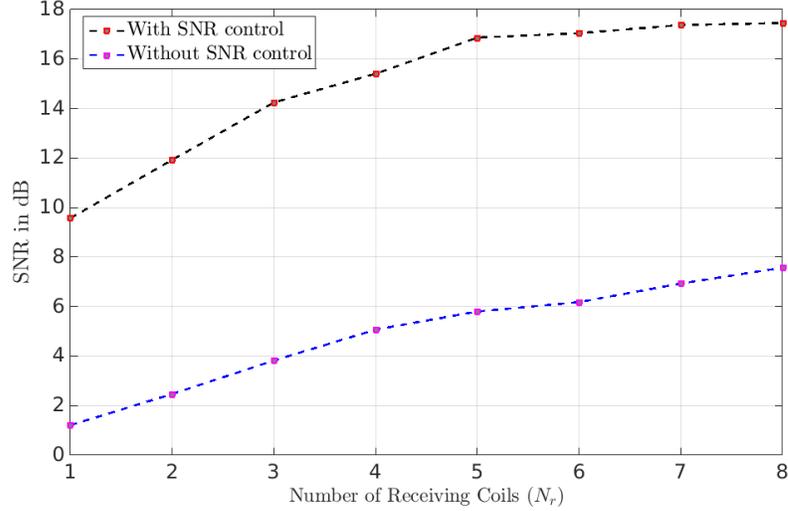
Fig. 3.6. $gSNR$ in two different RF pulse design methods, with and without SNR control. Different numbers of receive elements N_r (1-8) were applied in these simulations. Other parameters were $N_t = 8, AF_e = 2, T = 4.7\,ms$, NRMSE $\in [0.08, 0.10]$ in all simulations. The $gSNR$ increased when the number of receivers was increased in both methods, but the SNR-controlled method has greater improvement.

The $gSNR$ of the simulated MRI receive voltage, $\boldsymbol{V}$, was increased in both RF design methods (i.e., with SNR control and without SNR control). However, the SNR-controlled method increased the $gSNR$ substantially compared to the method without SNR control, especially when N_r was small.

We also observed that the intensity values of the receive voltage, $\boldsymbol{V}$, became more uniform when we used more receive coils during the image reception. In Fig. 3.7(a), we found that the output, $\boldsymbol{V}$, appeared dark at locations further away from the used receive coils, but this phenomenon completely disappeared in Fig. 3.7(c), suggesting that more receive coil elements could reduce the intensity homogeneity. There were also intensity variations when we used six receive coils for image reception (see Fig.

3.7(b)), but the variation became unnoticeable when eight receive coils were used. Fig. 3.7 illustrates that, when the number of receive elements was incremented from four to eight, the output, V, became more homogeneous, and the NRMSE between the simulated transverse magnetization and desired magnetization was equal to 0.1 for all simulations.
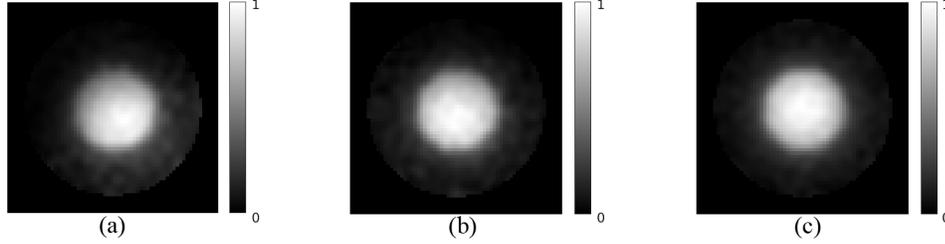
Fig. 3.7. The simulated MRI receive voltage, V, using different receive coil elements, under three different conditions. The MRI output became more homogeneous as the number of MRI receivers increased, although the excitation accuracy remained the same for all three conditions. (a) $N_t = 8, N_r = 4, AF_e = 2, T = 4.7\,ms, gSNR = 15.40\,dB$ (improved 203.75%), NRMSE = 0.10; (b) $N_t = 8, N_r = 6, AF_e = 2$, $T = 4.7\,ms, gSNR = 17.04\,dB$ (improved 175.73%), NRMSE = 0.10; (c) $N_t = 8, N_r = 8, AF_e = 2, T = 4.7\,ms, gSNR = 17.45\,dB$ (improved 130.52%), NRMSE = 0.10.

3.4.3   Maximum Excitation Error Tolerance

Fig. 3.8 plots the $gSNR$ associated with ϵ_{max} during the SNR optimization process. The desired excitation pattern in this simulation is shown in Fig. 3.3. An acceleration factor of 2 was chosen which corresponded to a $4.7\,ms$ RF pulse in the MIMO MRI model.

We further applied an interpolation method to show the increasing trend of $gSNR$ with respect to ϵ_{max}. The relationship trend in Fig. 3.8 shows that the $gSNR$ started to logarithmically increase when ϵ_{max} increased, suggesting that ϵ_{max} played a pivot role in controlling $gSNR$ values when it was small-valued. Furthermore, there was a

trade-off between the excitation errors and the resulting SNR. Note that ϵ_{max} can be used to optimize the excitation error tolerance and the SNR.
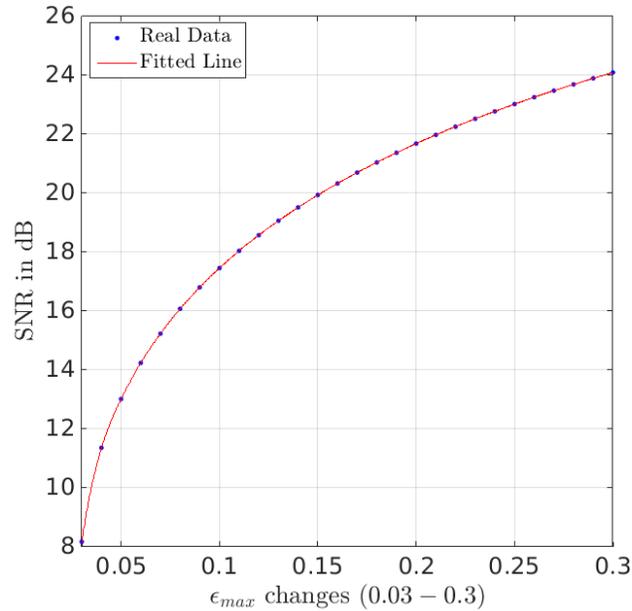
Fig. 3.8. $gSNR$ associated with ϵ_{max} changes. This simulation was implemented with the desired circular excitation pattern with $radius \approx 4.5cm$, $N_t = 8, N_r = 8, AF_e = 2, T = 4.7\,ms$. ϵ_{max} was increased from 0.03 to 0.3. Piecewise cubic interpolation method was applied to implement the regression. The $gSNR$ growed rapidly with small ϵ_{max}, but the gains decreased and became more difficult as ϵ_{max} increased. The $gSNR$ would not increase to infinity.

3.4.4 Different Excitation Pattern

Previous research showed that excitation patterns can be utilized to improve the receive performance in terms of g-factor [14], and this would further improve the performance in SNR. Fig. 3.9 shows results for a circular excitation pattern using various radius sizes. It can be concluded that $gSNR$ increased when the size of the excitation region increased. There were minimal bias effects in the center of both transverse plane magnetizations, and it was observed that the resulting excitation

patterns in both simulations remained homogeneous in the excited region. The applied ROI mask helped reduce the B_1 field inhomogeneity when the excitation region was close to the ROI. Additionally, the iterative method was effective in suppressing the aliasing effect outside the cylinder because the effect was accounted for in the optimization cost function and the constraint functions.

Fig. 3.10 illustrates the desired and simulated excitation patterns of the Shepp-Logan phantom, inverted to yield a larger region of appreciable magnetization. We applied an ROI mask to the edge of the desired excitation pattern. The SNR-controlled RF design method indicated higher excitation accuracy and improved $gSNR$ compared to the RF design method without SNR control. This simulation illustrated that the new multi-coil RF design method with SNR control could be utilized for arbitrary excitation patterns.

3.5   Discussion

Together, the results confirm that MIMO MRI is a novel and useful theoretical framework for MRI parallel RF transmission and reception. Through computer simulations, we verified the RF pulse design strategy for this specific model. This MIMO MRI framework constructed a signal processing model between the transmitters and receivers of the MRI system by using both parallel excitation and parallel reception. This signal processing model is analogous to techniques developed for MIMO communication.

This framework was built based on the *spatial domain* method [11] and the *Transmit SENSE* method [20]. These methods focus on transmit coils and other parameters that could significantly affect parallel excitation. However, our results confirm that it is worthwhile to investigate the relationship between transmit and receive coils in an MRI system. Our proposed RF pulse design method used RF pulses to optimize several quadratic cost functions comprised of $gSNR$, weighted error terms, and Tikhonov regularization terms, while maintaining the excitation errors in a con-
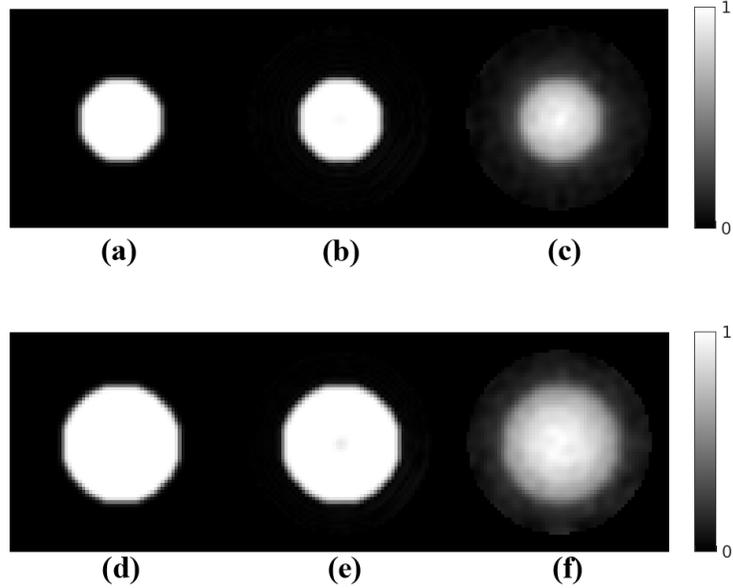
Fig. 3.9. Top row represents the desired circular excitation pattern with $radius^2 = 20\,cm^2$, $N_t = 8, N_r = 8, AF_e = 2$, $T = 4.7\,ms$, $gSNR = 13.02\,dB$ (improved 71.99%), NRMSE= 0.05; the second row represents the circular excitation pattern with $radius^2 = 40\,cm^2$, $N_t = 8, N_r = 8, AF_e = 2$, $T = 4.7\,ms$, $gSNR = 16.65\,dB$ (improved 73.08%), NRMSE = 0.05. (a) and (d) are the desired excitation patterns; (b) and (e) are the simulated excitation patterns; (c) and (f) are the simulated MRI receive voltages $\boldsymbol{V}$.

trollable scope. The Tikhonov regularization term was used to control the peak and integrated RF power. An ROI mask was incorporated in the model to instill high excitation accuracy at high acceleration factors since errors outside the ROI do not contribute to $gSNR$ and NRMSE.

In an ideal experiment, the ROI can be determined during the RF design procedure with pre-scan images obtained from sensitivity mapping, such as ASSET or ARC in a GE MRI system. Furthermore, our method compensated for the B_1 field inhomogeneity since eight transmit channels were used for parallel excitation. When all eight receive channels were used, the simulated transverse plane magnetization showed uniform transverse plane magnetization, though it may have induced excita-
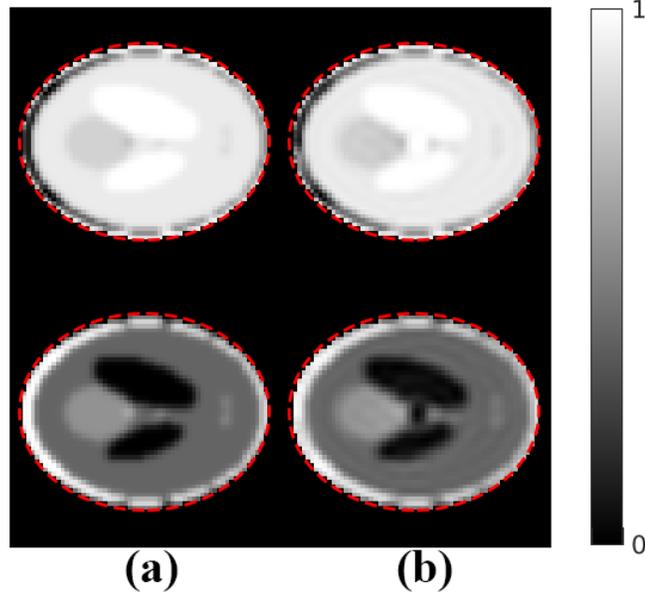
Fig. 3.10. Desired (a) and simulated (b) excited excitation pattern of the inverted Shepp-Logan phantom. The top row represents the M_{xy} magnetization, and the bottom row represents the M_z magnetization. $N_t = 8$, $N_r = 8$, $AF_e = 2$, $T = 4.7\,ms$, $gSNR = 16.05\,dB$ (improved 55.67%), NRMSE = 0.03. The red dashed ellipse indicates the applied ROI mask. The SNR-controlled method allowed for arbitrary excitation profile, and it indicated improved $gSNR$ compared to the method without SNR control.

tion errors outside the excited region. In addition to these advantages, our RF design method also allowed for SNR maximization, although there still existed a trade-off between $gSNR$ maximization and NRMSE minimization. These two cost functions must be balanced to achieve optimal system performance. When NRMSE = 0.12 and AF_e was increased from 2 to 6, the $gSNR$ in the SNR-controlled method improved by 18-130% as compared to the $gSNR$ obtained in the method without SNR control for a given circular excitation pattern (see Fig. 3.4).

Although most of the designed RF pulses were sub-optimal solutions (duality gap is not zero within KKT conditions), they still improved the $gSNR$ tremendously

and maintained excitation errors within the desired scope. Therefore, although these solutions were sub-optimal, they remained viable.

In the current study, the MIMO MRI model was described in the math form should be valid for arbitrary k-space trajectories and coil sensitivity profiles. This model worked successfully for tip angles ranging from $10°$ to $90°$, although the results presented here only used a tip angle equal to $90°$. However, this work cannot be applied on $180°$ refocusing pulses for spin echo sequences or inversion recovery sequences because small-tip-angle regime can lead to significant errors in excitation patterns. To overcome the limits in this regime, several approaches of large-tip-angle RF and gradient design have been proposed that apply or could apply to approximate the nonlinear Bloch equation [53, 54, 95]. The algorithm on large-tip-angle regime of MIMO MRI model needs to be researched on improving the inversion and refocusing at ultra-high fields.

3.6   Conclusions

Previous researchers have studied multi-coil transmission techniques, but there has been minimal research on multi-coil transmission and reception in MRI systems. Here, we considered an MRI system with a transceive coil array, where a number of coil elements are used for transmission and another number were used for reception. We termed this framework the MIMO MRI model. We proposed a novel RF pulse design method not only to minimize the difference between the desired magnetization pattern and the simulated pattern, but also to maximize the SNR in the MRI receive signal. Multiple simulations have been verified using the new RF pulse design method, and we further investigated the influence of acceleration factors, number of receive coil elements used, maximum excitation error tolerance, and different excitation patterns. We have demonstrated that our method improved the $gSNR$ by 18-130% as compared to the $gSNR$ obtained in the conventional pTx work ($AF_e = 2 - 6$), and simultaneously controlled the image difference between excitation error in a desired

scope (NRMSE = 0.12). The simulated transverse plane magnetization was uniform using the new RF design method, and there were minimal bias effects within the transverse magnetization. Research on large-tip angle simulations and different acceleration methods in parallel reception remain problems for future MIMO MRI investigation.

4. SAR MANAGEMENT

4.1 Introduction

Parallel Transmission (pTx) technique is majorly used in multi-coil MRI environment to reduce dielectric effects in obtained images [17–19]. The purpose of the pTx is to create a homogeneous RF field (i.e., B_1 field) or a uniform tip angle pattern in RF excitation by using multiple coils [20–26]. The major advantage of the pTx MRI system is to effectively mitigate the B_1 field inhomogeneity, but it also involves with other benefits, such as scan acceleration [11,54] and RF exposure control [68,69]. The pTx technique is becoming widespread and is shown to become effective in ultra-high field (≥ 7 T) [21–23, 33]. However, RF shimming is still considered to be the most commonly used technique to create the homogeneous B_1 field in high-field or UHF MRI [27–29].

RF exposure is one of the major issues in the pTx MRI system since more RF powers are involved in the multi-coil environment, and thus it is required to control the RF exposure and further control the internal temperature rise [35]. It is known that RF exposure will lead to the temperature rise, and higher temperature rise will likely cause irreversible tissue damage. RF exposure needs to be explicitly controlled, and thus a safety measurement is proposed to ensure the scan safety. SAR, which is defined as the dissipated power per unit mass, is used as the safety measurement to quantify the RF power deposition during the human scanning [36–38, 96]. Different guidelines on SAR limitations are proposed to ensure the MR safety, such as IEC standards [39], ICNIRP guidelines [40], and IEEE guidelines [41].

In the widely adopted IEC standards, there are two different SAR definitions, 10-g local SAR and global SAR. 10-g local SAR is defined as the averaged SAR of a 10 gram tissue volume, and the global SAR is defined as the averaged SAR of the entire

scanned body. Global SAR reflects the core temperature inside the human body, and increased global SAR value indicates the body heating inside MRI. Therefore, global SAR needs to be carefully monitored during the human scanning, using calorimetric method or pulse energy method [36]. Local SAR is difficult to measure in real-time, but it can be offline thresholded to avoid possible localized heating [37].

The pTx framework can be used to effectively control the RF exposure, i.e., control global/local SARs and/or temperature rise in the RF excitation. Prior researches have shown that global SAR [21, 97–99], local SAR [42, 68, 69, 100–105], and temperature rise [10] can be controlled within the pTx framework. One commonly used strategy was to apply regularization terms in the pTx RF pulse design procedure, to control the local and global SARs [18, 69]. These studies could further show the trade-off between different constraints (global/local SAR, excitation accuracy, etc.) during the RF excitation. There was an increasing trend to use a limited number of virtual observation points (VOPs) [10, 68, 69] to minimize the peak local SAR values or the temperature rise in each local region of the human body during the pTx process. This type of approach could control multiple constraints in the pTx RF pulse design, not only the peak local SARs, but also the peak/averaged RF power [100]. Apart from the offline SAR control method, online SAR monitoring is also another widely used technique in the multi-coil MRI. Some studies chose to monitor the RF pulses in each transmit channel [106, 107], and others used directional couplers [108–110] to monitor the transmitted and reflected RF pulses in real-time.

The goal of this chapter is to develop a pTx RF pulse design method that is modeled to constrain the peak local SARs using a compressed set of SAR matrices. We proposed an alternative SAR compression method, which used $k-$means clustering algorithm, for an upper-bounded estimation of peak local SARs in the RF excitation. The proposed SAR model compression method was shown to be superior because our generated SAR compression model was close to the actual SAR model and guaranteed to be safe by all means. Traditional VOP compression method, e.g., Lee's method [69], constructed the SAR compression model with a pre-defined overestimation bound, but

it did not provide control over the compression ratio. On the contrary, our $k-$means SAR compression model allows for the control over the number of generated clusters. When compared with Lee's method, the $k-$means SAR compression model showed a better performance by generating a narrower and more accurate overestimation bound. A preliminary version of this study has been reported in [111].

4.2 Theory

4.2.1 Basic pTx Concept and SAR Modeling

We consider an MRI system with an $N_t \times N_t$ transceive array and follow the general framework of *Transmit SENSE* for the description of multi-dimensional pTx pulse design [11, 20]. The RF pulses inducing the desired magnetization (denoted by $\boldsymbol{m}_{ref}$) can be generated by solving the magnetization matching problem

$$\boldsymbol{\mathcal{B}}_{ref} = \underset{\boldsymbol{\mathcal{B}}}{\mathrm{argmin}} \left(\left\| \tilde{\boldsymbol{S}}_A \boldsymbol{\mathcal{B}} - \boldsymbol{m}_{ref} \right\|_2^2 \right), \tag{4.1}$$

where $\tilde{\boldsymbol{S}}_A$ is the system matrix containing the transmit sensitivity maps and $\boldsymbol{\mathcal{B}}$ is the transmitted RF waveforms across N_t channels. Here, $\boldsymbol{\mathcal{B}} = [\boldsymbol{\mathcal{B}}(t_1), ... \boldsymbol{\mathcal{B}}(t_\mathcal{K})]^\top$ and $\boldsymbol{\mathcal{B}}(t_k) = [B_1(t_k), ..., B_{N_t}(t_k)]^\top$, where t_k refers to the kth time point in each RF pulse ($\mathcal{K}$ time points in total).

The local SAR of each location, $\boldsymbol{r}$, is obtained by averaging the SAR values found in its neighboring 10-g volume, which can be expressed as

$$SAR(\boldsymbol{r}) = \frac{1}{\mathcal{K}} \sum_{k=1}^{\mathcal{K}} \boldsymbol{\mathcal{B}}^H(t_k) \underbrace{\left[\frac{1}{\mathcal{V}} \int_{\mathcal{V}} \frac{\sigma(\boldsymbol{r})}{2\rho(\boldsymbol{r})} \boldsymbol{S}_E^H(\boldsymbol{r}) \boldsymbol{S}_E(\boldsymbol{r}) d\mathcal{V} \right]}_{\boldsymbol{Q}(\boldsymbol{r})} \boldsymbol{\mathcal{B}}(t_k), \tag{4.2}$$

where H denotes the conjugate transpose operator, σ refers to the electrical conductivity, ρ refers to the mass density of the tissue, $\mathcal{V}$ is the local volume containing 10-g of tissue, and $\boldsymbol{S}_E(\boldsymbol{r})$ denotes the electric field sensitivity at location $\boldsymbol{r}$.

Tissue density, conductivity, and electric field sensitivity maps are combined into a *Q-matrix* [67] in (4.2). This *Q-matrix* format is majorly used for offline SAR pre-

diction, but it can be extended for use in the online SAR monitoring. This expression can be written into a quadratic form

$$SAR(\boldsymbol{r}) = \frac{1}{\mathcal{K}} \boldsymbol{\mathcal{B}}^H \boldsymbol{Q_d}(\boldsymbol{r}) \boldsymbol{\mathcal{B}}, \tag{4.3}$$

where $\boldsymbol{Q_d}(\boldsymbol{r}) = diag(\boldsymbol{Q}(\boldsymbol{r}), ..., \boldsymbol{Q}(\boldsymbol{r})) \in \mathbb{C}^{(N_t \cdot \mathcal{K}) \times (N_t \cdot \mathcal{K})}$.

4.2.2 SNR Control in the MRI Receive Voltage

The SNR is defined as the ratio of the absolute value of the received NMR signal voltage (i.e., electromotive force) to the standard deviation of the noise voltage [71]. Mathematically, this is

$$|SNR(\boldsymbol{r})| = \frac{\left| M(\boldsymbol{r}) \sum_{q=1}^{N_t} \eta_q^*(\boldsymbol{r}) B_{1,q}^-(\boldsymbol{\mathcal{B}}, \boldsymbol{r}) \right|}{\sqrt{\sum_{j=1}^{N_t} \sum_{k=1}^{N_t} \eta_j^*(\boldsymbol{r}) \eta_k(\boldsymbol{r}) \Sigma_{jk}}}, \tag{4.4}$$

where $M(\boldsymbol{r})$ is the Bloch simulated transverse magnetization as defined in Eq.(1)-(2) in [71], $B_{1,q}^-(\boldsymbol{\mathcal{B}}, \boldsymbol{r})$ is defined as the B_1^- field from q^{th} receive coil at location $\boldsymbol{r}$, η_q is the receiver weight related to the qth receive coil, and Σ_{jk} represents the correlated ($j \neq k$) and uncorrelated ($j = k$) thermal noise between coils j and k. We assume Σ_{jk} is the noise covariance matrix induced by the additive Gaussian noises.

We further define $\boldsymbol{B}_1^-(\boldsymbol{\mathcal{B}}, \boldsymbol{r})$ as a column vector that contains the B_1^- fields from all receive coils at location $\boldsymbol{r}$ as

$$\boldsymbol{B}_1^-(\boldsymbol{\mathcal{B}}, \boldsymbol{r}) = \boldsymbol{F} \boldsymbol{D}(\boldsymbol{r}) \boldsymbol{\mathcal{B}}, \tag{4.5}$$

where $\boldsymbol{F} \in \mathbb{R}^{N_t \times N_t}$ is the receive coil selection matrix and $\boldsymbol{D}(\boldsymbol{r}) \in \mathbb{C}^{N_t \times (N_t \cdot \mathcal{K})}$ is defined as $\boldsymbol{D}(\boldsymbol{r}) = diag(\boldsymbol{A}_1(\boldsymbol{r}), ..., \boldsymbol{A}_\mathcal{K}(\boldsymbol{r}))$. The matrix $\boldsymbol{A}_k(\boldsymbol{r}) \in \mathbb{C}^{1 \times N_t}$ is defined as

$$\boldsymbol{A}_k(\boldsymbol{r}, q) = S_q'(\boldsymbol{r}) e^{i\gamma \Delta B_0(\boldsymbol{r})(T - t_k)}, \tag{4.6}$$

where T is the pulse length and $S_q'(\boldsymbol{r})$ is the receive sensitivity map of qth receive coil.

It is know that SNR is proportional to the B_1^- fields according to [88], and thus SNR should have a positive correlation with the complex-valued RF waveform. This

is because the amplitude of the transmitted RF pulses contributes to the strength of the B_1^- field. We define the B_1^- field to contain the receive coil pattern, $S_q'(\boldsymbol{r})$, and the RF waveform, $\boldsymbol{\mathcal{B}}$ [71]. The receive coil patterns contain the body/coil geometry and the electromagnetic parameters (e.g., permittivity and conductivity), which are similar but not the same as the transmit coil patterns, especially in high field and UHF MRI [88–90].

The discrete expression of the square of the SNR can be converted into a ratio of quadratic forms, and the optimal receiver weight η can be found by maximizing

$$|\boldsymbol{SNR}(\boldsymbol{\mathcal{B}},\boldsymbol{r})|^2 = \frac{|\boldsymbol{m}(\boldsymbol{r})|^2 \, \boldsymbol{\eta}(\boldsymbol{r})^{\mathrm{H}} \boldsymbol{B}_1^-(\boldsymbol{\mathcal{B}},\boldsymbol{r}) \boldsymbol{B}_1^-(\boldsymbol{\mathcal{B}},\boldsymbol{r})^{\mathrm{H}} \boldsymbol{\eta}(\boldsymbol{r})}{\boldsymbol{\eta}(\boldsymbol{r})^{\mathrm{H}} \boldsymbol{\Sigma} \boldsymbol{\eta}(\boldsymbol{r})}. \tag{4.7}$$

where $\boldsymbol{m}(\boldsymbol{r})$ is the discrete expression of $M(\boldsymbol{r})$. It should be close to the desired magnetization, $\boldsymbol{m}_{ref}(\boldsymbol{r})$, because the excitation error is minimized. We then defined global SNR as the averaged value of the SNR in the given ROI

$$gSNR(\boldsymbol{\mathcal{B}}) = \frac{1}{\mathcal{R}} \sum_{\boldsymbol{r}=1}^{\mathcal{R}} \boldsymbol{SNR}(\boldsymbol{\mathcal{B}},\boldsymbol{r}). \tag{4.8}$$

The expression in (4.7) is a ratio of two quadratic forms. We can determine the optimal SNR by finding the point when the gradient of $|\boldsymbol{SNR}(\boldsymbol{\mathcal{B}},\boldsymbol{r})|^2$ with respect to $\boldsymbol{\eta}(\boldsymbol{r})$ is zero (N_t equations). Solving for $\boldsymbol{\eta}$ yields the optimum receiver weights, the relevant equations can be found in [71].

4.3 Methods

4.3.1 RF Pulse Design with Multiple Constraints

Transmitted RF pulses in high field MRI need to be controlled to satisfy the global and local SAR constraints. Reducing global SAR values in single-coil or multi-coil MRI systems is straightforward, we could prolong the RF pulse duration, or increase TR, or reduce the resulting tip angle. However, minimizing local SAR values are often considered to be difficult. To guarantee that the designed RF pulses will satisfy local SAR constraints, we need to estimate the worst-case local SAR value of every

voxel in the image domain. This estimation is usually exhaustive and complex, thus we need to reduce the computational burden.

Here, the SAR compression model is introduced to monitor a few representative SAR matrices instead of all SAR matrices. The target becomes the determination of several positive semi-definite, Hermitian SAR matrices $\hat{\boldsymbol{Q}}_{\boldsymbol{d},1}, ..., \hat{\boldsymbol{Q}}_{\boldsymbol{d},\mathcal{N}}$, such that for any arbitrary $\boldsymbol{r} \in \{\boldsymbol{r}_1, ..., \boldsymbol{r}_\mathcal{R}\}$ there exists some $v \in \{1, ..., \mathcal{N}\}$ satisfying

$$\boldsymbol{\mathcal{B}}^H \boldsymbol{Q}_{\boldsymbol{d}}(\boldsymbol{r})\boldsymbol{\mathcal{B}} \leq \boldsymbol{\mathcal{B}}^H \hat{\boldsymbol{Q}}_{\boldsymbol{d},v}\boldsymbol{\mathcal{B}}, \ \forall \boldsymbol{\mathcal{B}} \in \mathbb{C}^{(N_t \cdot \mathcal{K}) \times 1}, \tag{4.9}$$

for any multi-channel RF pulse. The resulting SAR compression model should be accurate and guarantee to be safe by all means. To be more specific, the local SAR values obtained from the SAR compression model should not be smaller than the actual SAR values (i.e., no underestimation), and should not be vastly greater than the actual SAR values (i.e., controlled overestimation).

During the pTx RF pulse design procedure, instead of checking local SAR value of every voxel, we now only need to check the local SAR values of those representative SAR matrices. Here, different pTx RF pulse design methods are proposed, using the designed SAR compression model.

Power Back-off Approach

This approach only applies the transmit power constraint in the objective function of pTx RF pulse design problem. It is designed to back off the transmitted RF power if the resulting peak local SAR values are greater than the threshold values [112].

$$\boldsymbol{\mathcal{B}}_{opt}^{(1)} = \beta\boldsymbol{\mathcal{B}}_{ref}, \ \beta \in [0, 1] \tag{4.10}$$

$$\text{s.t.} \ \frac{1}{\mathcal{K}}\beta^2 \boldsymbol{\mathcal{B}}_{ref}^H \hat{\boldsymbol{Q}}_{\boldsymbol{d},v}\boldsymbol{\mathcal{B}}_{ref} \leq SAR_{max}, \forall v \in \{0, ..., \mathcal{N}\}. \tag{4.11}$$

Note that $\hat{\boldsymbol{Q}}_{\boldsymbol{d},0} = \mathcal{K}\boldsymbol{I}$, which represents the RF power constraint. Consequently, this approach will sacrifice the tip angle accuracy in the RF excitation.

Power and Local SAR Constraints

RF power and peak local SAR constraints are applied in this pTx RF pulse optimization problem

$$\boldsymbol{\mathcal{B}}_{opt}^{(2)} = \underset{\boldsymbol{\mathcal{B}}}{\operatorname{argmin}} \left|\left| \tilde{\boldsymbol{S}}_A \boldsymbol{\mathcal{B}} - \boldsymbol{m}_{ref} \right|\right|_2^2, \tag{4.12}$$

$$\text{s.t.} \ \frac{1}{\mathcal{K}} \boldsymbol{\mathcal{B}}^H \hat{\boldsymbol{Q}}_{\boldsymbol{d},v} \boldsymbol{\mathcal{B}} \leq SAR_{max}, \forall v \in \{0, ..., \mathcal{N}\}. \tag{4.13}$$

The expressions stated in (4.12) and (4.13) can be combined together if we set a Lagrange multiplier α_v for each peak local SAR constraint.

Power, SAR and SNR Constraints

We update equations in the previous method and include the $gSNR$ term described in (4.8). The resulting RF pulses, $\boldsymbol{\mathcal{B}}_{opt}^{(3)}$, can induce less dissipated RF power while maintaining the tip angle pattern and the SNR of the MRI receive voltage. The detailed constrained optimization is given as

$$\boldsymbol{\mathcal{B}}_{opt}^{(3)} = \underset{\boldsymbol{\mathcal{B}}}{\operatorname{argmin}} ||\boldsymbol{S}_A \boldsymbol{\mathcal{B}} - \boldsymbol{m}_{ref}||_2^2, \tag{4.14}$$

$$\text{s.t.} \ \frac{1}{\mathcal{K}} \boldsymbol{\mathcal{B}}^H \hat{\boldsymbol{Q}}_{\boldsymbol{d},v} \boldsymbol{\mathcal{B}} \leq SAR_{max}, \forall v, \tag{4.15}$$

$$\frac{||\boldsymbol{S}_A \boldsymbol{\mathcal{B}} - \boldsymbol{m}_{ref}||_2^2}{||\boldsymbol{m}_{ref}||_2^2} \leq diff_{max}, \tag{4.16}$$

$$gSNR(\boldsymbol{\mathcal{B}}) \geq SNR_{min}. \tag{4.17}$$

In this approach, we can combine (4.14) and (4.15) to set up the Lagrangian, then we can search for optimal Lagrange multipliers α_v to obtain pTx RF pulses that satisfy (4.16) and (4.17).

4.3.2   Model Compression in Local SAR Supervision

The SAR compression model requires to group voxels with similar SAR behaviours into a smaller set of clusters, and it selects one representative SAR matrix to represent each cluster. To ensure the local SAR safety, the representative SAR matrix, $\hat{\boldsymbol{Q}}_{d,v}$, of cluster v should be designed to have no underestimation but controlled overestimation, compared to the actual SAR model.

For no underestimation, the representative SAR matrix, $\hat{\boldsymbol{Q}}_{d,v}$, for each possible v should always satisfy (4.9). For controlled overestimation, we firstly need to find a core matrix (i.e., $\hat{\boldsymbol{Q}}'_{d,v}$) in each cluster. Each cluster's core matrix should represent highest local SAR value in this cluster. The SAR matrices around this core matrix are clustered together if the following inequality holds

$$\boldsymbol{\mathcal{B}}^H \hat{\boldsymbol{Q}}'_{d,v} \boldsymbol{\mathcal{B}} + \Omega \geq \boldsymbol{\mathcal{B}}^H \boldsymbol{Q}_d(\boldsymbol{r}) \boldsymbol{\mathcal{B}}, \forall \boldsymbol{r} \in C_v, \tag{4.18}$$

where C_v contains all voxels in cluster v and Ω is denoted as the overestimation bound. Ω needs to be as small as possible to have the minimized overestimation.

VOP Compression Algorithm

Lee et al. [69] demonstrated a greedy algorithm to construct VOPs in the pTx framework. It used a pre-defined overestimation bound as the input, and outputted the desired VOP segmentation which satisfied this upper bound. This algorithm had fast computation, reliable performance, but it had a loose overestimation bound and it did not have the control of the number of VOPs (i.e., the compression ratio).

The algorithm that constructed the VOPs can be described in Fig. 4.1. This algorithm firstly sorted all SAR matrices based on their dominant eigenvalues (i.e., $\lambda_0(\boldsymbol{r})$) in a descending order, and then looped over all SAR matrices to create the least number of VOPs that strictly satisfied the overestimation bound. Note that the overestimation bound, Ω, was chosen to determine the maximum peak local SAR difference between the compressed SAR model and the actual SAR model.
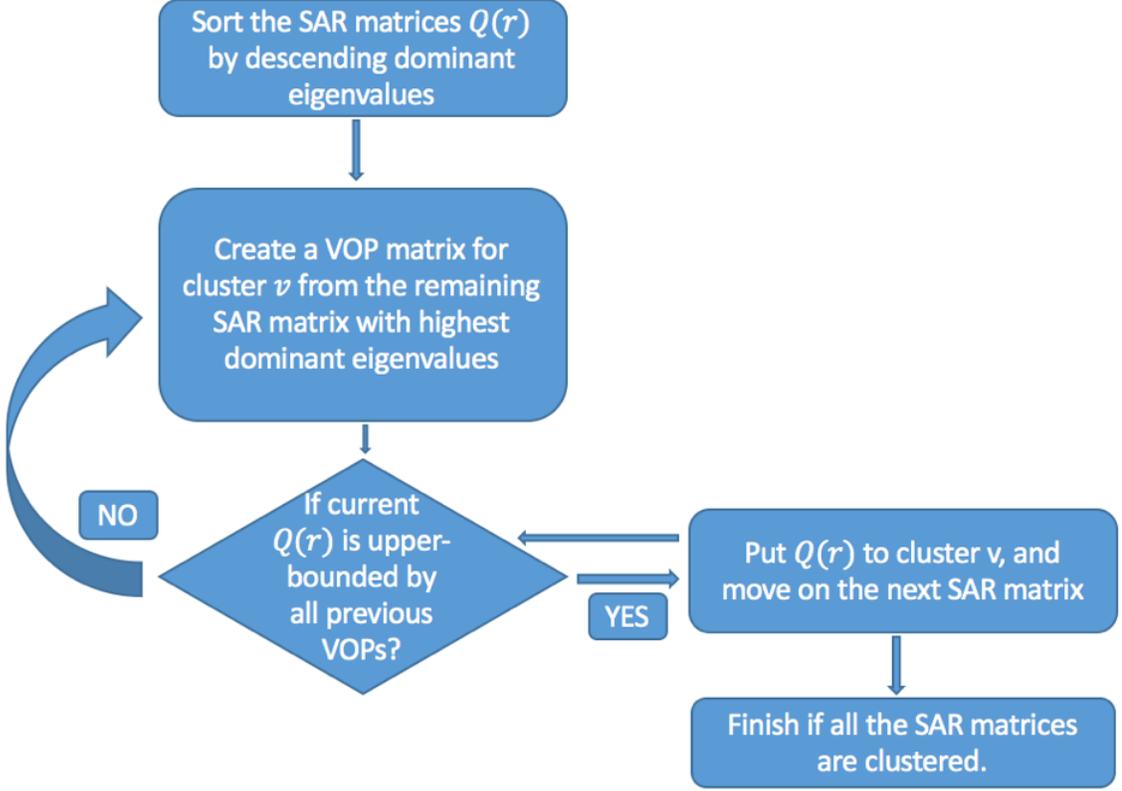
Fig. 4.1. The greedy algorithm that constructed VOPs in the pTx framework. The overestimation bound is the input of the model, and it determines the the correctness of the SAR compression model. The SAR compression is closer to the actual SAR model if the overestimation bound is smaller.

According to (4.9) and (4.18), the representative SAR matrix, $\hat{\boldsymbol{Q}}_{\boldsymbol{d},v}$, in cluster v has to satisfy the following inequality

$$\boldsymbol{\mathcal{B}}^H \boldsymbol{Q}_{\boldsymbol{d}}(\boldsymbol{r}) \boldsymbol{\mathcal{B}} \leq \boldsymbol{\mathcal{B}}^H \hat{\boldsymbol{Q}}_{\boldsymbol{d},v} \boldsymbol{\mathcal{B}} \leq \boldsymbol{\mathcal{B}}^H \hat{\boldsymbol{Q}}'_{\boldsymbol{d},v} \boldsymbol{\mathcal{B}} + \Omega, \forall \boldsymbol{\mathcal{B}}, \forall \boldsymbol{r} \in C_v. \tag{4.19}$$

The representative SAR matrix, $\hat{\boldsymbol{Q}}_{\boldsymbol{d},v}$, in each cluster v is defined as

$$\hat{\boldsymbol{Q}}_{\boldsymbol{d},v} = \hat{\boldsymbol{Q}}'_{\boldsymbol{d},v} + \boldsymbol{Z}_v, \tag{4.20}$$

where the $\boldsymbol{Z}_v$ matrix is used to determine the overestimation bound, Ω. The left side of (4.19) holds when $\hat{\boldsymbol{Q}}'_{d,v} + \boldsymbol{Z}_v - \boldsymbol{Q}_d(\boldsymbol{r})$ is positive semi-definite, indicating there is no underestimation. The right side of (4.19) automatically holds because the following inequality holds

$$\boldsymbol{\mathcal{B}}^H \hat{\boldsymbol{Q}}_{d,v} \boldsymbol{\mathcal{B}} - \boldsymbol{\mathcal{B}}^H \hat{\boldsymbol{Q}}'_{d,v} \boldsymbol{\mathcal{B}} \leq ||\boldsymbol{Z}_v||_2 \boldsymbol{\mathcal{B}}^H \boldsymbol{\mathcal{B}} \approx \Omega, \tag{4.21}$$

where the spectral norm of $\boldsymbol{Z}_v$ should be minimized to ensure the controlled overestimation. The spectral decomposition method described in [68] can be used to determine the matrix $\boldsymbol{Z}_v$. The overestimation rate $R = \max_v(||\boldsymbol{Z}_v||_2 / \max_{\boldsymbol{r}}(\lambda_0(\boldsymbol{r})))$, where the denominator is the maximum eigenvalue among the whole body SAR matrices.

$k-$Means Compression Algorithm

We here proposed an alternative method using the $k-$means clustering to implement the model compression with control over the number of clusters. It aimed to partition multiple SAR matrices into exactly k clusters in which each SAR matrix was put into the cluster with the nearest distance [113, 114]. In this application of $k-$means clustering, the distance between the centroid SAR matrix (i.e., $\hat{\boldsymbol{Q}}'_{d,v}$) and any arbitrary SAR matrix (i.e., $\boldsymbol{Q}_d(\boldsymbol{r})$) is quantified as the overestimation bound (i.e., $\boldsymbol{Z}'_v(\boldsymbol{r})$) between them

$$\boldsymbol{\mathcal{B}}^H \boldsymbol{Q}_d(\boldsymbol{r}) \boldsymbol{\mathcal{B}} - \boldsymbol{\mathcal{B}}^H \hat{\boldsymbol{Q}}'_{d,v} \boldsymbol{\mathcal{B}} \leq ||\boldsymbol{Z}'_v(\boldsymbol{r})||_2 \boldsymbol{\mathcal{B}}^H \boldsymbol{\mathcal{B}}, \forall \boldsymbol{\mathcal{B}}. \tag{4.22}$$

Generally, voxels in the image domain were clustered around the centroids which were obtained by minimizing the following objective function

$$\psi(l) = \sum_{v=1}^{k} (\max_{\boldsymbol{r}}(||\boldsymbol{Z}'_v(\boldsymbol{r})||_2)), \tag{4.23}$$

where l is the current iteration index. As the algorithm proceeds, $\psi(l) \leq \psi(l-1)$ holds for each iteration. The algorithm will end when $\psi(l) = \psi(l-1)$ (i.e., converged) or $l \geq L_{max}$ (L_{max} is the maximum number of iterations).

The detailed procedure followed a way to assign N local SAR matrices to exactly one of the k clusters such that the maximum distance between the selected SAR matrix and the corresponding cluster centroid was minimized. To assign N local SAR matrices into $\{C_1, ..., C_k\}$ clusters, the algorithm worked as follows:

1. Arbitrarily chose k initial centroids $\mathcal{C} = \{c_1, ..., c_k\}$. Ideally, we could use the $k-$means++ algorithm [115] to heuristically find the initial centroids, achieving faster convergence.

2. For each $v \in \{1, ..., k\}$, set the cluster C_v to be the set of voxels that were closer to c_v (i.e. smaller $||\mathbf{Z}'_v(\mathbf{r})||_2$) than they were to other centroids.

3. For each $v \in \{1, ..., k\}$, set the centroid c_v to be the SAR matrix in that cluster containing the minimum value of $\max_{\mathbf{r}} ||\mathbf{Z}'_v(\mathbf{r})||_2, \forall \mathbf{r} \in C_v$.

4. Repeated step 2 and 3 until cluster assignments stayed unchanged, or the maximum number of iterations (i.e., L_{max}) was reached.

To determine the degree of overestimation, we used the overestimation rate R to measure the SAR overestimation in the RF pulse design procedure. The overestimation rate for cluster v is

$$R'_v = \left(\max_{\mathbf{r}}(||\mathbf{Z}'_v(\mathbf{r})||_2) / \max_{\mathbf{r}}(\lambda_0(\mathbf{r})) \right), \tag{4.24}$$

since the inequality in (4.22) hold. The overestimation rate for the whole body model is $R = \max_v(R'_v)$.

4.4   Results

4.4.1   FDTD Simulation

The proposed pTx RF pulse design methods were conducted via a numerical head model and a 16-channel 7.0 T transceiver coil array [116]. The Duke head model was firstly loaded in the imaging domain, and FDTD simulation (Remcom XFdtd, PA, USA) was later implemented to extract the corresponding B_1 and E field maps from the transceiver coils. The FDTD mesh originally had $217 \times 251 \times 174$ voxels and each voxel is 1 mm^3. However, the $x - y$ plane of the imaging domain was later downsampled to have 128×128 voxels, to allow for fast computation. The surface of the Duke model and the 16-channel transceiver coil array are demonstrated in Fig. 4.2.
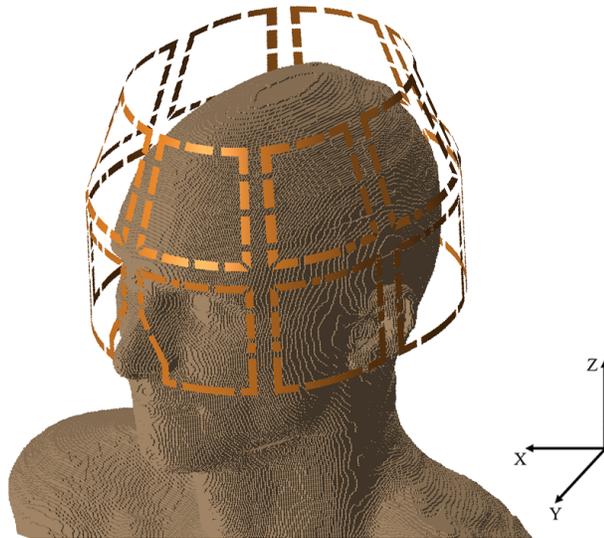
Fig. 4.2. The surface of the 3D numerical Duke model (from Virtual Family) placed in a 16-channel transceiver coil array. The eight coils in the lower ring of the loop array were used for the simulation.

All the simulations were conducted in different $x - y$ planes of the Duke model. Fig. 4.3 shows the original eight-channel B_1^+ field maps and the electric field after the FDTD simulation, since only the eight-channel transceiver coils in the lower ring

of the loop array were used for parallel RF transmission and reception. Transmit and receive sensitivity maps used in the pTx RF pulse design were obtained after the normalization. We also note that the local B_0 field inhomogeneities were not included in the FDTD simulation.
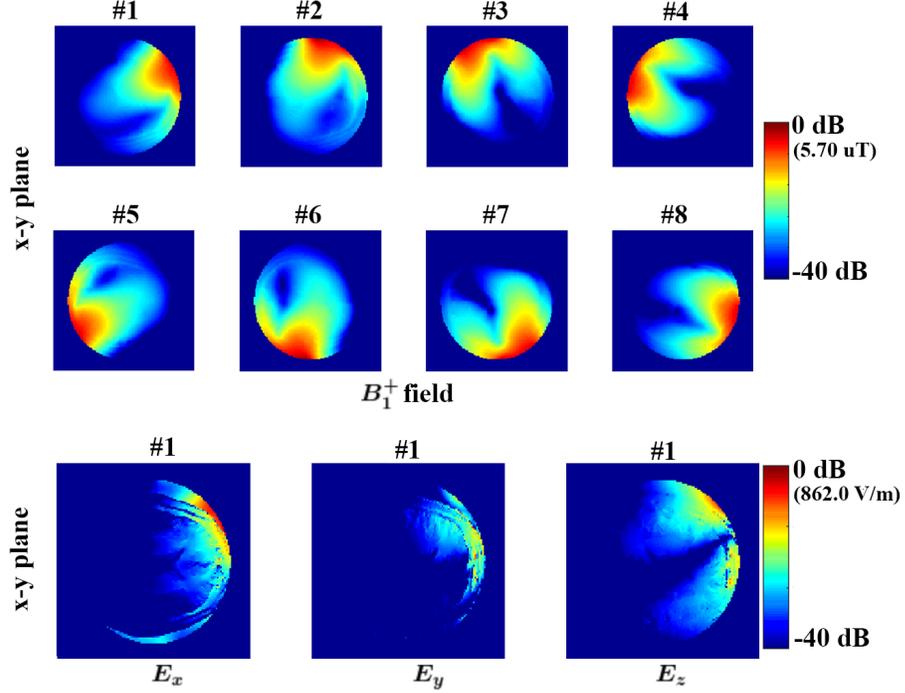
Fig. 4.3. Simulated $x - y$ plane ($z = 79\ mm$) of the B_1^+ fields in all 8 transmit elements with loaded Duke phantom, and simulated E_x, E_y, E_z fields in transmit channel #1. An ROI mask was applied to choose the interested region (see Fig. 4.4 (a)).

The proposed pTx RF pulse design methods should be conducted under the 7 T scanner parameters to ensure that human can tolerate the designed RF pulses: $100\ mT/m$ maximum gradient amplitude, $200\ T/m/s$ maximum slew rate, and 4 μs sampling period. The designed $k-$space trajectory was a single-shot, spiral-out excitation for all designs [94] and was under-sampled by a factor of 2. Therefore, the resulting RF pulse length was 4.7 ms, and this spiral $k-$space trajectory had a spatial resolution of 1.06 cm in this acceleration scheme.

The desired excitation pattern in this simulation, shown in Fig. 4.4(b), was a smoothed elliptic cylinder. The excitation pattern was convoluted with a Gaussian kernel to reduce the Gibbs ringing artifact. Fig. 4.4(c) shows the transverse plane magnetization after exciting the plane with a $(\pi/2)$ RF pulse. The applied ROI mask was a elliptical region shown in Figures 4.4(a) and (b).
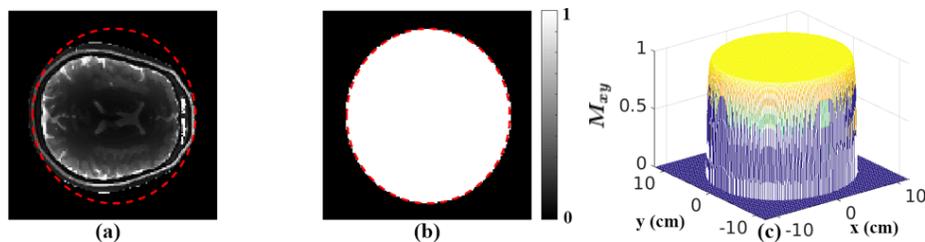
Fig. 4.4. Desired excitation pattern and transverse plane magnetization used in the simulations. (a) the red dashed ROI is a smoothed ellipse $(x_{radius} = 10\,cm,\ y_{radius} = 9.5\,cm)$ in the $z = 79\,mm$ plane of the Duke model; (b) the desired excitation pattern within the ROI; (c) the tip angle pattern of the desired excitation, which is a smoothed elliptic cylinder, whose peak was scaled to $\pi/2$, corresponding to 90° tip angle.

4.4.2   Local SAR Supervision

We performed two SAR compression models in 174 slices of the Duke head model, and we observed that the $k-$means compression model always showed a better performance than the greedy algorithm, although it takes a longer time to reach the convergence.

Greedy algorithm for model compression

We performed the greedy algorithm from Lee et al. [69] (i.e., Lee's method) for VOP compression with several over-estimating bounds, from 5% to 50% of the maximum eigenvalue in the whole imaging slice. For example, Lee's method captured

Table 4.1.
Comprehensive overview of the two SAR compression models in different slices of Duke head model in 7 T MRI.

Method	Body Region (# slice)	# of Voxels	Overestimation Rate (%)	# of VOPs	Compression Ratio
VOP Compression	Head (z=10mm)	11,891	10.00%	**26**	457.3
K-means Compression	Head (z=10mm)	11,891	**6.70%**	**26**	457.3
VOP Compression	Head (z=10mm)	11,891	**6.70%**	40	297.3
VOP Compression	Head (z=30mm)	12,019	5.00%	**60**	200.3
K-means Compression	Head (z=30mm)	12,019	**4.40%**	**60**	200.3
VOP Compression	Head (z=30mm)	12,019	**4.40%**	67	179.4
VOP Compression	Head (z=79mm)	11,423	3.00%	**121**	94.4
K-means Compression	Head (z=79mm)	11,423	**2.90%**	**121**	94.4
VOP Compression	Head (z=79mm)	11,423	**2.90%**	127	89.9
VOP Compression	Head (z=90mm)	11,293	5.00%	**73**	154.7
K-means Compression	Head (z=90mm)	11,293	**4.50%**	**73**	154.7
VOP Compression	Head (z=90mm)	11,293	**4.50%**	84	134.4

the local SAR estimations by compressing the voxels in the desired slice ($\sim 10^5$) to 36 clusters at the over-estimation bound $\Omega = 0.0975 \cdot \max_{\boldsymbol{r}}(\lambda_0(\boldsymbol{r}))$ in one chosen slice.

Fig. 4.5(a) and (b) shows the color-coded clustering result obtained in the selected slices ($z = 90\ mm$ and $z = 79\ mm$) of the Duke model in the 7 T MRI environment, and spectral norms of all VOPs in the clustering result. Each VOP's spectral norm equals to the worst-case local SAR value in that VOP for a unit excitation pulse, i.e., $\boldsymbol{\mathcal{B}}^H \boldsymbol{\mathcal{B}} = 1$.

$k-$Means algorithm for model compression

In the $k-$means clustering based SAR model compression, a local voxel is considered to be in a particular cluster if it is closer to that cluster's centroid than any other centroid. The resulting clusters from the $k-$means method are shown in Fig. 4.5 (c) and (d). This plot illustrates that 73 clusters were constructed in the $z = 90\ mm$ plane of the Duke model, with 4.5% overestimation rate; 30 clusters were constructed in the $z = 79\ mm$ plane of the Duke model, with 9.75% overestimation rate. Clusters

with high spectral norms (i.e. high peak local SAR values) are often near to the physical location of those transmit coils. These clusters can be easily identified and classified as significant clusters in the compressed SAR model.
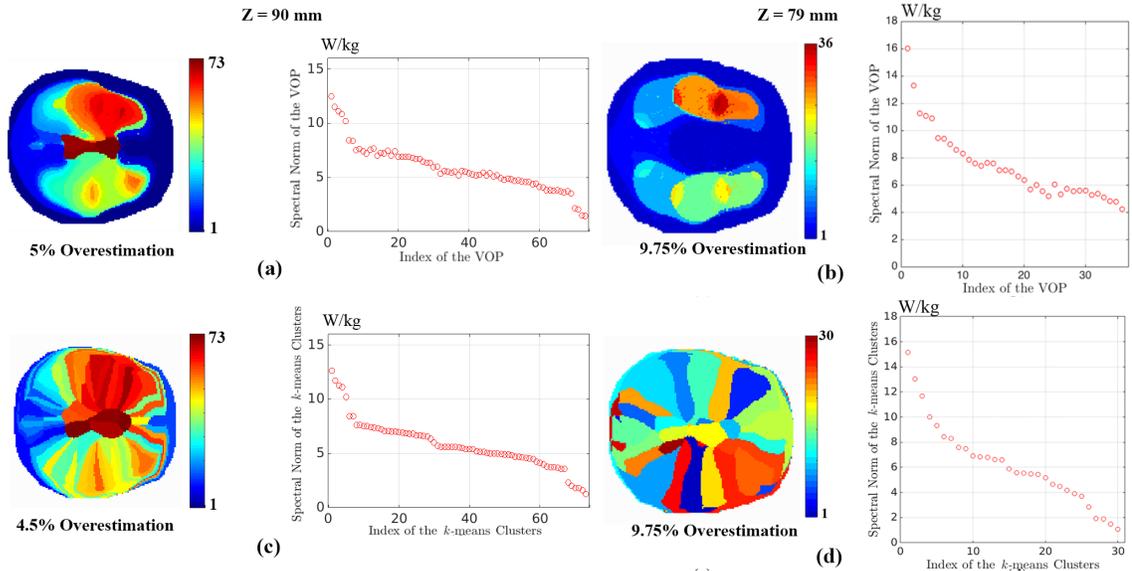
Fig. 4.5. (a) Total of 73 color-coded VOPs in the transverse plane at $z = 90\ mm$ of the Duke model, with 5% maximum overestimation rate; the spectral norm of each cluster refers to the peak local SAR value found in that cluster if use unit excitation pulses. (b) Total of 36 VOPs in the transverse plane at $z = 79\ mm$ of the Duke model, with 9.75% overestimation. (c) Total of 73 $k-$means clusters at $z = 90\ mm$ of the Duke model, with 4.5% overestimation. (d) Total of 30 $k-$means clusters at $z = 79\ mm$ of the Duke model, with 9.75% overestimation.

Fig. 4.6 compares the number of clusters and the overestimation rate in the $z = 79\ mm$ slice using Lee's method and the $k-$means clustering method. The two lines illustrate that the $k-$means method always generated fewer clusters than Lee's method when they have the same overestimation bound. The line plot demonstrates that the maximum overestimation rate of all clusters decreased as we increased the total number of clusters, and it would converge to $R = 0$ when k equals to the maximum number of voxels in the ROI. The $k-$means clustering algorithm chose random

initial centroids in the first step, and the total computation time could dramatically decrease if initial centroids are well-defined. The actual overestimation rate R in the $k-$means method became less than 5% when $k \geq 60$. Fig. 4.7 (a) and (b) further show the box-whisker plot of overestimation rates across all VOPs and $k-$means clusters in each of the clustering results. Fig. 4.8 demonstrates the difference between the two compressed SAR models and the actual SAR model in the $z = 79 \ mm$ plane of the Duke head.
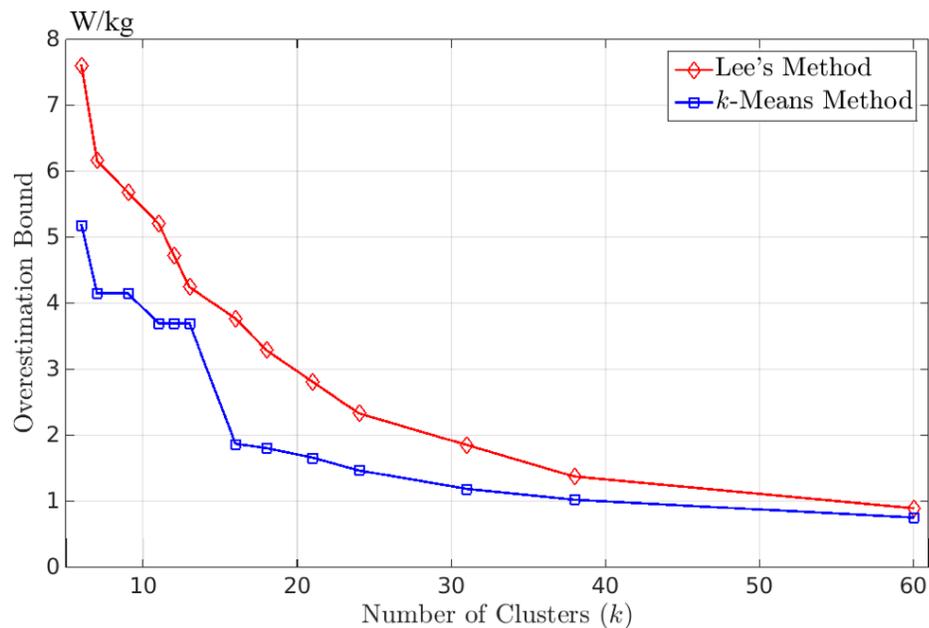
Fig. 4.6. Line plot of the maximum overestimation rate across all clusters along with the total number of clusters (k) in two SAR compression methods, Lee's method [69] and the $k-$means clustering method. The overestimation bound (W/kg) is the maximum peak local SAR difference between the compressed SAR model and the actual SAR model when unit excitation is applied. The actual overestimation rate R in the $k-$means method became less than 5% (overestimation bound ≤ 0.8) when $k \geq 60$.
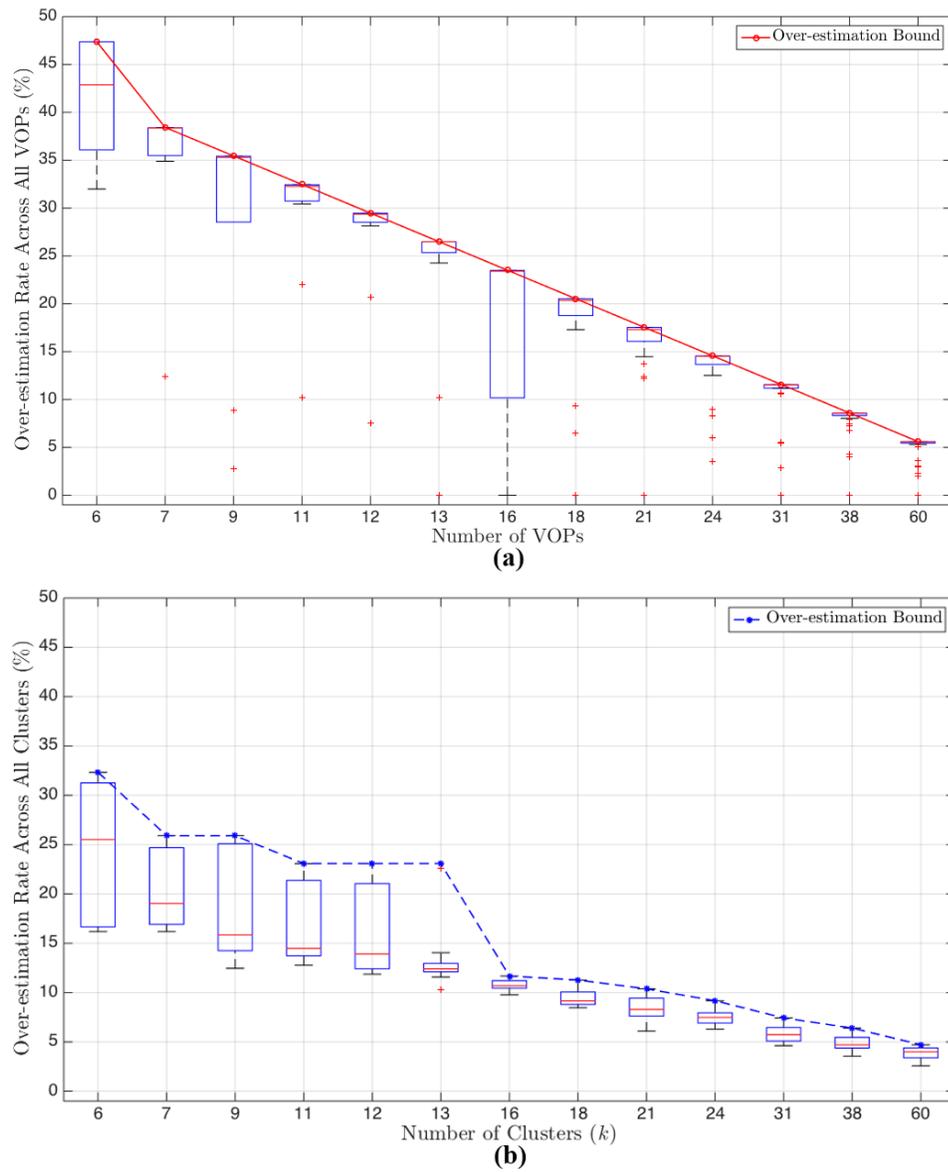
Fig. 4.7. Box-whisker plot of the over-estimation rates across all VOPs (a) and $k-$means clusters (b) in each segmentation assignment. The red solid line refers to the overestimation bound applied in each VOP assignment. The blue dash line refers to the overestimation bound found in each $k-$means clustering result.

4.4.3   Multi-Channel RF Pulse Design

In this simulation, we first investigated the basic pTx technology without any RF power, SNR, or SAR constraint in the multi-coil 7 T MRI system. The desired and
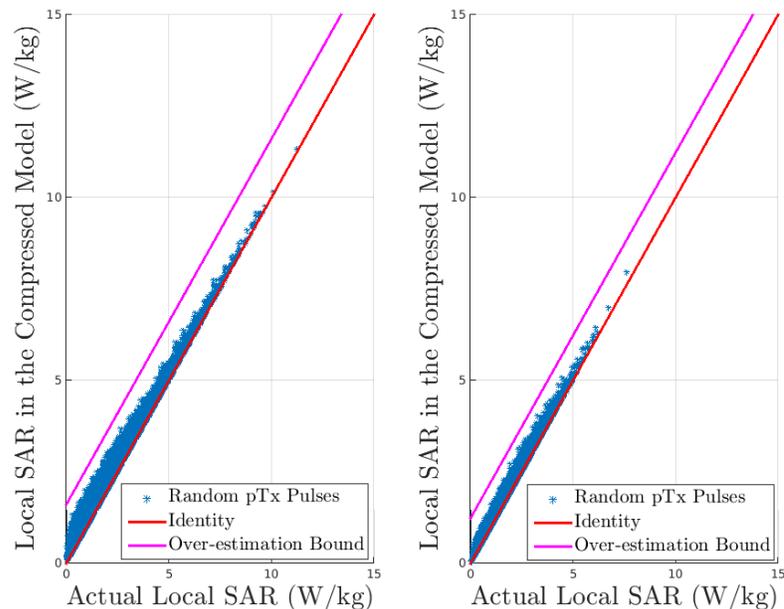
Fig. 4.8. Analysis of 30,000 random unit excitation pulses for local SAR prediction in the actual SAR model and the compressed SAR model in the transverse plane at $z = 79\ mm$ of the Duke model. (a) The compressed SAR model (34 VOPs) created by Lee's method, with 10% overestimation. (b) The compressed SAR model (34 clusters) created by the $k-$means model, with 7.7% overestimation.

simulated excitation patterns are shown in Fig. 4.9. This result provided the fundamental basis of the upcoming simulations. We compared the proposed constrained pTx methods with this *unconstrained approach* to quantify the percentage of the SAR reduction in each SAR-efficient pTx method.

The *unconstrained approach* was supposed to achieve the exact desired pattern during the parallel RF excitation, and Fig. 4.9 shows a relatively perfect excitation, with a NRMSE value of 0.003. Since there exists B_0 field inhomogeneity, some hypo-intense areas were observed near the physical coils during the parallel excitation.

Power Back-off Approach

The *power back-off approach* is commonly used in single-channel communication systems [112] to constrain the SAR values. The resulting peak SAR values are reduced by applying a power coefficient β. This coefficient back off the transmitted RF power in order to satisfy peak local SAR constraints.
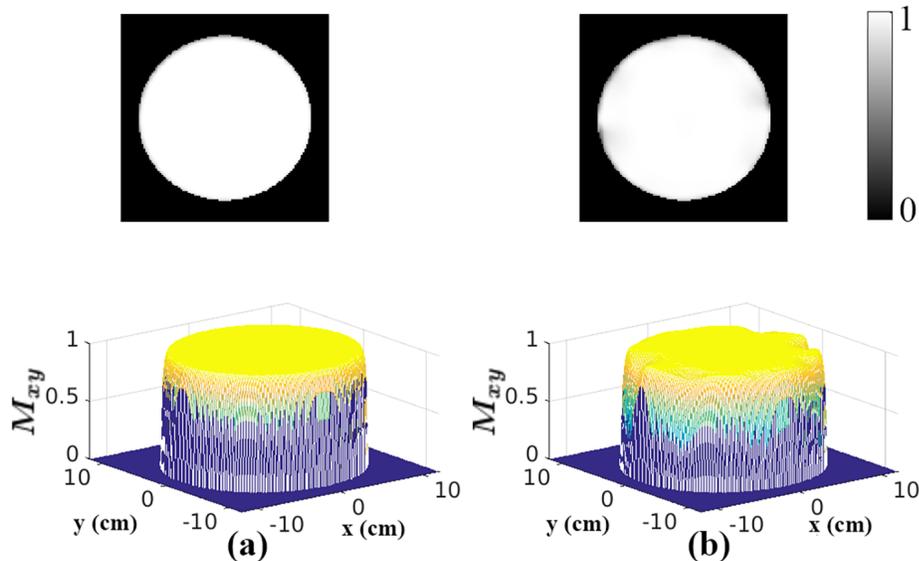
Fig. 4.9. The desired excitation pattern (a) and the simulated excitation pattern (b), whose peak value was scaled to $\pi/2$ tip angle. The RF pulse design procedure does not have any type of constraint. NRMSE = 0.003. $20 \cdot log(gSNR) = 19.18 \ dB$.

This *power back-off approach* achieved the SAR regulations by sacrificing the tip angle in the RF excitation. For example, if the $gSNR$ obtained from this approach is 90% of the $gSNR$ obtained from the *unconstrained approach*, then the peak local SAR would be reduced by 7.1% (see Fig. 4.10), compared to the *unconstrained approach*. However, the NRMSE between the desired and simulated excitation pattern would increase, and it can be concluded from Fig. 4.10(a) that the excitation pattern produced by the designed RF pulses (dash line) was shown to have a smooth excitation with a low penalty in the tip angle accuracy.
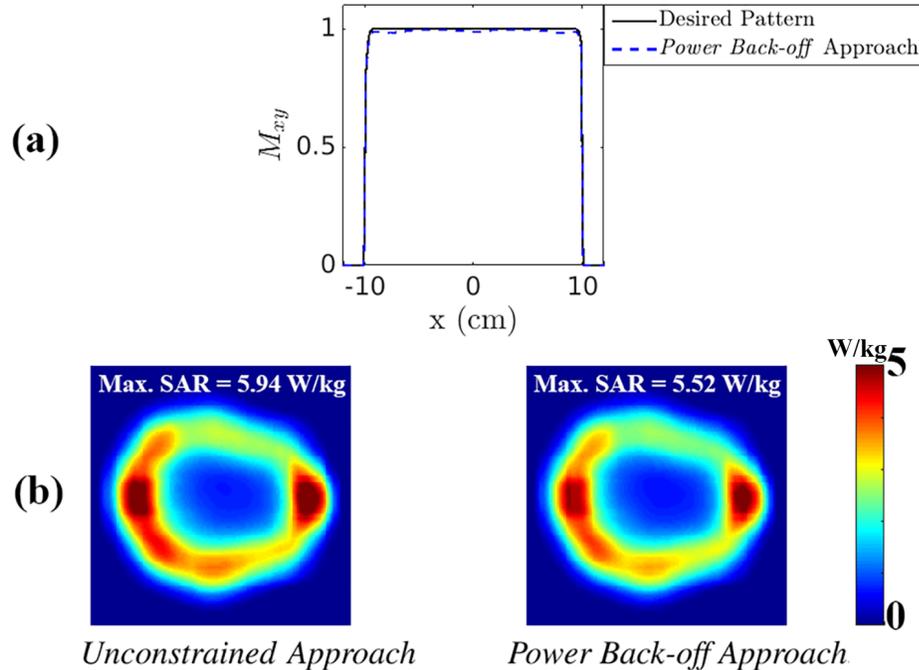
Fig. 4.10. (a) Two excitation patterns at $y = 0\ cm$, one is the desired excitation pattern (black solid line) and the other is Bloch simulated pattern produced by the designed RF pulses (blue dash line). NRMSE $= 0.036$. (b) The peak local SAR (W/kg) in the *power back-off* approach was reduced by 7.1%, compared to the *unconstrained approach*. The estimated SNR output in this approach is $20 \cdot log(gSNR) = 18.26$ dB, which is 90% of the $gSNR$ obtained from the *unconstrained approach*.

Power and Local SAR Constrained Approach

The pTx pulse design method offers a novel method to provide a large DOFs for controlling the local SAR reduction. This *power and local SAR constrained approach* can be classified as the approach that enforce peak local SAR constraints [9, 42, 104, 117].

In this approach, multiple peak local SAR constraints can be converted into Tikhonov regularization terms in the optimization problem. This is known as the Lagrangian-dual method [118]. The Lagrange coefficients associated with those regularization terms can be adjusted to control the peak local SAR values of those

generated VOPs or $k-$means clusters. In the SAR compression model, the representative SAR matrices with higher spectral norm are often considered to be vulnerable to the RF exposure and thus need to be thresholded with higher Lagrange multipliers, and SAR matrices with smaller spectral norm can be omitted with zero Lagrange multipliers. In the following simulation, 30 representative SAR matrices with 9.75% overestimation rate were applied in this pTx pulse design procedure.

Fig. 4.11 shows the bar plot of the peak local SAR reduction rate (compared to the *unconstrained approach*) associated with the number of non-zero Lagrange multipliers applied in this approach. The NRMSE bars indicate the excitation accuracy decreased when more non-zero Lagrange multipliers were introduced. The peak local SAR reduction rate is defined as the percentage of reduction compared to the *unconstrained approach*. During the pTx pulse design procedure, we firstly considered applying the non-zero multiplier to the SAR matrix with the largest spectral norm, and then moving to SAR matrices with smaller spectral norms. The red dash line plot of the peak local SAR reduction rate increased and eventually became steady. Some SAR matrices with smaller spectral norm might still be very vulnerable to the RF exposure, and their corresponding Lagrange multipliers should not be considered as zeros. The RF power constraint of this simulation was set with an empirical value.

Approach with Power, SAR, SNR Constraints

To evaluate the effect of RF pulse design method with RF power, SAR, and SNR constraints, we first fixed the α_0 coefficient associated with RF power to the previously prescribed empirical, and then we kept increasing those α_v coefficients that are associated with peak local SAR values of those SAR clusters. All excitation results should have a lower bound of $gSNR$ value.

Fig. 4.12 shows the relationship between NRMSE values, SAR reduction rates (compared to the *unconstrained approach*), and $gSNR$ values in this pTx RF design approach. We kept increasing α_v values to achieve the minimum peak local SAR
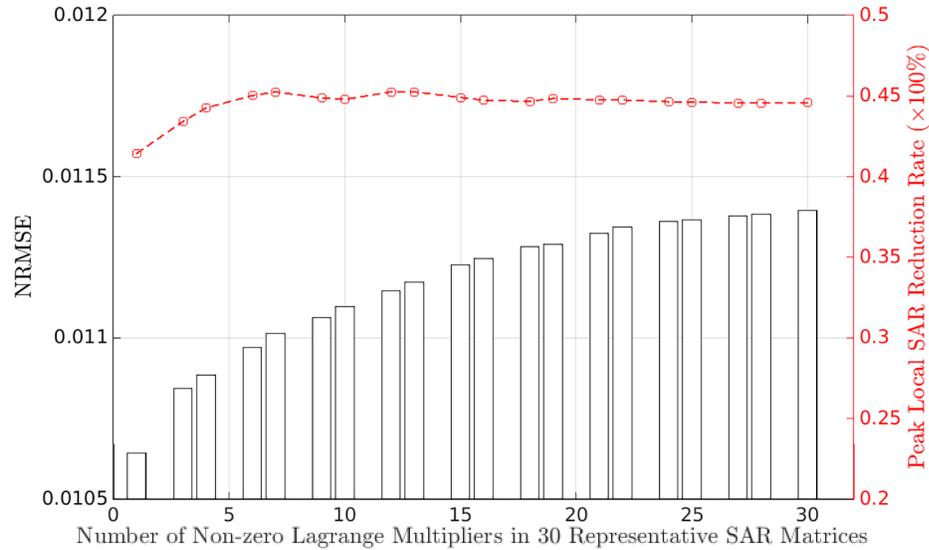
Fig. 4.11. This bar plot indicated the relationship between the NRMSE and the number of non-zero Lagrange multipliers applied in the *power and local SAR constrained approach*. The line plots show the relationship between the peak local SAR reduction rate (compared to the *unconstrained approach*) and the number of non-zero Lagrange multipliers (red dash line). This simulation was conducted with a fixed RF power coefficient $\alpha_0 = 5 \times 10^{-4}$, and the total number of representative SAR matrices is 30.

values, and the resulting $gSNR$ should be no smaller than the desired lower bound. The SAR reduction rate increased rapidly with smaller NRMSE values, but the gains decreased as NRMSE increased. The $gSNR$ dropped faster with respect to the excitation error tolerance. There existed a primary trade-off between the NRMSE/$gSNR$ and the SAR reduction rate. The results show that the peak local SAR was reduced by 44.6-54.2% when the constrained pTx approach had a lower bound of SNR equals to 90% of the $gSNR$ obtained in the *unconstrained approach*, with the NRMSE from 0.011-0.022.

Fig. 4.13 shows the trade-off between global/peak local SAR values and the excitation accuracy in this pTx approach. The resulting L-curves were obtained by varying α_v coefficients so that it always satisfied the NRMSE limit and the minimum $gSNR$ limit (90% of the $gSNR$ in the *unconstrained approach*). Both global SAR
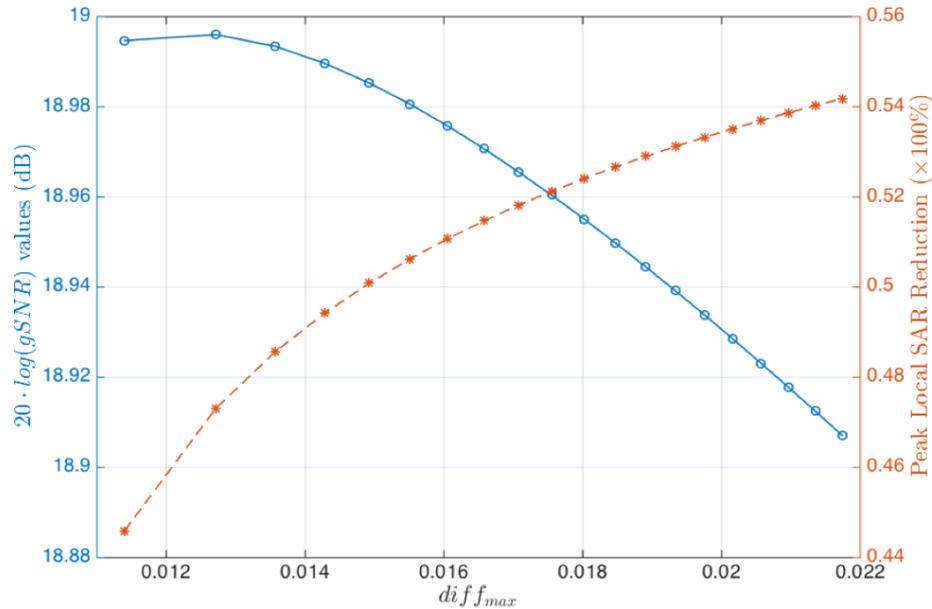
Fig. 4.12. Line plots illustrating the relationship between the NRMSE and the $20 \cdot log(gSNR)$ value (blue solid line); and the relationship between the NRMSE and the SAR reduction rate (compared to the *unconstrained approach*; red dash line). This simulation was conducted with a fixed RF power coefficient $\alpha_0 = 5 \times 10^{-4}$, and the total number of representative SAR matrices is 30.

(red curve) and peak local SAR (the blue curve) kept decreasing like L-curves with respect to the excitation error tolerance.

Fig. 4.14 plots the parallel RF excitation result and the SAR distribution map when the minimum $gSNR$ limit was achieved in this approach. However, a larger penalty than the *power back-off approach* on the tip angle pattern (NRMSE $= 0.059$) was induced in this excitation result. Two excitation patterns at $y = 0\ cm$ (one is the desired pattern, and the other is the Bloch simulated pattern using this approach) are shown in Fig. 4.14(a). Both raw SAR and 10-g SAR distribution maps of this approach are shown in Fig. 4.14(b), indicating that 58.7% of the peak local 10-g SAR was minimized in this constrained pTx approach.
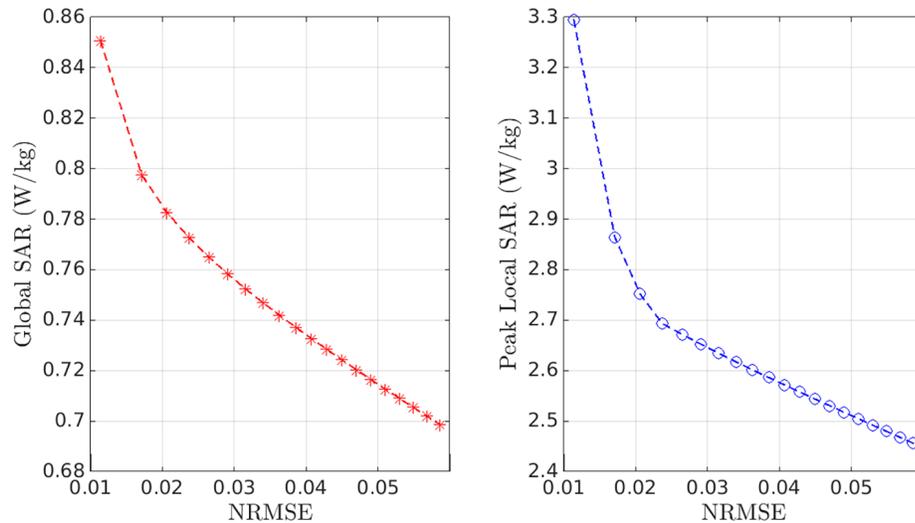
Fig. 4.13. Global/peak local SAR values are decreasing like L-curves with respect to the excitation error tolerance in this pTx RF pulse design approach. These curves are obtained by increasing Lagrange coefficients which control those peak local SAR values. The minimum $gSNR$ limit is $20 \cdot log(SNR_{min}) = 18.26 \ dB$ for all data points.

4.5 Discussion

Among different multi-channel RF pulse design approaches presented in this paper, the *unconstrained approach* provides the fundamental basis of the multi-channel RF pulse design method in parallel RF excitation [71]. Theoretically, the *unconstrained approach* generated the highest SNR among other approaches since it did not have any constraint on the transmitted RF power. The *power back-off approach* is the most common method that has been used in the field of wireless communication to achieve the regulatory SAR compliance [112]. However, this approach often induces a penalty in the tip angle accuracy. We achieved regulatory SAR limits by incorporating the transmit RF power constraint and local SAR constraints into the objective function of the pTx pulse design method. The resulting optimization problem can accommodate any arbitrary number of representative SAR matrices, and they enable the control over the RF power constraint and SAR constraints in either the whole body or in any local subvolume.
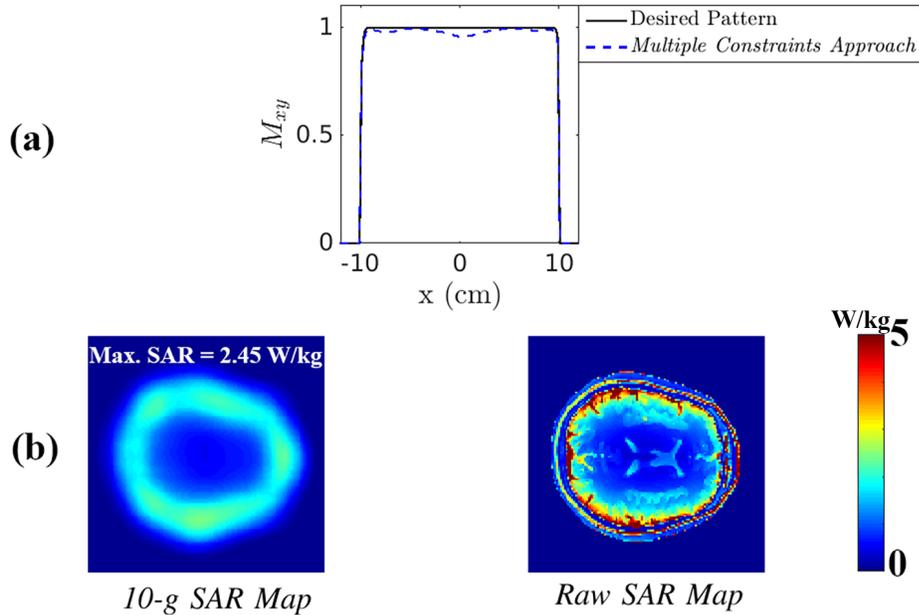
Fig. 4.14. (a) Two excitation patterns at $y = 0$ cm, one is the desired excitation pattern (black solid line) and the other is the Bloch simulation pattern produced by the designed RF pulses (blue dash line). NRMSE = 0.059. (b) The peak local SAR (W/kg) in this approach was reduced by 58.7%, compared to the *unconstrained approach*. The estimated SNR output in this approach is $20 \cdot log(gSNR) = 18.26$ dB.

We further provided an alternative method to implement the SAR model compression in the pTx pulse design method. The $k-$means clustering algorithm constructed a compressed set of representative SAR matrices to represent the actual SAR model. The resulting compressed SAR model is very close to the actual SAR model, and it is guaranteed to be safe under any circumstance. The results show that the $k-$means method, as compared to Lee's method [69], produced a smaller set of representative SAR matrices if under the same overestimation bound.

As for the speed and scope of the presented simulation work, the proposed approaches have shown their numerical robustness and efficiency. The $k-$means method is an NP-hard problem that usually takes less than an hour to obtain the clustered results. The computation time can be dramatically minimized if the initial centroids

are well-defined. A regular two-dimensional slice-selective RF pulse can be designed in a few seconds if only the transmit power constraint is applied, and the computation time can go up to a few minutes if less than 100 representative SAR matrices are introduced into the RF pulse design procedure. Furthermore, the pTx RF pulse design with larger number of global or peak local SAR constraints could be finished in a short time period if the computation can be done on the parallel computing system. VOP or $k-$means compression could substantially save the computation time in a parallel computing architecture. The time complexity reduction would potentially provide guidelines about real-time pTx RF pulse design in the future.

Our proposed $k-$means clustering algorithm has another potential benefit that allows for inner volume selection. The VOP compression method only selects representative SAR matrices to control peak local SARs in the human body, while the $k-$means clustering method is designed to cluster voxels with similar SAR behaviors. Generated clusters with higher values of spectral norm can be considered as clusters that are most vulnerable to the RF exposure. This advantage could be used for online local SAR monitoring in future studies. It is still an open question as to how to integrate the online SAR monitoring along with the offline SAR compression. Once this problem is solved, pTx technology will become the cutting-edge technology that can both effectively mitigate RF field inhomogeneity and explicitly control the RF exposure.

4.6   Conclusion

In this study, we proposed a pTx RF pulse design method which incorporated constraints on RF power, SAR, and the MRI receive voltage for a given human head model and a target excitation profile. The proposed pTx pulse design method jointly considered the RF power and SAR penalized RF pulse and the desired lower bound of SNR. It achieved the minimum peak local SAR while satisfying the lower bound of SNR and the excitation performance criterion. The results presented in

this chapter confirmed the effectiveness and the feasibility of different approaches on the multi-channel RF pulse design. We further demonstrated a new $k-$means clustering algorithm based SAR compression model in the pTx pulse design method. The $k-$means method constructed a compressed set of SAR matrices with a narrower and more accurate overestimation bound, as compared to the conventional VOP compression model.

5. MACHINE LEARNING BASED RF PULSE PREDICTION IN PTX MRI

5.1 Introduction

In this chapter, we proposed several machine learning methods to implement the pTx RF pulse design in UHF MRI, and they all saved a great amount of computation time in the design procedure. High field or UHF MRI systems can maintain high SNR in reconstructed MR images, but they also bring new unsolved issues in reality. On one hand, higher field strength allows for shorter image times, better image contrast, and more details in smaller regions to be visualized because of stronger detected MRI signals. On the other hand, inhomogeneous B_1 field, increased RF power deposition, and increased susceptibility artifacts are presented in high field and UHF MRI machines [27, 119].

RF pulses in high field or UHF MRI systems are always with high RF power to give rise to the excitation, but it further causes the transmit B_1^+ field to be spatially non-uniform. Consequently, the inhomogeneous transmit B_1^+ field will result in a non-uniform tip angle pattern and a spatially variant image contrast [120]. We know that pTx technology in high-field or UHF MRI machines offers the RF pulse designer extra DOFs to control the RF field, and it can be used to mitigate the inhomogeneity in magnetic fields. The pTx technique can be used to design multi-dimensional RF pulses on a patient-by-patient basis, and are referred to as tailored pTx RF pulse design. Many challenges include high RF exposure and high temperature rise in the human body can be controlled using this tailored method as these factors can be served as constraints in the pTx framework. Several advanced approaches have been proposed to optimize pTx RF pulses in high-field or UHF MRI environments [21–23, 33], but these methods usually generate local optimum solutions since there is

always a trade-off between the solution optimality and the computational time. The major problem of the pTx RF pulse design remains to be speeding up the computation without sacrificing the RF excitation performance.

Research on machine learning methods are fastly growing in recent years, and researchers started to apply machine learning methods in the pTx RF pulse design. Prior research [121] constructed a universal multi-dimensional RF pulses in multi-coil MRI by pre-optimizing them using a large set of human brains. Those pre-optimized RF pulses were applied in a 3D non-selective k_T-points excitation [23], and they were claimed to be applicable for any arbitrary human brain model. This method was considered to be one of the traditional machine learning methods, but its excitation performance was not as well as the tailored pTx approach. One recent research [122] used an advanced machine learning model, trained neural networks, for a tailored spokes RF pulse prediction, but it failed to have a constant performance in generating a homogeneous tip-angle pattern for any arbitrary case. Other approach [123] did not focus on the pTx RF pulse design, but predicted RF shim weights in multi-coil MRI using a trained ridge regression model. This method iteratively updated the optimal RF shims by merging the training and learning procedure together.

Here, we proposed two machine learning methods for the pTx RF pulse design and investigated the relations between optimized pTx RF pulses and transmit excitation profiles for which they were tailored. The two proposed methods are the kernelized ridge regression (KRR) based pTx RF pulse design, and the feedforward neural network (FNN) based pTx RF pulse design. Both designed algorithms can directly predict the pTx RF pulses using the extracted features from the given transmit sensitivity profiles. They both can finish the prediction quickly since these two methods only require the feature calculation and the RF pulse prediction, and both calculations can be finished in linear time.

5.2 Methods

5.2.1 Magnitude Least-Squares pTx Pulse Design

Patient-tailored, SAR-efficient pTx pulses can be generated by solving the following magnetization matching problem with Tikhonov regularization terms

$$\mathcal{B}_{opt}^{\dagger} = \underset{\mathcal{B}}{\operatorname{argmin}} \left(\left|\left| \tilde{\mathbf{S}}_A \mathcal{B} - \boldsymbol{m}_{ref} \right|\right|_2^2 + \frac{1}{\mathcal{K}} \sum_{i=0}^{\mathcal{N}} \theta_i \mathcal{B}^H \hat{\boldsymbol{Q}}_{\boldsymbol{d},v} \mathcal{B} \right),$$ (5.1)

where $\theta_i, i \in 0, ..., \mathcal{N}$ are the coefficients controlling the RF power and local SAR terms in the SAR observation clusters. As we discussed in the previous chapter, $\tilde{\mathbf{S}}_A$ is the diagonal matrix containing the transmit sensitivities and $\mathcal{B}$ is the transmitted RF waveforms across N_t channels. Here, $\mathcal{B} = [\mathcal{B}(t_1), ...\mathcal{B}(t_{\mathcal{K}})]^{\top}$ and $\mathcal{B}(t_k) = [B_1(t_k), ..., B_{N_t}(t_k)]^{\top}$, where t_k refers to the kth time point in each RF pulse ($\mathcal{K}$ time points in total).

The optimization problem in (5.1) is a conventional magnetization matching problem which can be solved using either CG iterations or pseudoinverse algorithms. As a result, a specific RF pulse vector is obtained for each transmit channel in multi-coil MRI.

5.2.2 Kernelized Ridge Regression Training Method

Based on the work proposed by Ianni et al. [123], we apply the KRR method to predict complex RF pulses for each coil and each slice from a $N_{feat} \times 1$ feature vector $\boldsymbol{f}$ (i.e., the input of the KRR learning model). The features include the central Fourier transform coefficient of the B_1^+ field map for each coil, the $z-$position of the slice in the head model, and the Fourier shape descriptor of the ROI mask contour. The vector $\boldsymbol{f}$ of the training dataset is normalized to have zero mean and unit standard deviation. The vector $\boldsymbol{f}$ of the testing dataset is normalized by the same scale. The first order cross-product and sign operator of the aforementioned features are computed and included in the feature vector, therefore this machine learning method is kernelized. The RF pulse $B_p(t_k)$ for coil p can be written as

$$B_p(t_k) = \boldsymbol{w}_{p,t_k}^H \boldsymbol{f}, \tag{5.2}$$

where the feature weight vector $\boldsymbol{w}_{p,t_k}$ relates each feature to t_k time point of coil p's transmitted RF pulse. The feature weight vector $\boldsymbol{w}_{p,t_k}$ is obtained by fitting the following linear model to the training dataset

$$\hat{\boldsymbol{w}}_{p,t_k} = \frac{1}{2}\underset{\boldsymbol{w}_{p,t_k}}{\operatorname{argmin}}||\boldsymbol{B}_{p,train}(t_k) - \boldsymbol{F}\boldsymbol{w}_{p,t_k}||_2^2 + \frac{1}{2}\zeta||\boldsymbol{w}_{p,t_k}||_2^2, \tag{5.3}$$

where the $N_{train} \times 1$ complex vector $\boldsymbol{B}_{p,train}(t_k)$ contains t_k time point of the optimal coil p's RF pulses for all the training slices, $\boldsymbol{F}$ is an $N_{train} \times N_{feat}$ matrix in which each row is $\boldsymbol{f}^\top$, and ζ is the Tikhonov regularization term.

Therefore, we can use the optimal training RF pulses that are solutions to (5.2) to solve for (5.3) for the feature weights. We further applied a 10-fold random cross validation to the training dataset, so that 90% of the training data were used for training and 10% training data were used for validation. The details about the steps of the training and testing procedures are summarized in Figure 5.1.

5.2.3 Neural Network Training Method

Artificial neural network becomes a hot topic in recent years, as it can be trained if matched input and output samples are provided. Recently, neural networks are used in pTx RF pulse design [122], where the input is the B_1^+ field map and output is the designed pTx RF pulse. In this section, we proposed trained feedforward neural networks to predict pTx RF pulses only from the key features extracted from the given B_1^+ field maps.

The designed neural network consisted of one input layer of 306 input nodes, one hidden layer with 10 nodes, and one output layer with 2 nodes. This three-layer FNN was applied to predict the magnitude and phase of the RF pulse at each time point and each coil. The reason we chose to train the RF pulse time point by time point is that we traded the time complexity for space complexity. We reduced the number of
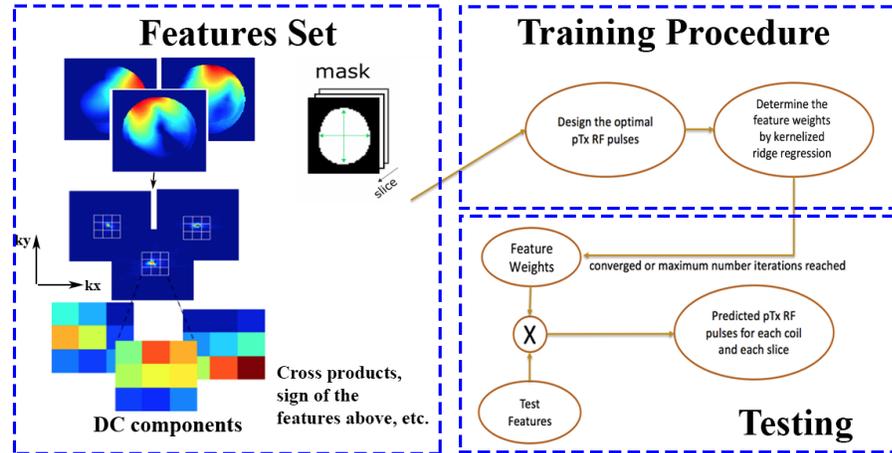
Fig. 5.1. (a) The pTx RF pulse prediction algorithm by kernelized ridge regression. The features of each slice included the DC components of the Fourier domain of the B_1^+ fields, tissue mask metrics, slices positions, etc. The first order cross-product terms and sign operators were applied for all features. (b) The training procedure fitted the training dataset into a regularized linear model to obtain the feature weights. The training dataset contained the SAR-efficient pTx RF pulses that were solutions to (5.2), and their corresponding feature vectors. (c) The testing procedure applied the learned feature weights to the testing slices' features to obtain the prediction of the pTx RF pulses for the testing slices.

layers and neurons in the network because it only had two neurons in the output layer. We later combined all trained networks together for an instantaneous prediction of the entire pTx RF pulse. The neural networks were implemented using MATLAB R2017a (The Mathworks, Natick, MA, US) to minimize the MSE between the neural network output pTx RF pulses and the trained pTx RF pulses. The training dataset contained all input features extracted from the training B_1^+ maps, and it was divided to have 80% training, 10% validation, and 10% for initial testing. Those computed neural networks were later tested with our produced testing dataset.
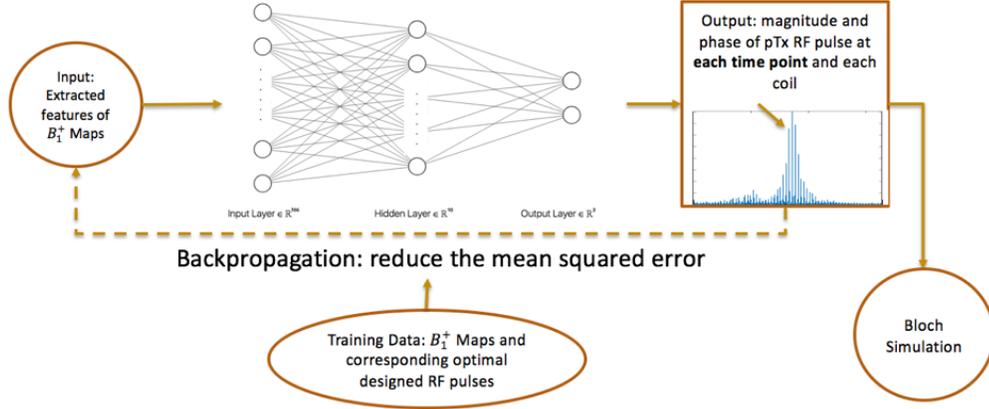
Fig. 5.2. The schematic diagram of the proposed FNN setup. The proposed FNN was used to predict the magnitude and phase of the pTx RF pulse at each time point and each coil. Full neural network for each prediction consisted of 306 input nodes, 10 hidden nodes, and 2 output nodes.

5.3 Results

5.3.1 FDTD Simulation and Feature Extraction

To learn the spatial variation of B_1^+ fields, a large number of training B_1^+ maps and optimal solutions need to be created. An initial full EM simulation was conducted using a commercial FDTD solver (Remcom, XFdtd, PA, USA) with a loaded Duke human model [124] (downsampled to 128×128 voxels) and an eight-channel 7.0 T transceiver coil array [116] (see Figure 5.3).

For the initial B_1^+ field of each coil, we need to perturb it to generate a large number of training data samples. We should first know that for a given ROI, we can separate the magnetic field in this ROI into two sub-fields, the local field and the background field [125]. The local field, $\hat{\boldsymbol{f}}_L$, reflects the magnetic response from the human tissues in this ROI. This field is majorly caused by the susceptibility distribution induced by tissue properties. The background field, $\hat{\boldsymbol{f}}_B$, reflects the magnetic response from the RF coil. This background field is the near field of the
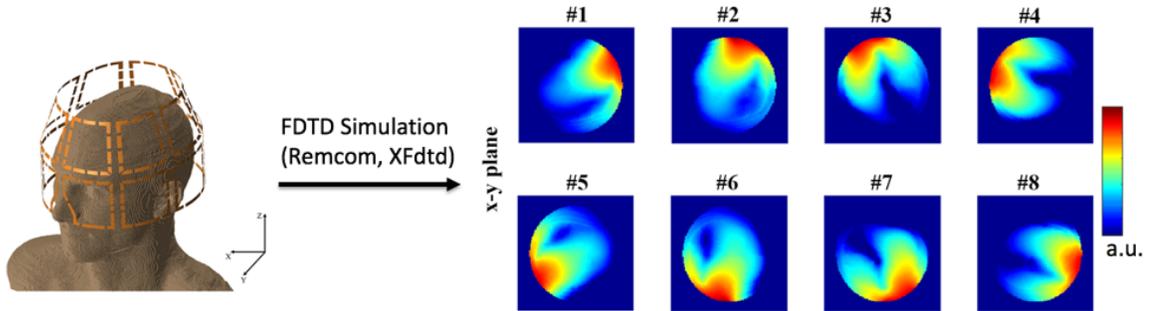
Fig. 5.3. The initial full EM simulation was conducted with the Duke human head (downsampled to 128×128 voxels) in an 8-channel transceive coil array (lower ring of the 16-channel array).

coil, which is inside the FOV but outside the ROI. The near field of the RF coil will likely produce some variations, and thus we can perturb the background field to obtain a large number of valid B_1^+ fields.

The algorithm of the field separation is similar to the projection theorem in the Hilbert space [125, 126]. The local field, $\hat{\boldsymbol{f}}_L$, should be orthogonal to the background field, $\hat{\boldsymbol{f}}_B$. To be more specific, for each unit dipole field in the local field, its inner product with any arbitrary background unit dipole field should be close to zero. We evaluated this projection algorithm by finding the maximum absolute value of the normalized inner product between two sub-fields, and the result should be driven to zero.

After the field separation, we then started to perturb the near field, or the background field of the total field. The background field was decomposed into its Fourier components. For each slice (174 slices in total), top 80% of Fourier coefficients were randomly perturbed 100 times according to a normal distribution (zero mean, 0.05 variance) in each of the transmit channels. A total dataset of $17, 400 \times 8$ B_1^+ fields and their corresponding optimal pTx RF pulses can be created (see Fig. 5.4). For the total of $17, 400 \times 8$ data samples, we used $15, 000 \times 8$ data samples for training and rest of them for testing. Note that there is no overlap in terms of the slice location

between training and testing data samples. In other words, training and testing data samples do not share the same local field.
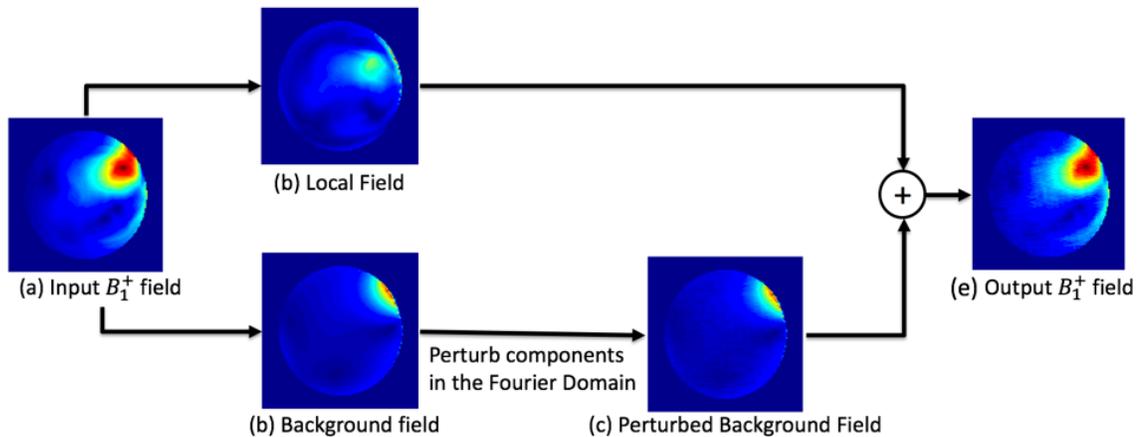
Fig. 5.4. The pipeline of the creation of the training and testing B_1^+ field map. (a) The initial B_1^+ field map obtained from the EM simulation in p^{th} transmit coil. (b) The local field extracted from the input B_1^+ field. (c) The background field extracted from the input B_1^+ field. (d) Perturbed background field after we randomly manipulate coefficients in the Fourier domain according to a normal distribution (zero mean, 0.05 variance). (e) The output B_1^+ field from the sum of the local field and the perturbed background field.

5.3.2 Comparison to Other Supervised Machine Learning Methods

Those pTx RF pulses generated by the KRR method and the FNN method were compared to pTx RF pulses obtained by the nearest neighbor (NN) and k-nearest neighbors (kNN) methods, the classic supervised learning methods in the machine learning research area. We used the Shepp-Logan phantom as the desired excitation pattern in this simulation. The NRMSE values between the desired Shepp-Logan pattern and the Bloch simulated pattern obtained using different methods were compared in the following simulations.

Fig. 5.5 shows best, median, and worst cases of the Bloch simulated excitation profiles in randomly selected training and testing slices (300 random slices of each dataset) across the direct design, KRR, NN, kNN ($k = 20$), and FNN methods. Fig. 5.6 shows the box plot of NRMSE distributions in training and testing slices, across all learning methods. The averaged NRMSE was 12.87% for the direct design methods, 16.09% for the KRR predictions, 13.98% for the NN predictions, 13.99% for the kNN predictions, and 14.72% for the FNN predictions. Therefore, the direct design method produced the most accurate excitation profile, followed by the NN, kNN, and FNN prediction methods. The direct design method produced the best result because we considered their results as the ground truth in the machine learning model. All the other machine learning methods were trying to obtain the results as well as the direct design methods, and it turned out they can maintain a good excitation accuracy in best and median cases. In worst cases, the machine learning methods still need improvements.

NN and kNN produced relatively good excitation profiles on average majorly because training and testing slices shared similar B_1^+ fields. Some slices in the training and testing dataset were physically close and they shared similar local fields. This is the major reason they have such good performances. The FNN method performed the third best, because it ran multiple epochs in the training process to verify the total loss was minimized in the training, validation, and testing cases. It selected the optimal neural network weights with the best validation performance. The KRR method, on the other hand, could only minimize the total loss in the training procedure, but could not guarantee to minimize the loss in the validation or testing procedure. Some KRR results had worse performance because there is a large variation in the phase data of the training pTx RF pulses. Although KRR maintained the excitation accuracy in best and median cases, it still generated a few worst and non-uniform excitation profiles as outliers.

Fig. 5.7 shows the estimated global and peak local SAR differences between the direct design method and the KRR learning method. For the training dataset,
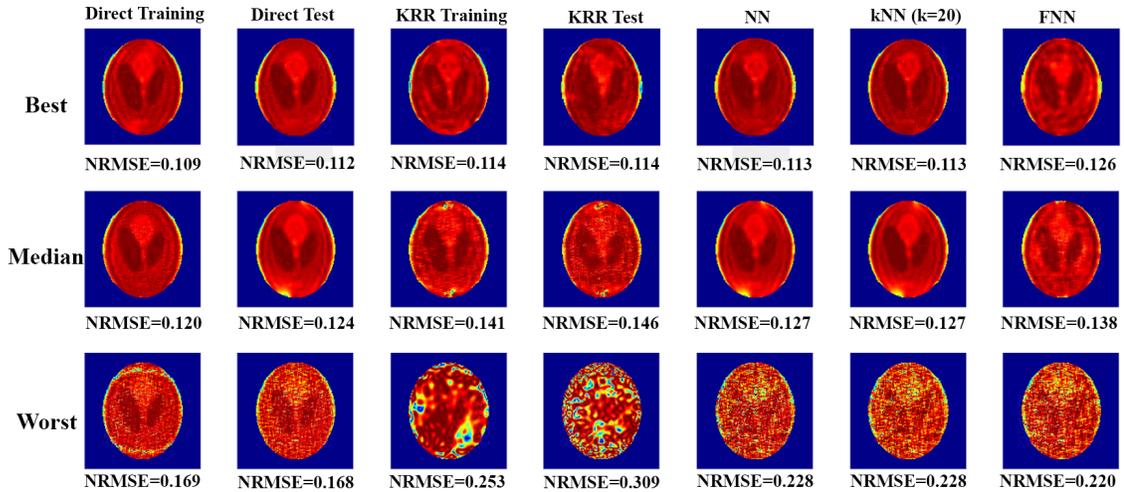
Fig. 5.5. Best, median, and worst simulated excitation patterns in training (213 random slices) and testing slices (240-300 random slices), for the direct design, KRR, NN, kNN, and FNN methods.

estimated global SARs in two methods demonstrated high positive correlation (i.e., $r = 0.989$, Pearson's correlation coefficient) with no statistically significant bias ($P = 0.68$, two-sided t-test), and the difference in peak local SARs was always within $[-7, +9]$ W/kg. As for the testing dataset, estimated global SARs in two methods also demonstrated high positive correlation (i.e., $r = 0.967$, Pearson's correlation coefficient) with no statistically significant bias ($P = 0.40$, two-sided t-test), and the difference in peak local SARs was between $[-4, +12]$ W/kg.

Fig. 5.8 demonstrates estimated global SAR and peak local SAR differences between the direct design method and other machine learning methods, including NN, kNN, and FNN methods. These estimated global SAR values between NN solutions and direct design solutions demonstrated high positive correlation (i.e., $r = 0.992$, Pearson's correlation coefficient), and the difference in peak local SARs was always within $[-2, +11]$ W/kg. For the kNN method, the Pearson's r value was 0.996, and the difference in peak local SARs was within $[-2, +3]$ W/kg. As for the FNN method, it sill showed a high positive correlation with the direct design method (i.e., $r = 0.949$).
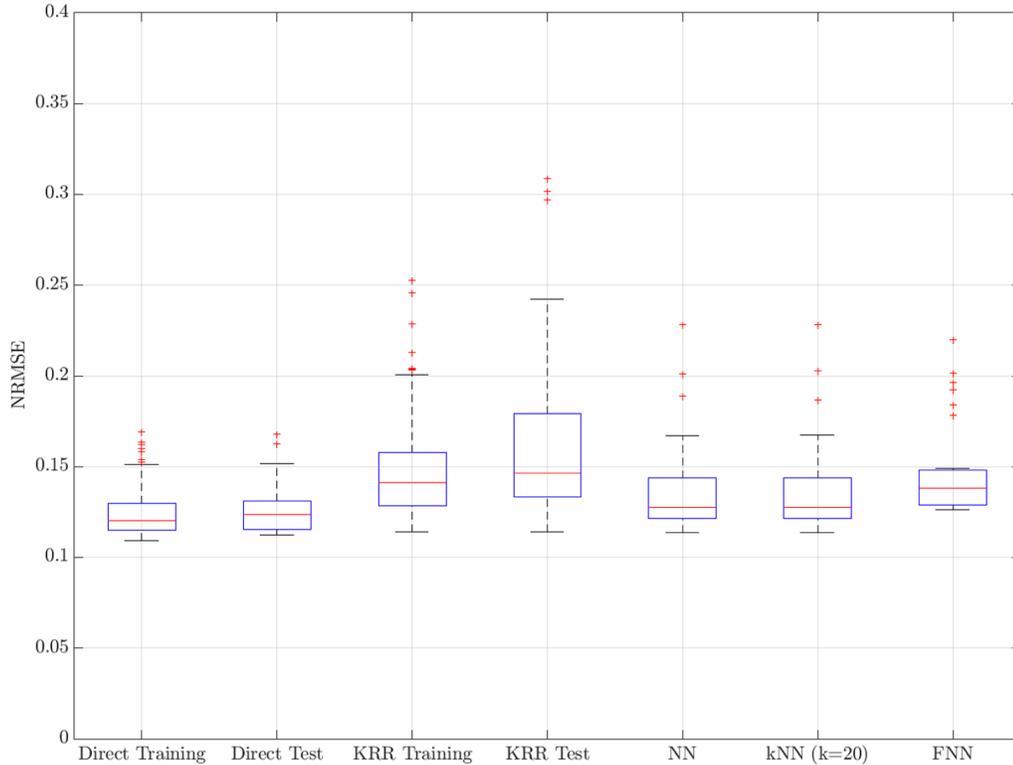
Fig. 5.6. The boxplot of the NRMSE values between Bloch simu-
lated excitation patterns and the desired simulated excitation pattern
across training (213 random slices) and testing slices (240-300 random
slices), for the direct design, KRR, NN, kNN, and FNN methods.

However, the range of differences in peak local SARs was wider than the former two
cases, which was in $[-2, +14]$ W/kg.

The major advantage of the machine learning based pTx RF pulse design is to
reduce the computational complexity in solving optimal pTx RF pulses. The machine
learning methods were implemented in MATLAB R2017a (The Mathworks, Natick,
MA, USA) using 12 core CPUs (2.60 GHz) to parallelize computations. The average
time to calculate the entire pTx RF pulse using the direct design method was $3 - 4$
mins, however the average prediction time for pTx RF pulses generated by KRR, NN,
kNN, and FNN methods were 0.033 s, 0.07 s, 0.04 s, and 0.055 s, respectively. The
training time for KRR and FNN methods were generally long, but the training was a
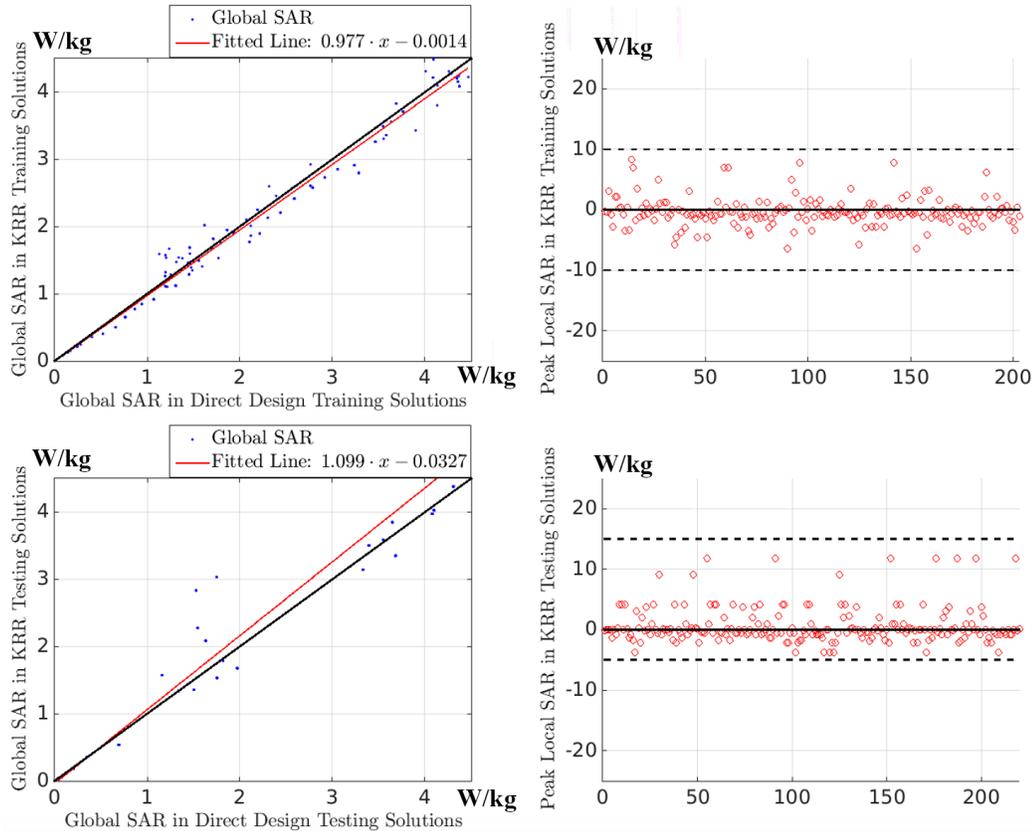
Fig. 5.7. The estimated global (left) and peak local SAR (right) difference between direct design solutions and KRR solutions, across training (203 random slices) and testing slices (220 random slices).

one-time computation. The current KRR training time for each time point and each coil was approximately 0.37 s (100 maximum iterations), and it took roughly 1.93 hours to finish the entire training for a 4.7 ms RF pulse and eight transmit coils. The current FNN training time for each time point and each coil was approximately $2 - 5$ min, and it took ~ 20 days to finish the entire training.

5.4 Discussion

All computational simulations presented in this chapter demonstrated that machine learning methods have the potential to predict SAR-efficient pTx RF pulses in
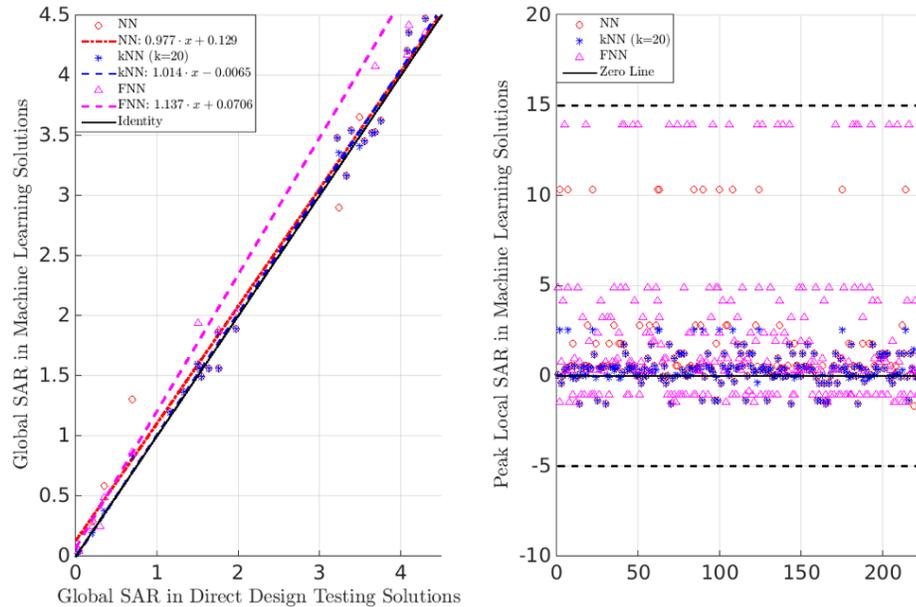
Fig. 5.8. The estimated global (left) and peak local SAR (right) difference between direct design solutions and NN/kNN/FNN machine learning solutions, across testing slices (225 random slices).

multi-coil and UHF MRI systems. The major advantage is that the computational burden in machine learning methods is much smaller than the direct design method. The computational time in the direct design method highly depends on the number of representative SAR matrices in the SAR compression model, and the time may vary from a few seconds to a couple of minutes. However, the computational time in machine learning methods remains constant for almost all cases. Once the training model is set up, the prediction time is predicable and is usually within the milliseconds level. The training time, on the other hand, is generally long, and it can be substantially minimized if we can use GPUs to parallelize computations. These machine learning methods could be extended to real-time training in order to maintain a big training library, which may require additional features for the RF pulse prediction, such as slice angles (i.e., the information about slice prescriptions) and slice locations (i.e., which part of the body).

Another advantage of the machine learning based pTx RF pulse design methods (KRR, NN, kNN, FNN, etc.) is that they require less information, compared to traditional pTx RF pulse design methods (i.e., the direct design method). They only need DC components of coils' B_1^+ field maps and slice-specific ROI masks. DC coefficients of a B_1^+ field can be obtained using one specific excitation pattern for each coil and each slice, and this scan time should be faster than the full B_1^+ mapping technique since we do not require extra phase-encoding lines (only need central Fourier transform coefficients). The ROI mask can be obtained using a gradient recalled echo pre-scan, which is commonly used in the receive sensitivity mapping scan for the parallel imaging preparation.

The major concern of machine learning based pTx RF pulse predictions in high-field or UHF MRI systems is the SAR control. As we summarized in Fig. 5.7 and Fig. 5.8, these machine learning methods generated diverse and random SAR values, although they tended to have high positive correlations with global SAR values obtained from the direct design method. In the direct design method, SAR and RF power constraints were considered as regularization terms, and thus those resulting pTx RF pulses were SAR-efficient. We then used these SAR-efficient pTx RF pulses as training data samples and applied aforementioned machine learning models. This was a sub-optimal solution to effectively control global and peak local SAR values because it did not guarantee that the predicted pTx RF pulses would meet SAR regulation constraints. One possible direction is to incorporate strict SAR constraints in the training model so that the resulting RF pulses are guaranteed to meet the constraints. However, this method might prolong the training time, and yet the optimal solution to this SAR problem is remained to be studied. The future goal is to bound peak local SARs in machine learning or deep learning methods without sacrificing the excitation accuracy or increasing the TR.

5.5   Conclusion

This chapter summarized different machine learning methods used in the tailored and slice-selective pTx RF pulse design. Two machine learning methods, KRR and FNN, were proposed to speed up the computation in the design procedure. KRR method learned the trained transform quickly, and it could maintain the excitation accuracy in best and median cases in the training/testing dataset. It still generated a few worst and non-uniform excitation profiles as outliers. FNN method usually took a longer time to train, but it maintained a good excitation accuracy and SAR behaviors in the testing slices. Traditional NN and kNN ($k = 20$) had a good performance majorly because training and testing slices shared similar B_1^+ fields. The major advantage of machine learning methods is that they do not require the full B_1^+ map as input, and they can save a considerable computation and pre-scan time.

6. DISCUSSION AND CONCLUSION

6.1 Contributions

This dissertation is focused on integrating the pTx technology into a transceiver coil array in high-field and UHF MRI systems, since higher field strength can be used to assist the improvement of enhanced imaging in the human body. High-field or UHF MRI systems will always lead to some distortions in the amplitude and phase data of the resonant magnetic field, which may cause some serious problems in the image acquisition. The pTx technology is therefore proposed to counteract this problem. The transmit coil array and the pTx technology combined take advantage of all transmit coils and generate an almost uniform tip angle pattern, therefore can reduce the distortions in reconstructed images into a negligible level. Further, they also provide extra DOFs in the RF pulse designer, which can be used to control some important imaging factors (SAR, temperature, etc.).

In this dissertation, we were especially interested in the optimal design of tailored pTx RF pulses, and we aimed to design the pTx RF pulses by including all possible constraints (SNR, SAR, NRMSE, etc.) into the design framework. The proposed methods were evaluated for various tip angle patterns and model parameters in simulations, but they have not been tested in real scanners yet. In practice, parallel transmit systems can be found in the recent commercial 3 T and 7 T MRI scanners, and two transmit channels are commonly used in these machines. However, optimal design and applications of pTx RF pulses with large number of coils still require further investigation.

The contributions of this dissertation can be classified into the following three sections.

- **MIMO MRI Model**: MIMO MRI is a novel and useful theoretical framework for MRI parallel RF transmission and reception. We have verified the pTx RF pulse design strategy for this specific model. The MIMO MRI framework constructed a signal processing model between transmitters and receivers of the MRI system by using both parallel excitation and parallel reception. This signal processing model was analogous to techniques developed for MIMO communication. It jointly optimized excitation errors and receive-side SNRs in the pTx RF pulse design procedure.

- **$k-$Means SAR Compression Model**: The proposed $k-$means clustering algorithm yielded a substantial gain in the SAR model compression, and showed an improved performance in the model accuracy. The computation time in the $k-$means clustering can be controlled by maximum iterations. The algorithm for all application cases can be converged using less than 100 iterations (< 1 hour). Although the computational time was a bit long, this was a one-time computation prior to the RF pulse design. The $k-$means clustering algorithm showed lower overestimation ratios, and allowed for inner volume selection. These attributes allowed the $k-$means SAR compression model to be used in real-time peak local SAR supervisions and possible use in the MR thermometry applications.

- **Machine Learning based pTx RF Pulse Prediction**: Several machine learning based pTx RF pulse prediction methods were proposed to obviate the the computational burden in tailored pTx RF pulse design in high-field or UHF MRI systems. All proposed machine learning methods could maintain a good excitation accuracy in best and median cases in the testing, but needed extra improvements to minimize the excitation error in worst cases. Traditional NN and KNN methods performed the best because of the existed similarity in training and testing data samples. KRR and FNN solutions showed relatively good performance in the excitation accuracy and SAR behaviors. Their SAR

values had high positive correlation with the SAR values in the optimal solutions, however they would need extra SAR control in order to meet strict SAR constraints. The biggest advantage of machine learning methods was that they could save a large amount of time in the RF pulse calculation. The resulting calculations required only some feature calculations and an application of the learned transform.

6.2   Limitations and Future Direction

Chapter 3 introduced the basic input-output model in the MRI system, which is referred to as MIMO MRI model. This model combined the parallel transmit and receive technology together in one framework. The MIMO MRI framework provides the fundamental basis for the multi-channel RF pulse design method, supporting the researches in **Chapter 4 and 5**. This model has the potential to overcome the majority of the aforementioned challenges in 3 T or 7 T MRI environments, and it further provides a guideline to solve the field inhomogeneity and image quality problems in UHF MRI machines ($\geq$ 7 T). Some major difficulties in this topic were 1) to produce a homogeneous excitation pattern while satisfying constraints of RF exposure, 2) to obtain a high SNR image after the image reconstruction in parallel reception. In the current stage of the MIMO MRI model, solutions of pTx RF pulse are sub-optimal. The next step is to construct an optimal pTx RF pulse design method to boost the SNR in the image reconstruction. The future direction of this research is to introduce acceleration schemes in the parallel reception into this model and investigate how excitation profiles and acceleration schemes affect $g-$factors and the image quality in the reconstructed image.

Chapter 4 introduced a novel $k-$means clustering algorithm to implement the SAR compression model in the pTx RF pulse design procedure. The biggest issue in the $k-$means model is the long computational time, and it needs extra improvements for fast converge. Future direction of this study is to ensure the local SAR safety by

all means in the pTx RF pulse design. To guarantee this, several safety conditions need to be considered:

- The SAR compression model needs to consider worst-case peak local SARs under any circumstance.

- The SAR compression model needs to consider the compression ratio and model accuracy in a way that they can be simultaneously controlled.

- To guarantee the SAR compression model can be used for any arbitrary subject, the SAR modeling should consider the inter-subject variability and possible EM simulation errors.

One major investigation focus of the SAR compression model in the future is to consider combining the offline SAR prediction with the real-time SAR monitoring. In this way, we could allow the pTx technology to be much more comprehensive and useful in high-field or UHF MRI machines.

Chapter 5 introduced several machine learning methods for pTx RF pulse design and validated their performance with the loaded Duke phantom to predict tailored SAR-efficient pTx RF pulses. The proposed KRR and FNN methods substantially reduce the required computational time down to the milliseconds level. However, the training time of those machine learning methods is still long. For example, although FNN method can maintain a good excitation accuracy, it appears to have an extremely long training process. This is because we traded the training time for memory because the testing station only has CPUs to parallelize computations. Future work of this study is to introduce wide and deep neural networks to train the entire pTx pulse in one time, therefore the training time can be substantially reduced. The biggest challenge in this research is to reduce the computational burden not only in the testing but also in the training procedure.

Another direction of this research is to build a big training library containing all kinds of variations. With this being stated, different human and phantom models,

different levels of noises, and different placements of the models need to be included in the training library. In general, we need to consider all possible model uncertainties in a way that the training library can handle all kinds of variations so that the prediction accuracy can be improved. The ultimate goal of this research is to build an online training library that can predict pTx RF pulses for any arbitrary B_1^+ field map and subject. Numerical simulation based training is beneficial because it can easily adapt all kinds of EM environments and different changes in B_1^+ field maps. Eventually, we target to have this technology implemented *in vivo*, but it might introduce new sources of variability. For example, subject's position may decrease the correctness of the prediction if they are not stayed in the center of the scanner. SAR management would be another significant concern in *in vivo* experiments. Current machine learning based pTx pulse design methods have loose control about global and peak SARs in the excitation, the final solutions are not guaranteed to meet strict SAR constraints. Therefore, future work of this research becomes finding a reliable method to control SAR values in predictions.

There is an increasing trend to apply different machine learning methods in neuroimaging and MRI field, since machine learning is a universal solution to many practical problems. Further, the model is easily trained once the matched input and output data samples can be provided. However, machine/deep learning aims for broad applicability but it will lose the prediction accuracy. One long-lasting objective of this future machine/deep learning based pTx technique is to reduce the overall variability and errors in the predictions, and this is also the fundamental goal of our proposed machine learning based pTx methods. We hope to see these proposed machine learning methods can be used to provide guidelines for new generations of machine/deep learning based pTx RF pulse design methods.

LIST OF REFERENCES

LIST OF REFERENCES

[1] B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde, "Functional Connectivity in the Motor Cortex of Resting Human Brain using Echo-planar MRI," *Magnetic Resonance in Medicine*, vol. 34, no. 4, pp. 537–541, 1995.

[2] D. Le Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat, "Diffusion Tensor Imaging: Concepts and Applications," *Journal of Magnetic Resonance Imaging*, vol. 13, no. 4, pp. 534–546, 2001.

[3] E. M. Haacke, S. Mittal, Z. Wu, J. Neelavalli, and Y.-C. Cheng, "Susceptibility-weighted Imaging: Technical aspects and Clinical Applications, Part I," *American Journal of Neuroradiology*, vol. 30, no. 1, pp. 19–30, 2009.

[4] S. Mittal, Z. Wu, J. Neelavalli, and E. M. Haacke, "Susceptibility-weighted Imaging: Technical Aspects and Clinical Applications, Part 2," *American Journal of Neuroradiology*, vol. 30, no. 2, pp. 232–252, 2009.

[5] Z. P. Liang and P. C. Lauterbur, *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective.* The Institute of Electrical and Electronics Engineers Press, 2000.

[6] B. P. Thomas, E. B. Welch, B. D. Niederhauser, W. O. Whetsell Jr, A. W. Anderson, J. C. Gore, M. J. Avison, and J. L. Creasy, "High-resolution 7T MRI of the Human Hippocampus In Vivo," *Journal of Magnetic Resonance Imaging*, vol. 28, no. 5, pp. 1266–1272, 2008.

[7] I. Tkáč, G. Öz, G. Adriany, K. Uğurbil, and R. Gruetter, "In Vivo 1H NMR Spectroscopy of the Human Brain at High Magnetic Fields: Metabolite Quantification at 4T vs. 7T," *Magnetic Resonance in Medicine*, vol. 62, no. 4, pp. 868–879, 2009.

[8] P. Vedrine, G. Aubert, F. Beaudet, J. Belorgey, C. Berriaud, P. Bredy, A. Donati, O. Dubois, G. Gilgrass, F. Juster *et al.*, "Iseult/INUMAC Whole Body 11.7 T MRI Magnet Status," *IEEE Transactions on Applied Superconductivity*, vol. 20, no. 3, pp. 696–701, 2010.

[9] C. M. Deniz, G. Carluccio, and C. Collins, "Parallel Transmission RF pulse Design With Strict Temperature Constraints," *NMR in Biomedicine*, vol. 30, no. 5, p. 3694, 2017.

[10] N. Boulant, X. Wu, G. Adriany, S. Schmitter, K. Uğurbil, and P.-F. Van de Moortele, "Direct Control of the Temperature Rise in Parallel Transmission by Means of Temperature Virtual Observation Points: Simulations at 10.5 Tesla," *Magnetic Resonance in Medicine*, vol. 75, no. 1, pp. 249–256, 2016.

[11] W. Grissom, C. Y. Yip, Z. Zhang, V. A. Stenger, J. A. Fessler, and D. C. Noll, "Spatial Domain Method for The Design of RF Pulses in Multi-coil Parallel Excitation," *Magnetic Resonance in Medicine*, vol. 56, no. 3, pp. 620–629, 2006.

[12] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger *et al.*, "SENSE: Sensitivity Encoding for Fast MRI," *Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–962, 1999.

[13] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA)," *Magnetic Resonance in Medicine*, vol. 47, no. 6, pp. 1202–1210, 2002.

[14] Z. Cao, X. Yan, and W. A. Grissom, "Parallel Transmit Excitation Pulses for Shuttered Echo Planar Imaging," in *Proceedings 27th Scientific Meeting, International Society for Magnetic Resonance in Medicine, Paris*, May 2018, p. 16.

[15] E. N. Yeh, C. A. McKenzie, M. A. Ohliger, and D. K. Sodickson, "Parallel Magnetic Resonance Imaging With Adaptive Radius in k-Space (PARS): Constrained Image Reconstruction using k-Space Locality in Radiofrequency Coil Encoded Data," *Magnetic Resonance in Medicine*, vol. 53, no. 6, pp. 1383–1392, 2005.

[16] R. G. Pinkerton, G. C. McKinnon, and R. S. Menon, "SENSE Optimization of A Transceive Surface Coil Array for MRI at 4 T," *Magnetic Resonance in Medicine*, vol. 56, no. 3, pp. 630–636, Sep 2006.

[17] J. Filos, J. Hajnal, and D. Larkman, *Parallel Transmission MRI.* Imperial College, London, 2008.

[18] Y. Zhu, "Parallel Excitation With An Array of Transmit Coils," *Magnetic Resonance in Medicine*, vol. 51, no. 4, pp. 775–784, 2004.

[19] K. Setsompop, L. L. Wald, V. Alagappan, B. Gagoski, F. Hebrank, U. Fontius, F. Schmitt, and E. Adalsteinsson, "Parallel RF Transmission With Eight Channels at 3 Tesla," *Magnetic Resonance in Medicine*, vol. 56, no. 5, pp. 1163–1171, 2006.

[20] U. Katscher, P. Börnert, C. Leussler, and J. S. Van Den Brink, "Transmit SENSE," *Magnetic Resonance in Medicine*, vol. 49, no. 1, pp. 144–150, 2003.

[21] K. Setsompop, L. Wald, V. Alagappan, B. Gagoski, and E. Adalsteinsson, "Magnitude Least Squares Optimization for Parallel Radio Frequency Excitation Design Demonstrated at 7 Tesla with Eight Channels," *Magnetic Resonance in Medicine*, vol. 59, no. 4, pp. 908–915, 2008.

[22] K. Setsompop, V. Alagappan, B. Gagoski, T. Witzel, J. Polimeni, A. Potthast, F. Hebrank, U. Fontius, F. Schmitt, L. L. Wald *et al.*, "Slice-selective RF Pulses for In-vivo B Inhomogeneity Mitigation at 7 Tesla Using Parallel RF Excitation with a 16-element Coil," *Magnetic Resonance in Medicine*, vol. 60, no. 6, pp. 1422–1432, 2008.

[23] M. Cloos, N. Boulant, M. Luong, G. Ferrand, E. Giacomini, D. Le Bihan, and A. Amadon, "kT-points: Short Three-dimensional Tailored RF Pulses for Flip-angle Homogenization over an Extended Volume," *Magnetic Resonance in Medicine*, vol. 67, no. 1, pp. 72–80, 2012.

[24] S. J. Malik, F. Padormo, A. N. Price, and J. V. Hajnal, "Spatially Resolved Extended Phase Graphs: Modeling and Design of Multipulse Sequences with Parallel Transmission," *Magnetic Resonance in Medicine*, vol. 68, no. 5, pp. 1481–1494, 2012.

[25] S. J. Malik, A. Beqiri, F. Padormo, and J. V. Hajnal, "Direct Signal Control of the Steady-state Response of 3D-FSE Sequences," *Magnetic Resonance in Medicine*, vol. 73, no. 3, pp. 951–963, 2015.

[26] I. Y. Chun, S. Noh, D. J. Love, T. M. Talavage, S. Beckley, and S. J. Kisner, "Mean Squared Error Based Excitation Pattern Design for Parallel Transmit and Receive SENSE MRI Image Reconstruction," *IEEE Transactions on Computational Imaging*, vol. 2, no. 4, pp. 424–439, Dec 2016.

[27] D. I. Hoult, "Sensitivity and Power Deposition in a High-Field Imaging Experiment," *Journal of Magnetic Resonance Imaging*, vol. 12, no. 1, pp. 46–67, 2000.

[28] C. A. Van den Berg, B. Van den Bergen, J. B. Van de Kamer, B. W. Raaymakers, H. Kroeze, L. W. Bartels, and J. J. Lagendijk, "Simultaneous B Homogenization and Specific Absorption Rate Hotspot Suppression Using A Magnetic Resonance Phased Array Transmit Coil," *Magnetic Resonance in Medicine*, vol. 57, no. 3, pp. 577–586, 2007.

[29] J. Vaughan, G. Adriany, C. Snyder, J. Tian, T. Thiel, L. Bolinger, H. Liu, L. DelaBarre, and K. Ugurbil, "Efficient High-Frequency Body Coil for High-Field MRI," *Magnetic Resonance in Medicine*, vol. 52, no. 4, pp. 851–859, 2004.

[30] P. B. Roemer, W. A. Edelstein, C. E. Hayes, S. P. Souza, and O. M. Mueller, "The NMR Phased Array," *Magnetic Resonance in Medicine*, vol. 16, no. 2, pp. 192–225, 1990.

[31] D. K. Sodickson and W. J. Manning, "Simultaneous Acquisition of Spatial Harmonics (SMASH): Fast Imaging With Radiofrequency Coil Arrays," *Magnetic Resonance in Medicine*, vol. 38, no. 4, pp. 591–603, 1997.

[32] R. G. Pinkerton, E. A. Barberi, and R. S. Menon, "Transceive Surface Coil Array for Magnetic Resonance Imaging of The Human Brain at 4 T," *Magnetic Resonance in Medicine*, vol. 54, no. 2, pp. 499–503, Aug 2005.

[33] G. Adriany, V. de Moortele, F. Wiesinger, S. Moeller, J. P. Strupp, P. Andersen, C. Snyder, X. Zhang, W. Chen, K. P. Pruessmann *et al.*, "Transmit and Receive Transmission Line Arrays for 7 Tesla Parallel Imaging," *Magnetic Resonance in Medicine*, vol. 53, no. 2, pp. 434–445, 2005.

[34] N. I. Avdievich and H. P. Hetherington, "4 T Actively Detunable Transmit/Receive Transverse Electromagnetic Coil and 4-Channel Receive-Only Phased Array for 1H Human Brain Studies," *Magnetic Resonance in Medicine*, vol. 52, no. 6, pp. 1459–1464, Dec 2004.

[35] Y. Zhu, L. Alon, C. M. Deniz, R. Brown, and D. K. Sodickson, "System and SAR Characterization in Parallel RF Transmission," *Magnetic Resonance in Medicine*, vol. 67, no. 5, pp. 1367–1378, 2012.

[36] H. Homann, *SAR Prediction and SAR Management for Parallel Transmit MRI.* KIT Scientific Publishing, 2012, vol. 16.

[37] H. Homann, I. Graesslin, H. Eggers, K. Nehrke, P. Vernickel, U. Katscher, O. Dössel, and P. Börnert, "Local SAR Management by RF Shimming: A Simulation Study With Multiple Human Body Models," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 25, no. 3, pp. 193–204, 2012.

[38] I. Graesslin, H. Homann, S. Biederer, P. Börnert, K. Nehrke, P. Vernickel, G. Mens, P. Harvey, and U. Katscher, "A Specific Absorption Rate Prediction Concept for Parallel Transmission MR," *Magnetic Resonance in Medicine*, vol. 68, no. 5, pp. 1664–1674, 2012.

[39] International Electrotechnical Commission, "International Standard, Medical Equipment IEC 60601-2-33: Particular Requirements for the Safety of Magnetic Resonance Equipment," 2010.

[40] International Commission on Non-Ionizing Radiation (ICNIRP), "Guidelines for Limiting Exposure to Time-Varying Electric, Magnetic, and Electromagnetic Fields (Up to 300 GHz)," *Helath Physics*, vol. 74, no. 4, pp. 494–522, 1998.

[41] IEEE Standards Coordinating Committee, *IEEE Standard for Safety Levels With Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3kHz to 300 GHz.* IEEE Standard C95.3-2002, 2008.

[42] B. Guérin, M. Gebhardt, S. Cauley, E. Adalsteinsson, and L. L. Wald, "Local SAR, Global SAR, Transmitter Power and Excitation Accuracy Trade-offs in Low Flip-angle Parallel Transmit Pulse Design," *Magnetic Resonance in Medicine*, vol. 71, no. 4, p. 1446, 2014.

[43] M. S. Vinding, B. Guérin, T. Vosegaard, and N. C. Nielsen, "Local SAR, Global SAR, and Power-constrained Large-flip-angle Pulses with Optimal Control and Virtual Observation Points," *Magnetic Resonance in Medicine*, vol. 77, no. 1, pp. 374–384, 2017.

[44] F. Zhao, "Methods for MRI RF Pulse Design and Image Reconstruction," Ph.D. dissertation, The University of Michigan, 2014.

[45] E. M. Haacke, R. W. Brown, M. R. Thompson, R. Venkatesan *et al.*, *Magnetic Resonance Imaging: Physical Principles and Sequence Design.* Wiley-Liss New York:, 1999, vol. 82.

[46] F. Bloch, "Nuclear Induction," *Physical Review*, vol. 70, no. 7-8, p. 460, 1946.

[47] J. Pauly, D. Nishimura, and A. Macovski, "A Linear Class of Large-Tip-Angle Selective Excitation Pulses," *Journal of Magnetic Resonance*, vol. 82, no. 3, pp. 571–587, 1989.

[48] C. Y. Yip, J. A. Fessler, and D. C. Noll, "Iterative RF Pulse Design for Multidimensional, Small-tip-angle Selective Excitation," *Magnetic Resonance in Medicine*, vol. 54, no. 4, pp. 908–917, 2005.

[49] J. A. Fessler and B. P. Sutton, "Nonuniform Fast Fourier Transforms using Min-max Interpolation," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 560–574, 2003.

[50] B. P. Sutton, D. C. Noll, and J. A. Fessler, "Fast, Iterative Image Reconstruction for MRI in the Presence of Field Inhomogeneities," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 178–188, 2003.

[51] D. C. Noll, C. H. Meyer, J. M. Pauly, D. G. Nishimura, and A. Macovski, "A Homogeneity Correction Method for Magnetic Resonance Imaging with Time-varying Gradients," *IEEE Transactions on Medical Imaging*, vol. 10, no. 4, pp. 629–637, 1991.

[52] J. Pauly, D. Nishimura, and A. Macovski, "A k-Space Analysis of Small-tip-angle Excitation," *Journal of Magnetic Resonance*, vol. 213, no. 2, pp. 544–557, 2011.

[53] D. Xu, K. F. King, Y. Zhu, G. C. McKinnon, and Z.-P. Liang, "Designing Multichannel, Multidimensional, Arbitrary Flip Angle RF pulses using an Optimal Control Approach," *Magnetic Resonance in Medicine*, vol. 59, no. 3, pp. 547–560, 2008.

[54] W. A. Grissom, D. Xu, A. B. Kerr, J. A. Fessler, and D. C. Noll, "Fast Large-tip-angle Multidimensional and Parallel RF Pulse Design in MRI," *IEEE Transactions on Medical Imaging*, vol. 28, no. 10, pp. 1548–1559, 2009.

[55] C. Y. Yip, W. A. Grissom, J. A. Fessler, and D. C. Noll, "Joint Design of Trajectory and RF Pulses for Parallel Excitation," *Magnetic Resonance in Medicine*, vol. 58, no. 3, pp. 598–604, 2007.

[56] A. C. Zelinski, L. L. Wald, K. Setsompop, V. K. Goyal, and E. Adalsteinsson, "Sparsity-enforced Slice-selective MRI RF Excitation Pulse Design," *IEEE Transactions on Medical Imaging*, vol. 27, no. 9, pp. 1213–1229, 2008.

[57] C. Ma, D. Xu, K. F. King, and Z.-P. Liang, "Joint Design of Spoke Trajectories and RF Pulses for Parallel Excitation," *Magnetic Resonance in Medicine*, vol. 65, no. 4, pp. 973–985, 2011.

[58] D. Yoon, J. A. Fessler, A. C. Gilbert, and D. C. Noll, "Fast Joint Design Method for Parallel Excitation Radiofrequency Pulse and Gradient Waveforms Considering Off-resonance," *Magnetic Resonance in Medicine*, vol. 68, no. 1, pp. 278–285, 2012.

[59] S. Saekho, F. E. Boada, D. C. Noll, and V. A. Stenger, "Small Tip Angle Three-dimensional Tailored Radiofrequency Slab-select Pulse for Reduced B1 Inhomogeneity at 3 T," *Magnetic Resonance in Medicine*, vol. 53, no. 2, pp. 479–484, 2005.

[60] S. Saekho, C. Y. Yip, D. C. Noll, F. E. Boada, and V. A. Stenger, "Fast-kz Three-dimensional Tailored Radiofrequency Pulse for Reduced B1 Inhomogeneity," *Magnetic Resonance in Medicine*, vol. 55, no. 4, pp. 719–724, 2006.

[61] C. M. Collins, Z. Wang, W. Mao, J. Fang, W. Liu, and M. B. Smith, "Array-optimized Composite Pulse for Excellent Whole-brain Homogeneity in High-field MRI," *Magnetic Resonance in Medicine*, vol. 57, no. 3, pp. 470–474, 2007.

[62] C. Yang, W. Deng, V. Alagappan, L. L. Wald, and V. A. Stenger, "Four-dimensional Spectral-spatial RF Pulses for Simultaneous Correction of B1+ Inhomogeneity and Susceptibility Artifacts in T2*-weighted MRI," *Magnetic Resonance in Medicine*, vol. 64, no. 1, pp. 1–8, 2010.

[63] W. Deng, C. Yang, V. Alagappan, L. L. Wald, F. E. Boada, and V. A. Stenger, "Simultaneous Z-shim Method for Reducing Susceptibility Artifacts with Multiple Transmitters," *Magnetic Resonance in Medicine*, vol. 61, no. 2, pp. 255–259, 2009.

[64] J. R. Burt, S. L. Zimmerman, I. R. Kamel, M. Halushka, and D. A. Bluemke, "Myocardial T1 Mapping: Techniques and Potential Applications," *Radiographics*, vol. 34, no. 2, pp. 377–395, 2014.

[65] S. Giri, Y. C. Chung, A. Merchant, G. Mihai, S. Rajagopalan, S. V. Raman, and O. P. Simonetti, "T2 Quantification for Improved Detection of Myocardial Edema," *Journal of Cardiovascular Magnetic Resonance*, vol. 11, no. 1, p. 56, 2009.

[66] S. A. Hamlin, T. S. Henry, B. P. Little, S. Lerakis, and A. E. Stillman, "Mapping the Future of Cardiac MR Imaging: Case-based Review of T1 and T2 Mapping Techniques," *Radiographics*, vol. 34, no. 6, pp. 1594–1611, 2014.

[67] F. Bardati, A. Borrani, A. Gerardino, and G. Lovisolo, "SAR Optimization in a Phased Array Radiofrequency Hyperthermia System," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 12, pp. 1201–1207, 1995.

[68] G. Eichfelder and M. Gebhardt, "Local Specific Absorption Rate Control for Parallel Transmission by Virtual Observation Points," *Magnetic Resonance in Medicine*, vol. 66, no. 5, pp. 1468–1476, 2011.

[69] J. Lee, M. Gebhardt, L. L. Wald, and E. Adalsteinsson, "Local SAR in Parallel Transmission Pulse Design," *Magnetic Resonance in Medicine*, vol. 67, no. 6, pp. 1566–1578, 2012.

[70] X. Mao, D. Love, J. Rispoli, and T. Talavage, "Multiple-Input Multiple-Output (MIMO) MRI: An Efficient Pulse Design Algorithm to Combine Parallel Excitation and Parallel Imaging," in *Proceedings of 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2017, pp. 909–913.

[71] X. Mao, N. Vike, T. Talavage, J. Rispoli, and D. Love, "Multiple-Input Multiple-Output (MIMO) MRI: Combining Parallel Excitation and Parallel Reception for Enhanced Imaging ," *IEEE Transactions on Computational Imaging*, 2019, doi:10.1109/TCI.2019.2904882.

[72] B. Park, A. G. Webb, and C. M. Collins, "A Method to Separate Conservative and Magnetically-induced Electric Fields in Calculations for MRI and MRS in Electrically-small Samples," *Journal of Magnetic Resonance*, vol. 199, no. 2, pp. 233–237, 2009.

[73] C. M. Collins and Z. Wang, "Calculation of Radiofrequency Electromagnetic Fields and Their Effects in MRI of Human Subjects," *Magnetic Resonance in Medicine*, vol. 65, no. 5, pp. 1470–1482, 2011.

[74] P. A. Bottomley, R. W. Redington, W. A. Edelstein, and J. F. Schenck, "Estimating Radiofrequency Power Deposition in Body NMR Imaging," *Magnetic Resonance in Medicine*, vol. 2, no. 4, pp. 336–349, 1985.

[75] C. Gabriel, S. Gabriel, and y. E. Corthout, "The Dielectric Properties of Biological Tissues: I. Literature Survey," *Physics in Medicine & Biology*, vol. 41, no. 11, p. 2231, 1996.

[76] C. E. Hayes, W. A. Edelstein, J. F. Schenck, O. M. Mueller, and M. Eash, "An Efficient, Highly Homogeneous Radiofrequency Coil for Whole-body NMR Imaging at 1.5 T," *Journal of Magnetic Resonance*, vol. 63, no. 3, pp. 622–628, 1985.

[77] F. G. Shellock, *Magnetic Resonance Procedures: Health Effects and Safety.* Crc Press, 2000.

[78] National Electrical Manufacturers Association, "Characterization of the Specific Absorption raRe for Magnetic Resonance Imaging Systems," *Standards Publication/No. MS*, vol. 8, 1993.

[79] P. T. Hardy and K. M. Weil, "A Review of Thermal MR Injuries," *Radiologic Technology*, vol. 81, no. 6, pp. 606–609, 2010.

[80] F. G. Shellock, "Radiofrequency Energy-induced Heating during MR Procedures: A Review," *Journal of Magnetic Resonance Imaging*, vol. 12, no. 1, pp. 30–36, 2000.

[81] A. Deshmane, V. Gulani, M. A. Griswold, and N. Seiberlich, "Parallel MR Imaging," *Journal of Magnetic Resonance Imaging*, vol. 36, no. 1, pp. 55–72, 2012.

[82] P. Ullmann, S. Junge, M. Wick, F. Seifert, W. Ruhm, and J. Hennig, "Experimental Analysis of Parallel Excitation Using Dedicated Coil Setups and Simultaneous RF Transmission on Multiple Channels," *Magnetic Resonance in Medicine*, vol. 54, no. 4, pp. 994–1001, 2005.

[83] R. Lattanzi, D. K. Sodickson, A. K. Grant, and Y. Zhu, "Electrodynamic Constraints on Homogeneity and Radiofrequency Power Deposition in Multiple Coil Excitations," *Magnetic Resonance in Medicine*, vol. 61, no. 2, pp. 315–334, Feb 2009.

[84] K. Kurpad, E. Boskamp, and S. Wright, "A Parallel Transmit Volume Coil With Independent Control of Currents on The Array Elements," in *Proceedings of the 13th Annual Meeting of ISMRM, Miami Beach, USA*, May 2005, p. 16.

[85] A. C. Zelinski, L. M. Angelone, V. K. Goyal, G. Bonmassar, E. Adalsteinsson, and L. L. Wald, "Specific Absorption Rate Studies of The Parallel Transmission of Inner-volume Excitations at 7T," *Journal of Magnetic Resonance Imaging*, vol. 28, no. 4, pp. 1005–1018, 2008.

[86] D. C. Noll, W. A. Grissom, and C. Y. Yip, "RF Pulse Design for Parallel Excitation," in *Proceedings of the 14th Annual Meeting of ISMRM, Seattle, USA*, 2006.

[87] H. She, R. R. Chen, D. Liang, Y. Chang, and L. Ying, "Image Reconstruction from Phased-array MRI Data Based on Multichannel Blind Deconvolution," in *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2010)*. IEEE, 2010, pp. 760–763.

[88] D. Hoult, "The Principle of Reciprocity in Signal Strength Calculations - A Mathematical Guide," *Concepts in Magnetic Resonance*, vol. 12, no. 4, pp. 173–187, Jan 2000.

[89] D. I. Hoult and R. Richards, "The Signal-to-Noise Ratio of the Nuclear Magnetic Resonance Experiment," *Journal of Magnetic Resonance*, vol. 24, no. 1, pp. 71–85, Oct 1976.

[90] M. V. Vaidya, C. M. Collins, D. K. Sodickson, R. Brown, G. C. Wiggins, and R. Lattanzi, "Dependence of B1- and B1+ Field Patterns of Surface Coils on the Electrical Properties of the Sample and the MR Operating Frequency," *Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering*, Feb 2016.

[91] J. A. Fessler, "Model-based Image Reconstruction for MRI," *IEEE Signal Processing Magazine*, vol. 27, no. 4, pp. 81–89, 2010.

[92] S. M. Wright and L. L. Wald, "Theory and Application of Array Coils in MR Spectroscopy," *NMR in Biomedicine*, vol. 10, no. 8, pp. 394–410, Dec 1997.

[93] R. F. Harrington and J. L. Harrington, *Field Computation by Moment Methods*. Oxford University Press, 1996.

[94] M. Lustig, S. J. Kim, and J. M. Pauly, "A Fast Method for Designing Time-optimal Gradient Waveforms for Arbitrary-space Trajectories," *IEEE Transactions on Medical Imaging*, vol. 27, no. 6, pp. 866–873, Jun 2008.

[95] Z. Cao, M. J. Donahue, J. Ma, and W. A. Grissom, "Joint Design of Large-Tip-Angle Parallel RF Pulses and Blipped Gradient Trajectories," *Magnetic resonance in medicine*, vol. 75, no. 3, pp. 1198–1208, 2016.

[96] B. M. Hochwald, D. J. Love, S. Yan, P. Fay, and J. M. Jin, "Incorporating Specific Absorption Rate Constraints Into Wireless Signal Design," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 126–133, 2014.

[97] B. Van den Bergen, C. A. T. Van den Berg, L. W. Bartels, and J. J. Lagendijk, "7 T Body MRI: B1 Shimming With Simultaneous SAR Reduction," *Physics in Medicine and Biology*, vol. 52, no. 17, p. 5429, 2007.

[98] C. M. Deniz, L. Alon, R. Brown, D. K. Sodickson, and Y. Zhu, "Specific Absorption Rate Benefits of Including Measured Electric Field Interactions in Parallel Excitation Pulse Design," *Magnetic Resonance in Medicine*, vol. 67, no. 1, pp. 164–174, 2012.

[99] H. Homann, I. Graesslin, K. Nehrke, C. Findeklee, O. Dössel, and P. Börnert, "Specific Absorption Rate Reduction in Parallel Transmission by K-space Adaptive Radiofrequency Pulse Design," *Magnetic Resonance in Medicine*, vol. 65, no. 2, pp. 350–357, 2011.

[100] D. O. Brunner and K. P. Pruessmann, "Optimal Design of Multiple-Channel RF Pulses Under Strict Power and SAR Constraints," *Magnetic Resonance in Medicine*, vol. 63, no. 5, pp. 1280–1291, 2010.

[101] A. Sbrizzi, H. Hoogduin, J. J. Lagendijk, P. Luijten, G. L. Sleijpen, and C. A. Van Den Berg, "Fast Design of Local N-Gram-Specific Absorption Rate–Optimized Radiofrequency Pulses for Parallel Transmit Systems," *Magnetic Resonance in Medicine*, vol. 67, no. 3, pp. 824–834, 2012.

[102] M. A. Cloos, M. Luong, G. Ferrand, A. Amadon, D. Le Bihan, and N. Boulant, "Local SAR Reduction in Parallel Excitation Based on Channel-dependent Tikhonov Parameters," *Journal of Magnetic Resonance Imaging*, vol. 32, no. 5, pp. 1209–1216, 2010.

[103] I. Graesslin, C. Steiding, B. Annighoefer, J. Weller, S. Biederer, D. Brunner, H. Homann, F. Schweser, U. Katscher, K. Pruessmann, and Others, "Local SAR Constrained Hotspot Reduction by Temporal Averaging," in *Proceedings of the 18th Scientific Meeting of ISMRM, Stockholm, Sweden*, 2010, p. 4932.

[104] A. Martin, E. Schiavi, Y. Eryaman, J. L. Herraiz, B. Gagoski, E. Adalsteinsson, L. L. Wald, and B. Guerin, "Parallel Transmission Pulse Design With Explicit Control for the Specific Absorption Rate in the Presence of Radiofrequency Errors," *Magnetic Resonance in Medicine*, vol. 75, no. 6, pp. 2493–2504, 2016.

[105] M. Pendse, R. Stara, M. Mehdi Khalighi, and B. Rutt, "IMPULSE: A Scalable Algorithm for Design of Minimum Specific Absorption Rate Parallel Transmit RF Pulses," *Magnetic Resonance in Medicine*, vol. 81, no. 4, pp. 2808–2822, 2019.

[106] I. Graesslin, S. Biederer, K. Falaggis, P. Vernickel, H. Dingemans, G. Mens, P. Roeschmann, C. Leussler, Z. Zhai, M. Morich *et al.*, "Real-time SAR Monitoring to Ensure Patient Safety for Parallel Transmission Systems," in *Proceedings of the 15th Annual Meeting of ISMRM, Berlin, Germany*, 2007, p. 1086.

[107] B. Gagoski, R. Gumbrecht, M. Hamm, K. Setsompop, B. Keil, J. Lee, K. Makhoul, A. Mareyam, K. Fujimoto, T. Witzel *et al.*, "Real-time RF Monitoring in A 7 T Parallel Transmit System," in *Proceedings of the 18th Annual Meeting of ISMRM, Stockholm, Sweden*, 2010, p. 781.

[108] I. Brote, S. Orzada, O. Kraff, S. Maderwald, H. Quick, P. Yazdanbakhsh, K. Solbach, K. Wicklow, A. Bahr, T. Bolz *et al.*, "A Multi-channel SAR Prediction and Online Monitoring System At 7T," in *Proceedings of the 17th Annual Meeting of ISMRM, Honolulu, Hawaii, USA*, 2009, p. 4788.

[109] M. Cloos, N. Boulant, M. Luong, G. Ferrand, D. Le Bihan, and A. Amadon, "Specific Absorption Rate Monitor for In-vivo Parallel Transmission At 7 Tesla," in *Proceedings of the 18th Annual Meeting of ISMRM, Stockholm, Sweden*, 2010, p. 3871.

[110] R. Gumbrecht, U. Fontius, H. Adolf, T. Benner, F. Schmitt, E. Adalsteinsson, L. Wald, and H. Fautz, "Online Local SAR Supervision for Transmit Arrays At 7T," in *Proceedings of the 21th Annual Meeting of ISMRM, Salt Lake City, USA*, 2013, p. 4420.

[111] X. Mao, J. Rispoli, and D. Love, "A k-Means Clustering Algorithm for MRI Virtual Observation Points Compression in Local SAR Supervision," in *Proceedings of the 27th Annual Meeting of ISMRM, Montreal, Canada*, 2019, p. 7026.

[112] D. Ying, D. J. Love, and B. M. Hochwald, "Closed-loop Precoding and Capacity Analysis for Multiple-Antenna Wireless Systems With User Radiation Exposure Constraints," *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5859–5870, 2015.

[113] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An Efficient k-Means Clustering Algorithm: Analysis and Implementation," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 7, pp. 881–892, 2002.

[114] S. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.

[115] D. Arthur and S. Vassilvitskii, "K-means++: The Advantages of Careful Seeding," in *Proceedings of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035.

[116] N. I. Avdievich, "Transceiver-phased Arrays for Human Brain Studies at 7 T," *Applied Magnetic Resonance*, vol. 41, no. 2-4, pp. 483–506, 2011.

[117] A. Zelinski, V. Goyal, L. Angelone, G. Bonmassar, L. L. Wald, and E. Adalsteinsson, "Designing RF Pulses With Optimal Specific Absorption Rate (SAR) Characteristics and Exploring Excitation Fidelity, SAR and Pulse Duration Tradeoffs," in *Proceedings of the 15th Annual Meeting of ISMRM, Berlin, Germany*, 2007, p. 1699.

[118] M. L. Fisher, "The Lagrangian Relaxation Method for Solving Integer Programming Problems," *Management Science*, vol. 27, no. 1, pp. 1–18, 1981.

[119] P. F. Van de Moortele, C. Akgun, G. Adriany, S. Moeller, J. Ritter, C. M. Collins, M. B. Smith, J. T. Vaughan, and K. Uğurbil, "B1 Destructive Interferences and Spatial Phase Patterns at 7 T with a Head Transceiver Array Coil," *Magnetic Resonance in Medicine*, vol. 54, no. 6, pp. 1503–1518, 2005.

[120] W. A. Grissom, L. Sacolick, and M. W. Vogel, "Improving High-field MRI using Parallel Excitation," *Imaging in Medicine*, vol. 2, no. 6, pp. 675–693, 2010.

[121] V. Gras, A. Vignaud, A. Amadon, D. Le Bihan, and N. Boulant, "Universal Pulses: A New Concept for Calibration-free Parallel Transmission," *Magnetic Resonance in Medicine*, vol. 77, no. 2, pp. 635–643, 2017.

[122] C. Mirfin, P. Glover, and R. Bowtell, "Optimisation of parallel transmission radiofrequency pulses using neural networks," in *Proceedings of the 26th Annual Meeting of ISMRM, Paris, France*, 2018, p. 3388.

[123] J. D. Ianni, Z. Cao, and W. A. Grissom, "Machine Learning RF Shimming: Prediction by Iteratively Projected Ridge Regression," *Magnetic Resonance in Medicine*, vol. 80, no. 5, pp. 1871–1881, 2018.

[124] A. Christ, W. Kainz, E. G. Hahn, K. Honegger, M. Zefferer, E. Neufeld, W. Rascher, R. Janka, W. Bautz, J. Chen *et al.*, "The Virtual Family Development of Surface-based Anatomical Models of Two Adults and Two Children for Dosimetric Simulations," *Physics in Medicine and Biology*, vol. 55, no. 2, p. N23, 2009.

[125] T. Liu, I. Khalidov, L. de Rochefort, P. Spincemaille, J. Liu, A. J. Tsiouris, and Y. Wang, "A Novel Background Field Removal Method for MRI Using Projection Onto Dipole Fields," *NMR in Biomedicine*, vol. 24, no. 9, pp. 1129–1136, 2011.

[126] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Prentice hall Upper Saddle River, NJ, 2000, vol. 1.