University of Alberta

Magnet Design and Optimization for an Integrated Linac-MRI System

by

Tony Tadic

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Medical Physics

Department of Physics

© Tony Tadic

Spring 2012
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission

ABSTRACT

A promising approach to the development of an integrated system for magnetic resonance imaging-guided radiotherapy involves the rigid coupling of a medical linear accelerator to a rotating biplanar magnet assembly. Through near real-time volumetric imaging during irradiation, this hybrid system will enable dynamic target localization and online treatment plan optimization. The work presented in this thesis was concerned with the design of non-axisymmetric yoked biplanar magnets that are suitable for use with this integrated system. Namely, the goal of this research was to develop and implement robust optimization algorithms for the design of novel permanent and superconducting magnet assemblies that are compact in size and provide an unobstructed beam path. A nonlinear iterative shape optimization scheme based on the finite element method was developed for the calculation of axisymmetric and non-axisymmetric pole piece surface contours that minimize the magnetic field inhomogeneity in a designated imaging volume. This method was applied to the theoretical design of a full-body 0.2 T biplanar permanent magnet system, for which the necessity of the optimized pole piece designs was demonstrated. The algorithm performance was evaluated and the sensitivity of the optimized designs was investigated. An iterative optimization scheme based on the finite element method was also developed for the calculation of minimum volume coil arrangements for superconducting magnets with magnetic materials. In particular, this method is well suited for the design of conduction-cooled cryogen-free magnet systems employing high-temperature superconducting coil configurations. The effectiveness of the proposed method was demonstrated with a uniform approach to the optimal design of several full-body 0.5 T superconducting biplanar magnet systems with bored yoke structures. These designs were subsequently evaluated with an emphasis on the relationship between the yoke geometries and magnet performance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

a_1, a_2, a_3, a_4	Finite element nodal integer labels
$a_{\ell m}, b_{\ell m}, c_{\ell m}$	Spherical harmonic expansion coefficients
a_j	Cross-sectional area of coil j
$\mathbf{A}$	Linear programming constraint matrix
A_x, A_y	Maximum dimensions of the imaging sample along the Cartesian directions
$\mathcal{A}$	Active constraint index set
$\boldsymbol{b}$	Linear programming constraint vector
b_{ci}	Axial component of the magnetic field at the target point with index i strictly due to free currents
b_{mi}	Axial component of the magnetic field at the target point with index i due to nonlinear magnetic materials
b_{ti}	Axial component of the target coil field at the target point with index i
b_{zi}	Total axial component of the magnetic field at the target point with index i
$\boldsymbol{B}$	Magnetic field or magnetic flux density vector
$\boldsymbol{B_0}$	Main static polarizing magnetic field
$\boldsymbol{B_1}$	Circularly polarized oscillating radio-frequency magnetic field
$\boldsymbol{B_{\mathrm{eff}}}$	Effective magnetic field during excitation
$\boldsymbol{B_{\mathrm{L}}}$, $\boldsymbol{B_{\mathrm{R}}}$	Left and right circularly polarized oscillating radio-frequency magnetic field
$\boldsymbol{B_{\mathrm{RF}}}$	Linearly polarized oscillating radio-frequency magnetic field
$\boldsymbol{B}'$	General time and spatially varying magnetic field

B_{peak}	Peak conductor field
B_s, B_ϕ, B_z	Cylindrical components of the magnetic field
B_x, B_y, B_z	Cartesian components of the magnetic field
δB	Absolute magnetic field variation or inhomogeneity
ΔB	Relative magnetic field variation or inhomogeneity
c_i	Optimization constraint function
$\boldsymbol{d}$	Optimization search direction
$\boldsymbol{D}$	Electric displacement vector
e	Elementary charge
$\boldsymbol{E}$	Electric field vector
E	Energy
ΔE	Energy difference
$\mathcal{E}$	Electromotive force
$\mathcal{E}$	Equality constraint index set
$\mathfrak{E}$	Complete elliptic integral of the second kind
f	General scalar function
$\boldsymbol{F}$	Force vector
$\mathcal{F}$	Fourier transform operator
$\mathbf{G}$	Search direction matrix parameter
$\boldsymbol{G}$	Linear magnetic field gradient vector
G_x, G_y, G_z	Cartesian components of the linear magnetic field gradient vector
ΔG_y	Increment in the phase-encoding gradient
$\mathcal{G}_1, \mathcal{G}_2$	Intermediate functions in the calculation of the magnetic field due to an ideal circular coil

$\hbar$	Planck's constant
$\boldsymbol{H}$	Auxiliary magnetic field
$\boldsymbol{H_m}$	Irrotational auxiliary magnetic field
$\boldsymbol{H_s}$	Solenoidal auxiliary magnetic field
H_x, H_y, H_z	Cartesian components of the auxiliary magnetic field
$\mathbf{I}$	Identity matrix
$\boldsymbol{I}$	Current vector
$\mathcal{I}$	Inequality constraint index set
$\boldsymbol{J}$	Current density vector
$\boldsymbol{J_M}, \boldsymbol{K_M}$	Effective or bound volume and surface current distributions for a magnetized object
$\boldsymbol{k}$	Spatial frequency vector
k_x, k_y, k_z	Cartesian components of the spatial frequency vector
$\Delta k_x, \Delta k_y, \Delta k_z$	Separation of k-space samples along the Cartesian directions
k_B	Boltzmann's constant
$\mathfrak{K}$	Complete elliptic integral of the first kind
l	Magnet assembly dimension
$L_{\mathscr{k}}^{\mathscr{p}}$	Volume coordinate with respect to vertex $\mathscr{k}$ in the finite element subdomain with index $\mathscr{p}$
$L_{\mathscr{k}i}^{\mathscr{p}}$	Volume coordinate of the local node i with respect to vertex $\mathscr{k}$ in the finite element subdomain with index $\mathscr{p}$
$\mathcal{L}$	Lagrangian function
$\mathfrak{Q}$	Differential operator
m_Ω	Number of nodes in a finite element subdomain of order n_Ω
$\mathfrak{m}$	Azimuthal quantum number
$\boldsymbol{M}$	Magnetic polarization or magnetization vector

$\boldsymbol{M_r}$	Remanent magnetization or remanance
$\boldsymbol{M_{xy}}$	Net transverse magnetization
$\boldsymbol{M_z}$	Net longitudinal magnetization
M_0	Magnetization amplitude
$\mathcal{M}_1, \mathcal{M}_2$	Intermediate functions in the calculation of the magnetic field due to a bulk circular coil
n	Dimension of optimization problem
n_Ω	Order of a finite element mesh
n_x, n_y	Number of image pixels along the Cartesian directions
$\widehat{\boldsymbol{n}}$	Outward unit normal vector
N_s	Total number of spins
N_t	Number of optimization target points
N_Ω	Total number of finite element subdomains
$\mathcal{N}_1, \mathcal{N}_2$	Intermediate functions in the calculation of the magnetic field due to a bulk circular coil
$\mathcal{O}$	Order of magnitude error function
$\boldsymbol{p}$	Optimization design vector
$\boldsymbol{p}^*$	Optimal solution vector
δp	Finite-difference perturbation interval
$\boldsymbol{P}$	Electric polarization
P_ℓ^m	Associated Legendre function of order ℓ and degree m
P_+, P_-	Spin-up and spin-down eigenstate probabilities
P_{50}, P_{80}, P_{90}	Design sensitivity statistical parameters
ΔP	Probability difference
$\mathcal{P}_i^q$	Lagrange interpolating polynomial of order q for the node with local index i

$\mathfrak{P}$	Feasible domain or set
$\mathfrak{P}^*$	Neighbourhood of a solution point
q	Electric charge
$\mathbf{Q}$	Real symmetric square matrix
$\mathcal{Q}$	Quadratic model function
$\boldsymbol{r}$	Position vector
$\boldsymbol{r}_i^{\mathscr{p}}$	Position of local node i in a curved finite element with index $\mathscr{p}$
$\mathfrak{r}$	Partial differential equation residual
R	Radius
$\mathfrak{R}$	Total finite element residual vector
$\mathfrak{R}^{\mathscr{p}}$	Finite element residual vector corresponding to the subdomain with index $\mathscr{p}$
$\overline{\mathfrak{R}}^{\mathscr{p}}$	Augmented finite element residual vector corresponding to the subdomain with index $\mathscr{p}$
$\mathfrak{R}_i^{\mathscr{p}}$	Finite element weighted residual corresponding to the weighting function $w_i^{\mathscr{p}}$
$\boldsymbol{\mathcal{R}}$	Vector rotation operator
s	Radial coordinate for cylindrical coordinates
s_j	Radius for coil j
$\mathscr{s}$	Measured electronic NMR signal
$\mathfrak{s}$	Spin quantum number
$\boldsymbol{S}$	Spin angular momentum vector
$\mathcal{S}$	General surface of integration
$\mathfrak{S}_i^{\mathscr{p}}$	Finite element basis function corresponding to node i in the subdomain with index $\mathscr{p}$
t	Time

$\boldsymbol{t}$	Linear programming auxiliary variable
T	Temperature
T_1	Longitudinal relaxation time constant
T_2	Transverse relaxation time constant
T_2'	Transverse relaxation time constant due to static magnetic field inhomogeneities
T_2^*	Total effective transverse relaxation time constant
TE	Echo time
TR	Repetition time
$\mathfrak{T}_i^{\mathscr{p}}$	Geometric shape transformation function corresponding to node i in the curved finite element subdomain with index $\mathscr{p}$
U	Potential energy
$\mathcal{U}_\Omega$	Finite element subdomain connectivity array
$\mathcal{U}_{\partial\Omega}$	Finite element boundary connectivity array
$\mathfrak{U}$	Uniform probability distribution
$\boldsymbol{v}$	Velocity
$V^{\mathscr{p}}$	Volume of the finite element subdomain with index $\mathscr{p}$
$V_{\mathscr{k}}^{\mathscr{p}}$	Volume of the internal tetrahedron formed by replacing vertex $\mathscr{k}$ with a given point lying within the finite element subdomain with index $\mathscr{p}$
$\mathcal{V}$	Volume
$\mathcal{V}_c$	Conductor volume
$w_i^{\mathscr{p}}$	Finite element weighting function associated with local node i within the subdomain with index $\mathscr{p}$
$\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}$	Cartesian unit vectors in the standard reference frame
$\hat{\boldsymbol{x}}_{\boldsymbol{\rho}}, \hat{\boldsymbol{y}}_{\boldsymbol{\rho}}, \hat{\boldsymbol{z}}_{\boldsymbol{\rho}}$	Cartesian unit vectors in the rotating reference frame
x, y, z	Cartesian coordinates in the standard reference frame

x_i, y_i, z_i	Cartesian coordinates of a point or finite element node with local index i
x_ρ, y_ρ, z_ρ	Cartesian coordinates in the rotating reference frame
$x_{5\mathrm{G}}, y_{5\mathrm{G}}, z_{5\mathrm{G}}$	Maximal extent of the 5 G magnetic field contours
$\tilde{x}$	Distorted Cartesian coordinate
Δx, Δy	Width of image pixels along the Cartesian directions
y_ℓ^m	Real spherical harmonic function of order ℓ and degree m
Y_ℓ^m	Ordinary spherical harmonic function of order ℓ and degree m
α	Optimization step size
$\|\alpha\rangle, \|\beta\rangle$	Spin-up and spin-down eigenstates
β	Search direction scalar parameter
χ_e, χ_m	Electric and magnetic susceptibilities
$\boldsymbol{\delta}$	Perturbation vector
δ_{ij}	Kronecker delta function
∇	Del operator
ϵ	Electric permittivity
ϵ_0	Electric permittivity of a vacuum
ϵ_r	Relative electric permittivity
ϕ	Azimuthal angle or phase
$\Delta\phi_{k_x}, \Delta\phi_{k_y}$	Differences in accumulated phase between successive frequency and phase encoding samples in the absence of magnetic field inhomogeneities
$\Delta\tilde{\phi}_{k_x}, \Delta\tilde{\phi}_{k_y}$	Differences in accumulated phase between successive frequency and phase encoding samples in the presence of magnetic field inhomogeneities
Φ	General unknown quantity in a partial differential equation

$\Phi^{\mathscr{p}}$	Approximating polynomial for an element or subdomain with index $\mathscr{p}$
$\Phi_i^{\mathscr{p}}$	Value of the finite element solution at the node with local index i in the subdomain with index $\mathscr{p}$
Φ_m	Scalar magnetic potential
$\Phi_{\mathcal{E}}$	Magnetic flux
$\widetilde{\Phi}$	Finite element trial function or continuous approximation over the entire domain
γ	Gyromagnetic ratio
γ_a	Rectangular coil aspect ratio
γ_r	Grid refinement factor
κ_1, κ_2	Relative field error tolerance and convergence factor
$\boldsymbol{\lambda}$	Lagrange multiplier vector
$\boldsymbol{\lambda}^*$	Optimal Lagrange multiplier vector
$\boldsymbol{\mu}$	Nuclear magnetic moment vector
μ	Magnetic permeability
μ_0	Magnetic permeability of a vacuum
μ_r	Relative magnetic permeability
μ_x, μ_y, μ_z	Cartesian components of the nuclear magnetic moment vector
ω	Angular frequency
ω_0	Larmor frequency of precession
ω_1	Effective Larmor frequency upon resonance
ω_{eff}	General effective Larmor frequency during excitation
ω_z	Central frequency for selective excitation pulse
ω'	Time and spatially varying component of the generalized Larmor frequency

$\Delta\omega_z$	Spectral width of spatially selective excitation pulse
Ω	Domain of a problem specified by a partial differential equation
$\Omega^{\mathscr{p}}$	Finite element subdomain with index $\mathscr{p}$
$\partial\Omega^{\mathscr{p}}$	Boundary of the finite element subdomain with index $\mathscr{p}$
$\boldsymbol{\Pi}$	Projection operator
$\boldsymbol{\psi}$	Quantum mechanical state vector
$\boldsymbol{\psi}_+, \boldsymbol{\psi}_-$	Spin-up and spin-down eigenstate vectors
θ	Polar angle
θ_{RF}	Flip angle
ρ	Electric charge density distribution
$\rho_{2\mathrm{D}}, \rho_{3\mathrm{D}}$	Two and three-dimensional effective spin density distributions
$\boldsymbol{\sigma}$	Design sensitivity perturbation vector
$\sigma_{\max}$	Maximum perturbation amplitude
$\boldsymbol{\varsigma}$	Linear programming cost vector
$\boldsymbol{\tau}$	Torque vector
τ_{PE}	Duration of phase-encoding gradient
τ_{FE}	Duration of frequency-encoding gradient
τ_{RF}	Duration of radio-frequency excitation pulse
τ_s	Duration of signal acquisition
$\Delta\tau_s$	Separation of consecutive samples during signal acquisition
ξ, η, ζ	Cartesian coordinates in a reference frame related to the standard frame by a curvilinear transformation

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
AGP	Axisymmetric grid parameterized
AISI	American Iron and Steel Institute
ART^2	Advanced real-time adaptive radiotherapy
BMA	Bored magnet assembly
CCI	Cross Cancer Institute
CMA	Laterally reduced compact magnet assembly
COM	Component object model
CS	C-shaped magnet design
CT	Computed tomography
CTV	Clinical target volume
DARS	Double annular ring shimmed
DSV	Diameter spherical volume
EBRT	External beam radiotherapy
EPID	Electronic portal imaging device
FE	Frequency encoding
FEM	Finite element method
GM	Gifford-McMahon
GMV	Global model volume
GRE	Gradient echo
GTV	Gross tumor volume

HB	H-box magnet design
HTS	High-temperature superconducting
IM	Internal margin
kV	Kilovoltage
LP	Linear programming
M4	Modified four-column magnet design
MR	Magnetic resonance
MRI	Magnetic resonance imaging
MV	Megavoltage
NAGP	Non-axisymmetric grid parameterized
NMR	Nuclear magnetic resonance
PDE	Partial differential equation
PB	Closed pill-box magnet design
PE	Phase encoding
PTV	Planning target volume
RF	Radio-frequency
ROI	Region of interest
S4	Standard four-column magnet design
SE	Spin echo
SM	Set-up margin
SNR	Signal to noise ratio
SARS	Single annular ring shimmed
UEMA	Uniformly enlarged magnet assembly
US	Ultrasonography

CHAPTER 1
INTRODUCTION

1.1 IMAGE-GUIDED RADIOTHERAPY

External beam radiotherapy (EBRT) refers to the treatment of cancer through the delivery of concentrated beams of ionizing radiation. This radiation is generally comprised of high-energy photons (x-rays) or charged particles (electrons, protons, or heavy ions), which are generated with a source that is external to the patient's body. The primary source of radiation in modern EBRT systems is the medical linear accelerator (linac), which is predominantly used for the generation of megavoltage (MV) photon beams for the radical treatment of deep-seated tumors.

The principal goal of EBRT is to deliver the highest possible dose to malignant tissues, while maximally sparing surrounding critical structures. This effectively results in the greatest therapeutic gain, as indicated by an increased probability of disease control, together with a reduced likelihood of normal tissue complications. In the current state of the art, this is achieved by means of the advanced practice of intensity modulated radiotherapy[1] (IMRT). By employing radiation fields with non-uniform intensities, IMRT enables the delivery of highly conformed dose distributions with sharp dose gradients.

The therapeutic effectiveness of EBRT and IMRT is critically related to the accuracy and precision to which the treatment can be delivered as planned. Ultimately, this is limited by the uncertainties that arise due to errors or variations in the positioning of the targeted tissues with respect to the treatment system. As EBRT is typically administered on a fractionated basis,

these variations occur both during treatment delivery (intrafractional) and over the course of the treatment schedule (interfractional). Positioning uncertainties related to the former typically arise from the anatomical changes associated with respiration, cardiac motion, swallowing, and bowel movements. On the other hand, interfractional uncertainties can arise from the irreproducibility of daily patient setup, mechanical deformation of the treatment system, and physiological changes such as tumor shrinkage, weight loss, and the filling status of internal organs.

The impact of positioning uncertainty is best understood in terms of the volumes and margins that are assigned to the targeted anatomy during the treatment planning process[2, 3] (Figure 1.1). The gross tumor volume (GTV)

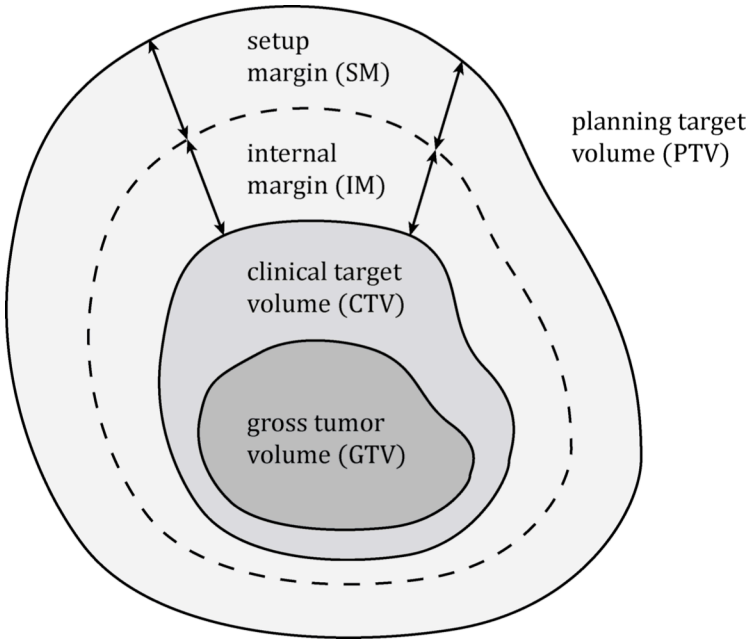


FIGURE 1.1: Schematic illustration of the treatment volumes defined during radiotherapy treatment planning. The internal margin (IM) and setup margin (SM) are appended to the clinical target volume (CTV) as a result of positioning uncertainties arising from intrafractional and interfractional variations in the patient anatomy, respectively.

is identified as the palpable and demonstrable extent of the cancerous tissue. The clinical target volume (CTV) is then defined to encompass the GTV and any suspected subclinical extension of the microscopic disease. Thus, it is this volume that must be delivered an adequate dose if the malignancy is to be eliminated. In order to ensure the requisite coverage of the CTV, the internal margin (IM) and setup margin (SM) are introduced. These margins act to increase the volume to which the prescribed dose is delivered, thereby compensating for any intrafractional and interfractional positioning uncertainties, respectively. Consequently, any healthy tissue within the resulting planning target volume (PTV) is unnecessarily irradiated, leading to an increased risk of treatment complications and secondary disease.

The widespread adoption of image-guided radiotherapy (IGRT) has helped improve EBRT treatment outcomes by enabling greater precision of the radiation delivery. Through frequent imaging in the treatment position, patient setup errors can be measured and changes in the shape, size, and location of the target volume can be detected. This information can then be used to make the necessary adjustments to maximize the geometric accuracy of the radiation delivery, ultimately leading to a reduction in positioning uncertainty and the associated margins introduced during treatment planning.

There are many devices capable of IGRT that have currently found mainstream clinical application. The most notable of these involve the on-board integration of an imaging device with the radiation delivery unit. A common such instrument is the electronic portal imaging device (EPID), which is generally positioned opposite a linear accelerator so as to utilize the MV treatment beam to obtain a two-dimensional (2D) projection image.[4] The entire system can then be rotated to obtain a series of projections that can be reconstructed into a volumetric MV cone-beam computed tomography (CT) data set.[5, 6] The on-board EPID has the advantage that it requires

minimal modification in order to be integrated with a conventional linear accelerator. In addition, the EPID can be directly used for dose measurements and therefore provides a useful quality assurance and treatment verification tool for IMRT.[4, 7, 8]

In an effort to obtain improved image quality and soft-tissue visualization, the on-board EPID can be augmented with an additional kilovoltage (kV) x-ray tube and flat-panel detector.[9, 10] This additional kV imaging system is typically mounted orthogonal to the treatment beam while maintaining the same axis of rotation as the linac gantry. Through rotation of the entire system, an additional cone-beam CT data set can be obtained with improved spatial resolution and soft-tissue contrast.

A completely integrated approach to volumetric x-ray based IGRT has been demonstrated with the TomoTherapy treatment system,[11] for which the capability of sequential helical IMRT has been combined with spiral megavoltage CT imaging. Both operations are performed in a continuous helical fashion as the patient couch is slowly translated through the bore, utilizing a single MV fan-beam source comprised of a rotating linear accelerator mounted on a ring-based gantry. This technique enables highly conformed dose distributions to be delivered and permits single-slice or volumetric MV imaging of the patient to be used for patient setup and treatment verification purposes.[12, 13]

Despite the ability to effectively reduce interfractional positioning uncertainties, the aforementioned devices are only capable of serial volumetric imaging. Specifically, the acquisition of three-dimensional (3D) data sets is restricted to immediately before or after delivery of the therapeutic radiation. In an effort to circumvent this limitation, stereoscopic x-ray imaging has been developed for intrafraction treatment guidance on the CyberKnife[14] and Novalis[15] stereotactic radiotherapy systems. Through

the cinematic acquisition of two orthogonal x-ray projections, these dedicated systems permit near real-time monitoring of the patient anatomy. This information can then be used to dynamically guide the treatment delivery, resulting in greater positional accuracy in the deposited dose.

Notwithstanding the many technological innovations established for x-ray based IGRT, all of these techniques suffer from two critical drawbacks. First, contrast in x-ray images is solely due to differences in tissue density and x-ray attenuation. As such, IGRT techniques based on this modality are of limited utility for sites with which bony structures cannot be used as reference objects.[16] Second, imaging with x-rays requires the transmission of additional radiation through the patient, resulting in the deposition of greater doses within healthy tissues and radiosensitive organs. Although the dose imparted by a single scan is significantly lower than the prescribed levels for treatment, the cumulative effect of frequent imaging may be harmful to the patient.[17]

Several other IGRT techniques are available that do not rely on radiographic imaging. One such modality is ultrasonography (US), for the effective visualization of soft-tissue anatomy that is unobstructed by gas or bone. In particular, US has been used to aid target localization prior to treatment of pelvic and abdominal malignancies.[18, 19] These devices require an operator to be present and therefore cannot be used during treatment delivery. Furthermore, inter-observer variations are a significant concern.[20] Peripheral solutions for IGRT also exist, which can involve radiographic,[21] electromagnetic,[22] or optical tracking[23] of implanted fiducial markers, radio-frequency transponders, or external surrogates for tumor motion, respectively. The major drawbacks of these techniques involve the difficulty of calibrating the detection systems for accurate target positioning, the inherent inability to visualize the actual target volume, and the invasive nature of fiducial implantation.

1.2 MAGNETIC RESONANCE IMAGING-GUIDED RADIOTHERAPY

In order to address the shortcomings of current IGRT technologies, several groups worldwide are now pursuing the integration of EBRT and magnetic resonance imaging (MRI) with the goal of achieving online target tracking and treatment guidance.[24–30] Due to the exquisite soft-tissue contrast attainable with MRI, it is becoming the preferred modality for the visualization and discrimination of anatomical structures such as tumors and organs at risk in cancer patients. MRI is capable of rapidly acquiring volumetric images with arbitrary orientations, permitting near real-time monitoring of target motion without interrupting radiation delivery. Furthermore, MRI does not employ ionizing radiation and is thus considered harmless to the patient being examined, provided the patient does not have any contraindications to strong static magnetic fields. Through a reduction of treatment errors and the associated margins related to patient positioning, daily anatomical variations, and intrafraction target motion, it is believed that improved treatment outcomes can be achieved with these hybrid systems.

Our group at the Cross Cancer Institute (CCI) in Edmonton, AB, Canada, is actively developing an integrated linac-MRI system[28–49] capable of performing advanced real-time adaptive radiotherapy (ART^2). The proposed design involves the rigid coupling of a 6 MV in-line side-coupled linear accelerator with a 0.2−0.5 T biplanar magnetic resonance (MR) imager, such that the two devices rotate in unison about the superior-inferior patient axis.

A rotating biplanar linac-MRI system permits two practical arrangements characterized by the relative orientation of the radiation source and magnetic field within the imaging volume (Figure 1.2). In the perpendicular configuration, the linac is mounted on the open end of the biplanar magnet such that the treatment beam is oriented perpendicular to the main magnetic

field. A prototype system comprised of a 6 MV linac and a 27.5 cm gap 0.2 T biplanar permanent magnet MRI scanner with a perpendicular configuration has been successfully constructed and tested to demonstrate the feasibility of the proposed concept.[29]

In the alternative parallel configuration, the linac is mounted exterior to the biplanar magnet with the treatment beam directed through the magnet poles and aligned parallel with the main magnetic field (Figure 1.2). It was recently shown that this arrangement leads to favourable dosimetric advantages. These include a decreased beam penumbra, no lateral shifting of the associated dose distributions, and a reduction of the hot and cold spots

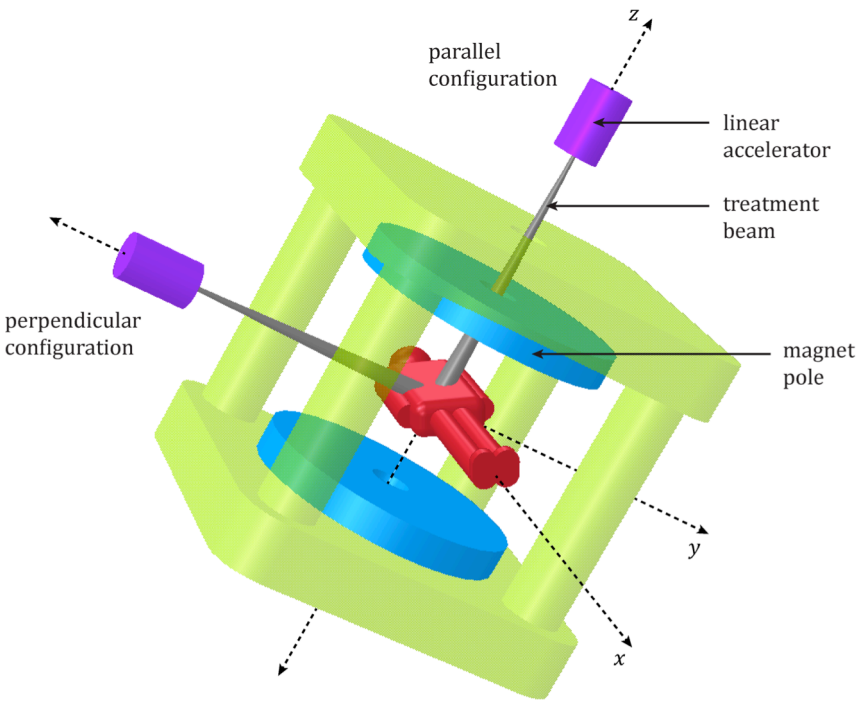


FIGURE 1.2: Parallel and perpendicular configurations for an integrated linac and biplanar magnet assembly. The two devices rotate in unison about the patient superior-inferior axis, such that the direction of the treatment beam and main magnetic field remain in a fixed relative orientation.

observed at tissue-air interfaces with the perpendicular configuration.[31, 46] The construction of a second prototype linac-MRI system with a parallel configuration is currently underway, for which a 6 MV linac will be coupled with a 0.5 T superconducting magnet system with a ferromagnetic yoke structure.

1.3 RESEARCH MOTIVATION

The rotating biplanar linac-MRI concept imposes two unique challenges involving the design of the MRI magnet assembly. The first is related to the requirement of rotation within a typical EBRT treatment vault. While maintaining a suitable pole gap for patient access, the lateral dimensions of the magnet assembly must be constrained. The second challenge is related to obtaining a parallel orientation of the linac with respect to the magnetic field. For a biplanar magnet employing a ferromagnetic yoke structure, a large hole needs to be bored from the magnet assembly in order to provide an unobstructed path for the treatment beam.

Both of the design challenges mentioned above involve a departure from the ideal configuration of the magnetic yoke structure and poles. This is realized by noting that a completely homogeneous magnetic field can only be obtained with a biplanar geometry if the magnet poles are infinite in extent. Constraining the pole dimensions and removing magnetic material from the region nearest isocenter is therefore expected to adversely impact the characteristics of the magnetic field in the imaging volume. Consequently, an investigation into the viability of a compact bored biplanar magnet for MRI was necessary.

Since the primary purpose of imaging with the hybrid linac-MRI system is to provide accurate geometrical information regarding the absolute shape and location of a target volume during treatment, obtaining distortion free

images through the reduction of magnetic field inhomogeneities is paramount. Therefore, it was the principal objective of this research to develop and implement novel computational methods for the optimal design of both permanent and superconducting biplanar magnets exhibiting uniform magnetic fields. In particular, a goal of this research was to obtain and evaluate theoretical magnet designs that are compact in size and permit an unobstructed beam path for a linac-MRI system with a parallel configuration.

1.4 THESIS OUTLINE

The structure of the thesis is as follows: Chapter 2 provides an overview of the basic principles of MRI, thereby establishing the necessary background relevant to nuclear magnetic resonance, image formation, and magnet design. Chapters 3 and 4 discuss the theory and techniques central to the work presented in this thesis. Specifically, Chapter 3 introduces classical electromagnetic theory, which leads to a description of the semi-analytic formulas used for the calculation of magnetic fields generated by circular coil systems. This is followed next by a theoretical description of the finite element method (FEM) and associated software. Chapter 4 introduces the basic theoretical concepts of mathematical optimization, which leads to a practical discussion of several algorithms that were incorporated into the methods developed in this thesis. Chapter 5 investigates the consequences of lateral size reduction for a 0.2 T biplanar permanent magnet assembly. This is followed by a description of an iterative pole piece shape optimization method based on field calculations with the FEM. Novel pole piece designs are obtained and compared, and the performance of the optimization algorithm is evaluated. Chapter 6 builds on the work in the preceding chapter by investigating the consequences imparted by inclusion of a large hole bored through the laterally reduced magnet assembly. Novel optimized pole piece designs are obtained and evaluated. Chapter 7 provides a

complete description of a robust optimization scheme for the calculation of minimum volume coil arrangements for homogenous superconducting magnet systems containing magnetic yoke structures. Chapter 8 applies the optimization method from the preceding chapter to a unified design approach for a collection of 0.5 T compact bored superconducting magnet assemblies. The resulting designs are evaluated and a candidate is identified for further optimization. Chapter 9 concludes this thesis with a summary of the presented work followed by a discussion regarding possible future extensions.

1.5 REFERENCES

[1] C. X. Yu, C. J. Amies, and M. Svatos, "Planning and delivery of intensity-modulated radiation therapy," *Med. Phys.* **35**(12), 5233−5241 (2008).

[2] International Commission on Radiation Units and Measurements, *ICRU Report 50: Prescribing, Recording and Reporting Photon Beam Therapy* (ICRU, Bethesda, 1993).

[3] International Commission on Radiation Units and Measurements, *ICRU Report 62: Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU Report 50)* (ICRU, Bethesda, 1999).

[4] J. Leong, "Use of digital fluoroscopy as an on-line verification device in radiation therapy," *Phys. Med. Biol.* **31**(9), 985−992 (1986).

[5] J. Pouliot, A. Bani-Hashemi, J. Chen, M. Svatos, F. Ghelmansarai, M. Mitschke, M. Aubin, P. Xia, O. Morin, K. Bucci, M. Roach, P. Hernandez, Z. R. Zheng, D. Hristov, and L. Verhey, "Low-dose megavoltage cone-beam CT for radiation therapy," *Int. J. Radiat. Oncol.* **61**(2), 552−560 (2005)

[6] O. Morin, A. Gillis, J. Chen, M. Aubin, M. K. Bucci, M. Roach, and J. Pouliot, "Megavoltage cone-beam CT: system description and clinical applications," *Med. Dosim.* **31**(1), 51−61 (2006).

[7] V. N. Hansen, P. M. Evans, and W. Swindell, "The application of transit dosimetry to precision radiotherapy," *Med. Phys.* **23**(5), 713−721 (1996).

[8] K. L. Pasma, B. J. Heijmen, M. Kroonwijk, and A. G. Visser, "Portal dose image (PDI) prediction for dosimetric treatment verification in radiotherapy. I. An algorithm for open beams," *Med. Phys.* **25**(6), 830−840 (1998).

[9] D. A. Jaffray, D. G. Drake, M. Moreau, A. A. Martinez, and J. W. Wong, "A radiographic and tomographic imaging system integrated into a medical linear accelerator for localization of bone and soft-tissue targets," *Int. J. Radiat. Oncol.* **45**(3), 773−789 (1999).

[10] D. A. Jaffray, J. H. Siewerdsen, J. W. Wong, and A. A. Martinez, "Flat-panel cone-beam computed tomography for image-guided radiation therapy," *Int. J. Radiat. Oncol.* **53**(5), 1337−1349 (2002).

[11] T. R. Mackie, J. Kapatoes, K. Ruchala, W. G. Lu, C. Wu, G. Olivera, L. Forrest, W. Tome, J. Welsh, R. Jeraj, P. Harari, P. Reckwerdt, B. Paliwal, M. Ritter, H. Keller, J. Fowler, and M. Mehta, "Image guidance for precise conformal radiotherapy," *Int. J. Radiat. Oncol.* **56**(1), 89−105 (2003).

[12] K. J. Ruchala, G. H. Olivera, E. A. Schloesser, and T. R. Mackie, "Megavoltage CT on a tomotherapy system," *Phys. Med. Biol.* **44**(10), 2597−2621 (1999).

[13] K. M. Langen, S. L. Meeks, D. O. Poole, T. H. Wagner, T. R. Willoughby, P. A. Kupelian, K. J. Ruchala, J. Haimerl, and G. H. Olivera, "The use of megavoltage CT (MVCT) images for dose recomputations," *Phys. Med. Biol.* **50**(18), 4259−4276 (2005).

[14] J. R. Adler, S. D. Chang, M. J. Murphy, J. Doty, P. Geis, S. L. Hancock, "The cyberknife: a frameless robotic system for radiosurgery," *Stereot. Funct. Neuros.* **69**(1-4), 124−128 (1997).

[15] T. D. Solberg, R. Fogg, M. T. Selch, A. A. F. De Salles, "Conformal radiosurgery using a dedicated linac and micro multileaf collimator," *Radiosurg.* **3**, 53−63 (2000).

[16] K. M. Langden and D. T. Jones, "Organ motion and its management," *Int. J. Radiat. Oncol.* **50**(1), 265−278 (2001).

[17] H. Shirato, M. Oita, K. Fujita, Y. Watanabe, and K. Miyasaka, "Feasibility of synchronization of real-time tumor-tracking radiotherapy and intensity-modulated radiotherapy from viewpoint of excessive dose from fluoroscopy," *Int. J. Radiat. Oncol.* **60**(1), 335−341 (2004).

[18] E. J. Holupka, I. D. Kaplan, E. C. Burdette, and G. K. Svensson, "Ultrasound image fusion for external beam radiotherapy for prostate cancer," *Int. J. Radiat. Oncol.* **35(**5), 975−984 (1996).

[19] J. Lattanzi, S. McNeeley, W. Pinover, E. Horwitz, I. Das, T. E. Schultheiss, and G. E. Hanks, "A comparison of daily CT localization to a daily ultrasound-based system in prostate cancer," *Int. J. Radiat. Oncol.* **43**(4), 719−725 (1999).

[20] K. M. Langen, J. Pouliot, C. Anezinos, M. Aubin, A. R. Gottschalk, I. C. Hsu, D. Lowther, Y. M. Liu, K. Shinohara, L. J. Verhey, V. Weinberg, M. Roach, "Evaluation of ultrasound-based prostate localization for image-guided radiotherapy," *Int. J. Radiat. Oncol.* **57**(3), 635−644 (2003).

[21] H. Shirato, S. Shimizu, K. Kitamura, T. Nishioka, K. Kagei, S. Hashimoto, H. Aoyama, T. Kunieda, N. Shinohara, H. Dosaka-Akita, K. Miyasaka, "Four-

dimensional treatment planning and fluoroscopic real-time tumor tracking radiotherapy for moving tumor," *Int. J. Radiat. Oncol.* **48**(2), 435–442 (2000).

[22] P. G. Seiler, H. Blattman, S. Kirsch, R. K. Muench, and C. Schilling, "A novel tracking technique for the continuous precise measurement of tumour positions in conformal radiotherapy," *Phys. Med. Biol.* **45**(9), N103–N110 (2000).

[23] J. D. P. Hoisak, K. E. Sixel, R. Tirona, P. Cheung, and J. P. Pignol, "Correlation of lung tumour motion with surrogates of respiration," *Med. Phys.* **31**(6), 1759 (2004).

[24] J. F. Dempsey, D. Benoit, J. R. Fitzsimmons, A. Haghighat, J. G. Li, D. A. Low, S. Mutic, J. R. Palta, H. E. Romeijn, and G. E. Sjoden, "A device for realtime 3D image-guided IMRT," *Int. J. Radiat. Oncol.* **63**(2), S202, (2005).

[25] J. Lagendijk, B. Raaymakers, U. van der Heide, J. Overweg, K. Brown, C. Bakker, A. Raaijmakers, M. Vulpen, J. Welleweerd, and I. Jurgenliemk-Schulz, "In room magnetic resonance imaging guided radiotherapy (MRIgRT)," *Med. Phys.* **32**(6), 2067 (2005).

[26] J. J. W. Lagendijk, B. W. Raaymakers, A. J. E. Raaijmakers, J. Overweg, K. J. Brown, E. M. Kerkhof, R. W. van der Put, B. Hardemark, M. van Vulpen, U. A. van der Heide, "MRI/linac integration," *Radiother. Oncol.* **86**(1), 25–29 (2008).

[27] B. W. Raaymakers, J. J. W. Lagendijk, J. Overweg, J. G. M. Kok, A. J. E. Raaijmakers, E. M. Kerkhof, R. W. van der Put, I. Meijsing, S. P. M. Crijns, F. Benedosso, M. van Vulpen, C. H. W. de Graaff, J. Allen, and K. J. Brown,

"Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept," *Phys. Med. Biol.* **54**(12), N229–N237 (2009).

[28] B. G. Fallone, M. Carlone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, K. Wachowicz, and C. Kirkby, "Development of a linac-MRI system for real-time ART," *Med. Phys.* **34**(6), 2547–2547 (2007).

[29] B. G. Fallone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, and S. Vidakovic, "First MR images obtained during megavoltage photon irradiation from a prototype integrated linac-MR system," *Med. Phys.* **36**(6), 2084–2088 (2009).

[30] B. G. Fallone, "Real-time MR-guided radiotherapy: integration of a low-field MR system," *Med. Phys.* **36**(6), 2774–2775 (2009).

[31] C. Kirkby, T. Stanescu, S. Rathee, M. Carlone, B. Murray, and B. G. Fallone, "Patient dosimetry for hybrid MRI-radiotherapy systems," *Med. Phys.* **35**(3), 1019–1027 (2008)

[32] C. Kirkby, T. Stanescu, and B. G. Fallone, "Magnetic field effects on the energy deposition spectra of MV photon radiation," *Phys. Med. Biol.* **54**(2), 243–247 (2009).

[33] B. Burke, M. Lamey, S. Rathee, B. Murray, and B. G. Fallone, "Radio frequency noise from clinical linear accelerators," *Phys. Med. Biol.* **54**(8), 2483–2492 (2009).

[34] B. Burke, B. G. Fallone, and S. Rathee, "Radiation induced currents in MRI RF coils: application to linac/MRI integration," *Phys. Med. Biol.* **55**(3), 735–746 (2010).

[35] M. Lamey, J. Yun, B. Burke, S. Rathee, and B. G. Fallone, "Radio frequency noise from an MLC: a feasibility study of the use of an MLC for linac-MR systems," *Phys. Med. Biol.* **55**(4), 981–994 (2010).

[36] M. Lamey, B. Burke, E. Blosser, S. Rathee, N. De Zanche, and B. G. Fallone, "Radio frequency shielding for a linac-MRI system," *Phys. Med. Biol.* **55**(4), 995–1006 (2010).

[37] J. St. Aubin, S. Steciw, and B. G. Fallone, "The design of a simulated in-line side-coupled 6 MV linear accelerator waveguide," *Med. Phys.* **37**(2), 466–476 (2010)

[38] K. Wachowicz, T. Stanescu, S. D. Thomas, and B. G. Fallone, "Implications of tissue magnetic susceptibility-related distortion on the rotating magnet in an MR-linac design," *Med. Phys.* **37**(4), 1714–1721 (2010).

[39] J. Yun, J. St. Aubin, S. Rathee, and B. G. Fallone, "Brushed permanent magnet DC MLC motor operation in an external magnetic field," *Med. Phys.* **37**(5), 2131–2134 (2010).

[40] J. St. Aubin, S. Steciw, C. Kirkby, and B. G. Fallone, "An integrated 6 MV linear accelerator model from electron gun to dose in a water tank," *Med. Phys.* **37**(5), 2279–2288 (2010)

[41] M. Lamey, S. Rathee, L. Johnson, M. Carlone, E. Blosser, and B. G. Fallone, "Radio frequency noise from the modulator of a linac," *IEEE Trans. Electromagn. C.* **52**(3), 530–536 (2010).

[42] J. St. Aubin, S. Steciw, and B. G. Fallone, "Effect of transverse magnetic fields on a simulated in-line 6 MV linac," *Phys. Med. Biol.* **55**(16), 4861–4869 (2010).

[43] J. St. Aubin, D. M. Santos, S. Steciw, and B. G. Fallone, "Effect of longitudinal magnetic fields on a simulated in-line 6 MV linac," *Med. Phys.* **37**(9), 4916−4923 (2010).

[44] J. St. Aubin, S. Steciw, and B. G. Fallone, "Waveguide detuning caused by transverse magnetic fields on a simulated in-line 6 MV linac," *Med. Phys.* **37**(9), 4751−4754 (2010)

[45] J. St. Aubin, S. Steciw, and B. G. Fallone, "Magnetic decoupling of the linac in a low field biplanar linac-MR system," *Med. Phys.* **37**(9), 4755−4761 (2010)

[46] C. Kirkby, B. Murray, S. Rathee, and B. G. Fallone, "Patient dosimetry in a linac-MR radiotherapy unit with a longitudinal magnetic field," *Med. Phys.* **37**(9), 4722−4732 (2010).

[47] T. Tadic and B. G. Fallone, "Three-dimensional non-axisymmetric pole piece shape optimization for biplanar permanent magnet MRI systems," *IEEE Trans. Magn.* **47**(1), 231−238 (2011).

[48] T. Tadic and B. G. Fallone, "Design and optimization of a novel bored biplanar permanent-magnet assembly for hybrid magnetic resonance imaging systems," *IEEE Trans. Magn.* **46**(12), 4052−4058 (2010).

[49] T. Tadic and B. G. Fallone, "Design and optimization of superconducting MRI magnet systems with magnetic materials," *IEEE Trans. Appl. Supercond.* (accepted for publication, 2012).

CHAPTER 2
MAGNETIC RESONANCE IMAGING

Magnetic resonance imaging (MRI) is a powerful non-invasive volumetric imaging technique that is widely regarded as one of the most significant advancements in medical technology in the 20th century. Through the exploitation of various tissue properties and unique contrast mechanisms, MRI can be used to generate high quality three-dimensional images for the detailed visualization of anatomical structures. Due to the physical mechanism underpinning signal formation, MRI is also capable of providing detailed metabolic and functional information of biological tissues in vivo. Consequently, MRI has become an indispensable clinical tool with a wide variety of applications involving the detection and treatment of disease.

2.1 NUCLEAR MAGNETIC RESONANCE

2.1.1 Microscopic Magnetism

Atomic nuclei primarily consist of neutrons and protons. These particles are collectively referred to as nucleons. The neutron is composed of two down quarks with charge $-e/3$ and one up quark with charge $+2e/3$. Thus, the total electric charge of the neutron is zero. On the other hand, the proton is composed of two up quarks and one down quark, resulting in a total electric charge of $+e$.

All quarks possess an intrinsic spin of $1/2$. In the familiar state for each nucleon, two of the quarks are in an antiparallel spin configuration. Hence, according to a simplified model, the presence of the third unpaired quark leads to a net intrinsic spin of $1/2$ for both nucleons.

TABLE 2.1: Typical values for the spin, gyromagnetic ratio, and natural abundance of various nuclei which exhibit a magnetic moment.

Isotope	Ground-state spin $\mathfrak{s}$	$\gamma/2\pi$ (MHz/T)	Natural abundance (%)	Abundance in human body (M)
^{1}H	1/2	42.577	99.99	88
^{17}O	5/2	–5.774	0.04	80×10^{-3}
^{19}F	1/2	40.078	~100	75×10^{-3}
^{23}Na	3/2	11.269	~100	16×10^{-3}
^{31}P	1/2	17.251	~100	4×10^{-6}

Atomic nuclei also possess the property of intrinsic spin, with a value for the nuclear spin quantum number $\mathfrak{s}$ determined by the configuration of the constituent nucleons. In particular, if the number of protons and neutrons are both even, the ground state nuclear spin is given by $\mathfrak{s} = 0$. Furthermore, an even number of nucleons leads to nuclei with integer spin, and an odd number leads to half-integer spin. The ground-state spins for several nuclear isotopes are provided[1, 2] in Table 2.1.

The presence of spin angular momentum as possessed by a particle gives rise to a nuclear magnetic moment $\boldsymbol{\mu}$ given as[3]

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

where the constant of proportionality γ, referred to as the gyromagnetic ratio, is a property of the nuclear species examined. In this equation, $\boldsymbol{S}$ is the spin angular momentum vector, related to the spin quantum number $\mathfrak{s}$ by

$$S^2 = \hbar^2 \mathfrak{s}(\mathfrak{s} + 1). \tag{2.2}$$

As evident in Eq. (2.1), the strength of the magnetic dipole moment (and extent of interaction with external magnetic fields) exhibited by a nucleus is directly related to the magnitude of γ. Consequently, those nuclei with large γ values are of special interest in nuclear magnetic resonance (NMR) experiments. In particular, the proton (or ^{1}H nucleus) is the dominant species targeted during magnetic resonance imaging of the human body, due to a large value of γ coupled with an overbearing abundance (as a constituent of H_20) in the human body (Table 2.1).

2.1.2 Spin Precession

The general quantum mechanical state $\boldsymbol{\psi}$ for a spin $1/2$ particle at rest may be expressed as the superposition

$$\boldsymbol{\psi} = a_+\boldsymbol{\psi}_+ + a_-\boldsymbol{\psi}_-, \tag{2.3}$$

where $\boldsymbol{\psi}_+$ and $\boldsymbol{\psi}_-$ represent two independent eigenstates corresponding to the azimuthal quantum numbers $\mathfrak{m} = 1/2$ and $-1/2$. These eigenstates are often referred to as the spin-up (parallel) state $|\alpha\rangle$ and the spin-down (anti-parallel) state $|\beta\rangle$, respectively.

If the particle is isolated in the absence of a magnetic field, the two independent eigenstates are degenerate. However, if the spin is immersed in a static uniform field $\boldsymbol{B_0} = B_0\hat{\boldsymbol{z}}$, the energies of the eigenstates become[1]

$$E_\pm = \mp\frac{\gamma B_0 \hbar}{2}, \tag{2.4}$$

in accordance with the general Zeeman effect. Hence, the energy of the eigenstate with a spin parallel to the external field is lower than that of a spin

anti-parallel to the field, as illustrated in Figure 2.1. Furthermore, the energy difference ΔE between the two states is given by

$$\Delta E = \gamma B_0 \hbar = \hbar \omega_0, \tag{2.5}$$

where the Larmor precession frequency ω_0 is defined by

$$\omega_0 = \gamma B_0. \tag{2.6}$$

The effect of the external field on the dynamics of the spin magnetic moment is of particular interest in NMR experiments. The torque $\boldsymbol{\tau}$ experienced by this spin is given by

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \boldsymbol{B_0}, \tag{2.7}$$

and the potential energy U associated with this torque is

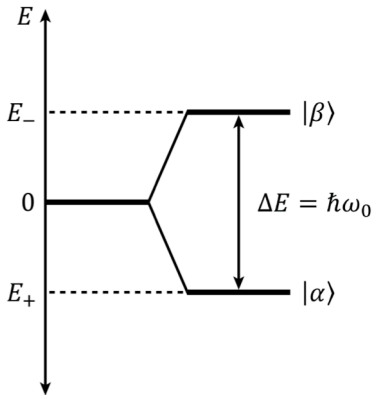


FIGURE 2.1: The nuclear Zeeman energy levels for a spin 1/2 particle in a magnetic field. The energy difference ΔE between the spin-up state $|\alpha\rangle$ and the spin-down state $|\beta\rangle$ is proportional to the magnitude of the magnetic field.

$$U = -\boldsymbol{\mu} \cdot \boldsymbol{B_0}. \tag{2.8}$$

A quantum mechanical treatment of the resulting spin behaviour yields an expectation value of the magnetic moment $\langle \boldsymbol{\mu}(t) \rangle$ defined as

$$\langle \boldsymbol{\mu}(t) \rangle = \langle \mu_x(t) \rangle \hat{\boldsymbol{x}} + \langle \mu_y(t) \rangle \hat{\boldsymbol{y}} + \langle \mu_z(t) \rangle \hat{\boldsymbol{z}}, \tag{2.9}$$

where the components are given by[1]

$$\langle \mu_x(t) \rangle = \frac{\gamma \hbar}{2} \sin\theta \cos(-\omega_0 t + \phi_0), \tag{2.10a}$$

$$\langle \mu_y(t) \rangle = \frac{\gamma \hbar}{2} \sin\theta \sin(-\omega_0 t + \phi_0), \tag{2.10b}$$

$$\langle \mu_z(t) \rangle = \frac{\gamma \hbar}{2} \cos\theta. \tag{2.10c}$$

Evidently, as illustrated in Figure 2.2, the magnetic moment (or spin) precesses about the magnetic field direction with a constant polar angle θ, initial phase ϕ_0, and azimuthal frequency $-\omega_0$ (in the right-hand sense). Moreover, the values[2] for the constants θ and ϕ_0 are determined by the initial relative magnitude and phases of the complex coefficients a_+ and a_- in the superposition in Eq. (2.3).

The spin motion described by Eqs. (2.10) is precisely that predicted by classical Newtonian mechanics. Consequently, it often suffices to refer to the classical interpretation of the spin system dynamics when describing the basic theory of NMR.

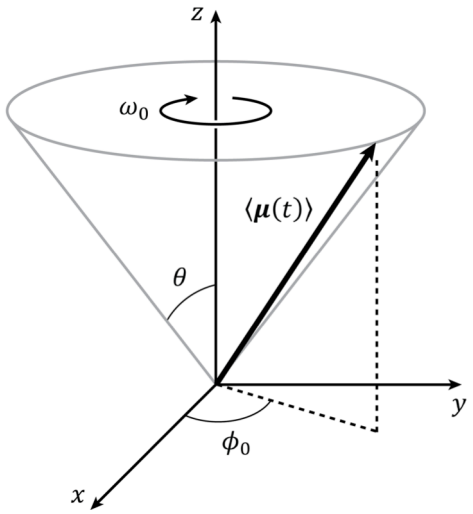


FIGURE 2.2: The expectation value $\langle \boldsymbol{\mu}(t) \rangle$ for the nuclear magnetic moment of a spin 1/2 particle precesses around the direction of the external magnetic field. The Larmor frequency of precession ω_0 is proportional to the magnitude of the applied field.

2.1.3 The Ensemble of Spins

Consider an ensemble of spins within a bulk object, such as protons within a sample of water. Further, imagine that the sample is segmented into small but finite volume segments, each containing trillions or more parent molecules and spins. The net magnetization $\boldsymbol{M}$ within a segment of volume $\mathcal{V}$ is then given by

$$\boldsymbol{M} = \frac{1}{\mathcal{V}} \sum_{i=1}^{N_s} \boldsymbol{\mu}_i, \tag{2.11}$$

where $\boldsymbol{\mu}_i$ is the magnetic moment of the ith spin and N_s is the total number of spins within $\mathcal{V}$.

Prior to application of an external magnetic field, the individual spins are randomly oriented within $\mathcal{V}$, yielding $\boldsymbol{M} = 0$ (Figure 2.3). However, once an external field $\boldsymbol{B_0} = B_0\hat{\boldsymbol{z}}$ is applied, each spin begins to precess about the field according to Eq. (2.5). Assuming the sample has a non-zero temperature T, thermal excitations lead to a biasing of the spin orientations towards a direction parallel with the magnetic field. According to Boltzmann statistics, the probabilities $P_{\pm}$ of the two eigenstates are

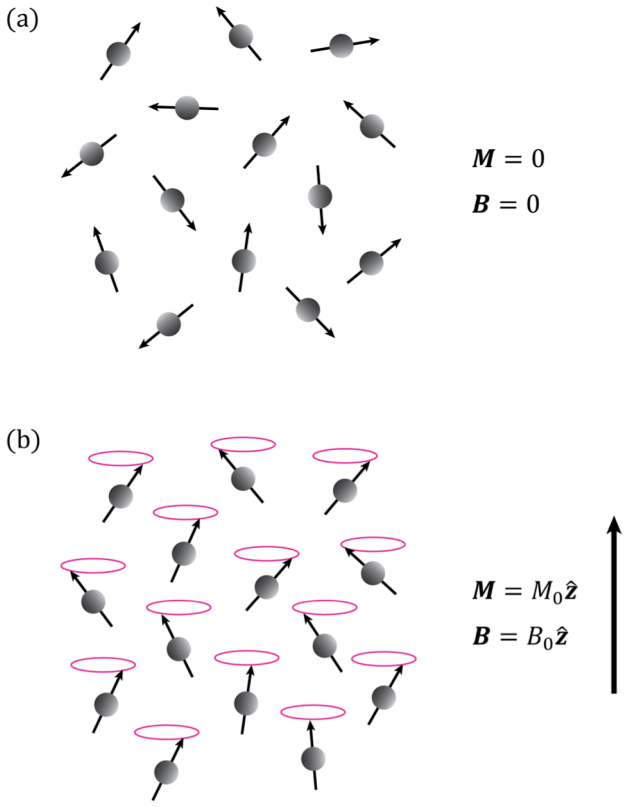


FIGURE 2.3: (a) Prior to the application of an externally applied magnetic field, the directions of the nuclear spins (indicated by the small arrows) are randomly oriented within the bulk object. (b) In the presence of an external magnetic field, the orientations of the nuclear spins are biased towards the direction parallel to the applied field. Coupled with the random thermal motion of the nuclei, this biasing results in a measurable net magnetization $\boldsymbol{M}$ in the bulk sample.

$$P_{\pm} = \frac{e^{\pm \Delta E/2k_B T}}{e^{\Delta E/2k_B T} + e^{-\Delta E/2k_B T}}, \tag{2.12}$$

where k_B is Boltzmann's constant. At temperatures roughly that of the human body (310 K), and for a practical field strength on the order of $B_0 = 1.0$ T, the energy difference ΔE is roughly five orders of magnitude smaller than the thermal energy $k_B T$. Consequently, the difference of probabilities is given by

$$\Delta P = P_+ - P_- \approx \frac{\Delta E}{2k_B T}. \tag{2.13}$$

This is an extremely small number on the order of 10^{-5}. However, considering the enormous number of spins N_s within a given voxel, a measurable net longitudinal magnetization $\boldsymbol{M_z} = M_0 \hat{\boldsymbol{z}}$ is developed at thermal equilibrium. Manipulation of this magnetization forms the ultimate basis of the NMR experiment, throughout which it is understood that the static $\boldsymbol{B_0}$ field remains active at all times.

2.1.4 Excitation of the Spin System

Typically, the diamagnetic magnetization of the sample examined is several orders of magnitude larger[1] than the extremely weak nuclear spin magnetization $\boldsymbol{M_z}$. Consequently, the direct study of $\boldsymbol{M_z}$ is impractical. To overcome this limitation, NMR exploits the fact that individual spins can undergo transitions between the parallel and antiparallel spin states via the absorption or release of photons with energy $\Delta E = \hbar \omega_0$. By triggering such transitions over the ensemble of spins, the direction of the net magnetization of the sample can be perturbed.

Consider the introduction of a left-circularly polarized time dependent magnetic field $\boldsymbol{B_1}(t)$ in addition to the static uniform field $\boldsymbol{B_0}$. If ω is the frequency of the field oscillation, then $\boldsymbol{B_1}(t)$ can be defined as

$$\boldsymbol{B_1}(t) = B_1(\cos\omega t\,\hat{\boldsymbol{x}} - \sin\omega t\,\hat{\boldsymbol{y}}), \tag{2.14}$$

where the unit vectors $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ define the laboratory frame of reference. In order to understand the effect of the combined field $\boldsymbol{B}(t) = \boldsymbol{B}_0 + \boldsymbol{B}_1(t)$ on the spin dynamics, the rotating frame of reference defined by the unit vectors $\left(\hat{\boldsymbol{x}}_{\boldsymbol{\rho}}, \hat{\boldsymbol{y}}_{\boldsymbol{\rho}}, \hat{\boldsymbol{z}}_{\boldsymbol{\rho}}\right)$ is introduced:

$$\hat{\boldsymbol{x}}_{\boldsymbol{\rho}} = \boldsymbol{\mathcal{R}}_{\hat{\boldsymbol{z}}}(-\omega t)\hat{\boldsymbol{x}} = \cos\omega t\,\hat{\boldsymbol{x}} - \sin\omega t\,\hat{\boldsymbol{y}}, \tag{2.15a}$$

$$\hat{\boldsymbol{y}}_{\boldsymbol{\rho}} = \boldsymbol{\mathcal{R}}_{\hat{\boldsymbol{z}}}(-\omega t)\hat{\boldsymbol{y}} = \sin\omega t\,\hat{\boldsymbol{x}} + \cos\omega t\,\hat{\boldsymbol{y}}, \tag{2.15b}$$

$$\hat{\boldsymbol{z}}_{\boldsymbol{\rho}} = \boldsymbol{\mathcal{R}}_{\hat{\boldsymbol{z}}}(-\omega t)\hat{\boldsymbol{z}} = \hat{\boldsymbol{z}}. \tag{2.15c}$$

In these equations, forward multiplication by the operator $\boldsymbol{\mathcal{R}_u}(\theta)$ defines a positive rotation through the angle θ (in the right-hand sense) about the vector $\boldsymbol{u}$. With this definition of the rotating frame, the unit vector $\hat{\boldsymbol{x}}_{\boldsymbol{\rho}}$ rotates in phase with the left-circularly polarized field $\boldsymbol{B_1}(t)$, such that:

$$\boldsymbol{B_1}(t) = B_1\hat{\boldsymbol{x}}_{\boldsymbol{\rho}}. \tag{2.16}$$

Consequently, both $\boldsymbol{B_0}$ and $\boldsymbol{B_1}(t)$ appear static when observed from the rotating frame, as illustrated in Figure 2.4.

A quantum mechanical calculation of the spin behaviour when subjected to the combined field $\boldsymbol{B}(t)$ yields a time dependent expectation value for the magnetic moment $\langle\boldsymbol{\mu}(t)\rangle_\rho$ given by[1]

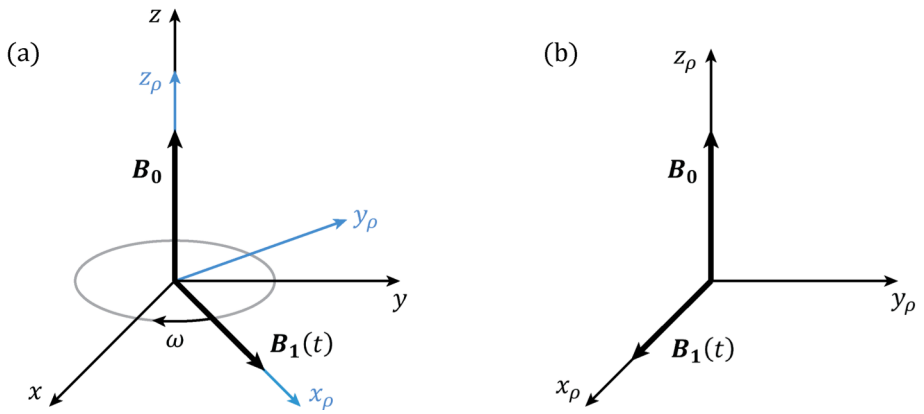


FIGURE 2.4: (a) The static magnetic field $\boldsymbol{B_0}$ and the left-circularly polarized time varying field $\boldsymbol{B_1}(t)$ are shown in the standard laboratory frame. The rotating frame defined by the unit vectors $(\boldsymbol{x_\rho}, \boldsymbol{y_\rho}, \boldsymbol{z_\rho})$ is highlighted in blue. (b) When viewed from the rotating frame, both magnetic fields $\boldsymbol{B_0}$ and $\boldsymbol{B_1}(t)$ appear static.

$$\langle \boldsymbol{\mu}(t) \rangle_\rho = \boldsymbol{\mathcal{R}}_{\boldsymbol{B_{\text{eff}}}}(-\omega_{\text{eff}} t) \langle \boldsymbol{\mu}(0) \rangle_\rho, \tag{2.17}$$

where $t = 0$ corresponds to the instant that $\boldsymbol{B_1}(t)$ is applied in addition to $\boldsymbol{B_0}$. In this equation, the effective field $\boldsymbol{B_{\text{eff}}}$ is defined in the rotating frame as

$$\boldsymbol{B_{\text{eff}}} = B_1 \hat{\boldsymbol{x}}_{\boldsymbol{\rho}} + \frac{\omega_0 - \omega}{\gamma} \hat{\boldsymbol{z}}_{\boldsymbol{\rho}}, \tag{2.18}$$

and the frequency ω_{eff} is given by

$$\omega_{\text{eff}} = \gamma B_{\text{eff}} = \sqrt{{\omega_1}^2 + (\omega_0 - \omega)^2}, \tag{2.19}$$

with

$$\omega_1 = \gamma B_1. \tag{2.20}$$

Therefore, the magnetic moment is rotated through the angle $-\omega_{\text{eff}} t$ about $\boldsymbol{B}_{\mathbf{eff}}$, as observed in the rotating frame. In particular, if the frequency of $\boldsymbol{B}_{\mathbf{1}}(t)$ is such that the magnetic resonance condition $\omega = \omega_0$ is satisfied, then the effective field $\boldsymbol{B}_{\mathbf{eff}}$ is coincident with the x_ρ axis and the magnetic moment is rotated by the flip angle θ_{RF} given by

$$\theta_{\text{RF}} = -\omega_1 \tau_{\text{RF}}, \tag{2.21}$$

where τ_{RF} is the duration for which $\boldsymbol{B}_{\mathbf{1}}(t)$ is applied. Since the resonant frequency ω_0 typically lies in the RF window of the electromagnetic spectrum (Table 2.1), the $\boldsymbol{B}_{\mathbf{1}}(t)$ field constitutes an RF electromagnetic pulse.

Since every member within an ensemble of polarized spins experiences the rotation described in Eq. (2.17), applying $\boldsymbol{B}_{\mathbf{1}}(t)$ to a bulk sample results in the same rotation of the net longitudinal magnetization $\boldsymbol{M}_{\boldsymbol{z}}$. Furthermore, if the duration τ_{RF} of $\boldsymbol{B}_{\mathbf{1}}(t)$ is chosen such that $\theta_{\text{RF}} = -\pi/2$, the net effect is a rotation of $\boldsymbol{M}_{\boldsymbol{z}}$ onto the xy plane, as illustrated in Figure 2.5. This process of generating a net transverse magnetization $\boldsymbol{M}_{\boldsymbol{xy}}$ via tipping of the individual spins is referred to as excitation of the nuclear system.

A left-circularly polarized time varying field, such as $\boldsymbol{B}_{\mathbf{1}}(t)$ defined in Eq. (2.14), would be difficult to generate on its own. Thus, a much simpler approach is often taken in practice. Consider a linearly polarized field $\boldsymbol{B}_{\mathbf{RF}}(t)$ oscillating along the x direction:

$$\boldsymbol{B}_{\mathbf{RF}}(t) = 2B_1 \cos \omega_0 t \, \hat{\boldsymbol{x}}. \tag{2.22}$$

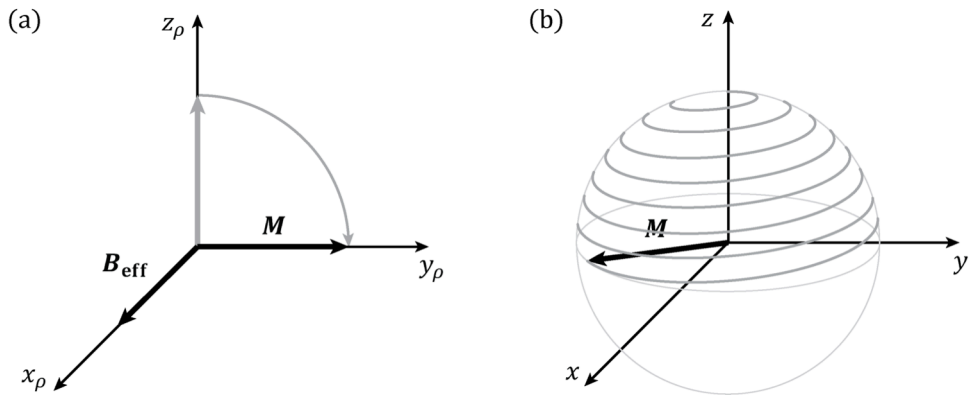


FIGURE 2.5: (a) When the magnetic resonance condition is satisfied, the effective field $\boldsymbol{B}_{\mathbf{eff}}$ is coincident with the x_ρ axis of the rotating frame. When viewed in this frame, the net magnetization $\boldsymbol{M}$ is tipped towards the transverse plane as it simply precesses about the direction of the effective field vector. (b) In the laboratory frame, the net magnetization spirals towards the transverse plane according to the illustrated path.

This field can be decomposed into left and right-circularly polarized components, denoted $\boldsymbol{B}_{\mathbf{L}}(t)$ and $\boldsymbol{B}_{\mathbf{R}}(t)$, according to the expression

$$\boldsymbol{B}_{\mathbf{RF}}(t) = \boldsymbol{B}_{\mathbf{L}}(t) + \boldsymbol{B}_{\mathbf{R}}(t), \tag{2.23}$$

where

$$\boldsymbol{B}_{\mathbf{L}}(t) = B_1(\cos\omega_0 t\,\hat{\boldsymbol{x}} - \sin\omega_0 t\,\hat{\boldsymbol{y}}), \tag{2.24a}$$

$$\boldsymbol{B}_{\mathbf{R}}(t) = B_1(\cos\omega_0 t\,\hat{\boldsymbol{x}} + \sin\omega_0 t\,\hat{\boldsymbol{y}}). \tag{2.24b}$$

Hence, $\boldsymbol{B}_{\mathbf{1}}(t)$ as defined in Eq. (2.14) is simply equal to $\boldsymbol{B}_{\mathbf{L}}(t)$. It can fortunately be shown that the non-resonant component $\boldsymbol{B}_{\mathbf{R}}(t)$ has a negligible effect on the motion of the spins in normal circumstances.[4] Therefore, the desired effect due to $\boldsymbol{B}_{\mathbf{1}}(t)$ can simply be achieved by applying

$\boldsymbol{B}_{\mathbf{RF}}(t)$ to the polarized sample. A simple coil that is capable of creating $\boldsymbol{B}_{\mathbf{RF}}(t)$ as defined in Eq. (2.22) is illustrated in Figure 2.6.

2.1.5 Relaxation

Upon excitation, the net magnetization $\boldsymbol{M}$ within a given voxel is perturbed from its previous orientation, generally resulting in nonzero transverse and longitudinal components. In addition to the external field, each of the spins contributing to $\boldsymbol{M}$ is subjected to a complex microscopic magnetic field, for which the magnitude and direction perpetually fluctuate due to the random and violent motion of the molecular surroundings. Consequently, these fluctuating microscopic fields perturb the constant angle of precession introduced in Eq. (2.9) as the individual spins exchange energy with the surrounding medium. Due to the anisotropic biasing of the spin directions resulting from the energy separation of the spin eigenstates, the individual magnetic moments tend towards the direction of the equilibrium

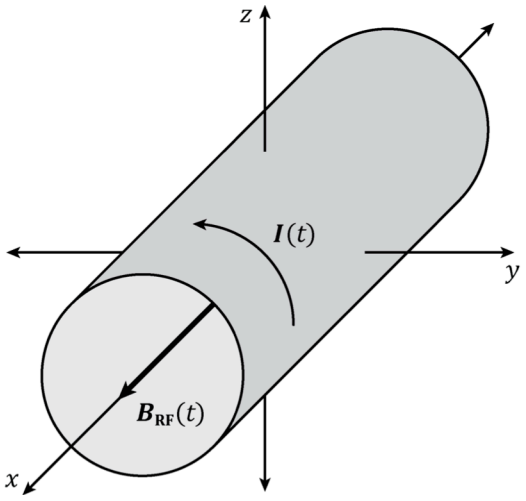


FIGURE 2.6: A simple solenoidal coil aligned with the x axis. If a sinusoidal time varying current $\boldsymbol{I}(t)$ is passed through the coil windings, then a linearly polarized magnetic field $\boldsymbol{B}_{\mathbf{RF}}(t)$ oscillating along the x direction is produced.

magnetization (see § 2.1.3). The macroscopic result of this effect can be summarized by the following differential equation for the longitudinal magnetization $M_z(\boldsymbol{r}, t)$:

$$\frac{dM_z(\boldsymbol{r}, t)}{dt} = \frac{1}{T_1(\boldsymbol{r})}[M_0(\boldsymbol{r}) - M_z(\boldsymbol{r}, t)], \tag{2.25}$$

where $M_0(\boldsymbol{r})$ is the magnitude of the equilibrium magnetization and $T_1(\boldsymbol{r})$ is the spatially dependant longitudinal (or spin-lattice) relaxation time. The solution to this equation is given by

$$M_z(\boldsymbol{r}, t) = M_z(\boldsymbol{r}, 0)e^{-t/T_1(\boldsymbol{r})} + M_0(\boldsymbol{r})\left[1 - e^{-t/T_1(\boldsymbol{r})}\right]. \tag{2.26}$$

where $t = 0$ corresponds to the moment immediately following excitation. Hence, $M_z(\boldsymbol{r}, t)$ experiences exponential regrowth towards the equilibrium magnetization, as illustrated in Figure 2.7. Typical values for T_1 in various human tissues[2] are provided in Table 2.2.

As the precessional frequency of an individual spin is directly related to the magnitude of the magnetic field it experiences (see § 2.1.2), the temporal and spatial fluctuations in the microscopic field lead to random local variations in the spin precessional frequencies. Consequently, the phase coherence of the excited spins contributing to the net transverse magnetization deteriorates over time. When combined with the thermal coupling of the spin motion with the surrounding medium, the dephasing of individual spins leads to a decay of the transverse magnetization $M_{xy}(\boldsymbol{r}, t)$ described by the differential equation

$$\frac{dM_{xy}(\boldsymbol{r}, t)}{dt} = -\frac{1}{T_2(\boldsymbol{r})}M_{xy}(\boldsymbol{r}, t), \tag{2.27}$$

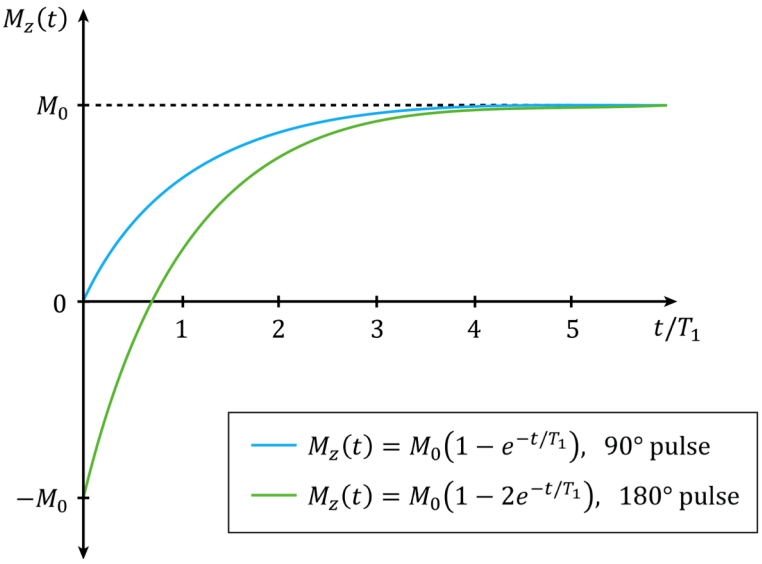


FIGURE 2.7: Relaxation curves for the net longitudinal magnetization following RF excitation pulses with 90° and 180° flip angles.

TABLE 2.2: Typical values for the longitudinal and transverse relaxation time constants in human tissues. The values given correspond to a field strength of $B_0 = 1.5\,\mathrm{T}$ and temperature of 37 °C.

Tissue	T_1 (ms)	T_2 (ms)
Gray matter	950	100
White matter	600	80
Muscle	900	50
Cerebrospinal Fluid	4500	2200
Fat	250	60
Arterial blood	1200	200
Venous blood	1200	100

where $T_2(\boldsymbol{r})$ is the spatially dependant transverse (or spin-spin) relaxation time, satisfying $T_2(\boldsymbol{r}) < T_1(\boldsymbol{r})$. The solution to this equation is given by

$$\boldsymbol{M_{xy}}(\boldsymbol{r}, t) = \boldsymbol{M_{xy}}(\boldsymbol{r}, 0)e^{-t/T_2(\boldsymbol{r})}. \tag{2.28}$$

Thus, in the presence of a longitudinal applied magnetic field, the transverse magnetization diminishes via exponential decay, as illustrated in Figure 2.8. Typical values for T_2 are provided in Table 2.2.

In practice, imperfections in the background field due to external sources, such as errors in the magnet design (see § 2.3), lead to additional static inhomogeneities that give rise to faster dephasing of $M_{xy}(\boldsymbol{r}, t)$. Consequently, the intrinsic relaxation time $T_2(\boldsymbol{r})$ in Eq. (2.28) can be replaced with a smaller effective time constant $T_2^*(\boldsymbol{r})$ given by

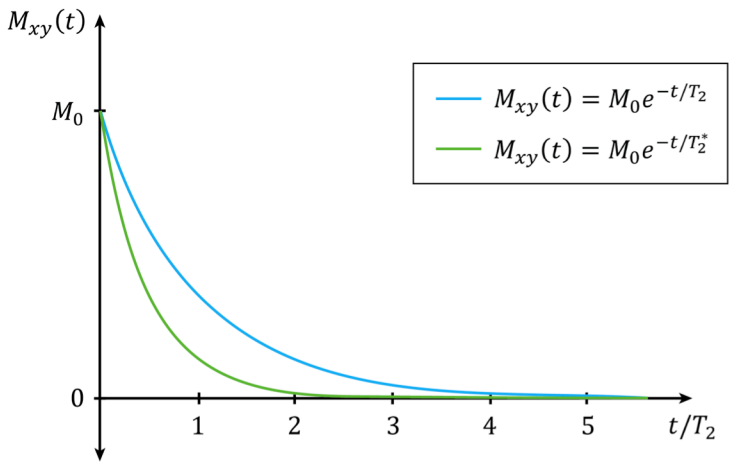


FIGURE 2.8: Relaxation curves for the transverse magnetization following an RF excitation pulse with a 90° flip angle. The magnitude of the transverse magnetization decays more rapidly in the presence of extrinsic magnetic field inhomogeneities, as characterized by the relaxation time constant $T_2^* < T_2$.

$$\frac{1}{T_2^*(\boldsymbol{r})} = \frac{1}{T_2(\boldsymbol{r})} + \frac{1}{T_2'(\boldsymbol{r})} \tag{2.29}$$

where $T_2'(\boldsymbol{r})$ is the relaxation time specifically associated with the static field inhomogeneities. Fortunately, certain imaging sequences can be employed to reverse the dephasing effects characterized by $T_2'(\boldsymbol{r})$ (see § 2.2.5). However, the loss of $M_{xy}(\boldsymbol{r}, t)$ due to intrinsic dephasing may only be recovered by allowing the system to reach complete thermal equilibrium and reinitiating the excitation process.

Differences in the density, microscopic structure, and unique molecular environment amongst various tissues in the human body give rise to spatially varying relaxation times, thereby providing a unique mechanism for generating image contrast that is not possible with other conventional imaging methods.

2.1.6 Signal Acquisition

Consider a generalized longitudinal magnetic field $\boldsymbol{B}(\boldsymbol{r}, t)$ given by

$$\boldsymbol{B}(\boldsymbol{r}, t) = \boldsymbol{B_0} + \boldsymbol{B}'(\boldsymbol{r}, t) = [B_0 + B'(\boldsymbol{r}, t)]\hat{\boldsymbol{z}}, \tag{2.30}$$

where $\boldsymbol{B}'(\boldsymbol{r}, t)$ represents a time and spatially varying component superimposed over the static uniform $\boldsymbol{B_0}$. Following excitation, the net transverse magnetization $\boldsymbol{M_{xy}}(\boldsymbol{r}, t)$ in the voxel located at the point $\boldsymbol{r} = (x, y, z)$ will precess according to

$$\begin{aligned} \boldsymbol{M_{xy}}(\boldsymbol{r}, t) = M_{xy}(\boldsymbol{r}, 0) e^{-t/T_2^*(\boldsymbol{r})} \times \\ \{\cos[\boldsymbol{\phi}(\boldsymbol{r}, t)]\, \hat{\boldsymbol{x}} + \sin[\boldsymbol{\phi}(\boldsymbol{r}, t)]\, \hat{\boldsymbol{y}}\}, \end{aligned} \tag{2.31}$$

with the motion of individual spins governed by Eq. (2.5). In this equation, $M_{xy}(\boldsymbol{r}, 0)$ is the transverse magnetization immediately following excitation, and the accumulated phase $\phi(\boldsymbol{r}, t)$ is given by

$$\phi(\boldsymbol{r}, t) = -\int_0^t \omega(\boldsymbol{r}, t')dt', \tag{2.32}$$

where the generalized Larmor frequency $\omega(\boldsymbol{r}, t)$ is

$$\omega(\boldsymbol{r}, t) = \gamma B(\boldsymbol{r}, t) = \omega_0 + \omega'(\boldsymbol{r}, t). \tag{2.33}$$

Similar to $\boldsymbol{B_1}(t)$ as defined in Eq. (2.14), a component of $\boldsymbol{M_{xy}}(\boldsymbol{r}, t)$ oscillates along the x direction. Therefore, the same RF coil apparatus used to generate $\boldsymbol{B_1}(t)$ can also be used to detect $\boldsymbol{M_{xy}}(\boldsymbol{r}, t)$. Faraday's law of electromagnetic induction states that the electromotive force $\mathcal{E}$ induced in the RF coil is given by[5]

$$\mathcal{E} = -\frac{d}{dt}\Phi_{\mathcal{E}}(t), \tag{2.34}$$

where $\Phi_{\mathcal{E}}(t)$ is the time dependent flux passing through the coil probe due to the rotating magnetization $\boldsymbol{M_{xy}}(\boldsymbol{r}, t)$. Thus, it can be shown that the measured electronic signal, denoted by the complex output $\mathcal{s}(t)$, is given by[2]

$$\mathcal{s}(t) = \int_{\mathcal{V}} \rho_{3\mathrm{D}}(\boldsymbol{r}) e^{-t/T_2^*(\boldsymbol{r})} e^{i\phi(\boldsymbol{r},t)}\, d\mathcal{V}'. \tag{2.35}$$

where $\mathcal{V}$ is the volume of the entire subject. The effective spin density $\rho_{3\mathrm{D}}(\boldsymbol{r})$ appearing in this equation is defined as

$$\rho_{3D}(\boldsymbol{r}) = \omega_0 \Lambda M_{xy}(\boldsymbol{r}, 0), \tag{2.36}$$

where the constant Λ entails all of the gain factors associated with the electronic detection system.

In a simple NMR spectroscopy experiment, information regarding the presence and relative abundances of various nuclear species within a given sample can be extracted via direct Fourier analysis of the measured signal $\mathcal{s}(t)$. However, in order to determine the spatial distribution of a particular species in one or more dimensions, a method within the NMR framework is required for encoding the individual spin locations.

2.2 IMAGE FORMATION

2.2.1 Linear Field Gradients and *k*-Space

Suppose the spatially and time varying longitudinal field $\boldsymbol{B}'(\boldsymbol{r}, t)$ appearing in Eq. (2.30) is defined as

$$\boldsymbol{B}'(\boldsymbol{r}, t) = [\boldsymbol{r} \cdot \boldsymbol{G}(t)]\hat{\boldsymbol{z}} = \left[xG_x(t) + yG_y(t) + zG_z(t)\right]\hat{\boldsymbol{z}}, \tag{2.37}$$

where $\boldsymbol{G}(t) = \nabla \cdot \boldsymbol{B}'(\boldsymbol{r}, t)$ is the spatially independent linear field gradient. With reference to Eq. (2.32), the phase accumulated by the spins within this field is given by

$$\phi(\boldsymbol{r}, t) = -\omega_0 t - \gamma \boldsymbol{r} \cdot \int_0^t \boldsymbol{G}(t')dt'. \tag{2.38}$$

By introducing the parameter $\boldsymbol{k}$ defined as[6]

$$\boldsymbol{k} = \gamma \int_0^t \boldsymbol{G}(t')dt', \tag{2.39}$$

the accumulated phase becomes

$$\phi(\boldsymbol{r}, t) = -\omega_0 t - 2\pi \boldsymbol{k} \cdot \boldsymbol{r}. \tag{2.40}$$

Thus, by imposing the linearly varying magnetic field $\boldsymbol{B}'(\boldsymbol{r}, t)$, the locations of the individual spins are encoded by their time-dependent phase. In particular, a one-to-one relationship is established between the spin position along any three-dimensional line and the phase (or frequency) of the spin precession.

If decay of the transverse magnetization is neglected, the acquired signal from Eq. (2.35) can be rewritten as

$$\mathcal{s}(\boldsymbol{k}) = \iiint_{-\infty}^{\infty} \rho_{3\mathrm{D}}(\boldsymbol{r}) e^{-i2\pi \boldsymbol{k}\cdot\boldsymbol{r}} dx\, dy\, dz. \tag{2.41}$$

where the integration limits have been extended over all space and it is assumed that the ω_0 frequency variation has been removed via demodulation. This equation exactly defines the three-dimensional Fourier transform of $\rho_{3\mathrm{D}}(\boldsymbol{r})$, denoted as

$$\mathcal{s}(\boldsymbol{k}) = \mathcal{F}\{\rho_{3\mathrm{D}}(\boldsymbol{r})\}, \tag{2.42}$$

where it is understood that $\boldsymbol{r}$ and $\boldsymbol{k}$ are conjugate variables. Hence, the three-dimensional distribution of the effective spin density can then be determined by the inverse Fourier transform of the measured MRI signal, according to

$$\rho_{3\mathrm{D}}(\boldsymbol{r}) = \mathcal{F}^{-1}\{\mathcal{s}(\boldsymbol{k})\} = \iiint_{-\infty}^{\infty} \mathcal{s}(\boldsymbol{k}) e^{i2\pi \boldsymbol{k}\cdot\boldsymbol{r}} dk_x\, dk_y\, dk_z. \tag{2.43}$$

Thus, the problem of determining $\rho_{3\mathrm{D}}(\boldsymbol{r})$ amounts to evaluating this integral over the domain of $\mathcal{s}(\boldsymbol{k})$, commonly referred to as k-space.

Due to practical time constraints, $\mathcal{s}(\boldsymbol{k})$ can only be determined at a finite number of points. Accordingly, a truncated and discretized version of $\mathcal{s}(\boldsymbol{k})$ is obtained in practice. Provided $\mathcal{s}(\boldsymbol{k})$ is sampled adequately, such that the Nyquist sampling criterion is satisfied (see § 2.2.6), faithful reconstruction of $\rho_{3\mathrm{D}}(\boldsymbol{r})$ may be possible by converting Eq. (2.43) into a discrete Fourier transform.

2.2.2 Slice Selection

If the excitation of spins is restricted to a single plane of finite thickness, then the task of spatial localization reduces to a two-dimensional (2D) problem. This can be accomplished by applying a linear field gradient along the direction orthogonal to the desired slice while simultaneously delivering a spatially selective excitation pulse, as illustrated in Figure 2.9.

In the presence of a slice-select gradient given by $\boldsymbol{G} = G_z\hat{\boldsymbol{z}}$, the spin precessional frequency defined in Eq. (2.33) becomes spatially dependant along the z direction according to

$$\omega(z) = \gamma B(z) = \omega_0 + \gamma z G_z. \tag{2.44}$$

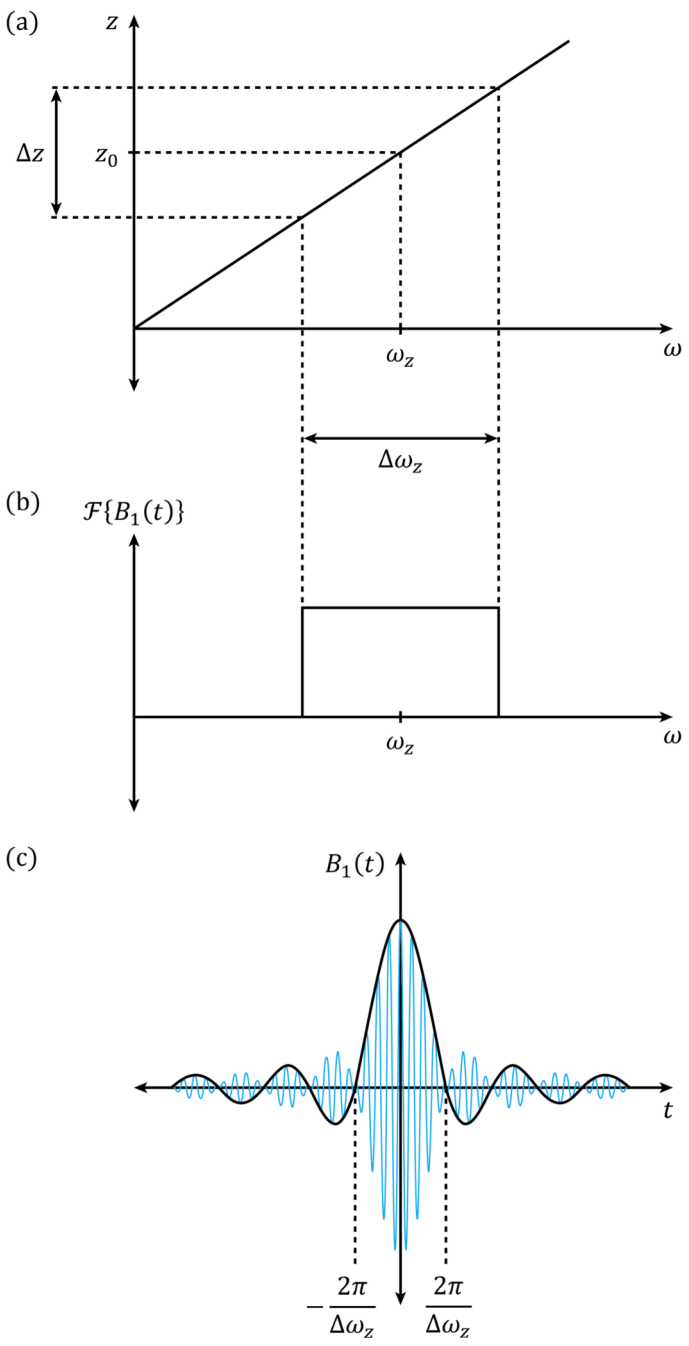


FIGURE 2.9: (a) In the presence of a linear field gradient $\boldsymbol{G} = G_z\hat{\boldsymbol{z}}$, the frequency of precession becomes linearly related to the spin locations along the z direction. (b) An RF excitation pulse with central frequency ω_z and spectral width $\Delta\omega_z$ may be used to selectively excite a slice of spins centered at z_0 and of thickness Δz. (c) The corresponding modulated RF excitation pulse is illustrated in the time domain.

Therefore, a slice of spins centered at z_0 and of thickness Δz can be uniformly excited by applying an excitation pulse possessing a rectangular frequency profile with center ω_z and width $\Delta\omega_z$ given by

$$\omega_z = \omega_0 + \gamma z_0 G_z, \tag{2.45a}$$

$$\Delta\omega_z = \gamma \Delta z G_z. \tag{2.45b}$$

With the transverse magnetization $M_{xy}(\boldsymbol{r}, 0)$ nonzero only within the slice of excited spins, Eq. (2.41) reduces to the 2D Fourier transform

$$\begin{aligned} \mathcal{s}(k_x, k_y) &= \iint_{-\infty}^{\infty} \rho_{2\mathrm{D}}(x, y) e^{-i2\pi(k_x x + k_y y)} dx\, dy, \\ &= \mathcal{F}\{\rho_{2\mathrm{D}}(x, y)\}, \end{aligned} \tag{2.46}$$

where the effective spin density within the slice is given by

$$\rho_{2\mathrm{D}}(x, y) = \int_{z_0 - \Delta z/2}^{z_0 + \Delta z/2} \rho_{3\mathrm{D}}(x, y, z) dz. \tag{2.47}$$

Hence, $\rho_{2\mathrm{D}}(x, y)$ can now be obtained via the 2D inverse Fourier transform:

$$\rho_{2\mathrm{D}}(x, y) = \mathcal{F}^{-1}\{\mathcal{s}(k_x, k_y)\}. \tag{2.48}$$

2.2.3 Frequency and Phase Encoding

Spatial localization of the spins within an excited slice can be achieved by manipulating the field gradients along the remaining x and y directions spanning the slice plane. In light of the integral defining $\boldsymbol{k}$ in Eq. (2.39), the manner in which these gradients are varied in time characterizes how the

MRI signal is sampled within the two-dimensional k-space spanned by k_x and k_y.

First consider the phase encoding (PE) gradient $\boldsymbol{G} = G_y\hat{\boldsymbol{z}}$ applied to a selectively excited slice for a length of time τ_{PE}. During the application of this gradient pulse, the accumulated phase within the slice becomes position dependant in accordance with Eq. (2.38). Therefore, after the PE gradient is removed, the value of k_y is given by

$$k_y = \gamma G_y \tau_{\mathrm{PE}}. \tag{2.49}$$

Similarly, consider next the application of a frequency encoding (FE) gradient $\boldsymbol{G} = G_x\hat{\boldsymbol{z}}$ for the length of time τ_{FE}, after which the value of k_x is given by

$$k_x = \gamma G_x \tau_{\mathrm{FE}}. \tag{2.50}$$

Thus, if the MRI signal is acquired following these gradient pulses, then a sample of $\mathcal{s}(k_x, k_y)$ is directly obtained at a single point. Furthermore, if the MRI signal is continuously acquired during the time interval for which the FE gradient is active, then samples of $\mathcal{s}(k_x, k_y)$ are obtained along a complete line segment within k-space defined by

$$\boldsymbol{k} = \begin{bmatrix} k_x \\ k_y \end{bmatrix} = \begin{bmatrix} 0 \\ \gamma G_y \tau_{\mathrm{PE}} \end{bmatrix} + t \begin{bmatrix} \gamma G_x \\ 0 \end{bmatrix}, \quad 0 \le t \le \tau_{\mathrm{FE}}, \tag{2.51}$$

where $t = 0$ corresponds to the moment when the FE gradient is initially applied. Since sampling of the MRI signal along this segment is performed in a consecutive fashion, the spatial direction associated with the FE gradient is commonly referred to as the readout direction.

By iterating selective slice excitation followed by the spatial encoding of the spins via the pulsed gradient method discussed above, a framework for the complete sampling of $\mathcal{s}(k_x, k_y)$ is established. In the following subsections, two common imaging sequences that are based on this framework are described.

2.2.4 The Gradient Echo Method

The gradient echo (GRE) method is a complete procedure for generating magnetic resonance (MR) images based on the Fourier transform approach introduced in the previous sections (see § 2.2.1). A key feature of this method is the use of a bipolar FE gradient along the readout direction. The first (dephasing) lobe of this gradient pulse acts to dephase the transverse magnetization in accordance with Eq. (2.38), which in turn leads to a reduction of the MRI signal amplitude. The second (rephasing) lobe of the FE gradient is of opposite polarity, thereby acting to refocus the dephasing induced during the prior lobe. When the general gradient echo condition is satisfied:

$$\int_0^{\boldsymbol{t}} G_{\boldsymbol{x}}(t')dt' = 0, \tag{2.52}$$

then the net accumulated phase due to the FE gradient is nulled for all spins. Consequently, the transverse magnetization is partially recovered and the signal amplitude peaks, forming what is known as a gradient echo. The precise moment at which this occurs is taken to be the echo time, denoted $t = TE$.

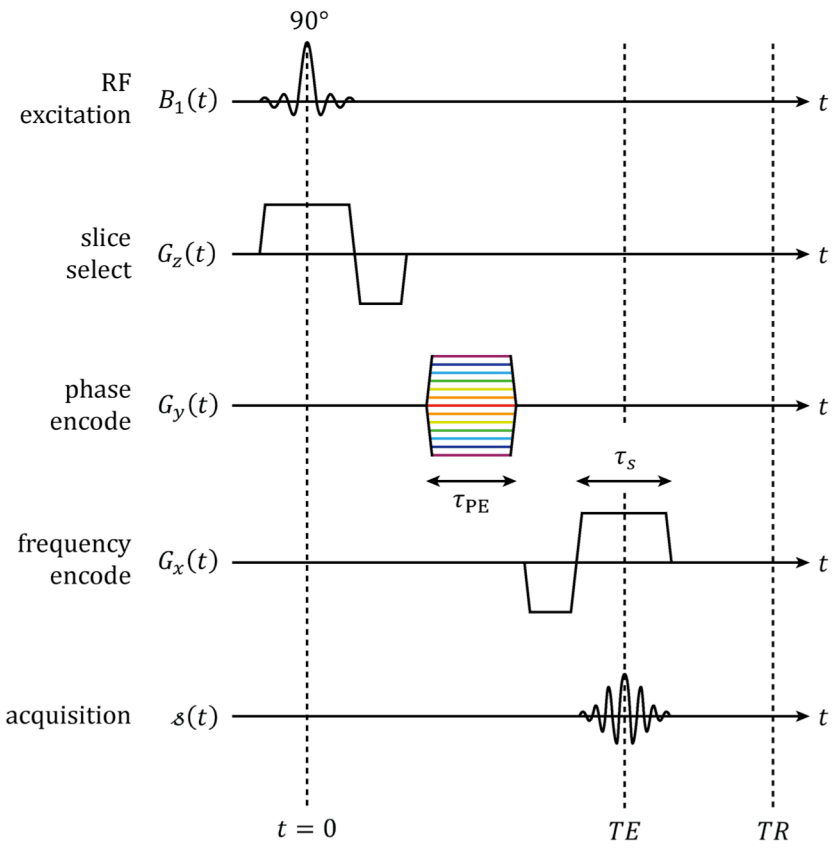


FIGURE 2.10: The pulse sequence diagram for a basic gradient echo imaging sequence. After the repetition time $t = TR$ has elapsed, the sequence is repeated with the magnitude of the phase encode gradient incremented along the coloured steps shown.

The essential steps of a basic GRE imaging sequence are summarized in the pulse sequence timing diagram provided in Figure 2.10. The corresponding k-space trajectory is illustrated in Figure 2.11. An initial net transverse magnetization is generated within a desired slice through the use of a spatially selective RF excitation pulse (see § 2.2.2). This is followed by a PE gradient of fixed duration τ_{PE}, such that k_y is given by Eq. (2.49). The bipolar FE gradient described above is subsequently applied, with acquisition of the MRI signal taking place during the second gradient lobe of duration τ_s. Since the dephasing lobe of the FE gradient has a negative amplitude, the resulting

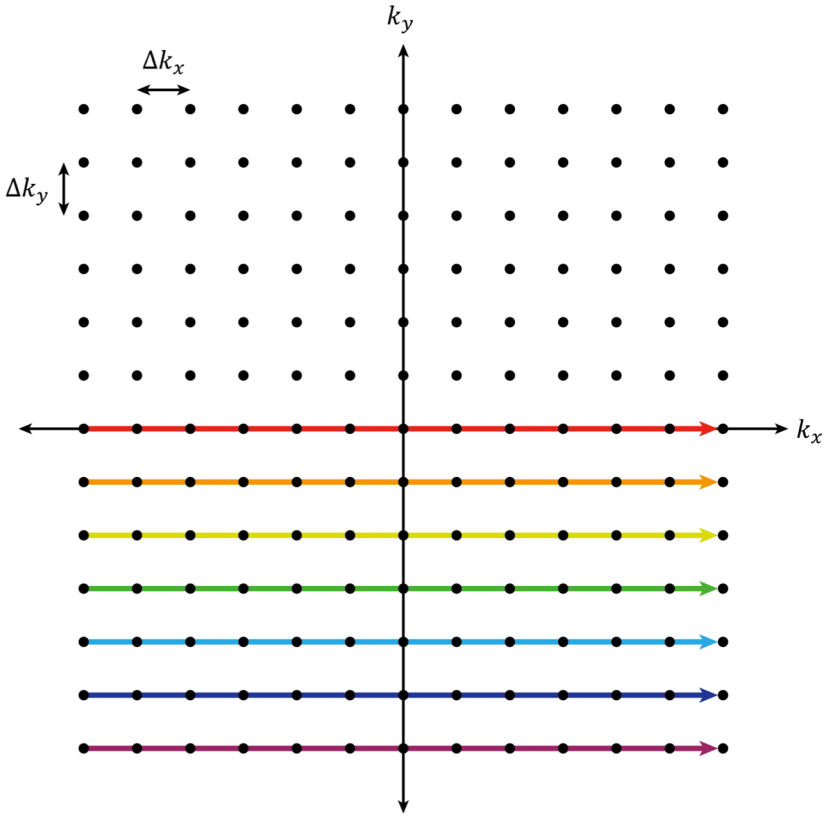


FIGURE 2.11: The k-space trajectory for a basic gradient echo imaging sequence. The illustrated dots represent the points at which the MRI signal is sampled. Each coloured line segment represents the portion of the trajectory associated with a single phase encoding step. The line segment arrows indicate the directions along which the samples are obtained during signal acquisition. This k-space trajectory is also applicable for a basic spin echo imaging sequence.

starting point of the associated k-space trajectory is specified by a negative k_x value in accordance with Eq. (2.39). The MRI signal is then consecutively sampled along the symmetric k_x-directed line segment, with the gradient echo forming at the moment Eq. (2.52) is satisfied and the line $k_x = 0$ is passed. After the repetition time $t = TR$ has elapsed, the entire sequence is repeated with the amplitude of the PE gradient incrementally adjusted, thereby specifying a different k_y value and corresponding k_x-directed line

segment in the k-space trajectory. Proceeding in this manner, $\mathcal{s}(k_x, k_y)$ is determined at a number of discrete grid points within k-space, permitting reconstruction of the estimated spin density via the discrete inverse Fourier transform.

Although the formation of a gradient echo involves reversing the phase induced by the first lobe of the FE gradient, the transverse magnetization still decays throughout the sequence due to the independent effects of both intrinsic and extrinsic magnetic field inhomogeneities. Consequently, the MRI signal in the GRE method experiences rapid T_2^* decay following each excitation.

2.2.5 The Spin Echo Method

The spin echo (SE) method is another basic imaging procedure similar to the GRE method, with the exception of an additional RF refocusing pulse that is delivered prior to application of the phase and frequency encoding gradients. By rotating all of the spins within the excited slice through an angle $\theta_{\mathrm{RF}} = \pi$, the refocusing pulse acts to reverse the polarity of any phase accumulated by that point. Hence, if the refocusing pulse is delivered at a time $t = \tau'$ following the initial excitation, then the net dephasing due to static magnetic field inhomogeneities is reversed at the moment $t = 2\tau'$. Similar to the gradient echo described in the previous section, this rephasing of the transverse magnetization leads to a peak in the MRI signal, forming what is known as a spin echo. Furthermore, to maximize the signal amplitude at the midpoint of the acquisition window, the timing of the refocusing pulse is typically specified such that the spin and gradient echoes coincide. Thus, the echo time is given by $t = 2\tau' = TE$.

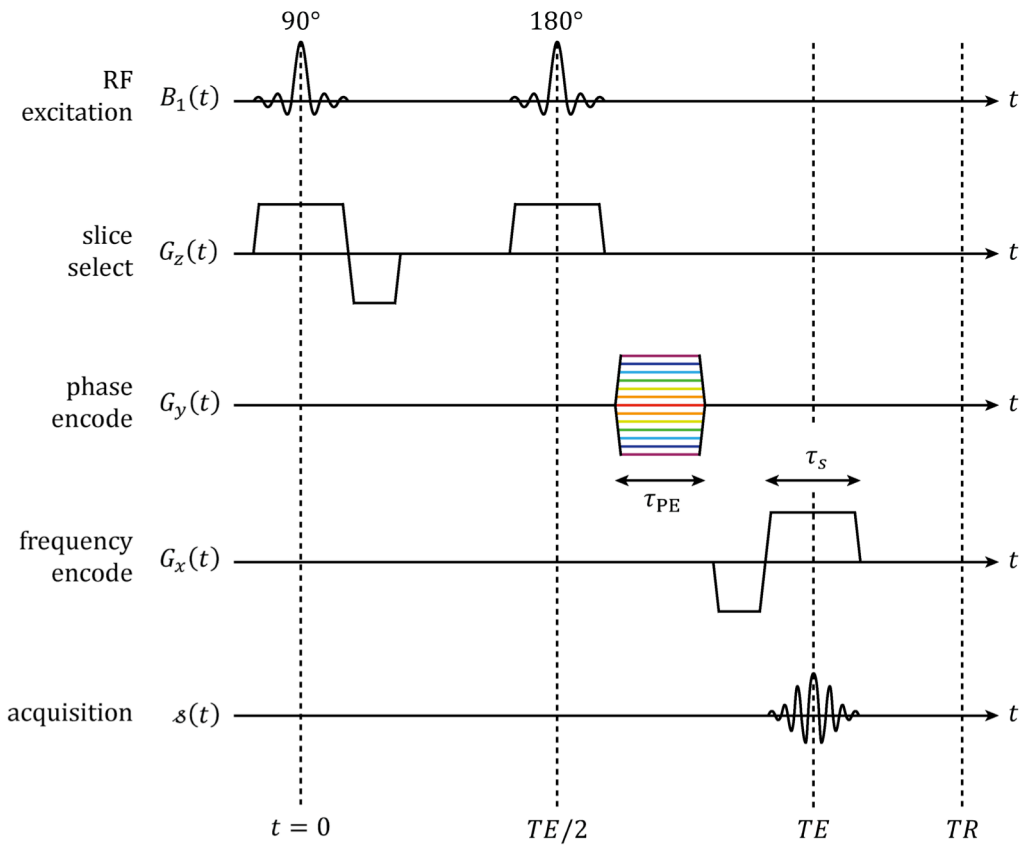


FIGURE 2.12: The pulse sequence diagram for a basic spin echo imaging sequence. After the repetition time $t = TR$ has elapsed, the sequence is repeated with the magnitude of the phase encode gradient incremented along the coloured steps shown.

The essential steps of a basic SE imaging sequence are illustrated in the pulse sequence timing diagram shown in Figure 2.12. As evident in Eq. (2.39), the RF refocusing pulse does not affect the k-space trajectory. Consequently, the GRE and SE methods are described by the same k-space diagram, shown in Figure 2.11.

As the dephasing effects due to static magnetic field inhomogeneities are reversed when a spin echo is formed, decay of the MRI signal in the SE method is characterized by the T_2 time constant (see § 2.1.5). This is

particularly advantageous in situations involving relatively large extrinsic field inhomogeneities, for which a small T_2' value leads to rapid signal loss. However, inclusion of the RF refocusing pulse requires additional time, resulting in longer scans than what may be achieved with other techniques such as the GRE method.

2.2.6 Nyquist Sampling Criteria

Sufficient sampling of the MRI signal $\mathcal{s}(k_x, k_y)$ must be performed if aliasing artifacts are to be avoided in the reconstructed image. This amounts to the well-known Nyquist sampling criterion, which requires that[2]

$$\Delta k_j \leq \frac{1}{A_j}, \quad \text{for } j = x, y, \tag{2.53}$$

where A_j represents the maximum dimension of the imaged object along the spatial direction j associated with the sample separation Δk_j.

For a fixed duration rectangular PE gradient pulse corresponding to k_y in Eq. (2.49), the sample separation Δk_y is given by

$$\Delta k_y = \gamma \Delta G_y \tau_{PE}, \tag{2.54}$$

where ΔG_y is the increment in the gradient amplitude between successive excitations (Figure 2.11). Similarly, for a fixed amplitude FE gradient pulse corresponding to k_x in Eq. (2.50), the sample separation Δk_x is given by

$$\Delta k_x = \gamma G_x \Delta \tau_s, \tag{2.55}$$

where $\Delta\tau_s$ is the time separation of the consecutively acquired samples. Hence, Eqs. (2.54) and (2.55) may be substituted into Eq. (2.53), yielding

$$\gamma G_x \Delta\tau_s A_x \leq 1, \tag{2.56a}$$

$$\gamma \Delta G_y \tau_{\mathrm{PE}} A_y \leq 1. \tag{2.56b}$$

Assuming that the object is resolved with n_j pixels of width Δj along the spatial direction j, such that $A_j = n_j \Delta j$, then the Nyquist criteria may be expressed as

$$\gamma G_x \Delta\tau_s n_x \Delta x \leq 1, \tag{2.57a}$$

$$\gamma \Delta G_y \tau_{\mathrm{PE}} n_y \Delta y \leq 1. \tag{2.57b}$$

2.2.7 Consequences of Magnetic Field Inhomogeneity

The principal assumption of spatial localization in MRI is the one-to-one linear relationship between the position of spins and their phase of precession. Specifically, if a known linear field gradient $\boldsymbol{G}$ is imposed on a perfectly uniform static magnetic field $\boldsymbol{B_0}$, then the individual spin locations may be discriminated based on their position-dependent phase given by Eq. (2.38). The basis for this approach is the fundamental property that the Larmor precessional frequency in Eq. (2.33) is directly proportional to the local magnetic field strength. Consequently, any macroscopic field variations in addition to the applied linear field gradient leads to the misregistration of spin positions, as their frequency of precession deviates from that assumed by the image reconstruction algorithm. This effect is manifested in the resultant images through geometric distortions, as the calculated spin density differs spatially from the true physical object.

The nature and severity of the geometric distortions in MRI are strongly dependent on the gradient parameters and the specific imaging sequence employed. This dependence can be analyzed by considering how the phase of precession at the point (x, y) within an excited slice evolves as the MRI signal is sampled along each orthogonal k-space direction. Specifically, for the gradient and spin echo methods in the absence of magnetic field inhomogeneities, the differences in accumulated phase $\Delta\phi_{k_x}(x, y)$ and $\Delta\phi_{k_y}(x, y)$ between consecutive frequency and phase encoding samples are given by

$$\Delta\phi_{k_x}(x, y) = x\Delta k_x = \gamma x G_x \Delta\tau_s, \tag{2.58a}$$

$$\Delta\phi_{k_y}(x, y) = y\Delta k_y = \gamma y \Delta G_y \tau_{\mathrm{PE}}. \tag{2.58a}$$

Similarly, in the presence of the additional magnetic field variation $\delta B(x, y)$, the phase differences $\Delta\tilde{\phi}_{k_x}(x, y)$ and $\Delta\tilde{\phi}_{k_y}(x, y)$ between adjacent samples along the frequency and phase encoding directions are given by

$$\Delta\tilde{\phi}_{k_x}(x, y) = \gamma[x G_x \Delta\tau_s + \delta B(x, y)\Delta\tau_s] = \Delta\phi_{k_x}(\tilde{x}, y), \tag{2.59a}$$

$$\Delta\tilde{\phi}_{k_y}(x, y) = \gamma y \Delta G_y \tau_{\mathrm{PE}} = \Delta\phi_{k_x}(\tilde{x}, y), \tag{2.59b}$$

where

$$\tilde{x} = x + \frac{\delta B(x, y)}{G_x}. \tag{2.60}$$

Hence, under the influence of the additional field variation $\delta B(x, y)$, the signal originating from the point (x, y) is incorrectly registered to the false position $(\tilde{x}, y)$. That is, the presence of $\delta B(x, y)$ directly leads to geometric distortions along the FE direction, whereas the PE direction remains

unaffected. Furthermore, it is evident from Eq. (2.60) that the extent of distortion is inversely proportional to the strength of the FE gradient. Thus, a reduction in this distortion can be always achieved by increasing G_x. However, according to the Nyquist criterion Eq. (2.57a), this would also necessitate a similar increase in the bandwidth of the acquisition hardware, as the sampling interval $\Delta\tau_s$ would have to be reduced in order to avoid aliasing.

In practice, nonuniformity in the static magnetic field may arise from various sources. Typically, the dominant culprit is the imperfect design or manufacturing of the primary magnet assembly generating the external magnetic field (see § 2.3). This may include errors in the winding or positioning of the magnet coils, magnetic impurities in the surrounding yoke structure and poles, or the shifting of components under the influence of magnetic forces. In addition, if the imaging subject is comprised of a heterogeneous distribution of magnetic susceptibilities, then the subject itself can induce significant variations in the local magnetic field.[7]

2.3 THE MAIN MAGNET

2.3.1 General Design Requirements

Magnetic resonance imaging critically relies on the availability of a strong static longitudinal magnetic field $\boldsymbol{B_0}$ for the primary purpose of magnetically polarizing the imaging subject. When immersed in this external magnetic field, the individual spin orientations are biased towards a direction parallel to $\boldsymbol{B_0}$, yielding a net longitudinal magnetization $\boldsymbol{M_z}$ that is subsequently manipulated to produce the encoded MRI signal used for image reconstruction. In the following discussion, the general requirements of the static magnetic field as related to the main magnet design are introduced.

The signal to noise ratio (SNR) for an imaging experiment is a key parameter describing the quality of the images produced. The SNR essentially quantifies the extent to which the signal generated by an object can be distinguished from the background noise. Maximizing the SNR is of central importance in MRI, as it leads to improved sensitivity while permitting greater resolution and faster scan times. The SNR in a magnetic resonance image is directly related to the magnetic field strength according to the approximate proportionality[2]

$$\text{SNR} \propto \begin{cases} B_0^{7/4}, & \text{for } B_0 < 0.5 \text{ T}, \\ B_0, & \text{for } B_0 \geq 0.5 \text{ T}. \end{cases} \tag{2.61}$$

Consequently, a primary requirement of the main magnet system is that the field produced should be as strong as possible (within the practical limits of modern magnet design). At present, typical values for clinical field strengths range from 0.1 to 3.0 T.

In order to facilitate the reliable diagnosis of disease or accurate treatment guidance, it is critical that reconstructed MR images form a faithful representation of the true patient anatomy. Consequently, obtaining geometric accuracy through the reduction of image distortions is an important concern in clinical MRI. As previously discussed, aberrations in the static magnetic field are directly translated into geometric distortions in MR images (see § 2.2.7). Hence, minimizing inhomogeneities in the static magnetic field is another primary requirement in the design of the main magnet system. For a practical scanner with a gradient strength of 25 mT/m, the geometric distortion along the FE direction given by Eq. (2.60) amounts to 4 cm/mT. Thus, for a 20×20 cm^2 field of view consisting of 128×128 pixels, the maximum permissible field variation would be approximately 40 μT if the maximum geometric distortion is to be less than the 1.6 mm

width of a single pixel. Evidently, image distortions in MRI are extremely sensitive to variations in the static magnetic field, leading to a required level of field uniformity that is extremely high. Accordingly, the relative magnetic field inhomogeneity ΔB over a region of interest (ROI) is typically quoted in terms of the parts per million (ppm) fractional variation given by:

$$\Delta B = \frac{B_{\max} - B_{\min}}{B_0} \cdot 10^6. \tag{2.62}$$

In this equation, $B_{\max}$ and $B_{\min}$ are the maximum and minimum field values within the ROI, commonly defined as a large diameter sphere at isocenter. The inhomogeneity at a given point in space can also be determined with the similar expression

$$\Delta B(\boldsymbol{r}) = \frac{B(\boldsymbol{r}) - B_0}{B_0} \cdot 10^6. \tag{2.63}$$

It often suffices to strictly consider the axial or longitudinal component of the static magnetic field when calculating these expressions, since the contribution of the transverse component in a homogenous magnet design can be considered negligible in comparison.[8] Namely, if the relative error in the axial field component is of order $\epsilon \ll 1$, then the relative error in the transverse field components will be of order $\epsilon^2 \ll \epsilon$.

Whole-body magnets capable of generating the strong uniform magnetic field required by MRI are typically based on either cylindrical or biplanar geometries, as illustrated in Figure 2.13. If these magnets were of infinite extent, then the magnetic field within the ROI would theoretically be perfectly uniform. Of course, due to practical constraints on the size and cost of these systems, the extent of the magnet geometry must be limited. Consequently, obtaining a level of uniformity suitable for medical imaging

requires the intelligent distribution of the magnet sources surrounding the ROI, which is generally determined with the aid of computer-based design optimization algorithms.

A common property of all magnetic fields is that they extend for infinite distances and permeate all space. For the simple magnet designs illustrated in Figure 2.13, the magnitude of the fringe magnetic field approximately decreases with the cube of the distance from the magnet sources. Due to the high field strength generated at the center of a typical MRI magnet, the strength of this fringe field in an unshielded system[2] may be in excess of 5 G (an order of magnitude larger than the Earth's magnetic field) at distances as large as 10 to 30 m. Consequently, these systems pose a significant risk of hazardous interaction with magnetic materials and sensitive electronic devices that are located in proximity to the imaging area.[9]

A general requirement for the installation of MRI systems is that the magnitude of the fringe magnetic field must be limited to less than 5 G at locations beyond the boundary of the imaging suite. Although this may be achieved by employing a suitable amount of magnetic material in the construction of the surrounding walls, the more common strategy is to integrate substantial magnetic shielding into the design of the main magnet system itself. In the approach of active shielding, the magnet design is supplemented with additional shielding coils surrounding the primary sources. The direction of the current in the shielding coils is chosen such that the magnetic field they generate is opposed to that produced by the primary sources, leading to a significant cancellation of the fringe magnetic field. In the alternative approach of passive shielding, a large magnetic yoke structure surrounding the primary sources is employed instead. This yoke structure is necessarily composed of magnetic material with high permeability (such as steel), thus forming a magnetic flux circuit that acts to confine the fringe fields produced by the magnet system.

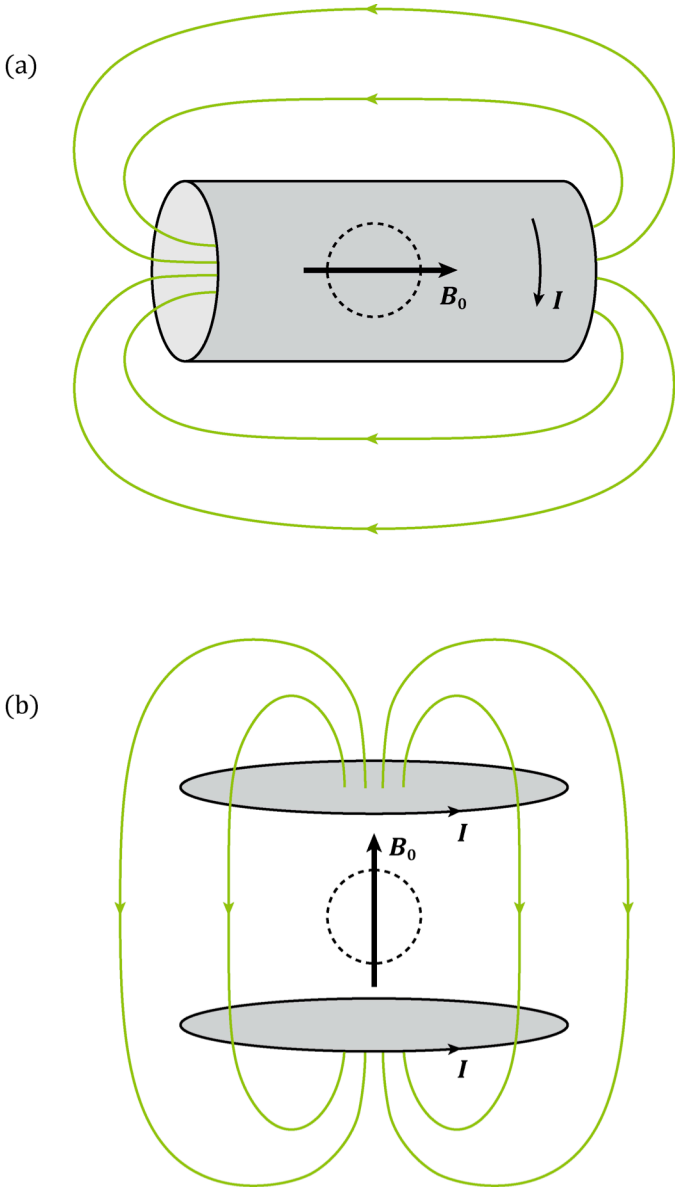


FIGURE 2.13: Full-body magnets for clinical MRI are generally based on (a) cylindrical or (b) biplanar geometries. The dotted circle indicates the typical location of the imaging volume for each geometry. Although the main magnet assembly is usually composed of several bulk coils (in the case of a superconducting magnet system), simplified current distributions associated with the illustrated orientations of the static magnetic field are provided. In addition, portions of the magnetic fringe fields are depicted in green.

2.3.2 Superconducting Magnet Systems

The vast majority of whole-body MRI magnet systems employ large coils constructed from superconducting material in order to generate the strong homogenous static magnetic field that is central to the imaging experiment. As of 2005, there had been over 15 000 superconducting MRI systems installed world-wide.[10] These systems are capable of greater field strengths, temporal stability, and field uniformity than their permanent magnet counterparts. These highly advantageous features permit faster imaging, greater resolution, and superior image quality. When cooled below their critical temperature, the superconducting coils used in these magnet systems adopt the property of zero electrical resistance. When functioning in this superconducting state, these systems are capable of supporting profoundly large and stable currents in a persistent mode of operation without the need of an external power supply. Due to the general desire for improved image quality in routine clinical imaging, there exists a continuous thrust to optimize the design of superconducting magnets with increased field strengths, maximal uniformity, and cost efficiency.[11–20]

At present, the most popular superconducting material used in MRI magnets is the metallic alloy of niobium titanium (NbTi). With a critical temperature of 9.3 K, this material must be cooled to cryogenic temperatures in order to support the high current densities and external magnetic fields that are characteristic of MRI applications. In general, this is achieved by submerging the coils in a liquid helium (LHe) bath, which is maintained at an ultra low temperature of 4.2 K within a large encapsulating cryostat vessel.

Due to the attractive possibility of eliminating the requirement of large and expensive cryogenic cooling systems, high-temperature superconducting materials are currently being explored for widespread use in MRI magnet systems.[21] In particular, of growing interest is the intermetallic compound of

magnesium diboride (MgB_2), for which the property of superconductivity was recently discovered.[22] With a relatively high critical temperature of 39 K, this material is suitable for use in a cryogen-free conduction-cooled environment.[23]

The majority of superconducting magnets for MRI are of cylindrical geometry and have field strengths in the range of 1.5 to 3.0 T. These systems typically employ active shielding, with the secondary shielding coils placed at a larger radius than the primary coils. As this shielding solution leads to a reduction in the imaging field strength per unit current, the coils in these systems are necessarily constructed from many windings and are required to carry especially large currents.

Due to growing concerns regarding patient claustrophobia and physician access, there has been a recent thrust towards superconducting magnets with biplanar geometries and open configurations. As a consequence of the large separation between the two magnet poles, these systems are generally limited to field strengths of 0.5 to 1.0 T. Moreover, biplanar systems often consist of large magnetic yoke structures to enhance the magnetic field strength in the region of interest, while aiding in the passive shielding of the fringe magnetic fields.

2.3.3 Permanent Magnet Systems

In dedicated applications for which the strength and complexity of superconducting systems is unnecessary and inconvenient, permanent magnet systems are becoming a simple and affordable alternative of growing popularity. These systems often possess biplanar geometries, consisting of large opposing magnetic poles composed of a rare-earth ferromagnetic material. The most common ferromagnetic material used in modern designs

is a neodymium-iron-boron compound ($Nd_2Fe_{14}B$), which typically leads to magnetic field strengths in the range of 0.1 to 0.3 T.

Due to the limited field strength attainable with a ferromagnetic source, permanent magnet systems normally involve the use of a large magnetic yoke structure in order to passively shield the stray magnetic fields. Furthermore, large pole plates constructed from a material of high permeability are generally fixed to the opposing surfaces of the magnetic sources. In addition to enhancing the magnetic field strength, the surfaces of these pole plates may be optimally contoured so as to assist in obtaining a suitably homogeneous magnetic field in the imaging volume.

Permanent magnet systems offer the technical advantage that they do not require a dedicated cooling system. This leads to simple and inexpensive maintenance, while alleviating the hazards associated with the risk of a quenching magnet. In addition, the fringe fields they exhibit are comparably weak, making planning and installation of these systems relatively easy. On the other hand, thermal instability of the ferromagnetic sources can result in temperature drift of the magnetic field strength, requiring careful calibration of the associated Larmor frequency. Furthermore, the level of field uniformity obtained in practice is often limited, due to unavoidable impurities in the ferromagnetic sources in addition to the magnetic influence of the surrounding yoke structure.

2.3.4 Shimming Strategies

Despite the quality of any theoretical magnet design, imperfections in the manufacturing process inevitably lead to residual field variations in the region of interest. The magnitude of this residual inhomogeneity is typically on the order of 10^2 to 10^3 ppm. Since medical imaging demands a much lower level of inhomogeneity on the order of a few ppm, a strategy for

measuring and eliminating the residual variations after manufacturing is required. The process that entails these final critical steps during the installation of a magnet system is referred to as shimming.[24–28]

The first step in the shimming process is the mapping of the magnetic field variation within the region of interest. This is typically performed by directly measuring the magnetic field at a collection of points located on the surface of a large diameter sphere or cylinder. An expansion of the magnetic field in terms of the complete basis of spherical harmonics is then calculated (see § 3.1.3). In this way, the contribution of each individual harmonic to the residual inhomogeneity can be determined based on the value of the associated expansion coefficient.

The next step of the shimming process is the determination of a magnetic source configuration that tends to eliminate the majority of the low order harmonics present in the previously calculated field expansion.[24] In a manner similar to the techniques of magnetic shielding, shimming methods can be classified as either active or passive, depending on whether the correction field is generated with current carrying coils or segments of magnetic material.

In the case of active shimming, the additional set of shim coils are typically arranged in a fixed configuration, with the geometry of each coil specifically designed to produce a magnetic field variation corresponding to a single harmonic in the field expansion. By independently manipulating the current in each of the coils, the corresponding component of the field inhomogeneity can be cancelled. In practice, the difficulty in realizing a pure harmonic shim increases with the complexity and order of the targeted harmonic. Specifically, any errors in the design, manufacturing, or positioning of a given shim system will lead to contamination of the correcting field with contributions of lower order than the targeted harmonic.[24] Although the

typical amplitudes of these contaminants are much smaller than that of the desired component, the lower-order terms become increasingly dominant as the origin is neared (see § 3.1.3).

In the alternative approach of passive shimming, a specific configuration of magnetic material is strategically placed in the region surrounding the imaging volume. Unlike the active method described above, modern techniques for passive shimming do not require independent targeting of orthogonal harmonics.[27, 28] Although possible, the generation of pure harmonic passive shims is hindered by considerable uncertainties in the shapes, locations, and magnetic properties of the individual segments comprising the shim. Rather, a total shim configuration as a whole is determined which perturbs the net magnetic field in a fashion that collectively leads to a reduction in the residual inhomogeneities in the region of interest.

Successful shimming of an MRI magnet typically requires several iterations of the measure and correct procedure described above. This systematic and experimental process is generally aided by the use of a computer, which performs the necessary decomposition of the magnetic field based on the measured data, followed by the calculation of the optimal adjustments to the correction system required to improve the homogeneity. Of course, the effectiveness of these shimming techniques in reaching acceptable levels of uniformity is limited by the initial inhomogeneities present after manufacturing. Namely, the ease with which a magnet may be shimmed is limited by the magnitude and complexity of the residual field variation. Ultimately, this leads to a significant constraint on the magnet design that competes with the additional requirement of a suitably large and accessible gap within the magnet for the subject being examined.

2.4 REFERENCES

[1] M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance*, 2nd ed. (Wiley, New York, 2008).

[2] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic Resonance Imaging: Physical Principles and Sequence Design* (Wiley, New York, 1999).

[3] D. J. Griffiths, *Introduction to Quantum Mechanics*, 2nd ed. (Prentice Hall, New Jersey, 2005).

[4] J. Jin, *Electromagnetic Analysis and Design in Magnetic Resonance Imaging* (CRC Press, Boca Raton, 1999).

[5] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).

[6] T. R. Brown, B. M. Kincaid, and K. Ugurbil, "NMR chemical shift imaging in three dimensions," *P. Natl. Acad. Sci. USA*, **79**(11), 3523−3526 (1982).

[7] J. F. Schenck, "The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds," *Med. Phys.* **23**(6), 815−850 (1996).

[8] H. Xu, S. M. Conolly, G. C. Scott, and A. Macovski, "Homogeneous magnet design using linear programming," *IEEE Trans. Magn.* **36**(2), 476−483 (2000).

[9] F. G. Shellock and J. V. Crues, "MR procedures: biologic effects, safety, and patient care," *Radiology*, **232**(3), 635−652 (2004).

[10] Y. Lvovsky and P. Jarvis, "Superconducting systems for MRI- present solutions and new trends," *IEEE Trans. Appl. Supercond.* **15**(2), 1317−1325 (2005).

[11] A. K. Kalafala, "A design approach for actively shielded magnetic resonance imaging magnets," *IEEE Trans. Magn.* **26**(3), no. 3, 1181−1188 (1990).

[12] S. Pissanetzky, "Structured coils for NMR applications," *IEEE Trans. Magn.* **28**(4), 1961−1968 (1992).

[13] V. Cavaliere, A. Formisano, R. Martone, G. Masullo, A. Matrone, and M. Primizia, "Design of split coil magnets for magnetic resonance imaging," *IEEE Trans. Appl. Supercond.* **10**(1), 759−762 (2000).

[14] V. Cavaliere, A. Formisano, R. Martone, and M. Primizia, "A genetic algorithm approach to the design of split coil magnets for MRI," *IEEE Trans. Appl. Supercond.* **10**(1), 1376−1379 (2000).

[15] S. Besio, S. Pittaluga, V. Punzo, and A. Trequattrini, "Elliptical coils for dedicated MRI magnets," *IEEE Trans. Appl. Supercond.* **18**(2), 908−911 (2008).

[16] G. Sinha, R. Sundararaman, and G. Singh, "Design concepts of optimized MRI magnet," *IEEE Trans. Magn.* **44**(10), 2351−2360 (2008).

[17] Q. Wang, G. Xu, Y. Dai, B. Zhao, L. Yan, and K. Kim, "Design of open high magnetic field MRI superconducting magnet with continuous current and genetic algorithm method," *IEEE Trans. Appl. Supercond.* **19**(3), 2289−2292 (2009).

[18] Q. M. Tieng, V. Vegh, and I. M. Brereton, "Minimum stored energy high-field MRI superconducting magnets," *IEEE Trans. Appl. Supercond.* **19**(4), 3645−3652 (2009).

[19] M. Sekino, H. Ohsaki, H. Wada, T. Hisatsune, O. Ozaki, and T. Kiyoshi, "Fabrication of an MRI model magnet with an off-centered distribution of homogeneous magnetic field zone," *IEEE Trans. Appl. Supercond.* **20**(3), 781−785 (2010).

[20] C. Wu, J. Guo, C. Chen, G. Yan, and C. Li, "Optimal design and test of main magnet in superconducting MRI," *IEEE Trans. Appl. Supercond.* **20**(3), 1810−1813 (2010).

[21] F. Steinmeyer, P. W. Retz, K. White, A. Lang, W. Stautner, P. N. Smith, and G. Gilgrass, "Towards the invisible cryogenic system for magnetic resonance imaging," *AIP Conf. Proc.* **613**, 1659−1666 (2002).

[22] J. Nagamatsu, N. Nakagawa, T. Muranaka, Y. Zenitani, and J. Akimitsu, "Superconductivity at 39 K in magnesium diboride," *Nature*, **410**, 63−64 (2001).

[23] M. Razeti, S. Angius, L. Bertora, D. Damiani, R. Marabotto, M. Modica, D. Nardelli, M. Perrella, and M. Tassisto, "Construction and operation of cryogen free MgB_2 magnets for open MRI systems," *IEEE Trans. Appl. Supercond.* **18**(2), 882−886 (2008).

[24] F. Roméo and D. I. Hoult, "Magnetic field profiling: analysis and correcting coil design," *Magnet. Reson. Med.*, **1**(1), 44−65 (1984).

[25] R. Vadovic, "Magnetic field correction using magnetized shims," *IEEE Trans. Magn.* **25**(4), 3133−3139 (1989).

[26] J. H. Battocletti, H. A. Kamal, T. J. Myers, and T. A. Knox, "Systematic passive shimming of a permanent magnet for P-31 NMR spectroscopy of bone mineral," *IEEE Trans. Magn.* **29**(3), 2139−2151 (1993).

[27] H. S. Lopez, F. Liu, E. Weber, and S. Crozier, "Passive shim design and a shimming approach for biplanar permanent magnetic resonance imaging magnets," *IEEE Trans. Magn.* **44**(3), 394−402 (2008).

[28] Y. Zhang, D. Xie, B. Bai, H. S. Yoon, and C. S. Koh, "A novel optimal design method of passive shimming for permanent MRI magnet," *IEEE Trans. Magn.* **44**(6), 1058−1061 (2008).

CHAPTER 3
THEORY AND TECHNIQUES I
COMPUTATIONAL ELECTROMAGNETICS

The generation of the strong uniform magnetic field required by magnetic resonance imaging (MRI) is primarily accomplished through the use of large magnetized pole pieces or immense superconducting coil configurations. In addition, MRI magnet assemblies often employ large magnetic yoke structures that act to confine the fringe magnetic fields while enhancing the field strength in the imaging volume. The design of these components to achieve the level of field uniformity necessary for medical imaging demands an understanding of how their dimensions and properties relate to the fields they produce. Consequently, this necessitates the use of accurate formulas and numerical methods for the calculation of the magnetic fields generated by these sources.

3.1 THEORY

3.1.1 Introduction

The electromagnetic force is one of the four fundamental interactions of nature described by the Standard Model of particle physics. The subject of electrodynamics is concerned with the characterization of this force in the classical regime, through an understanding of electromagnetic fields and their interactions with matter. The fundamental theory of electrodynamics can be elegantly summarized by Maxwell's equations in differential form:[5]

$$\nabla \cdot \boldsymbol{D} = \rho, \tag{3.1a}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{3.1b}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{3.1c}$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}. \tag{3.1d}$$

These equations describe the principle relationships between the charge density distribution ρ, the current density distribution $\boldsymbol{J}$, the electric field $\boldsymbol{E}$, and the magnetic field $\boldsymbol{B}$. The relations in Eq. (1) also refer to the electric displacement $\boldsymbol{D}$ and auxiliary magnetic field $\boldsymbol{H}$. In certain media, these quantities can be derived from the electric and magnetic fields via constitutive relationships of the general form

$$\boldsymbol{D} = \boldsymbol{D}(\boldsymbol{E}, \boldsymbol{B}), \tag{3.2a}$$

$$\boldsymbol{H} = \boldsymbol{H}(\boldsymbol{E}, \boldsymbol{B}). \tag{3.2b}$$

In many practical electromagnetic systems, the fields of interest are directly generated by prescribed configurations of charge or current. The electric field at a point $\boldsymbol{r}$ due to the charge density distribution ρ is explicitly given by the generalized form of Coulomb's law:

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0} \int_{\mathcal{V}} \rho(\boldsymbol{r}') \frac{\boldsymbol{r} - \boldsymbol{r}'}{\|\boldsymbol{r} - \boldsymbol{r}'\|^3} d\mathcal{V}', \tag{3.3}$$

where ϵ_0 is the permittivity of free space and $\mathcal{V}$ is any volume encompassing all of the charge. A similar expression for the magnetic field due to the current density distribution $\boldsymbol{J}$ is provided by the Biot-Savart law:

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} \boldsymbol{J}(\boldsymbol{r}') \times \frac{\boldsymbol{r} - \boldsymbol{r}'}{\|\boldsymbol{r} - \boldsymbol{r}'\|^3} d\mathcal{V}', \tag{3.4}$$

where μ_0 is the permeability of free space. The relationship between the field and source points $\boldsymbol{r}$ and $\boldsymbol{r}'$ is illustrated in Figure 3.1.

Electromagnetic fields were first characterized by their interaction with matter via the force they exert on charged particles. The fundamental equation describing this interaction is known as the Lorentz force law, which gives the force $\boldsymbol{F}$ acting on a particle with charge q and velocity $\boldsymbol{v}$ in the presence of an electric and magnetic field:

$$\boldsymbol{F} = q[\boldsymbol{E} + (\boldsymbol{v} \times \boldsymbol{B})]. \tag{3.5}$$

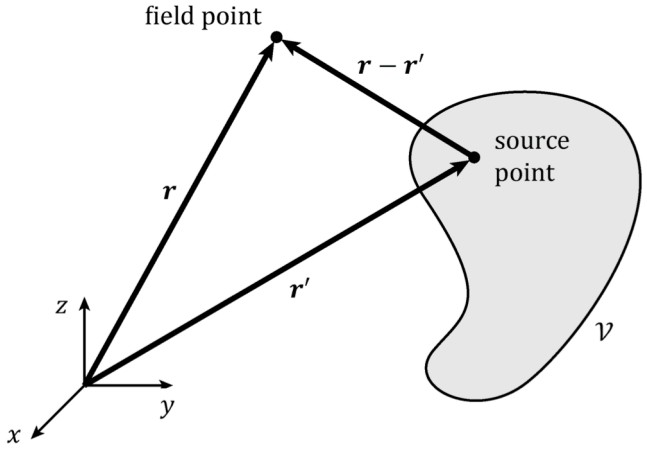


FIGURE 3.1: Relationship between the field and source points used in the evaluation of the electric and magnetic field integrals.

3.1.2 Electromagnetic Fields in Matter

Nearly all ordinary matter interacts with electromagnetic fields due to the intrinsic electric charge or spin angular momentum possessed by subatomic constituents. This is manifested in part by the polarization or alignment of atoms and molecules in the presence of an externally applied field. In bulk materials, this response may result in a net macroscopic electric polarization $\boldsymbol{P}$ or magnetization $\boldsymbol{M}$, defined as the electric and magnetic dipole moments per unit volume, respectively. The constitutive relations in Eq. (3.2) can then be expressed as

$$\boldsymbol{D} = \epsilon_0 \boldsymbol{E} + \boldsymbol{P}, \tag{3.6a}$$

$$\boldsymbol{H} = \frac{1}{\mu_0}\boldsymbol{B} - \boldsymbol{M}. \tag{3.6b}$$

The particular nature of the relationship from which $\boldsymbol{P}$ and $\boldsymbol{M}$ are derived from $\boldsymbol{E}$ and $\boldsymbol{B}$ depends on the specific properties of the associated medium. For isotropic linear media,

$$\boldsymbol{P} = \epsilon_0 \chi_e \boldsymbol{E}, \tag{3.7a}$$

$$\boldsymbol{M} = \chi_m \boldsymbol{H}, \tag{3.7b}$$

where the constants χ_e and χ_m are the electric and magnetic susceptibilities, respectively. Therefore,

$$\boldsymbol{D} = \epsilon \boldsymbol{E}, \tag{3.8a}$$

$$\boldsymbol{B} = \mu \boldsymbol{H}, \tag{3.8b}$$

where $\epsilon = \epsilon_0(1 + \chi_e)$ and $\mu = \mu_0(1 + \chi_m)$ are the permittivity and permeability of the medium, respectively.

Although χ_e is always a positive quantity, χ_m can be of any sign. Media containing atoms or molecules with unpaired electrons are typically paramagnetic and possess $\chi_m > 0$. The magnetization $\boldsymbol{M}$ induced in these materials is always oriented parallel to the externally applied $\boldsymbol{H}$. This is in contrast to diamagnetic materials with $\chi_m < 0$, for which the induced $\boldsymbol{M}$ is antiparallel to the applied $\boldsymbol{H}$.

In reality, the susceptibilities of most practical materials are not constant. Often they depend on the temperature of the medium, or on the frequency content of any time variance in the external applied field. Furthermore, the susceptibility may be a function of the applied field strength, such that the expressions in Eqs. (3.7) and (3.8) become nonlinear.

The precise spatial distribution of the magnetic field generated by a magnetized object is of great practical interest. The magnetization $\boldsymbol{M}$ in a particular region $\mathcal{V}$ with boundary $\mathcal{S}$ contributes an effective volume and surface current $\boldsymbol{J_M}$ and $\boldsymbol{K_M}$ given by[5]

$$\boldsymbol{J_M} = \nabla \times \boldsymbol{M} \tag{3.9a}$$

$$\boldsymbol{K_M} = \boldsymbol{M} \times \hat{\boldsymbol{n}} \tag{3.9b}$$

where $\hat{\boldsymbol{n}}$ is the outward normal unit vector on $\mathcal{S}$. The resulting magnetic field is then given by the Biot-Savart law in Eq. (3.4), yielding

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\left[\int_{\mathcal{V}} \boldsymbol{J_M}(\boldsymbol{r}') \times \frac{\boldsymbol{r}-\boldsymbol{r}'}{\|\boldsymbol{r}-\boldsymbol{r}'\|^3}\, d\mathcal{V}' + \int_{\mathcal{S}} \boldsymbol{K_M}(\boldsymbol{r}') \times \frac{\boldsymbol{r}-\boldsymbol{r}'}{\|\boldsymbol{r}-\boldsymbol{r}'\|^3}\, d\mathcal{S}'\right]. \tag{3.10}$$

3.1.3 The Laplace Equation

The calculation of magnetic fields from the explicit integrals in Eqs. (3.4) and (3.10) tends to be extremely cumbersome. Even in the case of trivial current or magnetization distributions, the integrals become very complicated and generally require the use numerical integration routines (see § 3.2). Moreover, if nonlinear magnetic materials are considered, evaluation of the field integral may be impossible.

An alternative approach to solving magnetostatics problems is derived from the scalar magnetic potential formulation of the auxiliary field $\boldsymbol{H}$. In a region absent of currents or time varying fields, Ampere's law from Eq. (3.1d) becomes

$$\nabla \times \boldsymbol{H} = 0. \tag{3.11}$$

Therefore, by the Helmholtz theorem,[6] $\boldsymbol{H}$ can be expressed as the gradient of a scalar magnetic potential Φ_m:

$$\boldsymbol{H} = -\nabla \Phi_m, \tag{3.12}$$

Assuming a constitutive relationship of the form in Eq. (3.8b), where the permeability $\mu = \mu(H)$ is nonlinear, Eq. (3.9) can be combined with Eq. (3.1b) to yield

$$-\nabla \cdot (\mu \nabla \Phi_m) = 0. \tag{3.13}$$

Directly solving this second order differential equation is in general very complicated. However, numerical techniques such as the finite element

method[7] are particularly well suited for obtaining its solution (see § 3.3). In the special case where μ is constant, Eq. (3.18) reduces to

$$\nabla^2 \Phi_m = 0. \tag{3.14}$$

A similar equation may also be obtained for the vector components of $\boldsymbol{H}$. Since the divergence of a curl is always zero, taking the divergence of Eq. (3.11) gives

$$\nabla^2 \boldsymbol{H} = \nabla^2 H_x \hat{\boldsymbol{\imath}} + \nabla^2 H_y \hat{\boldsymbol{\jmath}} + \nabla^2 H_z \hat{\boldsymbol{k}} = 0. \tag{3.15}$$

The vector components in this equation are independent, hence,

$$\nabla^2 H_x = \nabla^2 H_y = \nabla^2 H_z = 0. \tag{3.16}$$

Therefore, Φ_m, H_x, H_y, and H_z all satisfy the Laplace equation, which takes the form

$$\nabla^2 f = 0, \tag{3.17}$$

for a general scalar function f. In terms of spherical polar coordinates, Eq. (3.17) can be written as

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}(rf) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial f}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 f}{\partial \phi^2} = 0. \tag{3.18}$$

The general solution to this equation is then[5]

$$f(r,\theta,\phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \left[a_{\ell m} r^{\ell} + b_{\ell m} r^{-(\ell+1)}\right] Y_{\ell}^{m}(\theta,\phi), \tag{3.19}$$

where $Y_{\ell}^{m}(\theta,\phi)$ is the spherical harmonic function of order ℓ and degree m. By enforcing suitable boundary conditions on this solution, the expansion coefficients $a_{\ell m}$ and $b_{\ell m}$ can be determined. As evident in the summation in Eq. (3.19), the order ℓ may take on any nonnegative integer value, while the degree must satisfy $|m| \leq \ell$. The normalized functions $Y_{\ell}^{m}(\theta,\phi)$ are given by[5]

$$Y_{\ell}^{m}(\theta,\phi) = N_{\ell}^{m} P_{\ell}^{m}(\cos\theta) e^{im\phi}, \tag{3.20}$$

where $P_{\ell}^{m}(\cos\theta)$ is the associated Legendre polynomial of order ℓ and degree m. The normalization factor N_{ℓ}^{m} is defined as

$$N_{\ell}^{m} = \sqrt{\frac{(2\ell+1)}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}}, \tag{3.21}$$

such that the spherical harmonics satisfy the orthonormality relation

$$\int_0^{2\pi} \int_0^{\pi} Y_{\ell}^{m}(\theta,\phi) Y_{\ell'}^{m'}(\theta,\phi) \sin\theta \, d\theta \, d\phi = \delta_{\ell\ell'} \delta_{mm'}, \tag{3.22}$$

where δ_{ij} is the Kronecker delta function.

Within a region for which f is assumed regular, the solution given by Eq. (3.19) reduces to

$$f(r,\theta,\phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} r^{\ell} Y_{\ell}^{m}(\theta,\phi). \tag{3.23}$$

By exploiting the orthonormality property in Eq. (3.22), the expansion coefficients are given as

$$a_{\ell m} = \frac{1}{R^{\ell}} \int_0^{2\pi} \int_0^{\pi} f(R,\theta,\phi) Y_{\ell}^{m\dagger}(\theta,\phi) \sin\theta \, d\theta \, d\phi, \tag{3.24}$$

where R is the radius of a sphere on which the values of f are known, and the superscript † denotes the complex conjugate.

As evident in Eq. (3.20), the spherical harmonic basis functions are generally complex. Hence, it is often preferred to rewrite the expansion in Eq. (3.23) in terms of the all-real basis $\{y_{\ell}^{m}\}$:

$$f(r,\theta,\phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} c_{\ell m} \, r^{\ell} y_{\ell}^{m}(\theta,\phi), \tag{3.25}$$

where the normalized real spherical harmonic function $y_{\ell}^{m}(\theta,\phi)$ is defined according to

$$y_{\ell}^{m} = \begin{cases} \sqrt{2}\,\mathrm{Im}(Y_{\ell}^{m}) = \dfrac{1}{i\sqrt{2}}(Y_{\ell}^{-m} - (-1)^{m} Y_{\ell}^{m}), & \text{for } m < 0, \\ Y_{0}^{0}, & \text{for } m = 0, \\ \sqrt{2}\,\mathrm{Re}(Y_{\ell}^{m}) = \dfrac{1}{\sqrt{2}}(Y_{\ell}^{m} + (-1)^{m} Y_{\ell}^{-m}), & \text{for } m > 0, \end{cases} \tag{3.26a}$$

which may be expressed as

$$y_\ell^m = \begin{cases} \sqrt{2}\, N_\ell^{|m|} P_\ell^{|m|}(\cos\theta)\sin|m|\phi\,, & \text{for } m < 0, \\ N_\ell^0 P_\ell^0(\cos\theta), & \text{for } m = 0, \\ \sqrt{2}\, N_\ell^m P_\ell^m(\cos\theta)\cos m\phi\,, & \text{for } m > 0. \end{cases} \tag{3.26b}$$

The real spherical harmonics are also orthonormal on the surface of a sphere and hence the expansion coefficients $c_{\ell m}$ from Eq. (3.25) are given as

$$c_{\ell m} = \frac{1}{R^\ell} \int_0^{2\pi} \int_0^{\pi} f(R,\theta,\phi) y_\ell^m(\theta,\phi) \sin\theta \, d\theta \, d\phi. \tag{3.27}$$

The magnetic field in the imaging volume of an MRI system satisfies Eq. (3.17) and thus can be expressed in terms of the expansion in Eq. (3.25). Accordingly, the coefficients $c_{\ell m}$ are typically calculated during the shimming procedure for these systems, as the active correction coils used to improve the field uniformity generate field patterns described by the real harmonics y_ℓ^m defined in Eq. (3.26) (see § 2.3.4).

3.2 SEMI-ANALYTICAL METHODS

3.2.1 Ideal Circular Coils

The Biot-Savart integral in Eq. (3.4) provides a direct method for the calculation of the magnetic field generated by simple current distributions. Consider an ideal current loop with zero cross-sectional area, coaxial with the z axis, positioned at $z = z_0$, and with radius s_0 (Figure 3.2). The magnetic field at the point (s, ϕ, z) in cylindrical coordinates, due to a current I passing through the loop, is given by[8]

$$\boldsymbol{B}(s,\phi,z) = B_s(s,\phi,z)\hat{\boldsymbol{s}} + B_\phi(s,\phi,z)\hat{\boldsymbol{\phi}} + B_z(s,\phi,z)\hat{\boldsymbol{z}}, \tag{3.28}$$

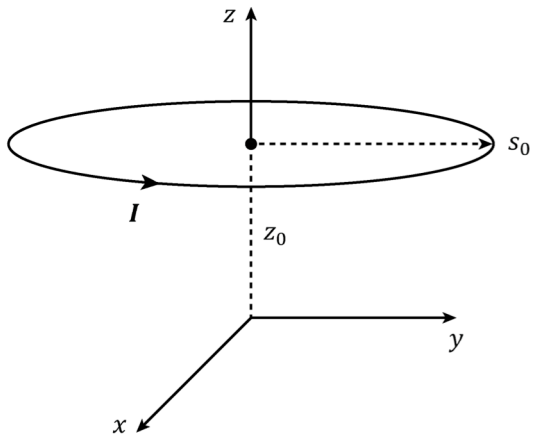


FIGURE 3.2: Geometry of an ideal current loop centered on the z axis.

where

$$B_s(s,\phi,z) = \frac{\mu_0 I}{2\pi s}\mathcal{G}_1(s,s_0,z-z_0), \tag{3.29a}$$

$$B_\phi(s,\phi,z) = 0, \tag{3.29b}$$

$$B_z(s,\phi,z) = \frac{\mu_0 I}{2\pi}\mathcal{G}_2(s,s_0,z-z_0). \tag{3.29c}$$

The intermediate functions $\mathcal{G}_1(t_1,t_2,t_3)$ and $\mathcal{G}_2(t_1,t_2,t_3)$ are given by

$$\mathcal{G}_1(t_1,t_2,t_3) = \frac{1}{\sqrt{(t_1+t_2)^2+{t_3}^2}} \times \left[-\mathfrak{K}(u) + \frac{{t_1}^2+{t_2}^2+{t_3}^2}{(t_1-t_2)^2+{t_3}^2}\mathfrak{E}(u)\right], \tag{3.30a}$$

$$\mathcal{G}_2(t_1, t_2, t_3) = \frac{1}{\sqrt{(t_1 + t_2)^2 + z^2}} \times \left[\mathfrak{K}(u) - \frac{{t_1}^2 - {t_2}^2 + {t_3}^2}{(t_1 - t_2)^2 + {t_3}^2}\mathfrak{E}(u)\right], \tag{3.30b}$$

where the parameter u, satisfying $0 \leq u \leq 1$, is defined as

$$u = \sqrt{\frac{4t_1 t_2}{(t_1 + t_2)^2 + {t_3}^2}}, \tag{3.31}$$

and where $\mathfrak{K}(u)$ and $\mathfrak{E}(u)$ denote the complete elliptic integrals of the first and second kind, defined by[6]

$$\mathfrak{K}(u) = \int_0^1 \left[\left(1 - \beta'^2\right)\left(1 - u^2\beta'^2\right)\right]^{-1/2} d\beta', \tag{3.32a}$$

$$\mathfrak{E}(u) = \int_0^1 \left(1 - \beta'^2\right)^{-1/2}\left(1 - u^2\beta'^2\right)^{1/2} d\beta'. \tag{3.32b}$$

The built-in function `ellipke` in MATLAB[9] can be used for the calculation of the integrals $\mathfrak{K}(u)$ and $\mathfrak{E}(u)$ via the arithmetic-geometric mean method,[10] thereby making rapid calculation of Eq. (3.28) practical.

3.2.2 Real Circular Coils with Rectangular Cross-sections

In order to achieve the highest possible accuracy when calculating the magnetic field due to a real physical coil, the finite cross-sectional area of the coil must be taken into account. This is necessary in the optimization of superconducting coil configurations used in MRI, for which the generation of

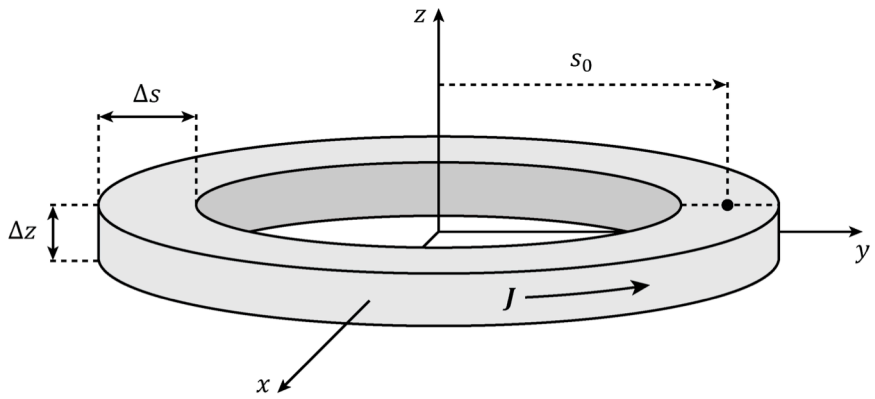


FIGURE 3.3: Geometry of a bulk coil with rectangular cross-section centered on the z axis.

a uniform magnetic field within a large imaging volume is a critical design goal (see Chapter 7).

Consider a bulk circular coil lying in the xy plane and coaxial with the z axis, as illustrated in Figure 3.3. Suppose the bottom and top surfaces of the coil are located at $z = -\Delta z/2$ and $z = \Delta z/2$, respectively. In addition, the inner and outer radii of the coil are $s = s_0 - \Delta s/2$ and $s = s_0 + \Delta s/2$. The magnetic field at the point (x, y, z), due to a uniform current density J across the coil cross section, is given by[11]

$$\boldsymbol{B}(x,y,z) = B_x(x,y,z)\hat{\boldsymbol{x}} + B_y(x,y,z)\hat{\boldsymbol{y}} + B_z(x,y,z)\hat{\boldsymbol{z}}, \tag{3.33}$$

where the field components are expressed as

$$B_x(x,y,z) = \tag{3.34a}$$
$$-\frac{\mu_0 J x}{2\pi s}\left[\mathcal{N}_1\left(s_0+\frac{\Delta s}{2},-\frac{\Delta z}{2},s,z\right)-\mathcal{N}_1\left(s_0-\frac{\Delta s}{2},-\frac{\Delta z}{2},s,z\right)\right.$$
$$\left.-\mathcal{N}_1\left(s_0+\frac{\Delta s}{2},\frac{\Delta z}{2},s,z\right)+\mathcal{N}_1\left(s_0-\frac{\Delta s}{2},\frac{\Delta z}{2},s,z\right)\right],$$

$$B_y(x,y,z) = \tag{3.34b}$$
$$-\frac{\mu_0 J y}{2\pi s}\left[\mathcal{N}_1\left(s_0+\frac{\Delta s}{2},-\frac{\Delta z}{2},s,z\right)-\mathcal{N}_1\left(s_0-\frac{\Delta s}{2},-\frac{\Delta z}{2},s,z\right)\right.$$
$$\left.-\mathcal{N}_1\left(s_0+\frac{\Delta s}{2},\frac{\Delta z}{2},s,z\right)+\mathcal{N}_1\left(s_0-\frac{\Delta s}{2},\frac{\Delta z}{2},s,z\right)\right],$$

$$B_z(x,y,z) = \tag{3.34c}$$
$$\frac{\mu_0 J}{2\pi s}\left[\mathcal{N}_2\left(s_0+\frac{\Delta s}{2},\frac{\Delta z}{2},s,z\right)-\mathcal{N}_2\left(s_0-\frac{\Delta s}{2},\frac{\Delta z}{2},s,z\right)\right.$$
$$\left.-\mathcal{N}_2\left(s_0+\frac{\Delta s}{2},-\frac{\Delta z}{2},s,z\right)+\mathcal{N}_2\left(s_0-\frac{\Delta s}{2},-\frac{\Delta z}{2},s,z\right)\right].$$

The intermediate functions $\mathcal{N}_1(t_1,t_2,t_3,t_4)$ and $\mathcal{N}_2(t_1,t_2,t_3,t_4)$ are defined as

$$\mathcal{N}_1(t_1,t_2,t_3,t_4) = \int_0^{\pi} \cos\beta' \left[\sqrt{(\gamma_1)^2+(\gamma_2)^2}\right. \tag{3.35a}$$
$$\left.+\gamma_3 \ln\left(\gamma_1+\sqrt{(\gamma_1)^2+(\gamma_2)^2}\right)\right] d\beta'$$

$$\mathcal{N}_2(t_1,t_2,t_3,t_4) = \int_0^{\pi} \cos\beta' \left[\mathcal{M}_1(t_1,t_2,t_3,t_4,\beta')\right. \tag{3.35b}$$
$$+\gamma_3\gamma_4\,\mathcal{M}_2(t_1,t_2,t_3,t_4,\beta')$$
$$\left.+\frac{1}{2}\gamma_5\sqrt{(\gamma_1)^2+(\gamma_2)^2}\right] d\beta'$$

where

$$\mathcal{M}_1(t_1, t_2, t_3, t_4, \beta') \tag{3.36a}$$
$$= 2\gamma_3\gamma_5 \ln\left[\gamma_1 + \sqrt{(\gamma_1)^2 + (\gamma_2)^2}\right]$$
$$+ \frac{1}{4}\{3[(\gamma_3)^2 - (\gamma_4)^2] - [(\gamma_1)^2 - (\gamma_3)^2]\} \times$$
$$\ln\left[\frac{\sqrt{(\gamma_1)^2 + (\gamma_2)^2} - \gamma_5}{\sqrt{(\gamma_1)^2 + (\gamma_2)^2} + \gamma_5}\right],$$

$$\mathcal{M}_2(t_1, t_2, t_3, t_4, \beta') \tag{3.36b}$$
$$= -2 \arctan\left[\frac{\gamma_1\left(\gamma_1 + \sqrt{(\gamma_1)^2 + (\gamma_2)^2}\right) + (\gamma_4)^2}{\gamma_4\gamma_5}\right]$$
$$+ \arctan\left[\frac{\gamma_5 + \gamma_1 + \sqrt{(\gamma_1)^2 + (\gamma_2)^2}}{\gamma_4}\right]$$
$$- \arctan\left[\frac{-\gamma_5 + \gamma_1 + \sqrt{(\gamma_1)^2 + (\gamma_2)^2}}{\gamma_4}\right],$$

and

$$\gamma_1 = t_1 - t_3 \cos\beta', \tag{3.37a}$$
$$(\gamma_2)^2 = {t_3}^2 \sin^2\beta' + (t_2 - t_4)^2, \tag{3.37b}$$
$$\gamma_3 = t_3 \cos\beta', \tag{3.37c}$$
$$\gamma_4 = t_3 \sin\beta', \tag{3.37d}$$
$$\gamma_5 = t_2 - t_4. \tag{3.37e}$$

The built-in function `quadgk` in MATLAB can be used to rapidly calculate the integrals defining $\mathcal{N}_1(t_1, t_2, t_3, t_4)$ and $\mathcal{N}_2(t_1, t_2, t_3, t_4)$ in Eq. (3.35). This routine employs adaptive Gauss-Kronrod quadrature and is especially well suited for obtaining high accuracy integral approximations.[12]

3.3 THE FINITE ELEMENT METHOD

3.3.1 Introduction

The finite element method[7,13] (FEM) is a numerical technique for calculating approximate solutions to partial differential equations of the general form:

$$\mathfrak{L}[\Phi(\boldsymbol{r})] = f(\boldsymbol{r}), \quad \boldsymbol{r} \in \Omega, \tag{3.38}$$

where $\mathfrak{L}$ is a differential operator defined on the domain Ω, f is the excitation or source function, and Φ is the unknown quantity. Together with suitable conditions enforced on the boundary of Ω, the problem in Eq. (3.38) constitutes a boundary value problem. Through discretization of the continuous domain Ω, the FEM provides a procedure for converting the mathematically posed problem of Eq. (3.38) into a set of linear equations with a finite number of unknowns. The resulting approximate problem is easily solved with the aid of a computer, whereas the solution to the original problem can only be obtained via mathematical manipulation. Furthermore, the accuracy of the FEM approximation is limited in theory only by the extent of domain discretization, which is in turn only constrained by computing speed and memory availability. Consequently, the FEM offers an attractive strategy for solving large arduous boundary value problems with complicated geometries.

There are four major steps involved with any FEM implementation. The first and often most difficult step is discretization of the solution domain. This consists of replacing the continuous domain Ω of the original problem by a finite number of subdomains in which the solution is to be approximated by an interpolating polynomial with unknown coefficients. The second step is then the specification of the interpolating functions, also known as the finite

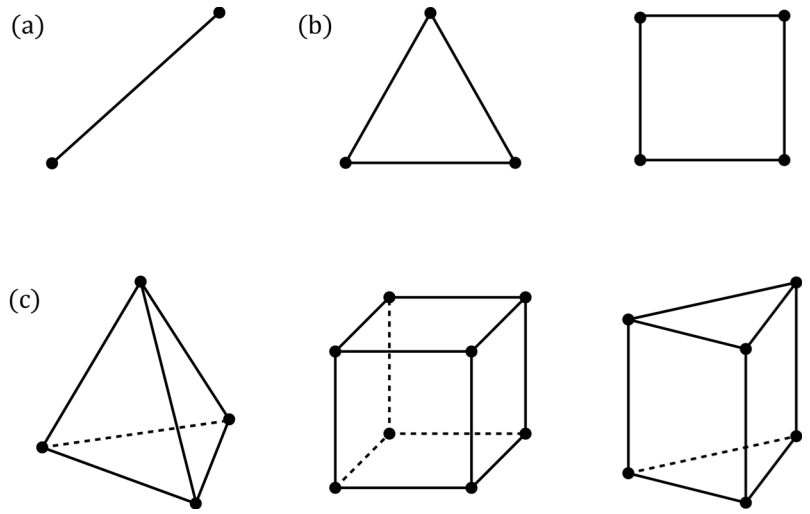


FIGURE 3.4: Common finite element shapes used in the discretization of (a) one, (b) two, and (c) three-dimensional geometries.

element basis functions, which are eventually combined to form the trial function serving as the approximate global solution. In the third step, the trial function is applied to the original problem formulation, leading to the generation and assembly of a large system of linear equations. Lastly, the fourth and most time consuming step amounts to applying a numerical routine to efficiently solve the system of equations, thereby obtaining the final approximation and concluding the procedure.

3.3.2 Domain Discretization

The first step of the FEM consists of partitioning the continuous domain Ω of the given boundary value problem with a finite set of N_Ω polytopal subdomains or elements, each denoted $\Omega^{\mathscr{p}}$ ($\mathscr{p} = 1, 2, \ldots, N_\Omega$). The resulting collection of elements constitutes what is commonly referred to as the finite element mesh. Common polytopes employed in the discretization of one, two, and three-dimensional problems are illustrated in Figure 3.4. Note that

triangular or tetrahedral elements are the most widely used, as they are especially effective at conforming to complex geometries and have also been shown to yield the greatest solution accuracy.[7] In a proper discretization, the elements are connected only at their vertices, such that there is neither overlap nor gaps amongst adjacent elements. Furthermore, narrow elements with large dihedral angles are usually avoided, as they compromise the local solution accuracy while leading to poor conditioning of the final system of equations.[7] Hence, in order to address these requirements, a Delaunay triangulation algorithm is commonly employed for the automatic generation of the finite element mesh.

3.3.3 Finite Element Basis Functions

Once the domain has been appropriately partitioned, the unknown quantity Φ is approximated by an interpolating function that is defined within each of the finite elements. Specifically, for a finite element mesh of order n_Ω, a piecewise polynomial $\Phi^{\mathscr{p}}$ of order n_Ω with compact support $\Omega^{\mathscr{p}}$ is defined within each element so that the trial function $\widetilde{\Phi}$ forms a continuous approximation of Φ over the entire solution domain:

$$\Phi(\boldsymbol{r}) \approx \widetilde{\Phi}(\boldsymbol{r}) = \sum_{\mathscr{p}=1}^{N_\Omega} \Phi^{\mathscr{p}}(\boldsymbol{r}), \quad \boldsymbol{r} = (x, y, z) \in \Omega. \tag{3.39}$$

In order to obtain solution isotropy, each interpolating polynomial $\Phi^{\mathscr{p}}$ must be complete in the sense that it must contain all mixed terms of the coordinate variables x, y, and z. For a polynomial of order n_Ω, this amounts to a total of $m_\Omega = (1/6)(n_\Omega + 1)(n_\Omega + 2)(n_\Omega + 3)$ terms.[7] Consequently, the m_Ω unknown coefficients of $\Phi^{\mathscr{p}}$ are uniquely determined only if the interpolating polynomial is enforced at m_Ω distinct points. These points,

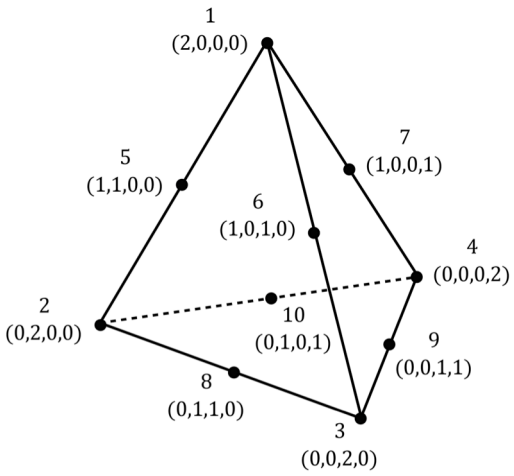


FIGURE 3.5: Distribution and labeling of nodes for a second-order tetrahedral finite element. Local node numbers and volume coordinate indices are shown.

commonly referred to as the nodes, are typically placed in a regular fashion within $\Omega^{\mathcal{p}}$, as illustrated in Figure 3.5 for a tetrahedral element with $n_\Omega = 2$.

The resulting system of equations can be solved for the unknown coefficients in terms of the coordinates and solution values at the m_Ω nodal points, denoted (x_i, y_i, z_i) and $\Phi_i^{\mathcal{p}}$ for the ith node, respectively. Once this is accomplished, the interpolating polynomial may be written in the form

$$\Phi^{\mathcal{p}}(x,y,z) = \sum_{i=1}^{m_\Omega} \Phi_i^{\mathcal{p}} \mathfrak{S}_i^{\mathcal{p}}(x,y,z), \tag{3.40}$$

where $\mathfrak{S}_i^{\mathcal{p}}(x,y,z)$ denotes the finite element basis function in $\Omega^{\mathcal{p}}$ corresponding to the ith node. These basis functions are also polynomials of order n_Ω, with the special property:

$$\mathfrak{S}_i^{\mathcal{p}}\left(x_j, y_j, z_j\right) = \delta_{ij}. \tag{3.41}$$

Accordingly, $\mathfrak{S}_i^{\mathscr{p}}$ is called a Lagrange basis function and the subdomain $\Omega^{\mathscr{p}}$ within which it is defined is referred to as a Lagrange finite element.

For convenience, each node within $\Omega^{\mathscr{p}}$ is labeled with a collection of four integers (a_1, a_2, a_3, a_4), such that $a_1 + a_2 + a_3 + a_4 = n_\Omega$. For the ith node located at (x_i, y_i, z_i), the integer labels are defined as

$$a_1 = n_\Omega L_1^{\mathscr{p}}(x_i, y_i, z_i), \tag{3.42a}$$

$$a_2 = n_\Omega L_2^{\mathscr{p}}(x_i, y_i, z_i), \tag{3.42b}$$

$$a_3 = n_\Omega L_3^{\mathscr{p}}(x_i, y_i, z_i), \tag{3.42c}$$

$$a_4 = n_\Omega L_4^{\mathscr{p}}(x_i, y_i, z_i). \tag{3.42d}$$

In these equations, $L_{\mathscr{k}}^{\mathscr{p}}(x, y, z)$ is the volume coordinate of the point (x, y, z) with respect to the $\mathscr{k}$th vertex of the element $\Omega^{\mathscr{p}}$, given by

$$L_{\mathscr{k}}^{\mathscr{p}}(x, y, z) = \frac{V_{\mathscr{k}}^{\mathscr{p}}(x, y, z)}{V^{\mathscr{p}}}, \quad \mathscr{k} = 1, 2, 3, 4, \tag{3.43}$$

where $V^{\mathscr{p}}$ is the volume of $\Omega^{\mathscr{p}}$ and $V_{\mathscr{k}}^{\mathscr{p}}(x, y, z)$ is the volume of the tetrahedron formed by replacing the $\mathscr{k}$th vertex of $\Omega^{\mathscr{p}}$ with the point located at (x, y, z). In this way, a combination of any three volume coordinates uniquely defines a point within the given element. The resulting nodal labels determined from Eq. (3.42) for a tetrahedral element with $n_\Omega = 2$ are shown in Figure 3.5.

With the nodal labels and volume coordinates defined above, a convenient generating formula for the basis functions $\mathfrak{S}_i^{\mathscr{p}}$ of arbitrary order may be used:[13]

$$\mathfrak{S}_i^{\mathcal{P}}(x,y,z) = \mathcal{P}_i^{a_1}\left(L_1^{\mathcal{P}}\right)\mathcal{P}_i^{a_2}\left(L_2^{\mathcal{P}}\right)\mathcal{P}_i^{a_3}\left(L_3^{\mathcal{P}}\right)\mathcal{P}_i^{a_4}\left(L_4^{\mathcal{P}}\right), \tag{3.44}$$

where $L_{\hbar}^{\mathcal{P}} = L_{\hbar}^{\mathcal{P}}(x,y,z)$ and $\mathcal{P}_i^{q}\left(L_{\hbar}^{\mathcal{P}}\right)$ is the Lagrange interpolating polynomial of order q, for the ith node, defined as

$$\mathcal{P}_i^{q}\left(L_{\hbar}^{\mathcal{P}}\right) = \prod_{\substack{\hbar=1 \\ \hbar\neq i}}^{q} \frac{\left(L_{\hbar}^{\mathcal{P}} - L_{\hbar\hbar}^{\mathcal{P}}\right)}{\left(L_{\hbar i}^{\mathcal{P}} - L_{\hbar\hbar}^{\mathcal{P}}\right)}, \tag{3.45}$$

where $L_{\hbar i}^{\mathcal{P}} = L_{\hbar}^{\mathcal{P}}(x_i, y_i, z_i)$. It can be verified that $\mathfrak{S}_i^{\mathcal{P}}(x,y,z)$ given by Eq. (3.44) indeed satisfies the property of Eq. (3.43).

It is now possible to address the relationship between the order of the FEM discretization error and the mesh element density. Consider the Taylor series expansion of the solution Φ about the point $\boldsymbol{r}$:

$$\Phi(\boldsymbol{r}+\boldsymbol{\delta}) = \sum_{q=0}^{n_\Omega} \frac{1}{q!}(\boldsymbol{\delta}\cdot\nabla)^q \Phi(\boldsymbol{r}) + \mathcal{O}(\delta^{n_\Omega+1}), \tag{3.46}$$

where $\mathcal{O}(\delta^{n_\Omega+1})$ represents the error encountered when terminating this expansion at a polynomial of order n_Ω. Therefore, if δ represents a measure of the size of the element $\Omega^{\mathcal{P}}$, the local discretization error in the FEM approximation is of order $\mathcal{O}(\delta^{n_\Omega+1})$ in the best-case scenario. Furthermore, it is well known from Taylor's theorem that the error term in Eq. (3.46) is related to the magnitude of $\nabla^{n_\Omega+1}\Phi$ within the domain of the approximation. Thus, refinement of the mesh in regions where Φ is expected to have rapid spatial variations is critical for obtaining an accurate FEM solution.

In reality, the geometry of practical problems often involves complex shapes with curved edges and surfaces. Consequently, the use of mesh elements

strictly possessing straight edges leads to unavoidable errors due to the nonconformity of the elements to the true problem geometry. In order to alleviate this shortcoming, the polytopal elements illustrated in Figure 3.4 can be mapped into distorted forms with curvilinear surfaces.[13] In a manner similar to Eq. (3.40), this transformation can be expressed as

$$\boldsymbol{r} = \sum_{i=1}^{m} \boldsymbol{r}_i^{\mathscr{p}} \mathfrak{T}_i^{\mathscr{p}}(\xi, \eta, \zeta), \quad \boldsymbol{r} = (x, y, z) \in \Omega^{\mathscr{p}}, \tag{3.47}$$

where $\boldsymbol{r}_i^{\mathscr{p}} = (x_i, y_i, z_i)$ denotes the position of the ith node of the curved element in xyz-space and $\mathfrak{T}_i^{\mathscr{p}}(\xi, \eta, \zeta)$ represents the shape function responsible for the transformation of the parent polytope defined in $\xi\eta\zeta$-space. This relationship between the parent and corresponding mapped element is illustrated in Figure 3.6. It can be shown that the shape functions in Eq. (3.47) are given by[7]

$$\mathfrak{T}_i^{\mathscr{p}}(\xi, \eta, \zeta) = \mathcal{P}_i^{a_1}(1 - \xi - \eta - \zeta)\mathcal{P}_i^{a_2}(\xi)\mathcal{P}_i^{a_3}(\eta)\mathcal{P}_i^{a_4}(\zeta), \tag{3.48}$$

where the nodal label (a_1, a_2, a_3, a_4) and Lagrange interpolating polynomial $\mathcal{P}_i^{q}$ are defined as before. If the order of the shape functions is taken to be equal to the order of the basis functions, the resulting elements defined by Eq. (3.47) are called isoparametric elements.

Most of the nodes generated by the discretization will simultaneously belong to several elements (see § 3.3.2). Consequently, each node is assigned the local index i within each associated element, in addition to a global index specifying its position within the complete problem geometry. These indices are then related via a connectivity array denoted $\mathcal{U}_\Omega(i, \mathscr{p})$, which gives the

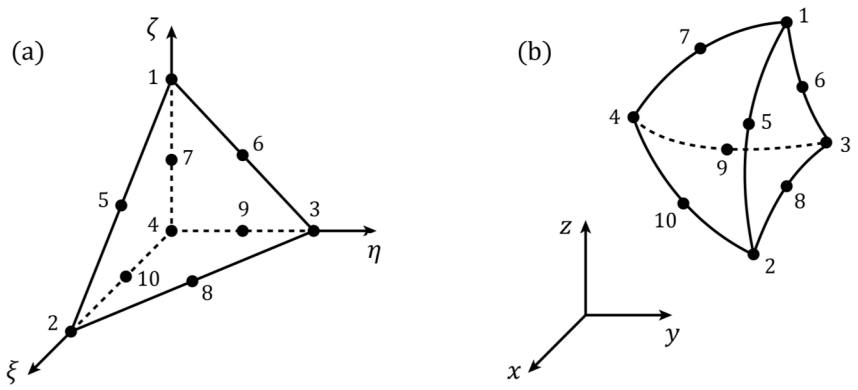


FIGURE 3.6: (a) Illustration of a second-order tetrahedral parent element with straight edges in $\xi\eta\zeta$-space. (b) The transformed element with curved edges in xyz-space. The local node indices are shown for both elements.

global number of the node possessing local index i in element $\Omega^{\mathscr{p}}$. In a similar manner, connectivity arrays denoted $\mathcal{U}_{\partial\Omega}(i,\mathscr{p})$ are constructed to relate local and global indices on specific interfaces and boundary segments on which boundary conditions of a particular type are enforced. These arrays are necessary for an efficient computer implementation of the FEM, eventually serving as a critical tool during assembly of the final system of linear equations (see § 3.3.5 and 3.3.6).

3.3.4 Galerkin's Method of Weighted Residuals

At the core of the FEM is the conversion of the original boundary value problem into a system of linear equations with a finite number of unknowns. This can be accomplished by interpreting the governing equation of Eq. (3.38) in terms of the residual $\mathfrak{r}$ defined as

$$\mathfrak{r}(\Phi) = \mathfrak{L}[\Phi] - f. \tag{3.49}$$

Thus, any Φ that is a solution of the original boundary value problem must result in $\mathfrak{r}(\Phi) = 0$. In contrast, an approximation of the exact solution $\widetilde{\Phi}$, such as the one constructed with Eqs. (3.39) and (3.40), would result in $\mathfrak{r}\big(\widetilde{\Phi}\big) \neq 0$. Hence, the problem of seeking an approximate solution of Eq. (3.38) can now be regarded as determining the approximation $\widetilde{\Phi}$ that results in the least value of $\mathfrak{r}$ at all points in Ω. Mathematically, this amounts to an attempt to satisfy the condition[14]

$$\mathfrak{R}_i^{\mathscr{p}} = \int_{\Omega} w_i^{\mathscr{p}} \mathfrak{r}\big(\widetilde{\Phi}\big)\, d\Omega' = 0, \tag{3.50}$$

where $\mathfrak{R}_i^{\mathscr{p}}$ is the weighted residual associated with the arbitrary weighting function denoted $w_i^{\mathscr{p}}$. Accordingly, this approximation procedure is commonly referred to as the method of weighted residuals. In particular, if a set of weighting functions $\{w_i^{\mathscr{p}}\}$ are selected such that $w_i^{\mathscr{p}} = \mathfrak{S}_i^{\mathscr{p}}$, then the Galerkin formulation of the weighted residuals method is obtained.[14] Hence, within each finite element $\Omega^{\mathscr{p}}$, a collection of weighted residuals are generated:

$$\mathfrak{R}_i^{\mathscr{p}} = \int_{\Omega^{\mathscr{p}}} \mathfrak{S}_i^{\mathscr{p}} \left(\mathfrak{L}\big[\widetilde{\Phi}(\boldsymbol{r})\big] - f(\boldsymbol{r}) \right) d\Omega' = 0, \quad \begin{cases} i = 1, 2, \dots, m_{\Omega}, \\ \mathscr{p} = 1, 2, \dots, N_{\Omega}. \end{cases} \tag{3.51}$$

As will be demonstrated in the following sections, this collection of weighted residuals actually represents a system of linear equations, for which the solution is comprised of the unknown nodal values of $\widetilde{\Phi}$.

3.3.5 Magnetostatics Formulation without Currents

The FEM is an effective tool for solving boundary value problems in magnetostatics characterized by the absence of free electrical current. In particular, this situation is applicable to the design of permanent magnet systems for MRI, for which calculation of the magnetic fields produced within and surrounding complex nonlinear magnetic structures is required (see Chapter 5). The formulation of this problem begins with the generalized constitutive relationship given by

$$\boldsymbol{B} = \mu \boldsymbol{H} + \mu_0 \boldsymbol{M_r}. \tag{3.52}$$

In ferromagnetic regions we have $\mu = \mu_0$ and Eq. (3.6b) is recovered. Similarly, in regions consisting of nonlinear magnetic media it is often assumed that $\boldsymbol{M_r} = 0$, and hence Eq. (3.8b) is obtained. Due to the absence of currents, $\boldsymbol{H}$ can be defined as the gradient of a scalar magnetic potential Φ_m (see § 3.1.3):

$$\boldsymbol{H} = -\nabla \Phi_m. \tag{3.53}$$

This result can be combined with the divergence of $\boldsymbol{B}$ in Eq. (3.1d) to obtain the second order partial differential equation (PDE):

$$-\nabla \cdot (\mu \nabla \Phi_m - \mu_0 \boldsymbol{M_r}) = 0. \tag{3.54}$$

By applying Galerkin's method of weighted residuals to this equation, Eq. (3.51) becomes

$$\Re_i^{\mathscr{p}} = \int_{\Omega^{\mathscr{p}}} \mathbb{G}_i^{\mathscr{p}} [\nabla \cdot (\mu \nabla \Phi_m - \mu_0 \boldsymbol{M_r})] \, d\Omega', \tag{3.55}$$

where it is implied that $i = 1, 2, \ldots, m_\Omega$ and $\mathscr{p} = 1, 2, \ldots, N_\Omega$. Hence, by Green's theorem,[6]

$$\Re_i^{\mathscr{p}} = \int_{\Omega^{\mathscr{p}}} \nabla \mathbb{G}_i^{\mathscr{p}} \cdot (\mu \nabla \Phi_m) \, d\Omega' - \int_{\Omega^{\mathscr{p}}} \nabla \mathbb{G}_i^{\mathscr{p}} \cdot (\mu_0 \boldsymbol{M_r}) \, d\Omega' + \int_{\partial \Omega^{\mathscr{p}}} \mathbb{G}_i^{\mathscr{p}} \boldsymbol{B} \cdot \hat{\boldsymbol{n}} \, d\Gamma', \tag{3.56}$$

where $\partial\Omega^{\mathscr{p}}$ is the boundary of $\Omega^{\mathscr{p}}$. By replacing Φ_m with the FEM approximation $\widetilde{\Phi}$ from Eqs. (3.39) and (3.40), the weighted residual in Eq. (3.56) becomes

$$\Re_i^{\mathscr{p}} = \sum_{j=1}^{m} \Phi_j^{\mathscr{p}} \int_{\Omega^{\mathscr{p}}} \nabla \mathbb{G}_i^{\mathscr{p}} \cdot \left(\mu \nabla \mathbb{G}_j^{\mathscr{p}}\right) d\Omega' - \int_{\Omega^{\mathscr{p}}} \nabla \mathbb{G}_i^{\mathscr{p}} \cdot (\mu_0 \boldsymbol{M_r}) \, d\Omega' + \int_{\partial \Omega^{\mathscr{p}}} \mathbb{G}_i^{\mathscr{p}} \boldsymbol{B} \cdot \hat{\boldsymbol{n}} \, d\Gamma'. \tag{3.57}$$

This equation can be expressed in matrix form as

$$\boldsymbol{\Re}^{\mathscr{p}} = \mathbf{K}^{\mathscr{p}} \boldsymbol{\Phi}^{\mathscr{p}} - \boldsymbol{b}^{\mathscr{p}} - \boldsymbol{g}^{\mathscr{p}}, \tag{3.58}$$

where the elements of the $m_\Omega \times m_\Omega$ matrix $\mathbf{K}^{\mathscr{p}}$, and $m_\Omega \times 1$ vectors $\boldsymbol{b}^{\mathscr{p}}$ and $\boldsymbol{g}^{\mathscr{p}}$, are defined as

$$K_{ij}^{\mathcal{P}} = \int_{\Omega^{\mathcal{P}}} \nabla \mathbb{G}_i^{\mathcal{P}} \cdot \left(\mu \nabla \mathbb{G}_j^{\mathcal{P}}\right) d\Omega', \tag{3.59a}$$

$$b_i^{\mathcal{P}} = \int_{\Omega^{\mathcal{P}}} \nabla \mathbb{G}_i^{\mathcal{P}} \cdot (\mu_0 \boldsymbol{M_r})\, d\Omega', \tag{3.59b}$$

$$g_i^{\mathcal{P}} = -\int_{\partial\Omega^{\mathcal{P}}} \mathbb{G}_i^{\mathcal{P}} \boldsymbol{B} \cdot \hat{\boldsymbol{n}}\, d\Gamma'. \tag{3.59c}$$

Summing the weighted residuals from all elements yields

$$\mathfrak{R} = \sum_{\mathcal{P}=1}^{N_\Omega} \overline{\mathfrak{R}}^{\mathcal{P}} = \sum_{\mathcal{P}=1}^{N_\Omega} \left(\overline{\mathbf{K}}^{\mathcal{P}} \overline{\boldsymbol{\Phi}}^{\mathcal{P}} - \overline{\boldsymbol{b}}^{\mathcal{P}} - \overline{\boldsymbol{g}}^{\mathcal{P}}\right), \tag{3.60}$$

$$= \mathbf{K}\boldsymbol{\Phi} - \boldsymbol{b} - \boldsymbol{g}, \tag{3.61}$$

where the augmented matrix $\overline{\mathbf{K}}^{\mathcal{P}}$ and vectors $\overline{\boldsymbol{\Phi}}^{\mathcal{P}}$, $\overline{\boldsymbol{b}}^{\mathcal{P}}$, and $\overline{\boldsymbol{g}}^{\mathcal{P}}$ are formed by expanding the associated elemental quantities of Eq. (3.58) through zero padding and using the relations[7]

$$\overline{K}^{\mathcal{P}}_{u_\Omega(i,\mathcal{P}),u_\Omega(j,\mathcal{P})} = K_{ij}^{\mathcal{P}}, \tag{3.62a}$$

$$\overline{\Phi}^{\mathcal{P}}_{u_\Omega(i,\mathcal{P})} = \Phi_i^{\mathcal{P}}, \tag{3.62b}$$

$$\overline{b}^{\mathcal{P}}_{u_\Omega(i,\mathcal{P})} = b_i^{\mathcal{P}}, \tag{3.62c}$$

$$\overline{g}^{\mathcal{P}}_{u_\Omega(i,\mathcal{P})} = g_i^{\mathcal{P}}. \tag{3.62d}$$

By setting the total residual $\mathfrak{R} = 0$ in accordance with Galerkin's method, the final matrix equation is obtained:

$$\mathbf{K}\boldsymbol{\Phi} = \boldsymbol{b} + \boldsymbol{g}. \tag{3.63}$$

Assuming that $\boldsymbol{M_r}$ is prescribed in the original problem, the elements of $\mathbf{K}$ and $\boldsymbol{b}$ can be obtained by direct integration of the formulae in Eqs. (3.59a) and (3.59b) using Gaussian quadrature rules.[7] However, the elements of $\boldsymbol{g}$ require close consideration of the imposed boundary conditions. From Eqs. (3.59c) and (3.60), g_i is given by

$$g_i = -\sum_{u}\left[\int_{\Gamma^u} \mathcal{S}_j^p(\boldsymbol{B}^+ - \boldsymbol{B}^-)\cdot\hat{\boldsymbol{n}}\, d\Gamma'\right] - \sum_{v}\left[\int_{\Gamma^v} \mathcal{S}_j^p \boldsymbol{B}\cdot\hat{\boldsymbol{n}}\, d\Gamma'\right], \tag{3.64}$$

where Γ^u and Γ^v represent internal interfaces and external boundary surface segments of the finite element mesh that are connected to the ith global node, respectively. Hence, p is such that $i = \mathcal{U}_\Omega(j,p)$ and therefore $\mathcal{S}_j^p$ represents an appropriate basis function on the surface of integration. Additionally, the $+$ and $-$ signs in this expression are used to label the quantities on the upper and lower sides of associated surface. The normal component of the magnetic field is required to be continuous across internal interfaces:

$$\hat{\boldsymbol{n}}\cdot(\boldsymbol{B}^+ - \boldsymbol{B}^-) = 0, \tag{3.65}$$

therefore, the integrals over Γ^u in Eq. (3.64) vanish. Furthermore, if Γ^v is an external boundary where the magnetic field is expected to be tangential to the boundary surface:

$$\hat{\boldsymbol{n}}\cdot\boldsymbol{B} = 0, \tag{3.66}$$

then the integral over $\Gamma^{\mathcal{V}}$ in Eq. (3.64) also vanishes. However, if the magnetic field is expected to be orthogonal to $\Gamma^{\mathcal{V}}$, then the potential on this surface is constant. Thus, it is possible to specify

$$\Phi_i = 0, \tag{3.67}$$

permitting g_i to be discarded and Φ_i directly removed as an unknown from the system of equations. Consequently, regardless of the boundary condition imposed, Eq. (3.63) simplifies to

$$\mathbf{K\Phi} = \boldsymbol{b}. \tag{3.68}$$

The AC/DC module of the commercially available software package COMSOL MULTIPHYSICS[15] can be used solve magnetostatics problems in the absence of free currents with the FEM. In doing so, COMSOL MULTIPHYSICS solves the linear system of Eq. (3.68) using a flexible generalized minimum residual method[16] (FGMRES) with a geometric multigrid preconditioner.[17]

3.3.6 Magnetostatics Formulation with Currents

In regions with conductors carrying electrical current, it is not possible to define $\boldsymbol{H}$ according to Eq. (3.53). Such conditions arise in the design of superconducting magnets for MRI, where large circular coils carry immense currents in order to generate the strong magnetic field used to uniformly polarize the object being examined (see Chapter 7). In this case, a reduced potential formulation can be employed, where $\boldsymbol{H}$ is separated into two components:[18]

$$\boldsymbol{H} = \boldsymbol{H_s} + \boldsymbol{H_m}, \tag{3.69}$$

where $\boldsymbol{H_s}$ is the solenoidal component directly generated by the current distribution and $\boldsymbol{H_m}$ is the irrotational component due to any magnetic materials. Consequently, $\boldsymbol{H_s}$ can be calculated directly with the Biot-Savart integral:

$$\boldsymbol{H_s}(\boldsymbol{r}) = \frac{1}{4\pi}\int_{\mathcal{V}} \boldsymbol{J}(\boldsymbol{r}') \times \frac{\boldsymbol{r}-\boldsymbol{r}'}{\|\boldsymbol{r}-\boldsymbol{r}'\|^3}\, d\mathcal{V}', \tag{3.70}$$

and $\boldsymbol{H_m}$ can be expressed as the gradient of a scalar magnetic potential as in Eqs. (3.12) and (3.53):

$$\boldsymbol{H_m} = -\nabla\Phi_m. \tag{3.71}$$

The benefit of this approach is that the reduced potential formulation of Eq. (3.65) only needs to be applied in the distinct regions of Ω containing the current carrying conductors. In all other regions, Eq. (3.53) holds as before.

According to the generalized constitutive relationship in Eq. (3.52), the magnetic field is given by

$$\boldsymbol{B} = \mu(-\nabla\Phi_m + \boldsymbol{H_s}), \tag{3.72}$$

where $\boldsymbol{M_r} = 0$ has been assumed. Combining this equation with the divergence of $\boldsymbol{B}$ in Eq. (3.1d) yields

$$\nabla\cdot[\mu(-\nabla\Phi_m + \boldsymbol{H_s})] = 0. \tag{3.73}$$

If the free currents are restricted to regions of uniform μ, then by noting that $\nabla\cdot(\mu\boldsymbol{H_s}) = 0$, Eq. (3.73) simplifies to

$$-\nabla \cdot (\mu \nabla \Phi_m) = 0. \tag{3.74}$$

This equation is identical to Eq. (3.54) with $\boldsymbol{M_r} = 0$, hence, the steps of Galerkin's method can be followed to yield the weighted residual matrix equation (see § 3.3.5)

$$\boldsymbol{\mathfrak{R}}^{\mathcal{p}} = \mathbf{K}^{\mathcal{p}} \boldsymbol{\Phi}^{\mathcal{p}} - \boldsymbol{g}^{\mathcal{p}}, \tag{3.75}$$

where the elements of the $m_\Omega \times m_\Omega$ matrix $\mathbf{K}^{\mathcal{p}}$ and $m_\Omega \times 1$ vector $\boldsymbol{g}^{\mathcal{p}}$ are given by

$$K_{ij}^{\mathcal{p}} = \int_{\Omega^{\mathcal{p}}} \nabla \mathcal{S}_i^{\mathcal{p}} \cdot \left(\mu \nabla \mathcal{S}_j^{\mathcal{p}} \right) d\Omega', \tag{3.76a}$$

$$g_i^{\mathcal{p}} = - \int_{\partial \Omega^{\mathcal{p}}} \mathcal{S}_i^{\mathcal{p}} (\mu \boldsymbol{H_m}) \cdot \hat{\boldsymbol{n}} \, d\Gamma'. \tag{3.76b}$$

Summing the weighted residuals from all elements yields

$$\boldsymbol{\mathfrak{R}} = \sum_{\mathcal{p}=1}^{N_\Omega} \overline{\boldsymbol{\mathfrak{R}}}^{\mathcal{p}} = \sum_{\mathcal{p}=1}^{N_\Omega} (\overline{\mathbf{K}}^{\mathcal{p}} \overline{\boldsymbol{\Phi}}^{\mathcal{p}} - \overline{\boldsymbol{g}}^{\mathcal{p}}), \tag{3.77}$$

$$= \mathbf{K}\boldsymbol{\Phi} - \boldsymbol{g}, \tag{3.78}$$

where the augmented quantities $\overline{\mathbf{K}}^{\mathcal{p}}$, $\overline{\boldsymbol{\Phi}}^{\mathcal{p}}$, and $\overline{\boldsymbol{g}}^{\mathcal{p}}$ are given by Eq. (3.62). By setting the total residual $\boldsymbol{\mathfrak{R}} = 0$ as before, the final matrix equation is obtained:

$$\mathbf{K}\boldsymbol{\Phi} = \boldsymbol{g}. \tag{3.79}$$

The elements of **K** can be obtained by direct integration of the formula in Eq. (3.76a) using Gaussian quadrature rules,[7] whereas the elements of $\boldsymbol{g}$ are determined by the imposed boundary conditions once again. In general, g_i has the form

$$g_i = -\sum_{u}\left[\int_{\Gamma^u} \mathfrak{S}_j^p(\mu^+ \boldsymbol{H}_m^+ - \mu^- \boldsymbol{H}_m^-)\cdot\hat{\boldsymbol{n}}\, d\Gamma'\right] \qquad (3.80)$$

$$-\sum_{v}\left[\int_{\Gamma^v} \mathfrak{S}_j^p(\mu \boldsymbol{H}_m)\cdot\hat{\boldsymbol{n}}\, d\Gamma'\right].$$

where p is such that $i = \mathcal{U}_\Omega(j,p)$ and therefore $\mathfrak{S}_j^p$ represents an appropriate basis function on the surface of integration. Therefore, on internal interfaces where the condition of Eq. (3.65) holds, the integrals over Γ^u become

$$\int_{\Gamma^u} \mathfrak{S}_j^p(\mu^+ \boldsymbol{H}_m^+ - \mu^- \boldsymbol{H}_m^-)\cdot\hat{\boldsymbol{n}}\, d\Gamma' \qquad (3.81)$$

$$= \int_{\Gamma^u} \mathfrak{S}_j^p(\mu^- \boldsymbol{H}_s^- - \mu^+ \boldsymbol{H}_s^+)\cdot\hat{\boldsymbol{n}}\, d\Gamma',$$

which may be obtained by direct integration using Eq. (3.70) (see § 3.2). Similarly, if Γ^v is an external boundary for which the tangential field condition of Eq. (3.66) holds, then the integral over Γ^v becomes

$$\int_{\Gamma^v} \mathfrak{S}_j^p(\mu \boldsymbol{H}_m)\cdot\hat{\boldsymbol{n}}\, d\Gamma' = -\int_{\Gamma^v} \mathfrak{S}_j^p(\mu \boldsymbol{H}_s)\cdot\hat{\boldsymbol{n}}\, d\Gamma'. \qquad (3.82)$$

Lastly, if $\boldsymbol{H_m}$ is expected to be orthogonal to Γ^{ν}, then it is possible to specify

$$\Phi_i = 0, \tag{3.83}$$

permitting g_i to be discarded and Φ_i once again directly removed as an unknown from the system of equations.

General magnetostatics problems involving current carrying conductors can be solved with the FEM using the commercially available software package OPERA-3D,[19] which employs the analysis module TOSCA to solve the linear system of Eq. (3.79) via an incomplete Cholesky decomposition and conjugate gradient method.[20]

3.4 REFERENCES

[1] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).

[2] D. R. Lide, ed. *CRC Handbook of Chemistry and Physics*, 67th ed. (CRC Press, Boca Raton, 1986).

[3] R. Skomski and J. M. D. Coey, *Permanent Magnetism* (Institute of Physics Publishing, Bristol, 1999).

[4] J. W. Fiepke, *Permanent Magnet Materials,* in: *ASM Handbook,* vol. 2, *Properties and Selection: Nonferrous Alloys and Special-Purpose Materials* (ASM International, Materials Park, 1990).

[5] D. J. Griffiths, *Introduction to Electrodynamics*, 3rd ed. (Prentice Hall, Upper Saddle River, 1999).

[6] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists,* 6th ed. (Elsevier, Burlington, 2005).

[7] J. Jin, *The Finite Element Method in Electromagnetics*, 2nd ed. (Wiley, New York, 2002).

[8] W. U. Wei, H. E. Yuan, M. A. Li-Zhen, H. Wen-Xue, and X. Jia-Wen, "Optimal design of a 7 T highly homogenous superconducting magnet for a Penning trap," *Chinese Phys. C.* **34**(7), 978–982 (2010).

[9] MATLAB, ver. 7.11 (The MathWorks, Natick, 2010).

[10] M. Abramowitz and I. A. Stegun, eds. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th ed. (United States Government Printing Office, Washington, D.C., 1972).

[11] L. K. Forbes, S. Crozier, and D. M. Doddrell, "Rapid computation of static fields produced by thick circular solenoids," *IEEE Trans. Magn.* **33**(5), 4405–4410 (1997).

[12] W. Gautschi, *Orthogonal Polynomials, Quadrature, and Approximation: Computational Methods and Software (in MATLAB)*, in: F. Marcellán and W. Van Assche, eds. *Orthogonal Polynomials and Special Functions: Computation and Applications,* (Springer, Berlin, 2006) pp. 1–77.

[13] O. C. Zienkiewicz, R. L. Taylor, and J. Z. Zhu, *The Finite Element Method: Its Basis and Fundamentals*, 6th ed. (Elsevier, Oxford, 2005).

[14] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods,* 3rd ed. (Springer, New York, 2008).

[15] COMSOL MULTIPHYSICS, ver. 3.4 (Comsol AB, Stockholm, 2007).

[16] Y. Saad, "A flexible inner-outer preconditioned GMRES algorithm," *SIAM J. Sci. Comput.* **14**(2), 461–469 (1993).

[17] W. Hackbusch, *Multi-grid Methods and Applications* (Springer, Berlin, 1985).

[18] J. Simkin and C. W. Trowbridge, "Three-dimensional nonlinear electromagnetic field computations, using scalar potentials," *IEE Proc. B.* **127**(6) 368–374 (1980).

[19] OPERA-3D, ver. 13 (Vector Fields, Oxford, 2009).

[20] J. A. Meijerink and H. A. van der Vorst, "An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix," *Math. Comput.* **31**(137) 148–162 (1977).

CHAPTER 4
THEORY AND TECHNIQUES II
NUMERICAL OPTIMIZATION

The topic of mathematical optimization deals with the minimization or maximization of a given function subject to constraints on its variables. It remains a major area of ongoing research, with significant efforts tracing back to the seminal work by Euler and Lagrange on the development of variational calculus in the 18th century. Optimization methods have widespread applications in nearly every field in which numerical information is processed, such as science, engineering, economics, and finance. In particular, mathematical optimization techniques play a central role in the optimal design of permanent and superconducting magnet systems for magnetic resonance imaging as presented in this thesis.

4.1 THEORY

4.1.1 Introduction

The general optimization problem may be mathematically expressed as

$$\text{minimize } f(\boldsymbol{p}), \tag{4.1a}$$

$$\text{subject to } \boldsymbol{p} \in \mathfrak{P}, \tag{4.1b}$$

where $f: \mathbb{R}^n \to \mathbb{R}$ is known as the objective function, $\boldsymbol{p} = [p_1, p_2, \dots, p_n]^T$ is the design vector with n independent design variables as elements, and $\mathfrak{P}$ is the feasible domain or feasible set. This general formulation also applies to maximization problems without loss of generality, since maximizing a function f is equivalent to minimizing the function $-f$.

The feasible domain $\mathfrak{P}$ is a subset of $\mathbb{R}^n$ that is defined by the particular conditions imposed on the design variables. Therefore, the problem in Eq. (4.1) is referred to as a constrained optimization problem and Eq. (4.1b) is known as the constraint. A point $\boldsymbol{p}$ which satisfies the constraint is said to be a feasible point. Typically, the feasible set is defined in terms of the constraint functions $c_i(\boldsymbol{p})$ as

$$\mathfrak{P} = \{\boldsymbol{p} | c_i(\boldsymbol{p}) = 0, i \in \mathcal{E}\} \cup \{\boldsymbol{p} | c_i(\boldsymbol{p}) \leq 0, i \in \mathcal{I}\}, \tag{4.2}$$

where $\mathcal{E}$ and $\mathcal{I}$ are the index sets for the equality and inequality constraints, respectively. In the special case where $\mathcal{E} = \mathcal{I} = \emptyset$, we have $\mathfrak{P} = \mathbb{R}^n$ and the problem in Eq. (4.1) is referred to as an unconstrained optimization problem (see § 4.2).

Optimization problems with the general form of Eq. (4.1) may be further classified according to the nature of the objective function, constraint functions, and the design vector. For example, if the objective and constraint functions are linear in $\boldsymbol{p}$, the problem is known as a linear programming (LP) problem (see § 4.3.2). If however at least one of the objective or constraint functions are nonlinear in $\boldsymbol{p}$, we have a nonlinear programming or nonlinear optimization problem.

A point $\boldsymbol{p}^*$ that is a solution to Eq. (4.1) is called a minimizer of the function f over $\mathfrak{P}$. In general, we ultimately wish to locate the global minimizer of Eq. (4.1), which in particular satisfies $f(\boldsymbol{p}^*) \leq f(\boldsymbol{p})$ for all $\boldsymbol{p} \in \mathfrak{P}$. However, this is extremely difficult in practice and often algorithms for solving Eq. (4.1) are only able to find a local minimizer, which satisfies $f(\boldsymbol{p}^*) \leq f(\boldsymbol{p})$ within a neighbourhood $\mathfrak{P}^*$ of $\boldsymbol{p}^*$.

When solving an optimization problem, we rarely have detailed prior knowledge of the overall shape and behaviour of the objective function. For example, the objective function may not be explicitly defined in terms of the variables, or the shear number of variables alone may complicate the problem to an incomprehensible extent. Consequently, optimization algorithms typically adopt an iterative approach. They begin with an initial starting point $\boldsymbol{p}^{(1)}$ and generate a sequence of iterates $\{\boldsymbol{p}^{(k)}\}$ that terminates when no further progress can be made or when an apparent solution has been reached. In general, this iterative procedure can be expressed as

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{p}^{(k)} + \alpha^{(k)}\boldsymbol{p}^{(k)}, \quad \mathrm{k} \geq 1, \tag{4.3}$$

where the vector $\boldsymbol{d}^{(k)}$ is the search direction, $\alpha^{(k)}$ is the scalar step size, and the superscript k denotes the iteration number. Algorithms therefore rely on the local information about f that is available at the current and previous iterations to generate search directions that yield future iterates exhibiting a decrease in the objective function value, that is:

$$f\left(\boldsymbol{p}^{(k+1)}\right) \leq f\left(\boldsymbol{p}^{(k)}\right). \tag{4.4}$$

We say a vector $\boldsymbol{d} \neq 0$ is a feasible direction at $\boldsymbol{p} \in \mathfrak{P}$ if there exists a nonzero α_0 such that $\boldsymbol{p} + \alpha\boldsymbol{d} \in \mathfrak{P}$ for all $\alpha \in [0, \alpha_0)$. By requiring the search direction $\boldsymbol{d}^{(k)}$ computed at each iteration to be a feasible direction, most algorithms ensure that the sequence of iterates $\{\boldsymbol{p}^{(k)}\}$ are feasible.

4.1.2 First-Order Necessary Conditions

The methods from which optimization algorithms calculate new search directions, or determine whether a given iterate is an acceptable solution, are

typically based on certain necessary and sufficient conditions regarding the gradient ∇f or Hessian $\nabla^2 f$ that can be established at a local minimizer $\boldsymbol{p}^*$.

A fundamental first-order necessary condition is that if $\boldsymbol{p}^*$ is a local minimizer of the function f over $\mathfrak{P}$, then for any feasible direction $\boldsymbol{d}$ at $\boldsymbol{p}^*$, we have[1]

$$\frac{\partial f}{\partial \boldsymbol{d}}(\boldsymbol{p}^*) = \boldsymbol{d}^T \nabla f(\boldsymbol{p}^*) \geq 0. \tag{4.5}$$

Essentially, this means that the rate of increase of f at $\boldsymbol{p}^*$ is nonnegative in any feasible direction at a local minimizer. If $\boldsymbol{p}^*$ is an interior point of $\mathfrak{P}$, or if the problem is unconstrained, then Eq. (4.5) simplifies to the familiar result[1]

$$\nabla f(\boldsymbol{p}^*) = 0. \tag{4.6}$$

Points that satisfy Eq. (4.6) are known as stationary points. Feasible directions that do not satisfy Eq. (4.5) are known as descent directions.

4.2 UNCONSTRAINED OPTIMIZATION METHODS

4.2.1 Introduction

Unconstrained optimization problems are a special class of problems in the form of Eq. (4.1) with no constraints present. Specifically, $\mathcal{E} = \mathcal{I} = \emptyset$ and we have

$$\text{minimize } f(\boldsymbol{p}), \tag{4.7a}$$

$$\text{subject to } \boldsymbol{p} \in \mathbb{R}^n. \tag{4.7b}$$

Solutions to Eq. (4.7) are strictly defined by the behaviour of f alone and therefore the methods that have been developed to solve this type of problem are significantly more straightforward and have a stronger theoretical foundation than those dealing with the more complicated class of general constrained optimization problems.

Unconstrained optimization methods are typically iterative and implement updates of the form in Eq. (4.3) in the search for a solution to Eq. (4.7). Specifically, the design vector is updated at the kth iteration according to

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{p}^{(k)} + \alpha^{(k)}\boldsymbol{d}^{(k)}, \quad k \geq 1, \tag{4.8}$$

with the search direction $\boldsymbol{d}^{(k)}$ often taking the general form

$$\boldsymbol{d}^{(k)} = -\mathbf{G}^{(k)^{-1}}\nabla f^{(k)} + \beta^{(k)}\boldsymbol{d}^{(k-1)}, \tag{4.9}$$

where $\nabla f^{(k)} = \nabla f\left(\boldsymbol{p}^{(k)}\right)$. The expressions defining the symmetric nonsingular matrix $\mathbf{G}^{(k)}$ and the scalar $\beta^{(k)}$ vary amongst the many established unconstrained optimization algorithms. If $\mathbf{G}^{(k)}$ is positive definite, $\beta^{(k)} \geq 0$, and $\boldsymbol{d}^{(k-1)^T}\nabla f^{(k)} < 0$, then $\boldsymbol{d}^{(k)}$ is a descent direction:

$$\boldsymbol{d}^{(k)^T}\nabla f^{(k)} = -\nabla f^{(k)^T}\mathbf{G}^{(k)^{-1}}\nabla f^{(k)} + \beta^{(k)}\boldsymbol{d}^{(k-1)^T}\nabla f^{(k)} < 0, \tag{4.10}$$

and the utility of Eq. (4.9) becomes apparent. In the following sections we review a selection of the most common methods for solving Eq. (4.7) and the associated choices for the definition of $\mathbf{G}^{(k)}$ and $\beta^{(k)}$.

4.2.2 The Steepest Descent Method

The steepest descent method[3] is the simplest first-order optimization method for solving problems in the form of Eq. (4.7). By choosing $\mathbf{G}^{(k)}$ and $\beta^{(k)}$ from Eq. (4.9) as

$$\mathbf{G}^{(k)} = \mathbf{I}, \tag{4.11a}$$

$$\beta^{(k)} = 0, \tag{4.11b}$$

where $\mathbf{I}$ is the identity matrix, the search direction for the steepest descent method simply becomes

$$\boldsymbol{d}^{(k)} = -\nabla f^{(k)}. \tag{4.12}$$

Motivation for the result in Eq. (4.12) follows from the fact that the gradient vector $\nabla f(\boldsymbol{p})$ points in the direction of maximum rate of increase of f at $\boldsymbol{p}$, or conversely, $-\nabla f(\boldsymbol{p})$ points in the direction of maximum rate of decrease. Therefore, this choice for the search direction is a reasonable starting point for a search for the minimizer of f. If we consider the linear approximation of f at $\boldsymbol{p}^{(k)}$ via the Taylor series expansion:

$$f\left(\boldsymbol{p}^{(k)} + \boldsymbol{\delta}\right) = f\left(\boldsymbol{p}^{(k)}\right) + \boldsymbol{\delta}^T \nabla f^{(k)} + \mathcal{O}(\delta^2), \tag{4.13}$$

where $\boldsymbol{\delta} = -\alpha \nabla f^{(k)}$, $\nabla f^{(k)} \neq 0$, then

$$f\left(\boldsymbol{p}^{(k)} + \boldsymbol{\delta}\right) = f\left(\boldsymbol{p}^{(k)}\right) - \alpha \left\|\nabla f^{(k)}\right\|^2 + \mathcal{O}(\alpha^2), \tag{4.14}$$

and for sufficiently small $\alpha > 0$ we have

$$f\left(\boldsymbol{p}^{(k)} + \boldsymbol{\delta}\right) < f\left(\boldsymbol{p}^{(k)}\right). \tag{4.15}$$

Thus, the steepest descent direction is always a descent direction when the gradient is nonzero, and so it is always possible to reduce the value of f by iteratively updating the design vector according to Eqs. (4.8) and (4.12) until $\nabla f = 0$ and convergence to a local minimizer of Eq. (4.7) is achieved. A rigorous treatment of the convergence properties of the steepest descent method can be found in the literature.[4]

It can be shown that if the step size $\alpha^{(k)}$ is chosen exactly so as to minimize f along $\boldsymbol{d}^{(k)}$ at each steepest descent iteration, then successive search directions will be orthogonal, that is $\boldsymbol{d}^{(k+1)^T}\boldsymbol{d}^{(k)} = 0$. This leads to a geometrical zig-zagging pattern for the sequence of iterates $\{\boldsymbol{p}^{(k)}\}$, which in turn results in a slow approach to the solution of many optimization problems in practice. However, selecting $\alpha^{(k)}$ in this fashion is computationally expensive and typically avoided in practice, thus alleviating this undesirable behaviour.

The steepest descent method is simple to implement and does not strongly depend on the smoothness of the objective function considered. Furthermore, it can easily be extended to a constrained optimization algorithm (see § 4.3.3), and thus it is frequently used as a starting point for solving many practical optimization problems.

4.2.3 The Conjugate Gradient Method

The conjugate gradient method is another first-order technique that has been proposed for solving nonlinear optimization problems.[5] Similar to the steepest descent method, the conjugate gradient method requires the calculation of the objective function gradient ∇f at each iteration. The

advantage of the conjugate gradient method is that the theoretical foundation for generating the search directions at each iteration is based on a quadratic model of the objective function. If we consider the Taylor series expansion of f at $\boldsymbol{p}^{(k)}$:

$$f\left(\boldsymbol{p}^{(k)}+\boldsymbol{\delta}\right)=f\left(\boldsymbol{p}^{(k)}\right)+\boldsymbol{\delta}^{T}\nabla f^{(k)}+\frac{1}{2}\boldsymbol{\delta}^{T}\nabla^{2}f^{(k)}\boldsymbol{\delta}+\mathcal{O}(\delta^{3}), \tag{4.16}$$

where $\nabla^2 f^{(k)} = \nabla^2 f\left(\boldsymbol{p}^{(k)}\right)$, we see that a quadratic model is generally a better approximation of f than the linearization in Eq. (4.13), for sufficiently small δ. Consequently, by tailoring a method to specifically perform well for quadratic functions, it may be possible to achieve improved convergence for general nonlinear functions as well.

In the conjugate gradient method, the search directions are chosen in an effort to obtain the property of **Q**-conjugacy, so that for all $j \neq k$, we have

$$\boldsymbol{d}^{(j)^{T}}\mathbf{Q}\boldsymbol{d}^{(k)}=0, \tag{4.17a}$$

where **Q** is real symmetric $n \times n$ matrix. This is achieved by choosing $\mathbf{G}^{(k)}$ from Eq. (4.9) as

$$\mathbf{G}^{(k)}=\mathbf{I}, \tag{4.18a}$$

with $\beta^{(k)}$ defined by either the Fletcher-Reeves formula:[5]

$$\beta^{(k)}=\frac{\nabla f^{(k)^{T}}\nabla f^{(k)}}{\nabla f^{(k-1)^{T}}\nabla f^{(k-1)}}, \tag{4.18b}$$

or the Polak-Ribière formula:[6]

$$\beta^{(k)} = \frac{\nabla f^{(k)^T}\left(\nabla f^{(k)} - \nabla f^{(k-1)}\right)}{\nabla f^{(k-1)^T}\nabla f^{(k-1)}}. \tag{4.18c}$$

In all cases, the conjugate gradient search directions take the form

$$\boldsymbol{d}^{(k)} = \begin{cases} -\nabla f^{(1)}, & \text{for } k = 1, \\ -\nabla f^{(k)} + \beta^{(k)}\boldsymbol{d}^{(k-1)}, & \text{for } k > 1. \end{cases} \tag{4.19}$$

The choice of using either Eq. (4.18a) or (4.18a) for computing $\beta^{(k)}$ is inconsequential for quadratic functions. However, the two variations can lead to noticeable differences in performance for general nonlinear functions. Although both approaches demonstrate unique advantages in certain cases, the Polak-Ribière formula is generally considered to be slightly more robust.[2]

Practically speaking, **Q**-conjugacy of the search directions can only be fully achieved for quadratic objective functions. In this special case, the conjugate gradient method exhibits the desirable property of finite termination, as the solution to Eq. (4.7) is obtained in at most n steps.

Although the conjugate gradient method often performs better than the steepest descent method, there is no guarantee that the search directions given by Eq. (4.19) will be descent directions.[2] Hence, conjugate gradient iterations may actually lead to an increase in f. To overcome this shortcoming, a modification is often implemented to restart the algorithm every n steps (or less). This acts to refresh the search direction calculation by discarding information from distant previous iterations that may no longer be beneficial.

4.3 CONSTRAINED OPTIMIZATION METHODS

4.3.1 Introduction

The most general class of optimization problems are those in which constraints are imposed on the design variable $\boldsymbol{p}$, so that the solution of the problem typically requires knowledge of how the objective function f behaves with respect to the topology of the feasible domain $\mathfrak{P}$. Mathematically, constrained optimization problems may be expressed as

$$\text{minimize } f(\boldsymbol{p}), \tag{4.20a}$$

$$\text{subject to } \boldsymbol{p} \in \mathfrak{P}. \tag{4.20b}$$

Methods which attempt to solve the problem in Eq. (4.20) are usually formulated from first and second-order optimality conditions. These methods are substantially more complicated than those dealing with unconstrained optimization, and thus in the following sections we restrict ourselves to the particularly special techniques of linear programming, as well as a basic projection method for extending the techniques of unconstrained optimization to problems with constraints.

4.3.2 Linear Programming

Linear programming (LP) refers to a branch of constrained optimization problems in which the objective function is linear and the constraints form a set of linear equations or inequalities. Historically speaking, LP methods have found many applications in science and economics. In particular, LP became especially popular during World War II, when it arose naturally in problems associated with the production and allocation of scarce resources. Consequently, the development of techniques tailored to this class of

constrained optimization problems has received great attention. Mathematically, an LP problem in standard form is expressed as

$$\text{minimize } f(\boldsymbol{p}) = \boldsymbol{\varsigma}^T \boldsymbol{p}, \tag{4.21a}$$

$$\text{subject to } \begin{cases} \mathbf{A}\boldsymbol{p} = \boldsymbol{b}, \\ \boldsymbol{p} \geq 0. \end{cases} \tag{4.21b}$$

where $\boldsymbol{\varsigma} \in \mathbb{R}^n$ is the cost vector, $\boldsymbol{b} \in \mathbb{R}^n$ is the constraint vector satisfying $\boldsymbol{b} \geq 0$, and $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the equality constraint matrix. With reference to Eq. (4.20), we see that the feasible domain $\mathfrak{P}$ is defined as

$$\mathfrak{P} = \{\boldsymbol{p} | \mathbf{A}\boldsymbol{p} = \boldsymbol{b}, \boldsymbol{p} \geq 0\}. \tag{4.22}$$

If the feasible set is nonempty and the objective function is bounded over $\mathfrak{P}$, then there must exist at least one solution to Eq. (4.21). Furthermore, since the objective function is linear, any solution to Eq. (4.21) is a global minimizer.

It is important to note that any LP problem may be expressed in the standard form of Eq. (4.21). Consider the seemingly more general expression:

$$\text{minimize } f(\boldsymbol{p}) = \boldsymbol{\varsigma}^T \boldsymbol{p}, \tag{4.23a}$$

$$\text{subject to } \begin{cases} \mathbf{A}_{\text{eq}}\boldsymbol{p} = \boldsymbol{b}_{\text{eq}}, \\ \mathbf{A}_{\text{ineq1}}\boldsymbol{p} \leq \boldsymbol{b}_{\text{ineq1}}, \\ \mathbf{A}_{\text{ineq2}}\boldsymbol{p} \geq \boldsymbol{b}_{\text{ineq2}}. \end{cases} \tag{4.23b}$$

By introducing auxiliary variables $\boldsymbol{t}_1, \boldsymbol{t}_2 \geq 0$, we may write

$$\mathbf{A}_{\text{ineq1}}\boldsymbol{p} + \boldsymbol{t}_1 = \boldsymbol{b}_{\text{ineq1}}, \tag{4.24a}$$

$$\mathbf{A}_{\text{ineq2}}\boldsymbol{p} - \boldsymbol{t}_2 = \boldsymbol{b}_{\text{ineq2}}. \tag{4.24b}$$

Furthermore, we can define $\boldsymbol{p} = \boldsymbol{p}^+ - \boldsymbol{p}^-$ where

$$\boldsymbol{p}^+ = \max(\boldsymbol{p}, 0) \geq 0, \tag{4.25a}$$

$$\boldsymbol{p}^- = \max(-\boldsymbol{p}, 0) \geq 0. \tag{4.25b}$$

Hence, by introducing the augmented quantities

$$\overline{\boldsymbol{p}} = \begin{bmatrix} \boldsymbol{p}^+ \\ \boldsymbol{p}^- \\ \boldsymbol{t}_1 \\ \boldsymbol{t}_2 \end{bmatrix}, \tag{4.26a}$$

$$\overline{\boldsymbol{\varsigma}} = \begin{bmatrix} \boldsymbol{\varsigma} \\ -\boldsymbol{\varsigma} \\ 0 \\ 0 \end{bmatrix}, \tag{4.26b}$$

$$\overline{\mathbf{A}} = \begin{bmatrix} \mathbf{A}_{\text{eq}} & -\mathbf{A}_{\text{eq}} & 0 & 0 \\ \mathbf{A}_{\text{ineq1}} & -\mathbf{A}_{\text{ineq1}} & \mathbf{I} & 0 \\ \mathbf{A}_{\text{ineq2}} & -\mathbf{A}_{\text{ineq2}} & 0 & \mathbf{I} \end{bmatrix}, \tag{4.26c}$$

$$\overline{\boldsymbol{b}} = \begin{bmatrix} \boldsymbol{b}_{\text{eq}} \\ \boldsymbol{b}_{\text{ineq1}} \\ \boldsymbol{b}_{\text{ineq2}} \end{bmatrix}, \tag{4.26d}$$

we may rewrite problem Eq. (4.23) in the equivalent standard form:

$$\text{minimize } \bar{f}(\overline{\boldsymbol{p}}) = \overline{\boldsymbol{\varsigma}}^T \overline{\boldsymbol{p}}, \tag{4.27a}$$

$$\text{subject to } \begin{cases} \overline{\mathbf{A}}\overline{\boldsymbol{p}} = \overline{\boldsymbol{b}}, \\ \overline{\boldsymbol{p}} \geq 0. \end{cases} \tag{4.27b}$$

An interesting feature of LP problems is that the constraint set $\mathfrak{P}$ defines a convex polytope in $\mathbb{R}^n$ (Figure 4.1). In fact, a fundamental theorem of LP is that if any solutions to Eq. (4.21) exist, then they must occur on the boundary of that polytope, and at least one solution exists at a vertex.[2] Thus, although $\mathfrak{P}$ may contain an infinite number of feasible points, a global minimizer of Eq. (4.21) must exist at one of the finite number of vertices of $\mathfrak{P}$. This important result forms the basis of all modern LP algorithms.

The well known simplex method is an especially elegant and straightforward approach to solving LP problems.[7, 8] This method is still commonly used and is the preferred choice for solving LP problems that do not contain a large number of variables. The simplex method is essentially an algebraic technique consisting of two major phases, as illustrated in Figure 4.1. In the first phase, an initial vertex of $\mathfrak{P}$ is located. Provided that this vertex is not the solution to the problem, linearity of f suggests that at least one edge containing the current vertex is a direction along which the value of f decreases. Therefore, in the second phase, such an edge is located and the algorithm proceeds to the next vertex on the opposite end. Proceeding in this manner, the simplex method will typically visit as many vertices of $\mathfrak{P}$ as necessary, until a solution is located. The algebraic details of this procedure are available in the literature.[4]

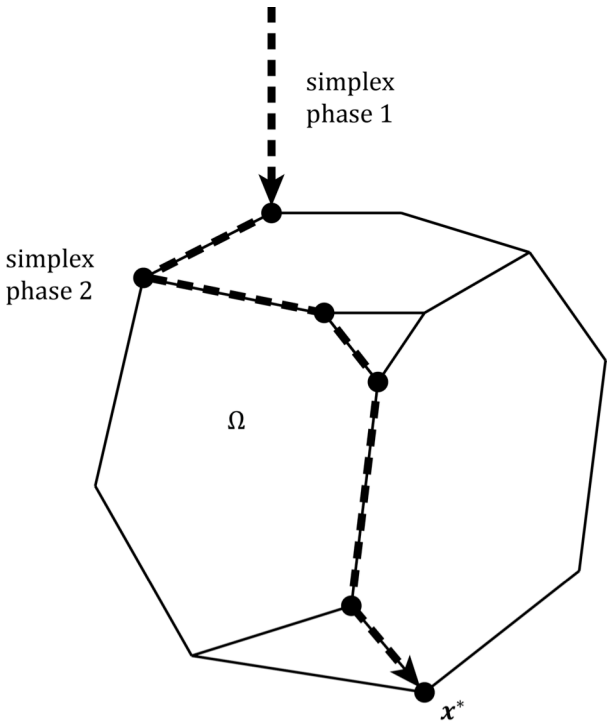


FIGURE 4.1: The two phases of the simplex method. In the first phase, an initial vertex of the feasible domain $\mathfrak{P}$ is located. In the second phase, the method steps through adjacent vertices until the minimum $\boldsymbol{p}^*$ is reached.

In certain special cases, there can be an iteration of the simplex method for which the objective function f is constant along all edges adjacent to the current vertex. Consequently, any step chosen by this method will leave the value of f unchanged. Iterations of this type are called degenerate steps. It is possible that several degenerate steps can occur in succession, causing the simplex method to arrive back at a vertex that was previously visited. Typically, the algorithm would then repeatedly cycle through the same degenerate steps and no further progress would be achieved. In practice, cycling may indeed be observed with large problems, however, modifications to the simplex method can be made which attempt to circumvent this behaviour.[4]

An obvious disadvantage of the simplex algorithm is the possibility that the method will visit many vertices of $\mathfrak{P}$ in a large-dimensional problem. In fact, examples can be provided for which the simplex method visits each of the 2^n vertices of the feasible domain before locating the solution.[9] Therefore, the simplex method is said to have exponential worst-case complexity.

An alternative method to solving LP problems was proposed by Karmarkar[10] in 1984. Karmarkar's projective algorithm is referred to as an interior-point method, since the algorithm traverses the interior of the feasible polytope $\mathfrak{P}$ when searching for the solution to Eq. (4.21), as opposed to the simplex method which traverses the boundary. Karmarkar's method essentially employs a modified Newton step that is projected onto a central region in $\mathfrak{P}$ that leads from the initial point to the solution vertex. This results in a decrease in f at each iteration, while approaching the solution from a more direct path than would be possible by following the boundary of $\mathfrak{P}$. By doing so, Karmarkar's method achieves polynomial complexity and performs much faster than the simplex method on large-scale LP problems. A review of the sophisticated details of Karmarkar's method is provided in the literature.[2]

After the introduction of Karmarkar's method, intense efforts were directed at developing more advanced interior-point techniques with better performance. The most notable such method is the primal-dual predictor-corrector algorithm[11] incorporated into the LIPSOL algorithm,[12] which has been implemented in many commercial software packages such as the OPTIMIZATION TOOLBOX of MATLAB.[13]

4.3.3 Projected Gradient Methods

Several methods were previously discussed for solving unconstrained optimization problems in the form of Eq. (4.7a). These techniques iteratively updated the design vector according to

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{p}^{(k)} + \alpha^{(k)}\boldsymbol{d}^{(k)}. \tag{4.28}$$

If this formula is directly applied to a constrained optimization problem, the sequence of iterates $\{\boldsymbol{p}^{(k)}\}$ may not satisfy the constraint Eq. (4.20b). However, through a simple modification of Eq. (4.28), the structure of unconstrained optimization methods may be extended to problems with constraints. The modified update formula is given as[14]

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{\Pi}\left(\boldsymbol{p}^{(k)} + \alpha^{(k)}\boldsymbol{d}^{(k)}\right), \tag{4.29}$$

where $\boldsymbol{\Pi}(\boldsymbol{p})$ is the projection of $\boldsymbol{p}$ onto $\mathfrak{P}$, and $\boldsymbol{\Pi}$: $\mathbb{R}^n \to \mathbb{R}^n$ is the projection operator defined according to

$$\|\boldsymbol{p} - \boldsymbol{\Pi}(\boldsymbol{p})\| = \min \|\boldsymbol{p} - \overline{\boldsymbol{p}}\|, \quad \overline{\boldsymbol{p}} \in \mathfrak{P}. \tag{4.30}$$

Hence, $\boldsymbol{\Pi}(\boldsymbol{p})$ is the closest point in $\mathfrak{P}$ to $\boldsymbol{p}$, with respect to the Euclidean norm. Therefore, provided $\boldsymbol{\Pi}(\boldsymbol{p})$ is well defined, the update formula in Eq. (4.29) may be used to generate a sequence of feasible iterates that satisfy Eq. (4.20b).

An explicit formula for $\boldsymbol{\Pi}(\boldsymbol{p})$ may not always be available. However, for the case of box constraints, where $\mathfrak{P}$ is defined as

$$\mathfrak{P} = \{\boldsymbol{p} | l_i \leq p_i \leq u_i, i = 1, \dots, n\}, \tag{4.31}$$

the projection $\boldsymbol{y} = \boldsymbol{\Pi}(\boldsymbol{p})$ can be defined according to[4]

$$y_i = \min\{u_i, \max\{l_i, p_i\}\} = \begin{cases} l_i, & \text{if } p_i < l_i, \\ p_i, & \text{if } l_i \le p_i \le u_i, \\ u_i, & \text{if } p_i > u_i. \end{cases} \tag{4.32}$$

If the steepest descent search direction in Eq. (4.12) is employed in Eq. (4.29), the update formula for the projected gradient algorithm becomes:

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{\Pi}\left(\boldsymbol{p}^{(k)} - \alpha^{(k)} \nabla f^{(k)}\right). \tag{4.33}$$

Similarly, if the conjugate gradient search direction in Eq. (4.19) is used, the update formula for the projected gradient algorithm becomes:

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{\Pi}\left(\boldsymbol{p}^{(k)} + \alpha^{(k)}\left[-\nabla f^{(k)} + \beta^{(k)} \boldsymbol{d}^{(k-1)}\right]\right). \tag{4.34}$$

There has been a great deal of research characterizing the convergence properties of projected gradient algorithms,[14–16] making this simple class of methods an attractive choice for solving the constrained optimization problem of Eq. (4.20). In particular, it can be shown that similar to the steepest descent and conjugate gradient methods, the iterates generated from Eqs. (4.33) and (4.34) will converge to a point satisfying the fundamental first-order necessary conditions[4] stated in Eq. (4.5).

4.4 REFERENCES

[1] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. (Wiley, New York, 1987).

[2] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. (Springer, New York, 2006).

[3] A. Cauchy, “Méthodes générales pour la résolution des systèmes d‘équations simultanées,” *C. R. Acad. Sci.* **25**, 536−538 (1847).

[4] E. K. P. Chong and S. H. Żak, *An Introduction to Optimization*, 3rd ed. (Wiley, New York, 2008).

[5] R. Fletcher and C.M. Reeves, “Function minimization by conjugate gradients,” *Comput. J.* **7** (2), 149−154 (1964).

[6] E. Polak, *Computational Methods in Optimization: A Unified Approach* (Academic Press, New York, 1971).

[7] L. Kantorovich, “Mathematical methods of organizing and planning production,” *Manag. Sci.* **6** (4), 366−422 (1960).

[8] G. B. Dantzig, *Linear Programming and Extensions* (Princeton University Press, Princeton, 1963).

[9] V. Klee and G. J. Minty, *How Good is the Simplex Algorithm?* in: O. Shisha, ed. *Inequalities III* (Academic Press, New York, 1972) pp. 159−175.

[10] N. Karmarkar, “A new polynomial-time algorithm for linear programming,” *Combinatorica*, **4** (4), 373−395 (1984).

[11] S. Mehrotra, “On the implementation of a primal-dual interior point method,” *SIAM J. Optim.* **2** (4), 575−601 (1992).

[12] Y. Zhang, “Solving large-scale linear programs by interior-point methods under the MATLAB environment,” *Optim. Method. Softw.* **10** (1), 1−31 (1998).

[13] MATLAB, ver. 7.11 (The MathWorks, Natick, 2010).

[14] A. A. Goldstein, "Convex programming in Hilbert space," *Bull. Amer. Math. Soc.* **70**(5), 709−710 (1964).

[15] J. C. Dunn, "Global and asymptotic convergence rate estimates for a class of projected gradient processes," *SIAM J. Control Optim.* **19**(3), 368−400 (1981).

[16] C. Wang and N. Xiu, "Convergence of the gradient projection method for generalized convex minimization," *Comput. Optim. Appl.* **16**(2), 111−120 (2000).

CHAPTER 5
NON-AXISYMMETRIC POLE PIECE SHAPE OPTIMIZATION FOR BIPLANAR PERMANENT MAGNET MRI SYSTEMS

A version of this chapter has been published. T. Tadic and B. G. Fallone, "Three-dimensional non-axisymmetric pole piece shape optimization for biplanar permanent magnet MRI systems," *IEEE Trans. Magn.*, **47** (1), 231–238 (2011).

5.1 INTRODUCTION

Magnetic field inhomogeneities in the imaging volume of an MRI system are directly translated into geometrical distortions in the final images produced (see § 2.2.7). Presently, the major limiting factor in the design of permanent magnet systems is the difficulty in obtaining a large diameter spherical volume (DSV) exhibiting a sufficiently homogenous magnetic field while maintaining suitable patient access and practical device dimensions.[1] Efforts in permanent magnet passive shimming have helped to cope with this issue,[2–6] however the effectiveness of these methods remains limited by the extent of the initial field inhomogeneities present after manufacturing.

There have been many attempts to improve the magnetic field uniformity of biplanar permanent magnet systems through the optimal design of the shape and configuration of the magnetic pole pieces.[1, 7–16] Generally, these efforts have focused on annular shimmed pole piece designs,[1, 7–11] or shape optimization of the pole piece surfaces.[12–16] The final pole piece geometries obtained and evaluated in these cases are all notably axisymmetric. As such, these designs are inherently limited in their ability to produce optimally

homogenous fields since biplanar assemblies typically have non-axisymmetric configurations. As biplanar magnet systems advance towards more compact geometries with reduced pole sizes, new design methods and pole piece geometries are sought that yield improved levels of field uniformity as compared to those currently provided.

Ultimately, field uniformity on the order of $10-50$ ppm over a 40 cm diameter spherical volume (DSV) is desired, which has been assumed throughout this work to translate into a specified requirement of better than 500 ppm at the design stage for a biplanar magnet system. When combined with the additional errors likely to be introduced during manufacturing (which may be of higher order and several times larger in magnitude), this level of inhomogeneity constitutes a practical estimate for the maximum limit that may be passively shimmed to a level suitable for medical imaging.

In this chapter, the finite element method (FEM) was applied to quantify the effects imparted by lateral size reduction of a full body 0.2 T biplanar permanent magnet assembly with a four-column yoke structure. The magnet geometry selected for this investigation was based on the 0.2 T biplanar permanent magnet utilized by our group in the construction of a prototype integrated linear accelerator (linac) and MRI system with a perpendicular configuration.[1, 29] Due to the unique requirement of this hybrid system to rotate within a typical radiotherapy vault, it was necessary that the overall lateral extent, and hence pole diameter, of the magnet assembly be constrained.

In the present work, an iterative optimization method based on the FEM was developed to calculate optimal pole piece surface contours that minimize the magnetic field variation within a specified region of interest. In an attempt to compensate for the non-axisymmetric magnetic field inhomogeneities that are characteristic of biplanar permanent magnets, a grid parameterization of

the pole piece surface was introduced so as to obtain optimized axisymmetric and non-axisymmetric pole piece designs for the laterally reduced compact magnet assembly (CMA). These designs were then compared against flat and single annular ring shimmed pole piece designs as found commonly in industry. In addition, the sensitivity of the magnetic field inhomogeneities to geometrical errors in the design surface for the nonaxisymmetric design was also explored, and statistical parameters quantifying this sensitivity were approximated.

5.2 METHODS

5.2.1 Magnet Assembly Specifications

In order to quantify the effects of reducing the size of a permanent magnet MRI system, the four-column biplanar magnet assembly shown in Figure 5.1 was considered. Using the rotating biplanar linac-MRI application[1, 29] as an example, the minimum pole separation required for rotation around a typical human patient was chosen as 70 cm. Due to the presence of pancake gradient coils mounted to the pole surfaces, this nominal pole separation would be further reduced by approximately $6-10$ cm. Table 5.1 provides the remaining dimensions labeled in Figure 5.1(b), which were approximately obtained through a uniformly scaled enlargement of the commercially available 27.5 cm gap permanent magnet system utilized by a prototype linac-MRI system described in the literature.[1,29] This structure is herein referred to as the uniformly enlarged magnet assembly (UEMA).

Due to a diagonal dimension l_1 of 2.862 m, the UEMA would provide a vertical clearance of 63.8 cm if rotated within a typical radiotherapy vault with 3.50 m height. In order to achieve a more practical clearance of 90 cm to allow room for additional peripheral equipment, the lateral dimension l_2 of the magnet assembly was reduced in order to maintain the minimum

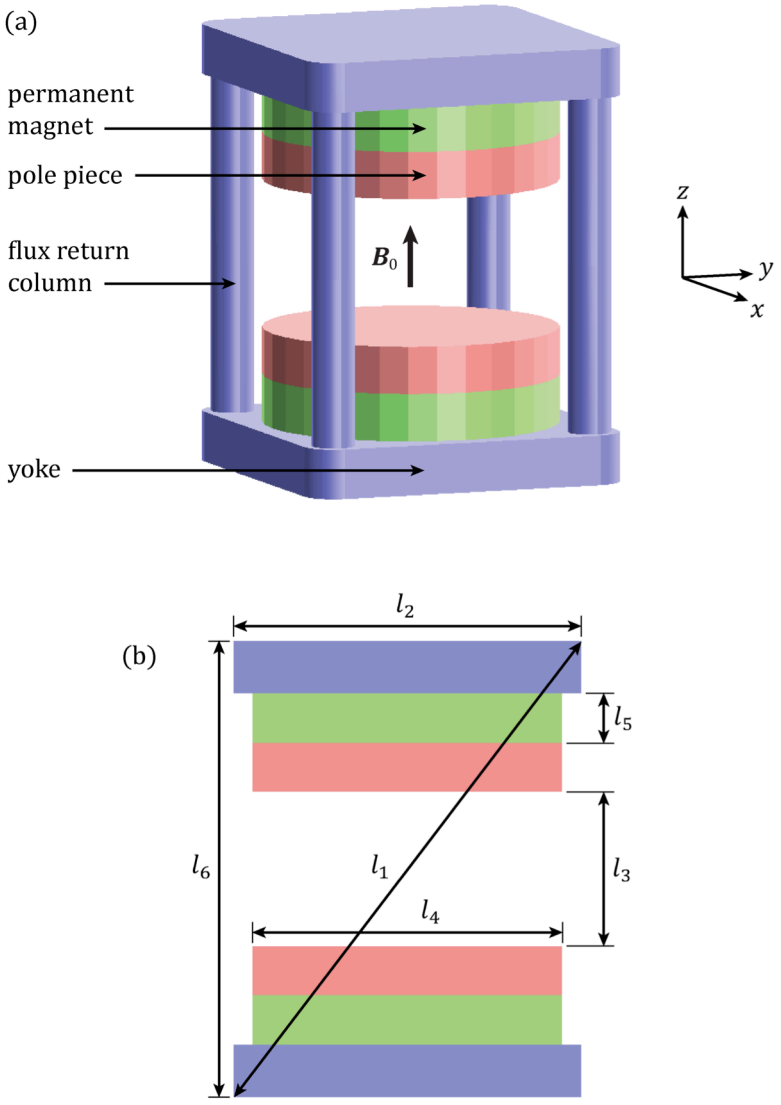


FIGURE 5.1: (a) A 3D schematic of the four-column biplanar permanent magnet assembly with the relative orientation of the Cartesian coordinate axes shown. (b) A cross-sectional view with dimension labels.

nominal pole separation l_3 of 70 cm. This effectively results in a reduction of the magnetic pole diameter l_4 from 175.5 to 139.8 cm. The remaining dimensions of this laterally reduced compact magnet assembly (CMA) are provided in Table 5.1.

TABLE 5.1: Dimensions of the permanent magnet assembly illustrated in Figure 5.1.

	Length (mm)	
Quantity	UEMA	CMA
l_1	2862	2600
l_2	1980	1578
l_3	700	700
l_4	1755	1398
l_5	235	235
l_6	2066	2066

The yoke and column structures of the magnet assemblies described are composed of AISI 1020 plain carbon steel, and the pole pieces are composed of a special Armco magnetic steel. The nonlinear magnetic properties for these materials are defined via the magnetization curves[22, 23] shown in Figure 5.2. The permanent magnet poles are constructed from a neodymium-iron-boron compound ($Nd_2Fe_{14}B$), with a homogenous remanent magnetic field of 1.005 T oriented in the z direction.[3]

For each magnet design, the quality of the magnetic field in a given volume of interest $\mathcal{V}$ was measured in terms of the parts per million (ppm) inhomogeneity ΔB defined as:

$$\Delta B = \frac{B_{\max} - B_{\min}}{B_0} \cdot 10^6, \tag{5.1}$$

where $B_{\max}$ and $B_{\min}$ are the maximum and minimum values for the magnetic field magnitude within $\mathcal{V}$, respectively, and B_0 is the magnetic field strength at isocenter. For completeness, the maximal extent (x_{5G}, y_{5G}, z_{5G}) of the 5 G magnetic field contours (relative to isocenter) was also calculated.

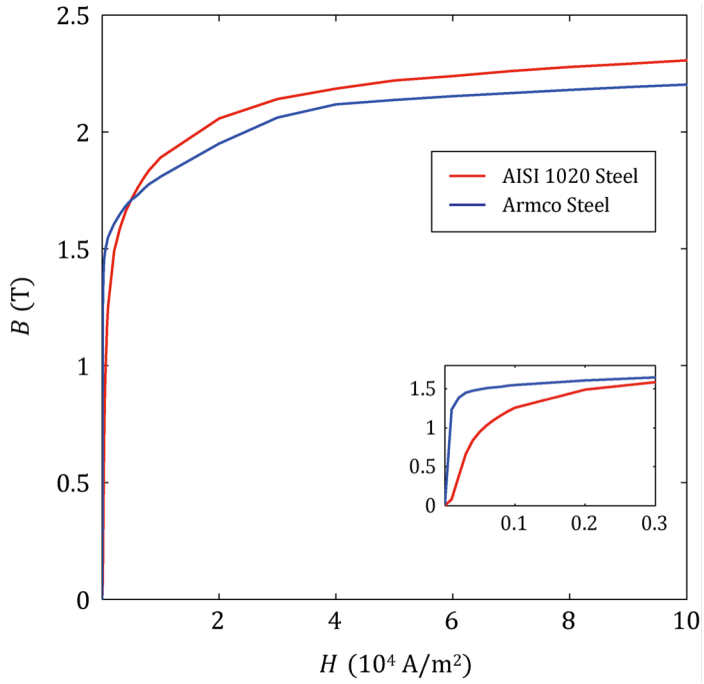


FIGURE 5.2: Nonlinear magnetization curves for AISI 1020 plain carbon steel and Armco magnetic steel.

5.2.2 The Finite Element Method

5.2.2.1 Problem Formulation

The FEM is a widely established technique for numerically solving discretized boundary value problems (see § 3.3). It is particularly well suited for problems involving complex geometries and its use in electromagnetic analysis has been extensively documented.[30] The required magnetic field calculations performed in this chapter were accomplished with the commercially available three-dimensional (3D) FEM software package COMSOL MULTIPHYSICS.[20] The details of the FEM implementation employed by COMSOL MULTIPHYSICS have been provided in § 3.3.5, for which the important

aspects are reviewed here. The FEM problem formulation begins with Maxwell's equations in a current free region:

$$\nabla \cdot \boldsymbol{B} = 0, \tag{5.2a}$$

$$\nabla \times \boldsymbol{H} = 0, \tag{5.2b}$$

where $\boldsymbol{B}$ is the magnetic field vector and $\boldsymbol{H}$ is the auxiliary magnetic field. The scalar magnetic potential Φ_m can then be defined as (see § 3.1.3):

$$\boldsymbol{H} = -\nabla \Phi_m. \tag{5.3}$$

In a region characterized by a magnetic permeability μ and remanent magnetization $\boldsymbol{M_r}$, the fields $\boldsymbol{B}$ and $\boldsymbol{H}$ are related through the constitutive relation:

$$\boldsymbol{B} = \mu \boldsymbol{H} + \mu_0 \boldsymbol{M_r}. \tag{5.4}$$

Combining Eqs. (5.2), (5.3), and (5.4) yields the second-order partial differential equation (PDE):

$$-\nabla \cdot (\mu \nabla \Phi_m - \mu_0 \boldsymbol{M_r}) = 0. \tag{5.5}$$

The magnetostatics application mode in the AC/DC module of COMSOL MULTIPHYSICS can be employed to numerically solve Eq. (5.5) for Φ_m at discrete points within the model geometry, subject to appropriate boundary conditions. Solution convergence is achieved when the specified relative tolerance of 10^{-6} exceeds the relative error estimate,[21] which is computed as a weighted L_2-norm of the estimated nodal error of Φ_m. The required fields $\boldsymbol{H}$ and $\boldsymbol{B}$ can then be calculated directly by Eqs. (5.3) and (5.4), respectively.

5.2.2.2 Model Description

The FEM model constructed in this work consists of the biplanar magnet assembly positioned at the center of a large cubic domain referred to as the global model volume (GMV). This domain was required in order to simulate the far field behaviour of the magnetic field within the vacated region surrounding the magnet. Although the real magnetic field is of infinite extent, the GMV must be truncated at a finite distance so as to limit the number of degrees of freedom in the FEM analysis. A suitable truncation distance of 5 m from isocenter was considered to have a negligible effect on the solution accuracy within a 40 cm DSV at isocenter. This was determined through a progressive enlargement of the GMV that was terminated when the associated perturbation of the field solution was dominated by the mesh discretization error addressed in a later discussion below.

To obtain the closed-loop behaviour characteristic of the magnetic field at far distances, the tangential field condition was applied at the outer boundaries of the GMV:

$$\hat{\boldsymbol{n}} \cdot \boldsymbol{B} = 0, \tag{5.6}$$

where $\hat{\boldsymbol{n}}$ is the unit normal vector at the particular boundary. This condition can also be applied on internal boundaries or planes across which both the geometry and $\boldsymbol{B}$ are symmetric upon reflection, effectively simplifying the model by only solving Eq. (5.5) explicitly on one side of the designated boundary. As such, this condition was also applied to the xz and yz planes of the biplanar magnet model. In a similar manner, the model may be further simplified by applying the zero potential condition

$$\Phi_m = 0, \tag{5.7}$$

on internal boundaries or planes across which the geometry and $\boldsymbol{B}$ are symmetric and antisymmetric upon reflection, respectively. This condition results in a magnetic field that is orthogonal to the specified boundary surface. As such, this condition was applied to the xy plane, resulting in a total reduction of the FEM model to just one eighth of the complete geometry.

The simplified FEM model was partitioned into approximately 1.4×10^5 isoparametric tetrahedral quadratic Lagrange elements (see § 3.3.2). This mesh was automatically generated with a Delaunay triangulation algorithm with a manual selection of parameters that controlled the maximum element size and maximum element growth rate within the various regions comprising the model geometry. An appropriate combination of these parameters was selected so as to obtain a suitably accurate mesh-independent solution that adequately resolved the field variation in regions with large gradients. This was determined through a progressive refinement of the mesh that was terminated when the field variation within a 40 cm DSV at isocenter was limited to below 1 μT (approximately 5 ppm) upon doubling the overall nodal density in the model. As a result, a maximum element dimension of 1.5 cm was specified within this region of interest.

An illustration of the simplified FEM model geometry and finite element discretization is provided in Figure 5.3. The depicted magnetic field corresponding to the solution of Eq. (5.5) was calculated in less than two minutes on a 2.7 GHz Intel Xeon quad-core PC workstation with 16 GB of RAM.

5.2.3 Pole Piece Design Parameterization

The uniformity of the magnetic field generated by a given magnet design is strongly influenced by the shape of the ferromagnetic pole pieces that support the permanent magnet sources. Hence, the contoured surfaces of the

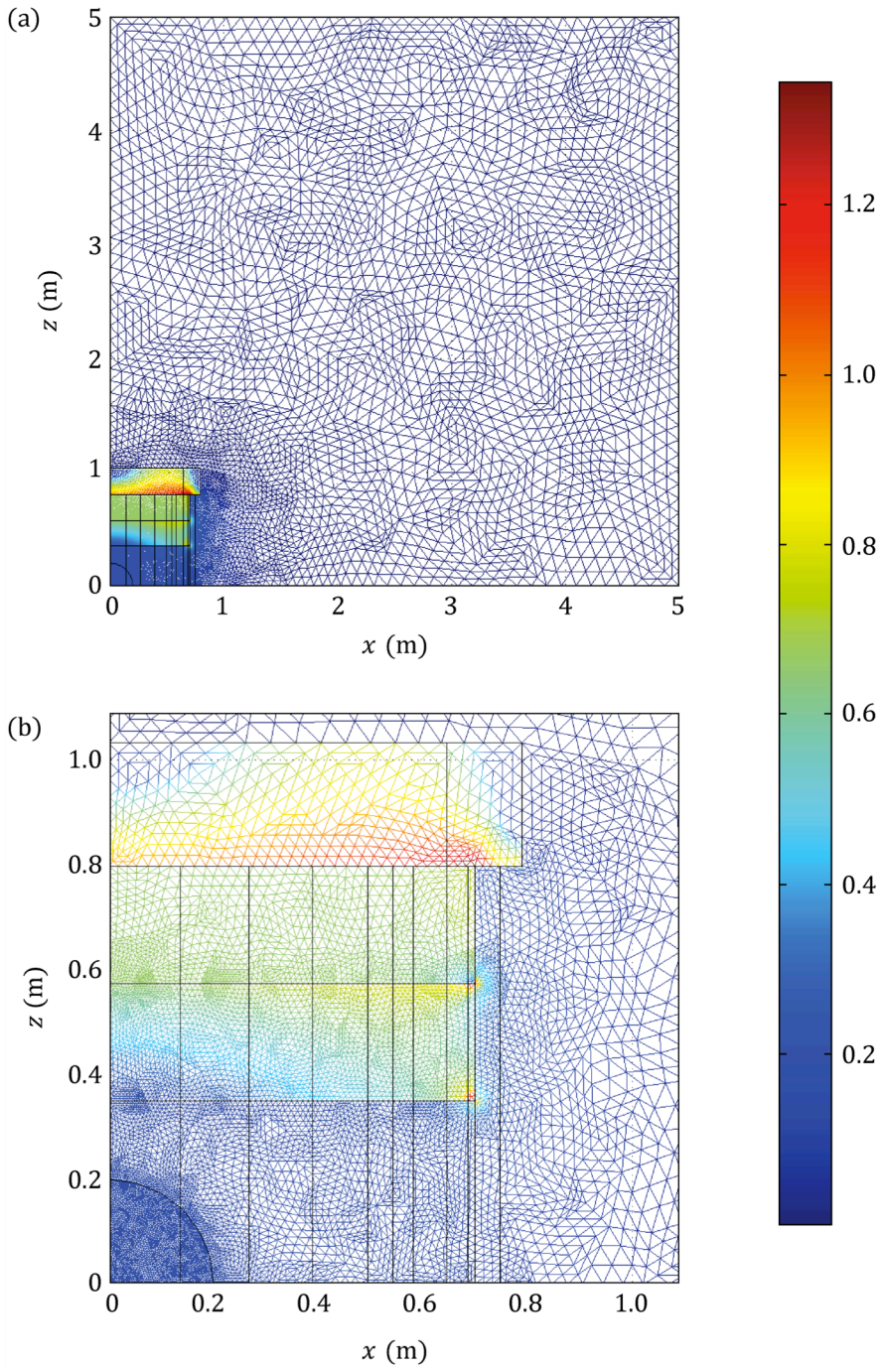


FIGURE 5.3: (a) A section of the simplified FEM model geometry corresponding to the CMA with a flat pole piece design. The magnetic field solution is overlaid with the FEM mesh. Field map values are displayed in units of T. (b) A close-up section illustrating a fine mesh within the magnet geometry and a 40 cm DSV at isocenter.

pole piece designs considered were optimized for maximum magnetic field uniformity in a 40 cm DSV. This was accomplished by mathematically optimizing a set of design variables that define the geometrical shape of the pole piece surface according to a specified parameterization. The details of this optimization method are provided in § 5.2.4. Pole pieces possessing a flat surface or a single annular ring shimmed (SARS) design were achieved with a simple parameterization of the pole piece cross-section. More complex axisymmetric and non-axisymmetric designs were generated from a direct two-dimensional (2D) grid parameterization of the pole piece surface. In all cases, the design variables were constrained such that the nominal pole separation l_3 remained at least 70 cm, with all other dimensions of the magnet structure held fixed.

5.2.3.1 Annular Shim Parameterization

The design variables for the SARS parameterizations were selected as the dimensions of the pole piece cross-section, as shown for an upper pole section in Figure 5.4. This parameterization results in three unique design variables, denoted p_i (for $i = 1, 2, 3$). The initial values specified prior to optimization were taken from uniform enlargement of the pole piece dimensions taken from the 27.5 cm gap commercial magnet system discussed in § 5.2.1. When considering a simple flat surface, the design variables were constrained such that the pole piece had a maximual uniform thickness of $p_1 = 218$ mm with $p_2 = p_3 = 0$.

5.2.3.2 Grid Parameterization

In addition to the flat and SARS designs, axisymmetric and non-axisymmetric grid parameterized pole pieces were obtained for the CMA. To this end, the open surface of the pole piece was parameterized with a distribution of 112 control points defined in the $r\theta$ plane, as shown in Figure 5.5. The thickness of the pole pieces at these locations was varied during the optimization

FIGURE 5.4: Design parameterization for optimization of the SARS pole piece designs. Only the upper pole piece is illustrated.

process, as the z coordinates of the control points were taken as the design variables. The pole surface was then generated based on a linear spline interpolation between the control points.

Due to symmetry in the magnet geometry, additional constraints were imposed on the control points in order to reduce the number of degrees of freedom in the optimization procedure. The z coordinates for the control points in quadrant II were obtained by reflecting the coordinates of the control points in quadrant I over the line $(x = 0)$. Similarly, the z coordinates for the points in quadrants III and IV were obtained by reflecting the coordinates of the points in quadrants I and II over the line $(y = 0)$. Additionally, the control points on the line $(r = 0)$ were constrained to have the same z coordinate, as they share the same location on the design surface.

When axisymmetric grid parameterized (AGP) designs were sought, the control points on the lines $(\theta = \pi/8, \pi/4, 3\pi/8, \text{and } \pi/2)$ were constrained to have the same z coordinates as the points on the line $(\theta = 0)$, resulting in a total of 7 unique design variables. When non-axisymmetric grid

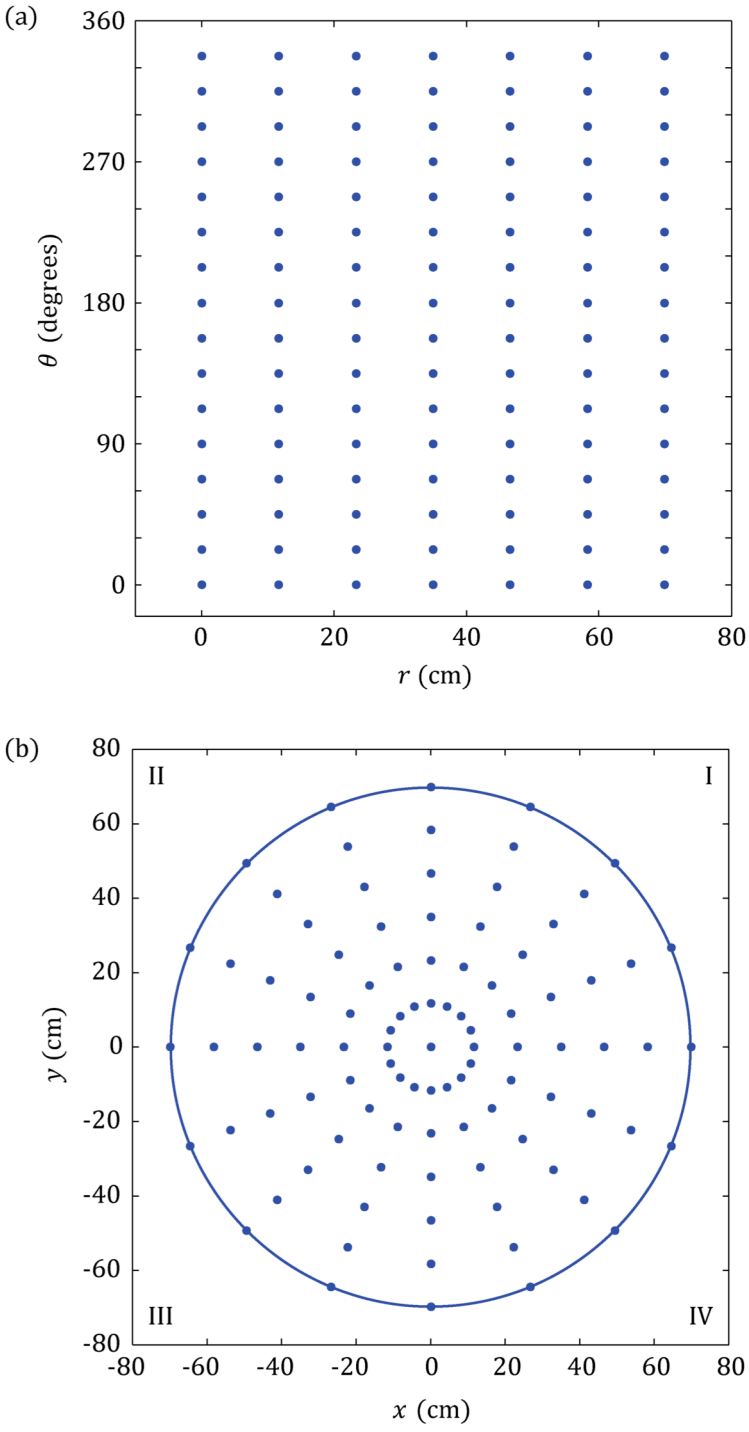


FIGURE 5.5: Control point distribution for the grid parameterization, plotted in (a) polar coordinates and (b) Cartesian coordinates.

parameterized designs (NAGP) were considered, the control points on the lines ($\theta = 3\pi/8$ and $\pi/2$) were constrained to have the same z coordinates as those on the lines ($\theta = \pi/8$ and 0), respectively, resulting in a total of 19 unique design variables.

Optimization of the AGP designs was conducted from an initial flat pole surface configuration, defined by $p_i = 350$ mm (for $i = 1, \ldots, 7$). In order to partially alleviate the burden of the increased number of degrees of freedom, the optimization of the NAGP designs was conducted as a two step procedure. As such, the NAGP optimization was initiated from a manually selected design state from within the evolution of the aforementioned AGP optimization.

5.2.4 Nonlinear Optimization

In order to determine the optimal parameters for the pole piece designs considered, the iterative shape optimization method summarized in Figure 5.6 was developed. This algorithm was implemented in the MATLAB[24] scripting environment and utilized macros from the COMSOL MULTIPHYSICS scripting language. This configuration permitted automatic iterative modification and analysis of the required FEM models, as COMSOL MULTIPHYSICS was directly integrated with the MATLAB application through a built-in interface.

Mathematically, the optimization problem can be expressed as:

$$\text{minimize } f(\boldsymbol{p}), \tag{5.8a}$$

$$\text{subject to } \boldsymbol{p} \in \mathfrak{P}, \tag{4.1b}$$

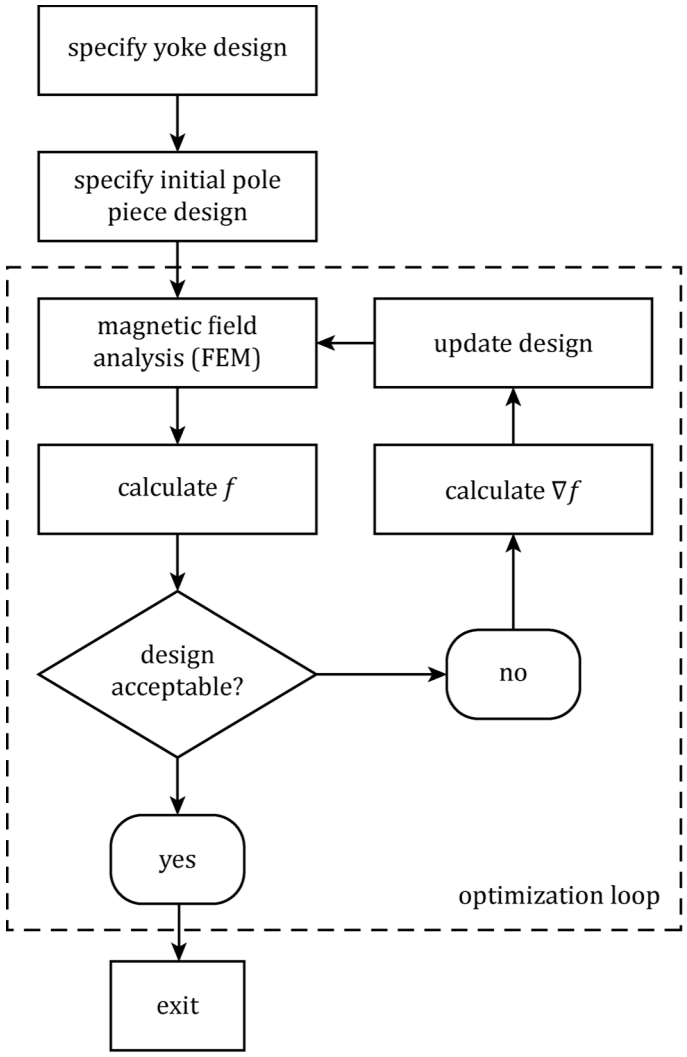


FIGURE 5.6: Iterative optimization flow diagram.

where f is the objective function and $\boldsymbol{p}$ is the design vector. The feasible domain $\mathfrak{P}$ is defined according to:

$$\mathfrak{P} = \{\boldsymbol{p} | l_i \leq p_i \leq u_i, i = 1, \ldots, n\}, \tag{5.9}$$

such that the elements of $\boldsymbol{p}$ are subject to the lower and upper limitations l_i and u_i, respectively. For n unique design variables, the elements of $\boldsymbol{p}$ are denoted p_i (for $i = 1, \ldots, n$).

The objective function is defined by the integral formula:

$$f(\boldsymbol{p}) = \int_{\mathcal{V}} [B(\boldsymbol{r}; \boldsymbol{p}) - B_0]^2 d\mathcal{V}', \tag{5.10}$$

where $B(\boldsymbol{r}; \boldsymbol{p})$ is the magnitude of the magnetic field at the point $\boldsymbol{r}$, implicitly due to the magnet design defined by $\boldsymbol{p}$. As such, the objective function provides a cumulative measure of the magnetic field inhomogeneity within the target region of interest $\mathcal{V}$. In the current implementation, $\mathcal{V}$ was taken as a 40 cm DSV at isocenter.

The gradient of the objective function ∇f with respect to the design variables is given by

$$\nabla f = \left[\frac{\partial f}{\partial p_1} \quad \cdots \quad \frac{\partial f}{\partial p_n} \right], \tag{5.11}$$

for which each of the partial derivatives was approximated using a forward finite-difference approximation:

$$\frac{\partial f}{\partial p_i} \approx \frac{f(p_i + \delta p) - f(p_i)}{\delta p}, \tag{5.12}$$

where the perturbation interval δp was chosen as 0.1 mm according to trial and error. Each of the objective function terms in the numerator of Eq. (5.12) corresponds to a particular design for which an independent FEM

analysis is performed and the integral in Eq. (5.10) over the entire volume $\mathcal{V}$ calculated.

During the optimization process, the design vector was updated at the kth iteration according to the projected gradient formula (see § 4.3.3):

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{\Pi}\left(\boldsymbol{p}^{(k)} + \alpha^{(k)}\boldsymbol{d}^{(k)}\right), \tag{5.13}$$

where $\alpha^{(k)}$ is the scalar step size, $\boldsymbol{d}^{(k)}$ is the search direction, and $\boldsymbol{\Pi}(\boldsymbol{p})$ denotes the projection of $\boldsymbol{p}$ onto the feasible domain $\mathfrak{P}$. Two versions of this optimization scheme were developed, as characterized by the calculation method specified for $\boldsymbol{d}^{(k)}$. In the first version, the steepest descent (SD) method[3] was selected, with $\boldsymbol{d}^{(k)}$ calculated according to (see § 4.2.2):

$$\boldsymbol{d}^{(k)} = -\nabla f^{(k)}. \tag{5.14}$$

This was a suitable starting point, as the SD method is a simple and robust first-order technique that only requires the calculation of ∇f at each iteration. In an attempt to improve the performance of the optimization procedure and obtain faster convergence, the second version employed the conjugate gradient (CG) method, with $\boldsymbol{d}^{(k)}$ given by (see § 4.2.3)

$$\boldsymbol{d}^{(k)} = \begin{cases} -\nabla f^{(1)}, & \text{for } k = 1, \\ -\nabla f^{(k)} + \beta^{(k)}\boldsymbol{d}^{(k-1)}, & \text{for } k > 1. \end{cases} \tag{5.15}$$

The parameter $\beta^{(k)}$ appearing in this expression was calculated according to the Polak-Ribière formula, [6] given as

$$\beta^{(k)} = \frac{\nabla f^{(k)^T}\left(\nabla f^{(k)} - \nabla f^{(k-1)}\right)}{\nabla f^{(k-1)^T}\nabla f^{(k-1)}}. \tag{5.16}$$

In an effort to refresh the CG method and discard information from distant previous iterations, the CG search direction was periodically reset to the SD search direction (i.e. $\beta^{(k)} = 0$). A reset frequency of seven iterations was found to be effective through trial and error.

In general, a solution to the optimization problem in Eq. (5.8a) is reached when convergence to a local minimum of f has been achieved. Since this point is characterized by a vanishing objective function gradient (see §), a sensible choice for the termination criteria would include the condition

$$\left|\nabla f^{(k)}\right| < \kappa, \tag{5.17}$$

where $\kappa > 0$ is a convergence parameter. The appropriate specification of this parameter, along with the selection of any additional conditions to be imposed, requires careful thought in addition to thorough experience with the algorithm performance. In order to avoid untimely termination, specification of automatic termination criteria was avoided and the optimization scheme was halted manually.

5.2.5 Design Sensitivity

The relatively complex surface topology of the NAGP pole piece would likely be subject to random errors during the manufacturing process. In order to simulate these errors and quantify the effects of geometry perturbations, a sensitivity analysis was performed on the optimal NAGP pole piece design obtained with the SD method described in the previous section. To this end, uniformly distributed random variations, denoted σ_i (for $i = 1, \ldots, n$), were

applied to each of the unique design variables of the non-axisymmetric parameterization, thus generating a random perturbation of the optimal NAGP design. These random variations are mathematically described as:

$$\sigma_i \sim \mathfrak{U}(-\sigma_{\max}, \sigma_{\max}), \tag{5.18}$$

where $\mathfrak{U}$ denotes the standard uniform probability distribution with the specified maximum variation amplitude $\sigma_{\max}$. The design vector for a perturbed optimized design $\boldsymbol{p}'$ was therefore calculated as:

$$\boldsymbol{p}' = \boldsymbol{p} + \boldsymbol{\sigma}. \tag{5.19}$$

where $\boldsymbol{\sigma}$ is the perturbation vector. Populations of 100 such perturbed designs were obtained for various maximum amplitudes in the range of 0.1 to 5 mm. For each perturbed design, the resulting field variation ΔB was calculated over 10, 20, 30, and 40 cm DSVs.

In order to evaluate the sensitivity of the optimized magnetic field inhomogeneity to the random variations imposed, the three statistical parameters P_{50}, P_{80}, and P_{90} were calculated for each population. These parameters are defined as the ΔB values that upper bound 50, 80, and 90% of the perturbed designs in a given population, respectively.

5.3 RESULTS AND DISCUSSION

5.3.1 Magnet Assembly Size Reduction

Table 5.2 compares the magnetic field at isocenter B_0 and calculated inhomogeneities for the UEMA and CMA. Flat and SARS pole piece designs were considered, where the width and depth of the peripheral ring of the

TABLE 5.2: Summary of optimization results and comparison of magnetic field properties for the magnet designs with flat and annular shimmed pole pieces.

	UEMA		CMA	
Quantity	Flat Pole Piece	SARS Pole Piece	Flat Pole Piece	SARS Pole Piece
Iterations	-	9	-	11
B_0 (T)	0.233	0.204	0.203	0.172
f (10^{-11} T^2·m^3)	81.54	0.018	1455	1.899
ΔB, 30 cm DSV (ppm)	1797	29	8774	321
ΔB, 40 cm DSV (ppm)	3134	100	15299	1030
x_{5G} (m)	2.7	2.7	2.4	2.4
y_{5G} (m)	2.7	2.7	2.4	2.4
z_{5G} (m)	3.0	3.0	2.8	2.8

SARS designs were optimized to minimize the cumulative field inhomogeneity f as measured over a 40 cm DSV.

The magnetic field inhomogeneity ΔB increased significantly upon reduction of the magnet assembly dimensions. In particular, ΔB increased by an order of magnitude for the SARS pole piece design, resulting in 321 and 1030 ppm over 30 and 40 cm DSVs, respectively. This effect can be visualized in the corresponding magnetic field maps in the xy and xz planes at isocenter provided in Figure 5.7. Assuming that successful post-manufacturing shimming of a magnet assembly requires better than 500 ppm in a given volume of interest, these results indicate that the CMA would be unsuitable for medical imaging within the desired 40 cm DSV.

As noted in § 5.2.1, lateral reduction of the biplanar magnet necessarily results in a smaller pole diameter and thus a decrease of permanent magnet

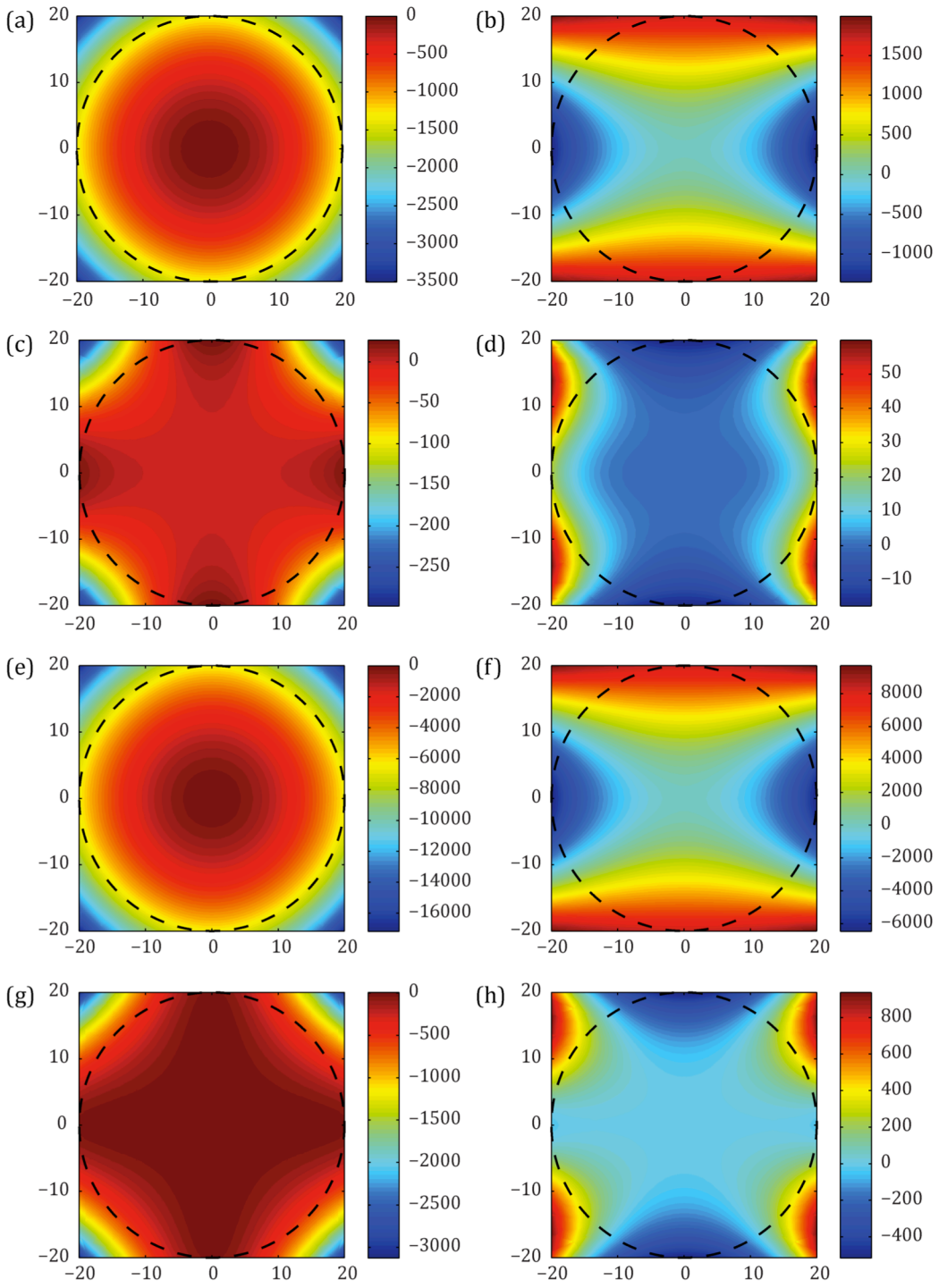


FIGURE 5.7: Magnetic field maps in the xy (first column) and xz (second column) planes at isocenter corresponding to the: (a), (b) flat and (c), (d) SARS designs for the UEMA; (e), (f) flat and (g), (h) SARS designs for the CMA. Axis dimensions are in cm. Map values are displayed in parts per million, measured relative to B_0 for each design. The dashed circles represent the boundary of a 40 cm DSV.

source material. Consequently, B_0 was reduced by at least 0.03 T for each pole piece design. The strength of the corresponding fringe magnetic fields also decreased, as the extent of the 5 G field contours was reduced from $(x_{5G}, y_{5G}, z_{5G}) = (2.7, 2.7, 3.0)$ to $(2.4, 2.4, 2.8)$ m.

5.3.2 Nonlinear Optimization

5.3.2.1 Design Comparison

Optimal axisymmetric and non-axisymmetric grid-parameterized pole piece designs were obtained for the CMA using the SD and CG versions of the optimization scheme described in § 5.2.4. These designs are visually compared with the flat and SARS pole piece designs in Figure 5.8. The calculated field strengths, objective function values, and residual field inhomogeneities for these designs are compared in Table 5.3.

As seen in Tables 5.2 and 5.3, the AGP pole piece designs resulted in improved field uniformity as compared to the SARS design. The residual field variation ΔB over a 30 cm DSV was improved to 191 and 187 ppm with the SD and CG variants, respectively. The value for the objective function f was also improved by a factor of approximately 2.7, despite nearly equal ΔB values over a 40 cm DSV. Hence, the advantages of the AGP designs diminish with increasing distance from isocenter, as both the SARS and AGP designs were inherently unable to compensate for the non-axisymmetric inhomogeneity arising from the four-column yoke geometry that dominates in larger regions of interest.

In contrast, the optimized NAGP pole piece designs clearly exhibited a surface topology with a non-axisymmetric shape variation. As such, these designs provided the greatest levels of field uniformity, now permitting post-manufacturing shimming and medical imaging over both 30 and 40 cm DSVs

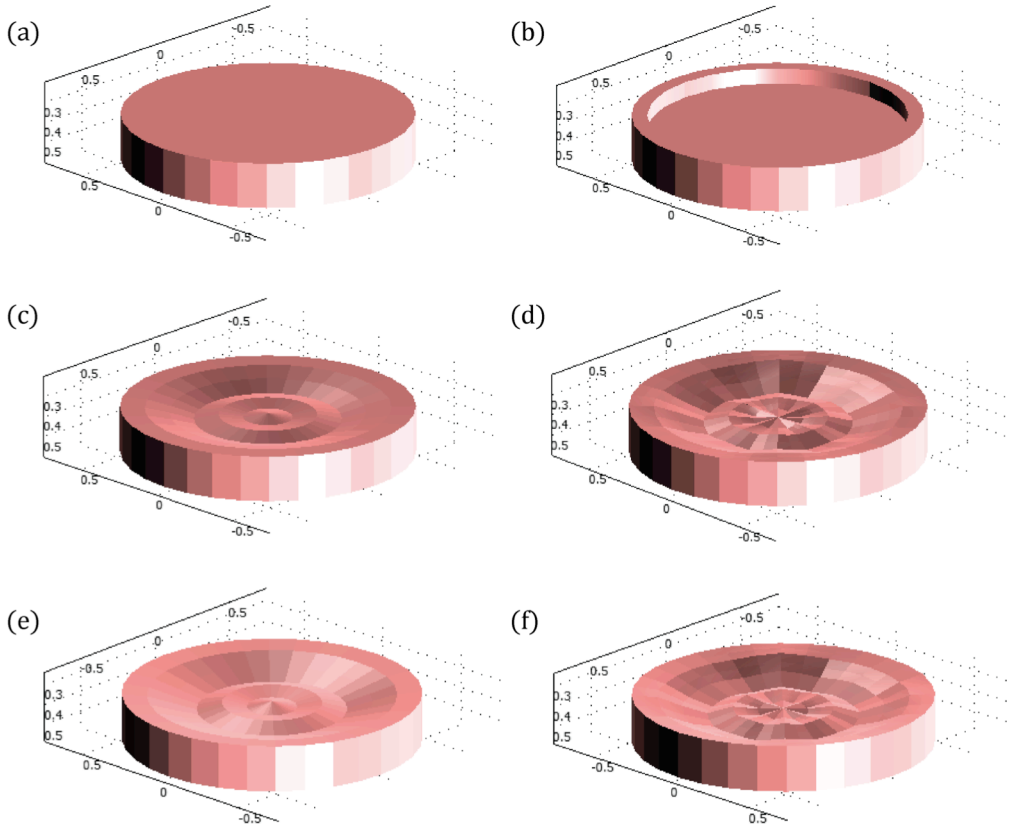


FIGURE 5.8: 3D visualizations of the (a) flat and (b) SARS pole pieces for the CMA, compared with the (c) AGP and (d) NAGP optimized designs obtained with the SD method, as well as the (e) AGP and (f) NAGP optimized designs obtained with the CG method. Axis dimensions are displayed in meters.

for the CMA. Specifically, ΔB values over a 40 cm DSV were improved to 279 and 237 ppm for the SD and CG variants, respectively, as compared to values exceeding 1000 ppm for the SARS and AGP designs. Furthermore, an additional order of magnitude reduction in the objective function was achieved, indicating a significant improvement throughout the 40 cm DSV. These results are visualized in the corresponding magnetic field maps in the xy and xz planes at isocenter provided in Figure 5.9.

TABLE 5.3: Summary of optimization results and comparison of magnetic field properties for the magnet designs with grid parameterized pole pieces.

	Steepest Descent Method		Conjugate Gradient Method	
Quantity	AGP Pole Piece	NAGP Pole Piece	AGP Pole Piece	NAGP Pole Piece
Iterations	720	683	256	358
B_0 (T)	0.177	0.178	0.179	0.175
f (10^{-11} T^2·m^3)	0.710	0.067	0.703	0.049
ΔB, 30 cm DSV (ppm)	191	51	187	37
ΔB, 40 cm DSV (ppm)	1001	279	1016	237
x_{5G} (m)	2.4	2.4	2.4	2.4
y_{5G} (m)	2.4	2.4	2.4	2.4
z_{5G} (m)	2.8	2.8	2.8	2.8

As discussed in § 5.2.3.2, the AGP and NAGP designs were based on the same geometrical parameterization used to generate the pole piece surfaces. Accordingly, the radial densities of the control points were equal, but the number of unique design variables was not. By increasing the number of design variables employed in the axisymmetric approach, the advantages of the NAGP design may be reduced. That said, the results indicate that the non-axisymmetric nature of the yoke geometry has a limiting effect on the uniformity of the residual magnetic field. With regards to the shape of the pole piece, this can only be abated through use of a non-axisymmetric design.

The field strengths resulting from the SARS, AGP, and NAGP designs were relatively similar, with values in the range of $0.172-0.179$ T. As mentioned in the previous section, the field strength is limited by the reduced dimensions of the permanent magnet source. Furthermore, the pole plate design was not seen to have a significant effect on the passive shielding

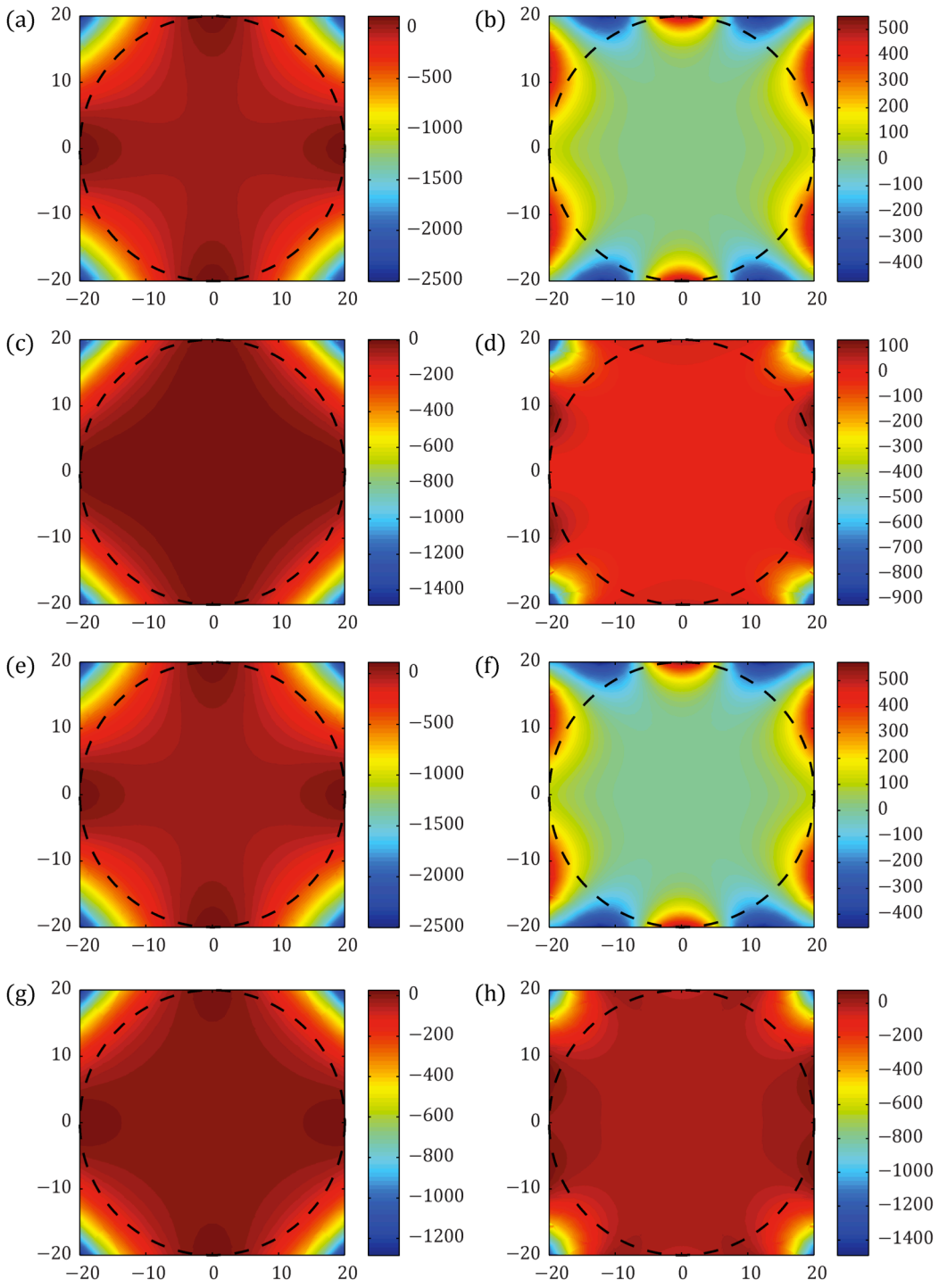


FIGURE 5.9: Magnetic field maps in the xy (first column) and xz (second column) planes at isocenter corresponding to the: (a), (b) AGP and (c), (d) NAGP designs obtained with the SD method; (e), (f) AGP and (g), (h) NAGP designs obtained with the CG method. Axis dimensions are in cm. Map values are displayed in parts per million, measured relative to B_0 for each design. The dashed circles represent the boundary of a 40 cm DSV.

performance of the magnet assembly. The fringe field extent (x_{5G}, y_{5G}, z_{5G}) resulting from the AGP and NAGP designs remained unchanged at (2.4, 2.4, 2.8) m. The corresponding 1, 5, and 10 G magnetic field contours are shown in Figure 5.10.

5.3.2.2 Algorithm Performance

The evolution of the objective function f and gradient norm $|\nabla f|$ is plotted for the SD and CG optimization schemes in Figures 5.11 and 5.12, respectively. As evident in these figures, optimization of the NAGP pole piece design was conducted as a two-step procedure, with the first step coinciding with a partial segment of the AGP optimization run. The CG version of the optimization scheme performed significantly better than the SD version, requiring considerably fewer iterations to arrive at AGP and NAGP pole piece designs exhibiting comparable levels of field uniformity. Specifically, the optimized AGP and NAGP designs were obtained with the CG method in 256 and 358 iterations, as compared to 720 and 683 iterations with the SD method, respectively.

An interesting feature of Figures 5.11(a) and 5.12(a) is that the evolution of f is comprised of several long plateaus with relatively slow improvement, separated by short segments exhibiting rapid reduction. This suggests that the optimization algorithms spend a considerable amount of time trapped near local minima, during which many computationally expensive iterations are performed with little or no improvement. This undesirable behaviour is likely attributed to the first-order nature of the SD and CG methods employed, for which the curvature information of f is not directly considered.

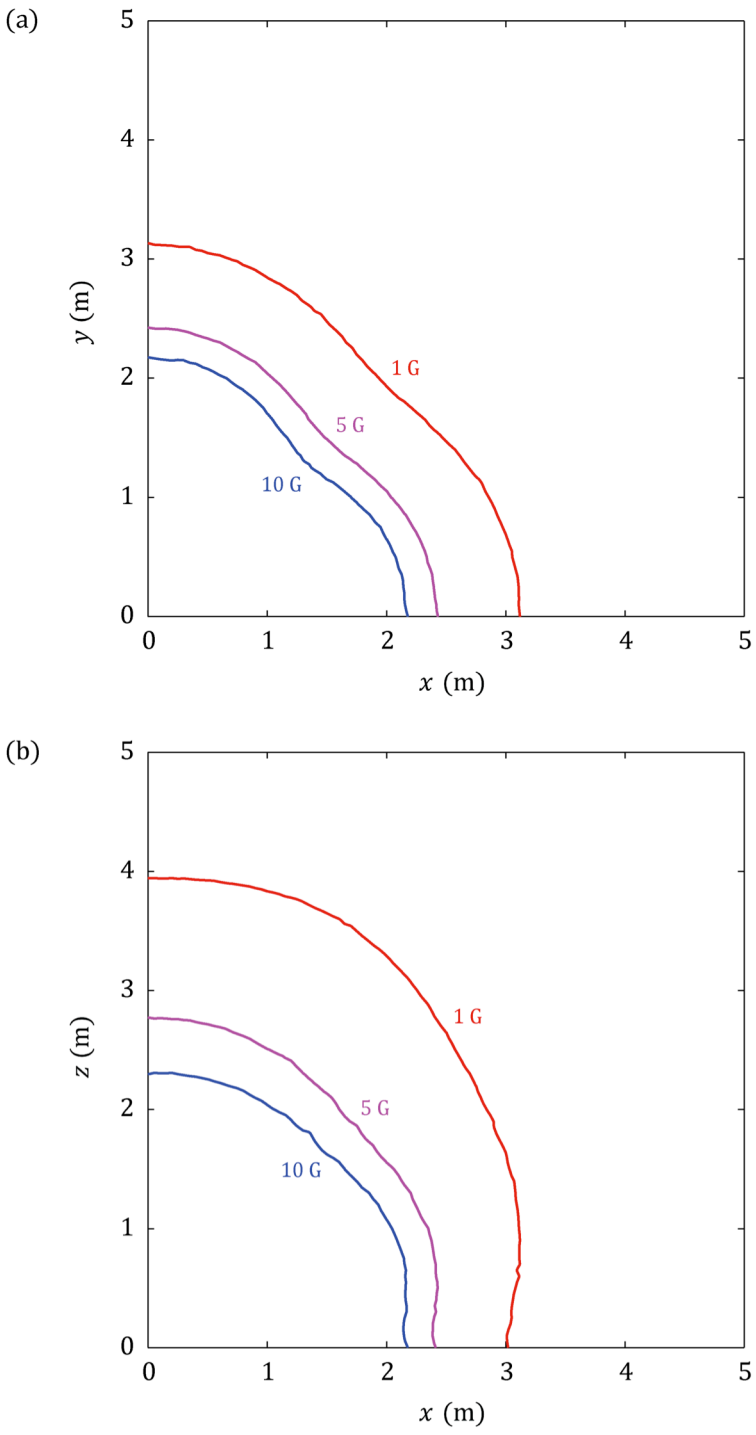


FIGURE 5.10: Plots of the 1, 5, and 10 G magnetic field contours projected onto the (a) xy and (b) xz planes for the CMA with the NAGP pole piece design obtained via the SD method.

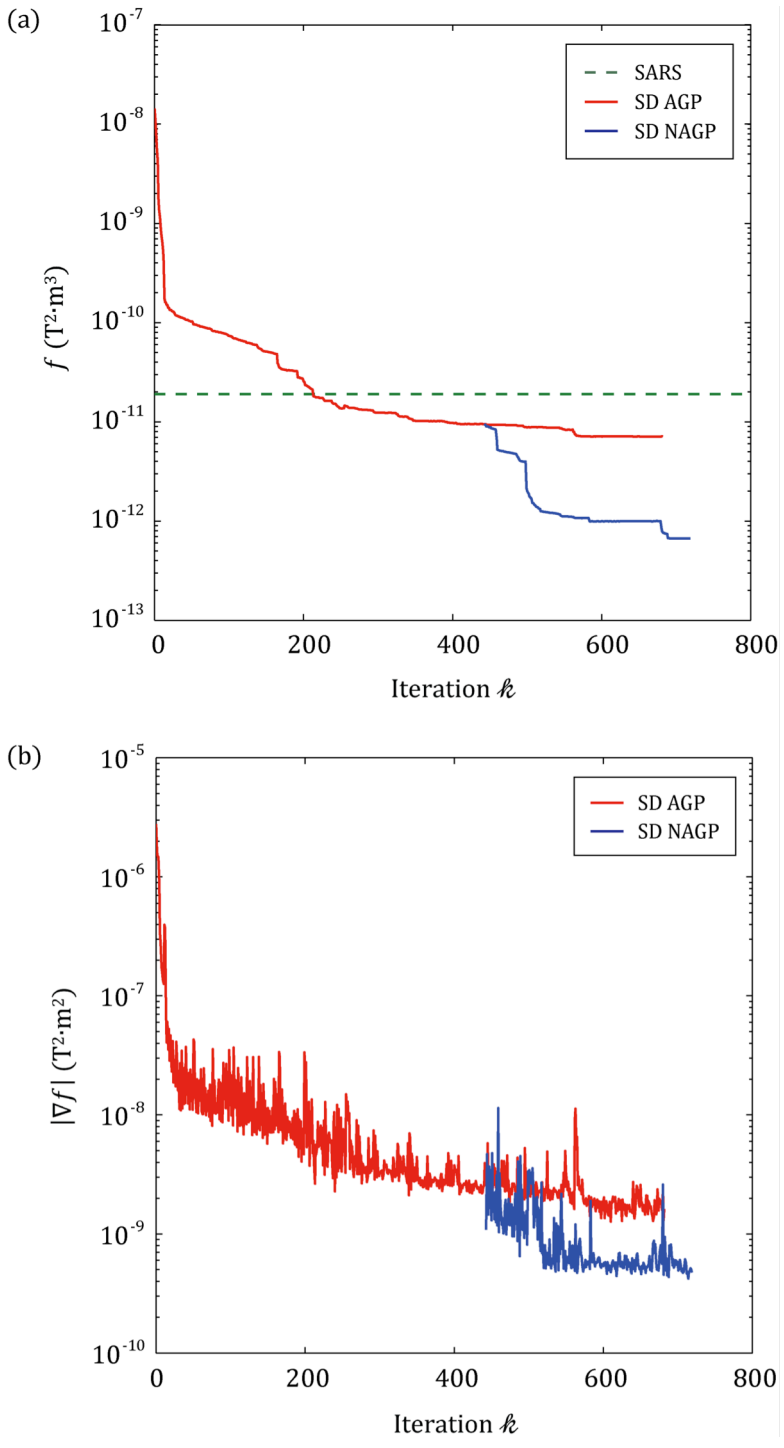


FIGURE 5.11: Evolution of the (a) objective function f and (b) gradient norm $|\nabla f|$ during optimization of the AGP and NAGP pole piece designs via the SD method.

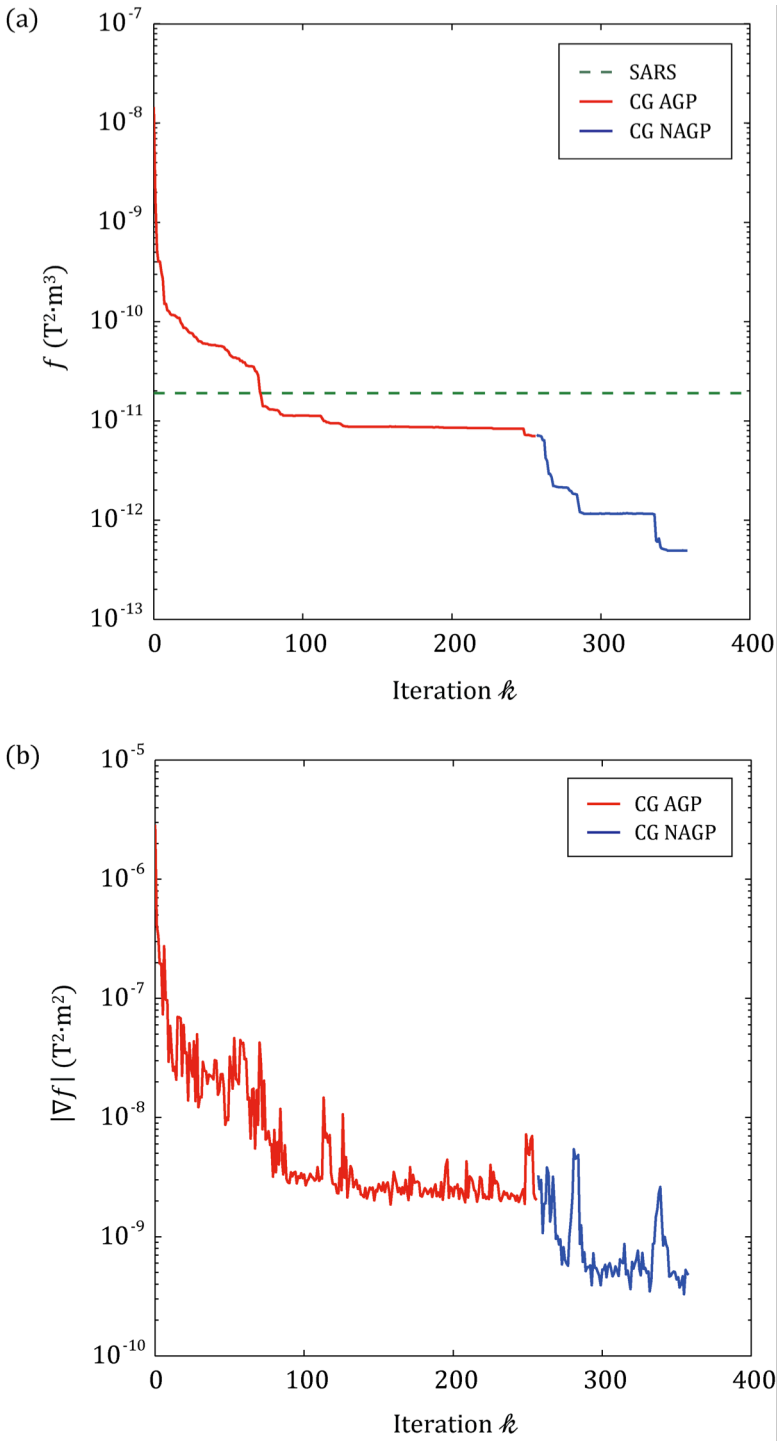


FIGURE 5.12: Evolution of the (a) objective function f and (b) gradient norm $|\nabla f|$ during optimization of the AGP and NAGP pole piece designs via the CG method.

The SD and CG methods are nonlinear optimization techniques that are designed to arrive at local solutions to a given nonlinear optimization problem. As such, the solutions obtained above are not guaranteed to be globally optimal. This is immediately evident from the results in Table 5.3, from which it can be seen that the NAGP designs obtained with the SD and CG methods correspond to different local minima of the objective function f. Confidence in the solutions above may further be established only through repeated execution of the optimization algorithm with alternative initial values for the set of design variables.

As noted in § 5.2.4, a required property of an optimal solution is the nulling of the objective function gradient ∇f. It is therefore an attractive feature of the optimization scheme developed here that the minimization of f is accompanied by a corresponding reduction of the gradient norm $|\nabla f|$ by approximately four orders of magnitude, as shown in Figures 5.11(b) and 5.12(b). Furthermore, the segments mentioned above during which f experiences rapid reduction are coincident with localized spikes in the gradient norm $|\nabla f|$, which is the expected behaviour for the SD and CG algorithms.

The quality of the search directions, and hence the performance of the optimization algorithm, is closely related to the accuracy of the objective function gradients used in Eqs. (5.14) and (5.15). Since the curve describing $|\nabla f|$ in Figures 5.11(b) and 5.12(b) is notably nonsmooth, it is important to consider the accuracy of the forward-difference approximation used to calculate ∇f. If δf denotes a bound on the absolute error of f, then the noise-level perturbation interval $\delta \check{p}$ can be approximated with[28]

$$\delta \check{p} \approx \frac{\delta f}{|\nabla f|}. \tag{5.20}$$

This interval represents a lower bound for the useful choice of the perturbation δp appearing in Eq. (5.12). Any perturbation $\delta p < \delta\check{p}$ results in $f(p_i + \delta p) - f(p_i) < \delta f$. In other words, the resulting change in f will be less than the noise-level (or error-level) defined by δf, thereby yielding a meaningless approximation of $\partial f / \partial p_i$.

It is immediately evident from Eq. (5.20) that in any practical implementation, the reliability of the forward-difference approximation decreases with the diminishing gradient $|\nabla f|$. As δf is ultimately limited by the accuracy of the FEM simulations, $\delta\check{p}$ becomes increasingly large. Therefore, the optimization algorithm will inherently suffer a loss of performance as a stationary point is approached. As such, this may be a contributing factor leading to the long plateaus of limited improvement observed in the evolution of f discussed above.

In order to estimate $\delta\check{p}$ at the point of termination of the NAGP optimization with the SD method, the absolute error δf was approximated. This was accomplished by progressively increasing the GMV truncation distance for the FEM model and calculating the resulting peak variation observed in f. This variation was random in nature and was attributed to the discretization error introduced by the unique mesh generated for each truncated model. Using this procedure, an estimate of $\delta f \approx 2 \times 10^{-15}$ T^2·m^3 was obtained. To complete the calculation of $\delta\check{p}$ with Eq. (5.20), a worst-case estimate of $|\nabla f| \approx 5 \times 10^{-10}$ T^2·m^2 was taken from the data graphed in Figure 5.11(a). Using these values, a noise-level perturbation of $\delta\check{p} \approx 0.004$ mm was estimated. As noted in § 5.2.4, a perturbation interval $\delta p = 0.1$ mm was specified for the finite-difference calculations performed in this work. Since this value is nearly two orders of magnitude larger than the estimate for $\delta\check{p}$, the aforementioned choice of δp was considered sufficiently large.

5.3.3 Design Sensitivity

The sensitivity analysis described in § 5.2.5 was performed on the optimal NAGP pole piece design obtained with the SD method. An example of the typical data obtained for a given population of randomly perturbed pole piece designs is provided in Figure 5.13, for which a maximum amplitude of $\sigma_{\max} = 2$ mm was specified. For each population obtained, the statistical parameters P_{50}, P_{80}, and P_{90} were calculated for 10, 20, 30, and 40 cm DSVs at isocenter. These results are plotted as a function of $\sigma_{\max}$ in Figures 5.14 and 5.15.

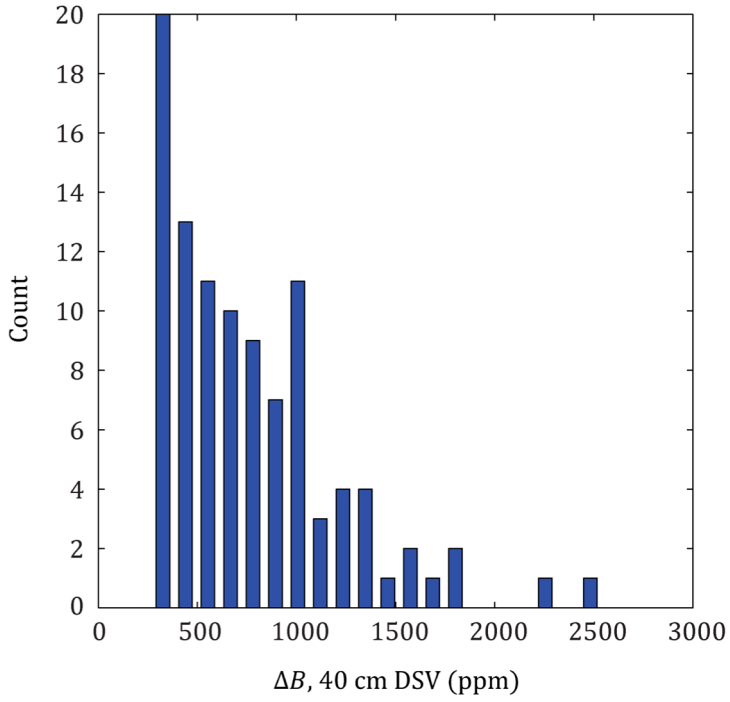


FIGURE 5.13: Histogram of magnetic field inhomogeneity values over a 40cm DSV for a population of 100 perturbed NAGP optimized designs with a maximum variation amplitude of $\sigma_{\max} = 2$ mm.

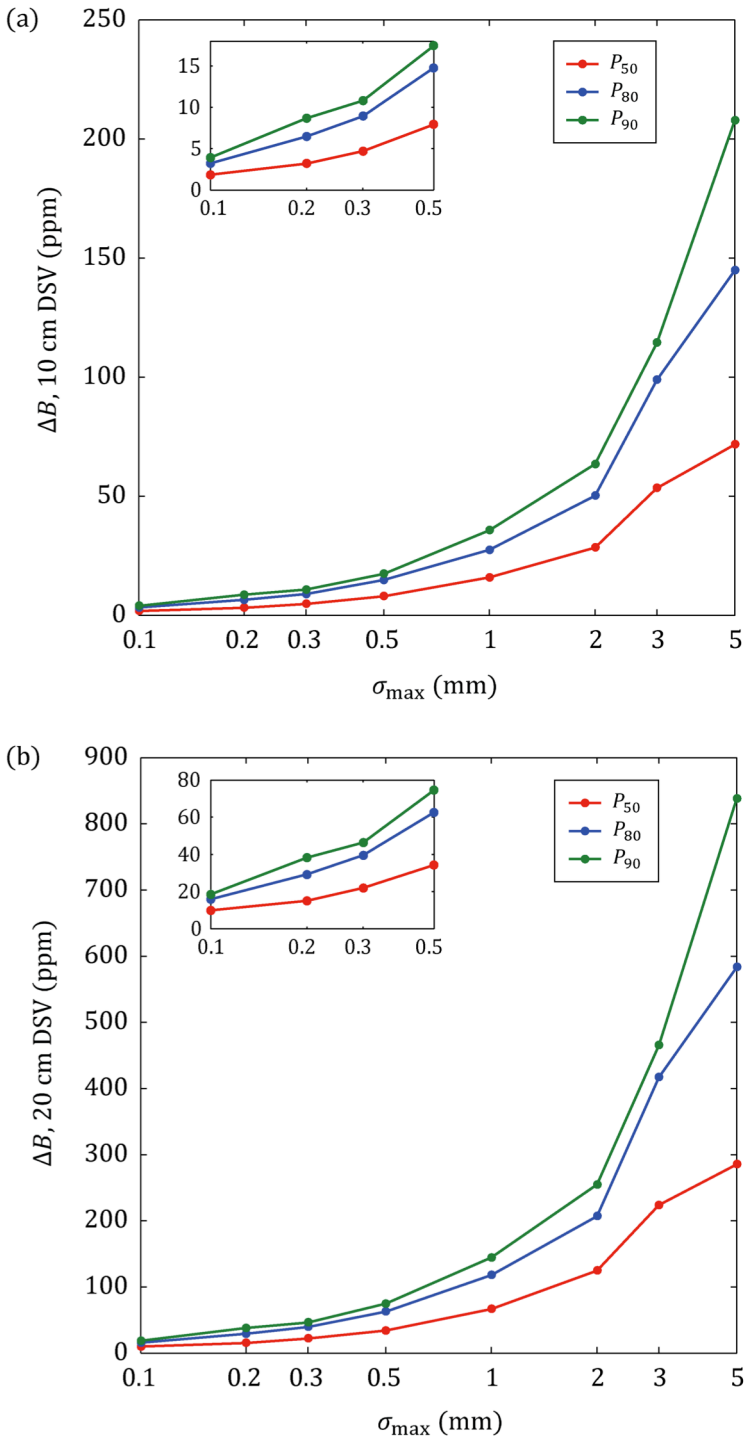


FIGURE 5.14: Plots of the statistical parameters P_{50}, P_{80}, and P_{90} as a function of the maximum variation amplitude σ_{max} for (a) 10 cm and (b) 20 cm DSVs.

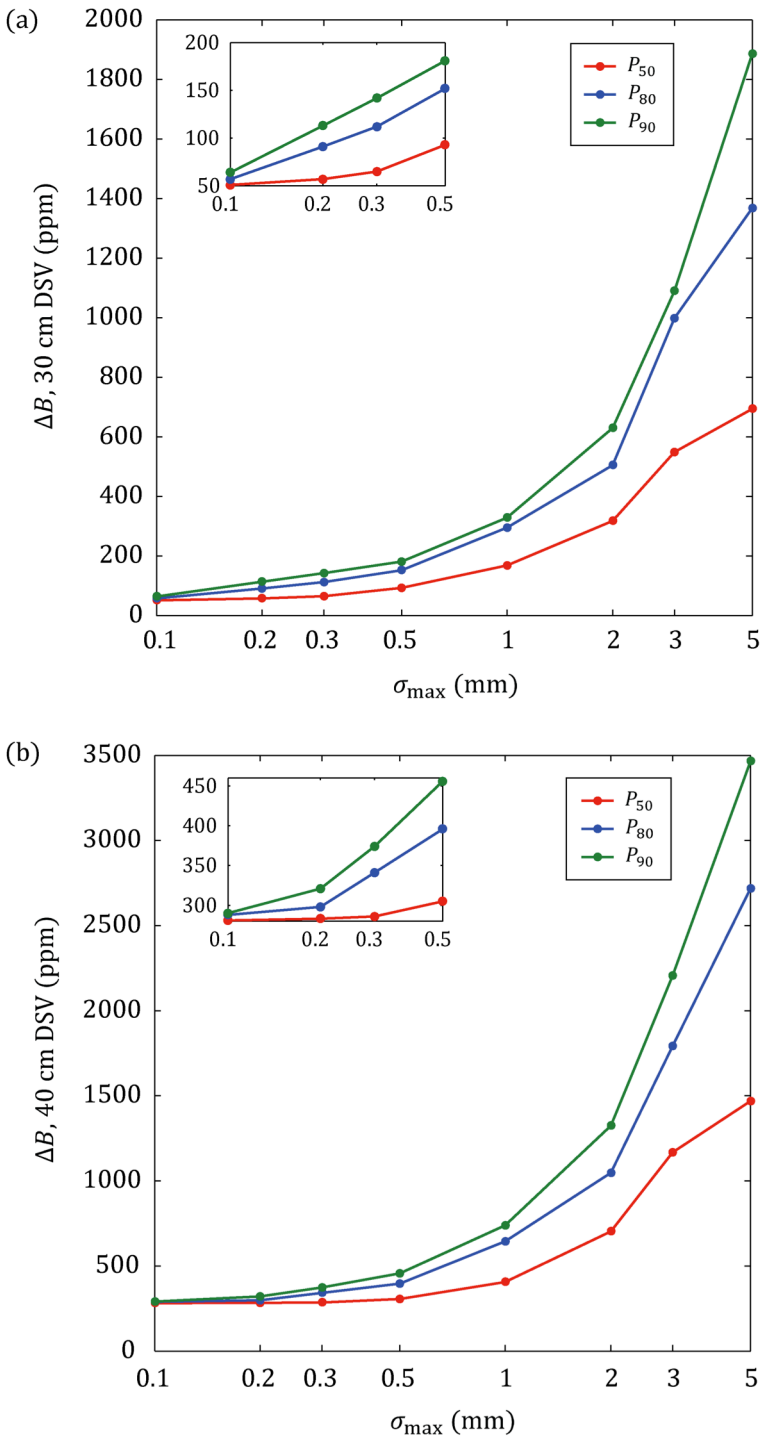


FIGURE 5.15: Plots of the statistical parameters P_{50}, P_{80}, and P_{90} as a function of the maximum variation amplitude σ_{max} for (a) 30 cm and (b) 40 cm DSVs.

The parameters P_{50}, P_{80}, and P_{90} rapidly increase with the diameter of the spherical volume for which they are calculated. Of course, this is the expected result as uniformity in the magnetic field is most easily achieved over smaller volumes.

As shown in Figures 5.14 and 5.15, the NAGP pole piece design is relatively stable for perturbation amplitudes of $\sigma_{\max} \leq 0.5$ mm. This is indicated by the fact that over 90% of the perturbed optimal designs corresponding to $\sigma_{\max} = 0.5$ mm provided inhomogeneity values better than 200 and 500 ppm over 30 and 40 cm DSVs, respectively.

The parameters calculated here and plots such as those provided in Figures 5.14 and 5.15 would be helpful in determining tolerance specifications for the manufacturing process of the NAGP optimized pole pieces. In particular, the results suggest that a tolerance on the surface geometry should be no greater than 0.5 mm if the field uniformity of the optimal design is to be preserved.

5.4 CONCLUSION

Through the numerical results, it was demonstrated that reducing the lateral extent of a biplanar permanent magnet MRI system with a four-column yoke structure leads to a degradation of the field uniformity in the imaging volume. Flat or annular shimmed pole piece designs were shown to be ineffective at compensating for the predominantly non-axisymmetric residual field inhomogeneity. In order to improve on the results obtained with these designs, a nonlinear pole-piece shape optimization method based on field calculations with the finite element method was developed and evaluated. This method was applied to determine axisymmetric and non-axisymmetric grid parameterized pole piece designs that were optimally contoured so as to minimize the magnetic field inhomogeneity in a 40 cm

DSV at isocenter. A subsequent analysis of these designs demonstrated that the NAGP pole piece design yielded superior levels of field uniformity through the compensation of the non-axisymmetric field variations induced by the four-column yoke configuration. As such, it was shown that this design was necessary for the CMA in order to permit successful passive shimming within a 40 cm DSV. In this way, the viability of novel compact magnet assemblies with significantly reduced pole dimensions was demonstrated, further permitting the development of unique hybrid medical devices such as an integrated linac-MRI system. Lastly, a sensitivity analysis was performed to quantify the inhomogeneity consequence of random geometrical errors on the optimized surface of the NAGP pole piece. The results of this analysis demonstrated that the uniformity over a 40 cm DSV obtained with this design was relatively stable for perturbations with a maximum amplitude that was less than 0.5 mm. Consequently, this value represents an appropriate upper limit for the specification of manufacturing tolerances on the design surface.

5.5 REFERENCES

[1] A. Podol'skii, "Development of permanent magnet assembly for MRI devices," *IEEE Trans. Magn.*, **34**(1), 248−252 (1998).

[2] F. Roméo and D. I. Hoult, "Magnetic field profiling: analysis and correcting coil design," *Magnet. Reson. Med.* **1**(1), 44−65 (1984).

[3] R. Vadovic, "Magnetic field correction using magnetized shims," *IEEE Trans. Magn.* **25**(4), 3133−3139 (1989).

[4] J. H. Battocletti, H. A. Kamal, T. J. Myers, and T. A. Knox, "Systematic passive shimming of a permanent magnet for P-31 NMR spectroscopy of bone mineral," *IEEE Trans. Magn.* **29**(3), 2139−2151 (1993).

[5] H. S. Lopez, F. Liu, E. Weber, and S. Crozier, "Passive shim design and a shimming approach for biplanar permanent magnetic resonance imaging magnets," *IEEE Trans. Magn.* **44**(3), 394−402 (2008).

[6] Y. Zhang, D. Xie, B. Bai, H. S. Yoon, and C. S. Koh, "A novel optimal design method of passive shimming for permanent MRI magnet," *IEEE Trans. Magn.* **44**(6), 1058−1061 (2008).

[7] T. Miyamoto, H. Sakurai, H. Takabayashi, and M. Aoki, "A development of a permanent magnet assembly for MRI devices using Nd-Fe-B material," *IEEE Trans. Magn.* **25**(5), 3907−3909 (1989).

[8] M. G. Abele and J. H. Jensen, "Hybrid pole pieces for permanent magnets," *J. Appl. Phys.*, **79**(8), 5199−5201 (1996).

[9] J. H. Jensen, "Optimization method for permanent-magnet structures," *IEEE Trans. Magn.* **35**(6), 4465−4472 (1999).

[10] A. Podol'skii, "Design procedure for permanent magnet assemblies with uniform magnetic fields for MRI devices," *IEEE Trans. Magn.* **36**(2), 484−490 (2000).

[11] X. Jiang, G. Shen, Y. Lai, and J.Tian, "Development of an open 0.3 T NdFeB MRI magnet," *IEEE Trans. Appl. Supercond.* **14**(2), 1621−1623 (2004).

[12] D.-H. Kim, B.-S. Kim, J.-H. Lee, W.-S. Nah, and I.-H. Park, "3-D optimal shape design of ferromagnetic pole in MRI magnet of open permanent-magnet type," *IEEE Trans. Appl. Supercond.* **12**(1), 1467−1470 (2002).

[13] Y. Yao, C. S. Koh, and D. Xie, "Three-dimensional optimal shape design of magnetic pole in permanent magnet assembly for MRI taking account of eddy

currents due to gradient coil field," *IEEE Trans. Magn.* **40**(2), 1164−1167 (2004).

[14] D.-H. Kim, J. K. Sykulski, and D. A. Lowther, "A novel scheme for material updating in source distribution optimization of magnetic devices using sensitivity analysis," *IEEE Trans. Magn.* **41**(5), 1752−1755 (2005).

[15] J. S. Ryu, Y. Yao, and C. S. Koh, "3-D optimal shape design of pole piece in permanent magnet MRI using parameterized nonlinear design sensitivity analysis," *IEEE Trans. Magn.* **42**(4), 1351−1354 (2006).

[16] M. G. Abele, W. Tsui, and H. Rusinek, "Methodology of pole piece design in permanent magnets," *J. Appl. Phys.* **99**(8), 08D903-1−08D903-3, (2006).

[17] B. G. Fallone, M. Carlone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, K. Wachowicz, and C. Kirkby, "Development of a linac-MRI system for real-time ART," *Med. Phys.* **34**(6), 2547−2547 (2007).

[18] B. G. Fallone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, and S. Vidakovic, "First MR images obtained during megavoltage photon irradiation from a prototype integrated linac-MR system," *Med. Phys.* **36**(6), 2084−2088 (2009).

[19] J. Jin, *The Finite Element Method in Electromagnetics*, 2nd ed. (Wiley, New York, 2002).

[20] COMSOL MULTIPHYSICS, ver. 3.4 (Comsol AB, Stockholm, 2007).

[21] *Comsol Multiphysics Reference Guide*, ver. 3.4 (Comsol AB, Stockholm, 2007).

[22] N. B. S. Gloria, M. C. L. Areiza, I. V. J. Miranda, and J. M. A. Rebello, "Development of a magnetic sensor for detection and sizing of internal pipeline corrosion defects," *NDT&E Int.* **42**(8), 669−677 (2009).

[23] D. M. Kohler, "Production and properties of grain-oriented commercially pure iron," *J. Appl. Phys.* **38**(3), 1176−1178 (1967).

[24] R. Skomski and J. M. D. Coey, *Permanent Magnetism* (Institute of Physics Publishing, Bristol, 1999).

[25] MATLAB, ver. 7.11 (The MathWorks, Natick, 2010).

[26] A. Cauchy, "Méthodes générales pour la résolution des systèmes d'équations simultanées," *C. R. Acad. Sci.* **25**, 536−538 (1847).

[27] E. Polak, *Computational Methods in Optimization: A Unified Approach* (Academic Press, New York, 1971).

[28] P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, "Computing forward-difference intervals for numerical optimization," *SIAM J. Sci. Stat. Comput.* **4**(2), 310−321 (1983).

CHAPTER 6
DESIGN AND OPTIMIZATION OF A NOVEL BORED BIPLANAR PERMANENT MAGNET ASSEMBLY FOR HYBRID MRI SYSTEMS

A version of this chapter has been published. T. Tadic and B. G. Fallone, "Design and optimization of a novel bored biplanar permanent magnet assembly for hybrid magnetic resonance imaging systems," *IEEE Trans. Magn.* **46**(12), 4052–4058 (2010).

6.1 INTRODUCTION

Due to the highly competitive nature of the current MRI industry, permanent magnet assemblies have gained widespread popularity in recent years. These devices are considered to have a more reasonable weight, size, and overall cost when compared against competing resistive and superconducting systems. Furthermore, permanent magnet assemblies can be found in widely varying geometrical configurations and offer greater flexibility for achieving compact and open designs tailored for dedicated applications.[1–6]

In this chapter, a design is proposed for a novel bored biplanar permanent magnet assembly for MRI. A unique feature of this design is the presence of a large cylindrical hole longitudinally bored through the entire yoke and pole structures. The existence of this bore permits the inclusion of additional therapeutic or diagnostic devices within or proximate to the magnet structure that might benefit from being directed along, or oriented parallel to, the main magnetic field of the MRI system. The bored magnet assembly

(BMA) disclosed in the present work is particularly well suited for integration with a medical linear accelerator (linac) in a parallel configuration, as currently being developed by our group.[7–30] A schematic diagram illustrating this configuration has been shown in Figure 1.2.

It was established in the foregoing chapter that the magnetic field distribution within the imaging volume of biplanar permanent magnets is strongly related to the configuration of the yoke and pole structures.[47] Accordingly, the finite element method (FEM) was employed in the present work to quantify the effects of the yoke bore on the magnetic field homogeneity in the imaging volume of the BMA. Furthermore, similar to the compact magnet assembly (CMA) investigated in Chapter 5, the overall dimensions of the BMA were constrained to permit rotation of this structure within a typical radiotherapy vault. Since it was shown that optimized pole pieces were required to yield suitable levels of uniformity with the CMA, the iterative FEM-based optimization method[47] that was developed in Chapter 5 was also applied to the BMA. Through a modification of the surface parameterizations introduced for the CMA, a novel double annular ring shim pole piece design was obtained, in addition to axisymmetric and non-axisymmetric grid parameterized (AGP and NAGP) designs. The magnetic field inhomogeneities corresponding to these designs were calculated and the results were compared against flat and single annular ring shimmed designs as found commonly in industry. The axial and transverse magnetic fields produced by the BMA with the NAGP optimized design were then examined within and proximate the large bore, in order to acquire an understanding of the unique magnetic field characteristics, particularly within regions that may potentially be occupied by magnetically sensitive devices. Lastly, a sensitivity analysis as introduced in the preceding chapter was performed to determine the sensitivity of the field uniformity to random geometrical variations in the design surface of the NAGP pole piece.

6.2 METHODS

6.2.1 Magnet Assembly Specifications

The bored four-column biplanar permanent magnet assembly considered in this work is shown in Figure 6.1. This magnet assembly is identical to the 0.2 T laterally reduced magnet assembly (CMA) analyzed in Chapter 5, with the exception of a large 20 cm diameter hole longitudinally bored through the yoke and poke structures along the z axis. Table 6.1 provides the values for the dimensions labeled in Figure 6.1(b). For a representative linac target to isocenter distance of 135 cm, a maximum radiation field size of 27×27 cm^2 at isocenter would be possible for this design.

The yoke and column structures of the bored magnet assembly (BMA) are composed of AISI 1020 plain carbon steel, and the pole pieces are composed of a special Armco magnetic steel. The magnetization curves[22, 23] defining the nonlinear properties of these materials were previously shown in Figure 5.2. The permanent magnet poles from which the magnetic field originates are constructed from a neodymium-iron-boron compound ($Nd_2Fe_{14}B$), with a homogenous remanent magnetic field of 1.005 T aligned in the z direction.[3]

In order to evaluate the quality of the magnetic field produced by the BMA, the parts per million (ppm) inhomogeneity ΔB of the magnetic field was introduced:

$$\Delta B = \frac{B_{\max} - B_{\min}}{B_0} \cdot 10^6, \tag{6.1}$$

where $B_{\max}$ and $B_{\min}$ are the maximum and minimum values for the magnetic field magnitude within the volume of interest $\mathcal{V}$, respectively, and B_0 is the magnetic field strength at isocenter. Additionally, the strength of

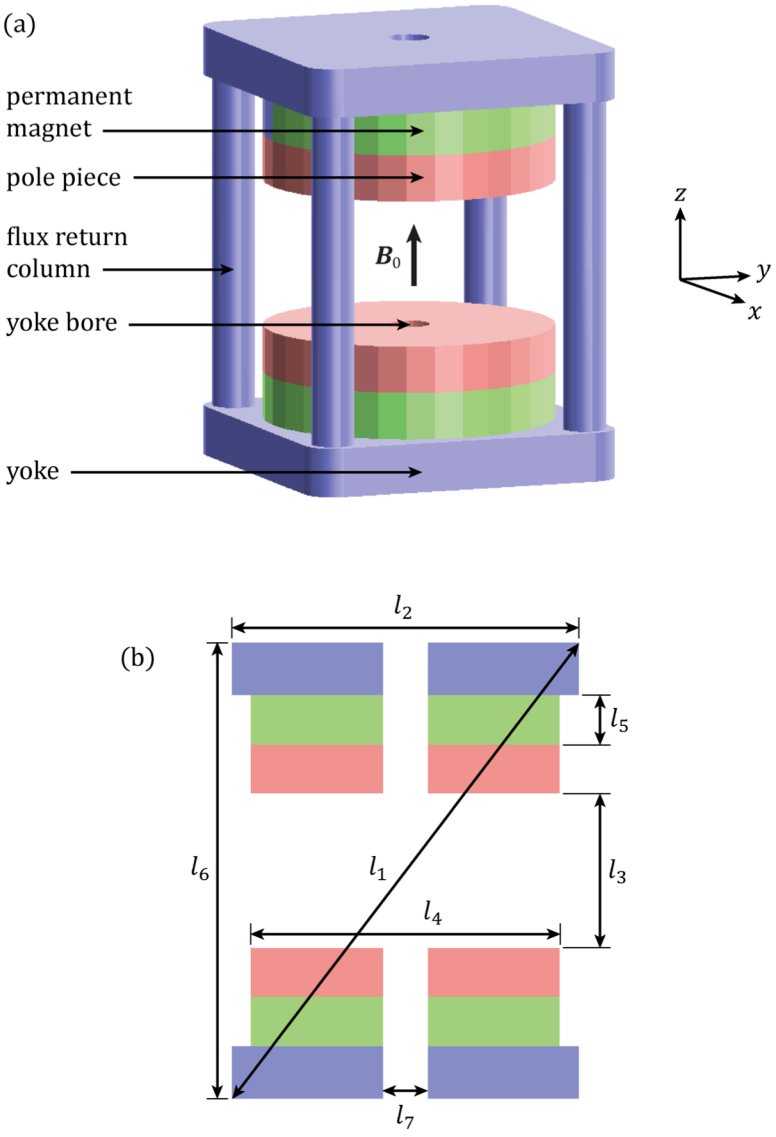


FIGURE 6.1: (a) A 3D schematic of the novel bored four-column biplanar permanent magnet assembly with the relative orientation of the Cartesian coordinate axes shown. (b) A cross-sectional view with dimension labels.

TABLE 6.1: Dimensions of the permanent magnet assembly illustrated in Figure 6.1.

Quantity	Length (mm)
l_1	2600
l_2	1578
l_3	700
l_4	1398
l_5	235
l_6	2066
l_7	200

the fringe magnetic fields produced by the magnet assembly were calculated in terms of the maximal extent (x_{5G}, y_{5G}, z_{5G}) of the 5 G magnetic field contours.

6.2.2 The Finite Element Method

The required magnetic field calculations performed in this chapter are accomplished with the commercially available three-dimensional (3D) FEM software package COMSOL MULTIPHYSICS.[20] Details regarding the FEM implementation employed by COMSOL MULTIPHYSICS have been provided in § 3.3.5.

6.2.2.1 Model Description

Due to similarity with the CMA investigated in Chapter 5, the FEM model geometry for the BMA was obtained through a simple modification of the CMA model already available. As such, the FEM model constructed in this work consists of the BMA positioned at the center of a large cubic global model volume (GMV) truncated at a distance of 5 m from isocenter. Through an application of the same boundary conditions described in the preceding

chapter, the FEM model was reduced to just one eighth of the complete geometry.

The simplified FEM model was partitioned into approximately 1.4×10^5 isoparametric tetrahedral quadratic Lagrange elements. As before, this resulted in a mesh-independent solution, as indicated by a field variation within a 40 cm DSV of no more than 1 μT (approximately 5 ppm) upon doubling the overall nodal density in the model.

The magnetostatics application mode in the AC/DC module of COMSOL MULTIPHYSICS was employed to numerically solve the governing equation in Eq. (3.54) for the scalar magnetic potential at discrete points within the model geometry, from which the magnetic field could be calculated through use of Eqs. (3.52) and (3.53). An illustration of the simplified model geometry and finite element discretization is provided in Figure 6.2. The depicted magnetic field corresponding to the FEM solution was calculated in less than two minutes on a 2.7 GHz Intel Xeon quad-core PC workstation with 16 GB of RAM.

6.2.3 Pole Piece Design Parameterization

Each pole piece design considered that exhibits a contoured open face was optimized for maximum magnetic field homogeneity in a 40 cm DSV. The geometric parameterizations used in defining the surface geometries of the pole pieces were modified from those described in Chapter 5, so as to account for the presence of the 20 cm diameter yoke bore. In all cases, the design variables were constrained such that the nominal pole separation l_3 remained at least 70 cm, with all other dimensions of the magnet structure held fixed.

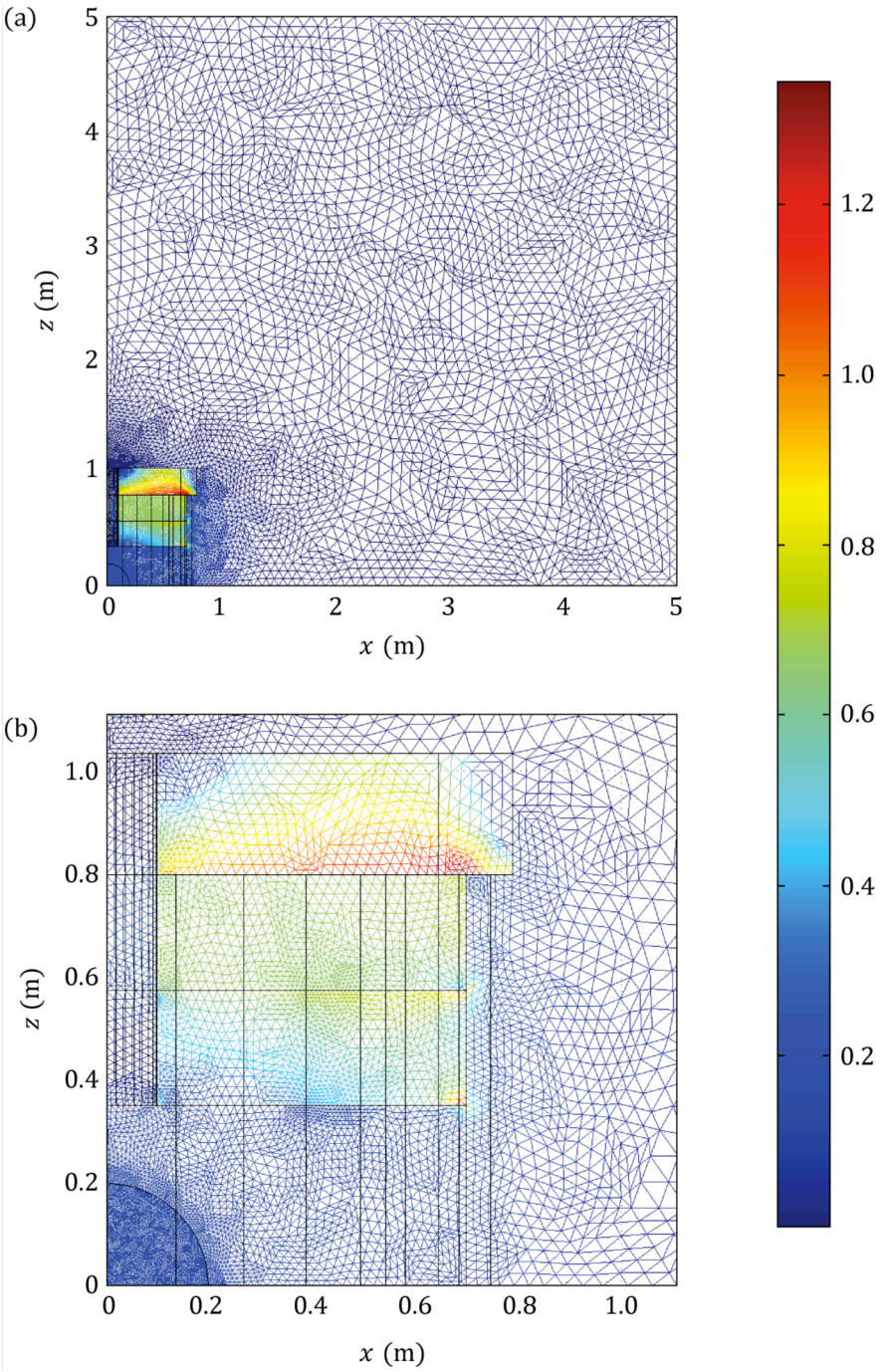


FIGURE 6.2: (a) A section of the simplified FEM model geometry corresponding to the BMA with a flat pole piece design. The magnetic field solution is overlaid with the FEM mesh and displayed in units of T. (b) A close-up section illustrating the fine mesh resolving the magnet geometry and 40 cm DSV at isocenter.

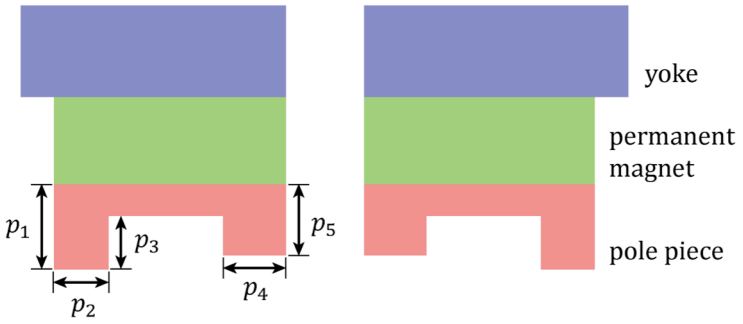


FIGURE 6.3: Design parameterization for optimization of the SARS and DARS pole piece designs. Only the upper pole piece is illustrated.

6.2.3.1 Annular Shim Parameterization

The design variables for the annular ring shim parameterizations were selected as the dimensions of the pole piece cross-section, as shown for an upper pole section in Figure 5.4. This parameterization resulted in 5 unique

design variables for the double annular ring shimmed (DARS) design, denoted p_i (for $i = 1, 2, \ldots, 5$). By enforcing the constraints $p_4 = 0$ and $p_5 = p_1 - p_3$, the parameterization of the SARS design as defined by 3 unique design variables was obtained. Lastly, in order to consider a simple flat surface, the remaining design variables were constrained such that the pole piece had a maximal uniform thickness of $p_1 = 218$ mm with $p_2 = p_3 = 0$.

6.2.3.2 Grid Parameterization

In addition to the flat, SARS, and DARS designs, axisymmetric and non-axisymmetric grid parameterized pole pieces were obtained for the BMA. To this end, the open surface of the pole piece was parameterized with 112 control points distributed on a grid in the $r\theta$ plane, as shown in Figure 6.4. The z coordinates of the control points were taken as the design variables, at

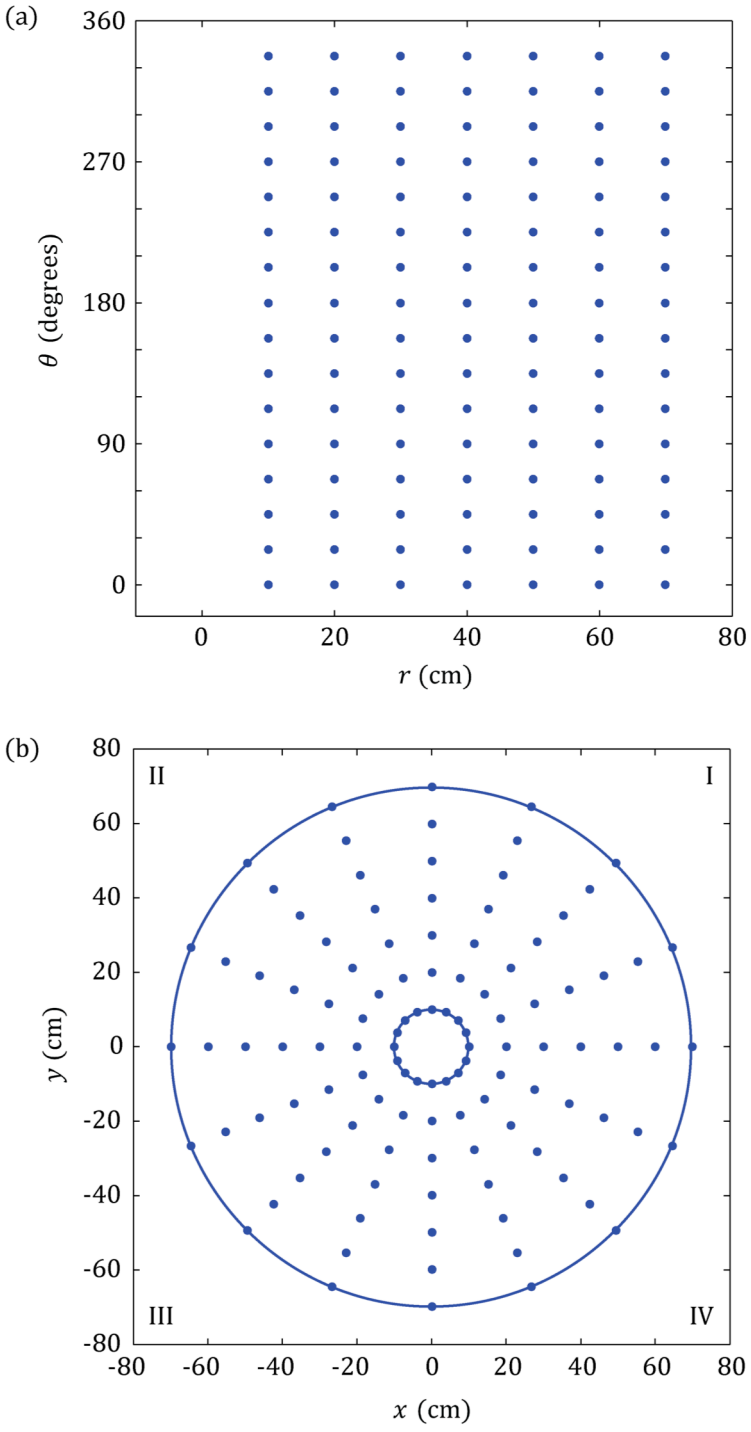


FIGURE 6.4: Control point distribution for the grid parameterization, plotted in (a) polar coordinates and (b) Cartesian coordinates.

which locations the thickness of the pole piece was varied during optimization. The complete pole piece surface was then generated based on a linear spline interpolation between the control points.

In a manner similar to that employed in § 5.2.3.2, symmetry in the pole piece surface allowed for the reduction in the number of unique degrees of freedom in the optimization procedure. To this end, the same constraints were applied in this work, with the exception that the control points on the line ($r = 0$) were not constrained to share the same z coordinate, as they no longer reside in the same location on the design surface. In this way, a total 7 and 21 unique design variables were obtained for the AGP and NAGP designs, respectively.

6.2.4 Nonlinear Optimization

Details of the iterative optimization method employed in this work were presented in Chapter 5. This algorithm was programmed in the MATLAB[24] scripting environment and utilized macros from the COMSOL MULTIPHYSICS scripting language for the automatic generation and analysis of the required models. A flow diagram summarizing the optimization process is provided in Figure 5.6.

Mathematically, the optimization problem can be expressed as:

$$\text{minimize } f(\boldsymbol{p}), \tag{6.2a}$$

$$\text{subject to } \boldsymbol{p} \in \mathfrak{P}, \tag{6.2b}$$

where f is the objective function, $\boldsymbol{p}$ is the design vector, and $\mathfrak{P}$ is the feasible domain defined according to Eq. (5.9).

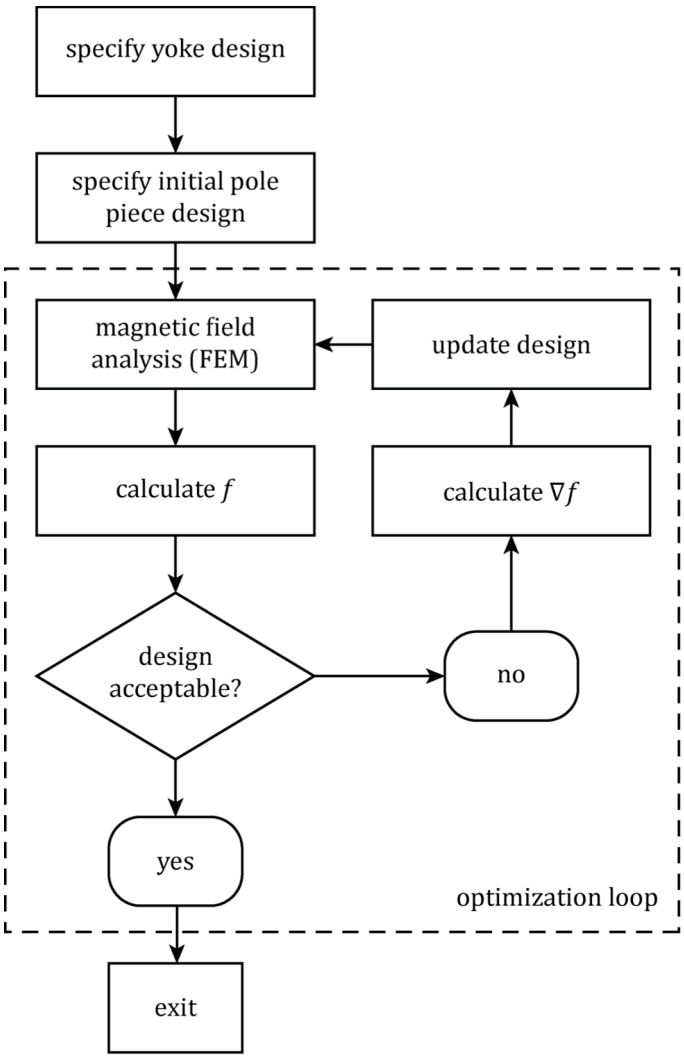


FIGURE 6.5: Iterative optimization flow diagram.

The objective function is defined by the integral formula:

$$f = \int_{\mathcal{V}} [B(\boldsymbol{r};\boldsymbol{p}) - B_0]^2 d\mathcal{V}', \tag{6.3}$$

where $B(\boldsymbol{r};\boldsymbol{p})$ is the magnitude of the magnetic field at the point $\boldsymbol{r}$, implicitly due to the magnet design defined by $\boldsymbol{p}$. As before, the target region of interest $\mathcal{V}$ is taken to be a 40 cm DSV at isocenter.

In order to calculate the gradient of the objective function ∇f with respect to the design variables, the forward finite-difference approximation in Eq. (5.12) was used. A perturbation interval of $\delta p = 0.1$ mm was specified for this calculation, as determined by trial and error. This value was shown to be reasonable based on an estimate of the noise-level perturbation $\delta\check{p}$ conducted previously in § 5.3.2.2.

During the optimization process, the design vector was updated at the kth iteration according to the projected gradient formula (see § 4.3.3):

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{\Pi}\left(\boldsymbol{p}^{(k)} + \alpha^{(k)}\boldsymbol{d}^{(k)}\right), \tag{6.4}$$

where $\alpha^{(k)}$ is the scalar step size, $\boldsymbol{d}^{(k)}$ is the search direction, and $\boldsymbol{\Pi}(\boldsymbol{p})$ denotes the projection of $\boldsymbol{p}$ onto the feasible domain $\mathfrak{P}$. The conjugate gradient (CG) version of this optimization method as described in § 5.2.4 was not implemented at the time of this work. Consequently, $\boldsymbol{d}^{(k)}$ was calculated according the steepest descent (SD) method[3] via Eq. (4.12).

6.2.5 Yoke Bore Magnetic Field Analysis

The primary motivation for considering a large hole vacated from the yoke and pole structures was to permit the inclusion of additional devices within or proximate the magnet assembly that would benefit from an orientation parallel to the main magnetic field. Therefore, it was of particular interest to explore the characteristics and behaviour of the magnetic field in these regions. To this end, the magnetic field produced by the BMA with the

optimal NAGP pole piece design was examined in terms of its axial B_z and transverse $B_\perp$ components within and proximate the 20 cm diameter yoke bore. Specifically, linear profiles of these field components, passing through the yoke bore and oriented in the z direction, were obtained for a selection of positions along the x axis.

6.2.6 Design Sensitivity

As discussed in Chapter 5, the NAGP pole piece design would likely be subject to random errors during the manufacturing process. In order to simulate these errors and quantify the effects of perturbations to the design surface, a sensitivity analysis was performed on the optimal NAGP pole piece design. The details of this procedure were provided in § 5.2.5.

Populations consisting of 100 perturbed designs were obtained for various maximum amplitudes $\sigma_{\max}$ in the range of 0.1 to 5 mm. For each perturbed design, the residual field variation ΔB was calculated over 10, 20, 30, and 40 cm DSVs. The sensitivity of the optimized design was then evaluated in terms of the parameters P_{50}, P_{80}, and P_{90}, that are defined as the ΔB values that upper bound 50, 80, and 90% of the perturbed designs in a given population, respectively.

6.3 RESULTS AND DISCUSSION

6.3.1 Yoke Bore Insertion

Table 6.2 compares the magnetic field properties for the magnet assemblies with and without the 20 cm bore. Flat and SARS pole pieces were considered, where the width and depth of the peripheral ring of the SARS designs were optimized to minimize the cumulative field inhomogeneity f as measured over a 40 cm DSV.

TABLE 6.2: Summary of optimization results and comparison of magnetic field properties for the magnet designs with flat and annular shimmed pole pieces.

	CMA		BMA	
Quantity	Flat Pole Piece	SARS Pole Piece	Flat Pole Piece	SARS Pole Piece
Iterations	-	11	-	75
B_0 (T)	0.203	0.172	0.197	0.157
f (10^{-11} T^2·m^3)	1455	1.899	647.1	19.51
ΔB, 30 cm DSV (ppm)	8774	321	7130	1355
ΔB, 40 cm DSV (ppm)	15299	1030	24978	4474
x_{5G} (m)	2.4	2.4	2.4	2.4
y_{5G} (m)	2.4	2.4	2.4	2.4
z_{5G} (m)	2.8	2.8	2.8	2.7

The presence of the yoke bore resulted in a dramatic increase in the magnetic field inhomogeneity ΔB for the SARS pole piece design. In particular, ΔB for this design increased by greater than a factor of four, resulting in 1355 and 4474 ppm over 30 and 40 cm DSVs, respectively. In contrast, an interesting effect was observed for the flat pole piece design, for which the field variation within a 30 cm DSV actually improved from 8774 to 7130 ppm upon inclusion of the yoke bore. The value for f also decreased by greater than a factor of two for this design, despite a near doubling of ΔB to 24978 ppm within a 40 cm DSV. As evident in the magnetic field maps provided in Figure 6.6, the yoke bore results in a smoothing of the characteristic peak near isocenter in the field distribution for the flat pole piece. However, this smoothing effect is of limited benefit within larger volumes, due to the rapid drop off in the field magnitude as the yoke bore is neared along the z direction. Furthermore, this effect is not observed with the SARS pole piece, as a peak in the field distribution near isocenter was likely eliminated through optimization of the pole surface. In any case, based

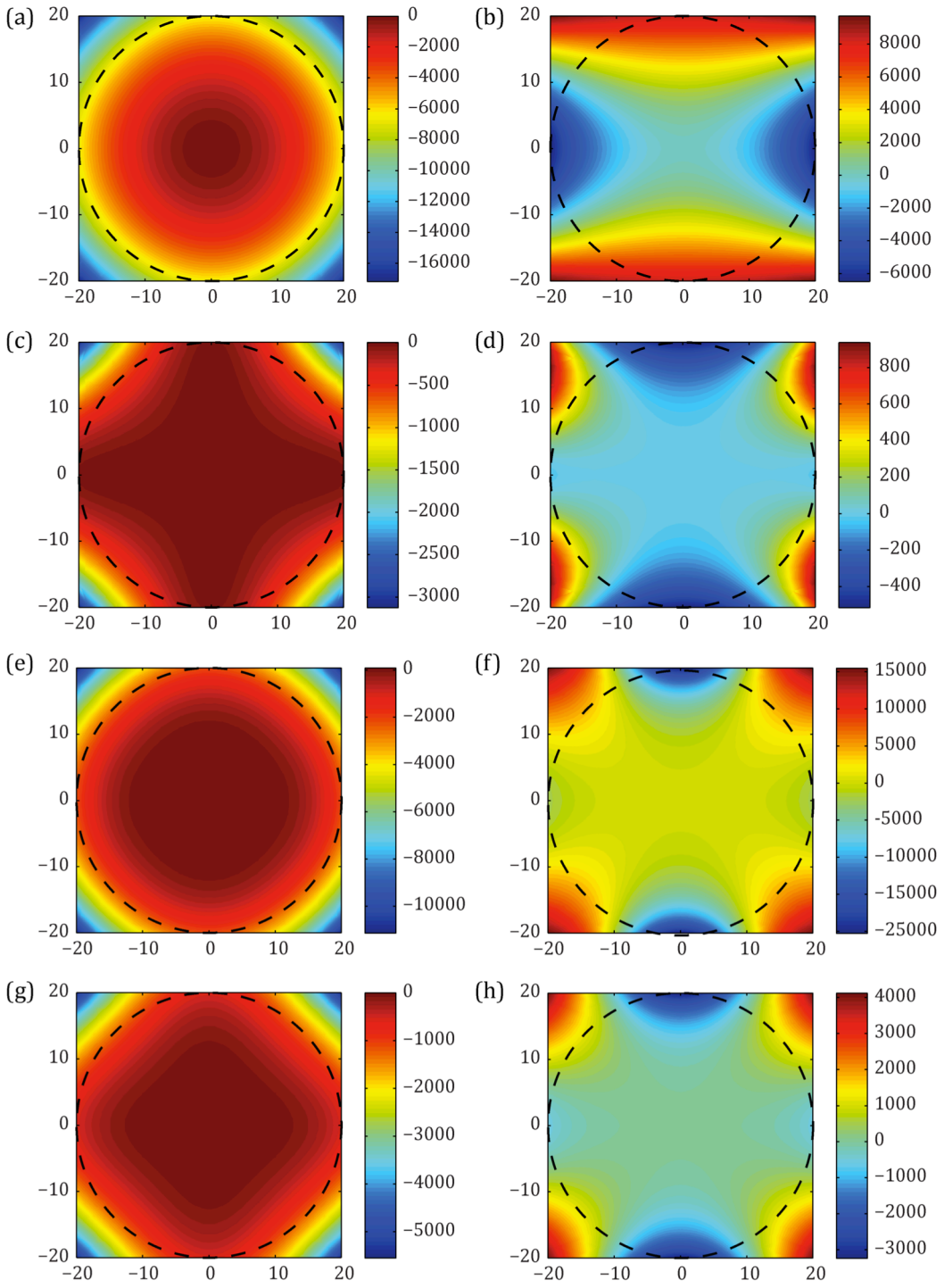


FIGURE 6.6: Magnetic field maps in the xy (first column) and xz (second column) planes at isocenter corresponding to the: (a), (b) flat and (c), (d) SARS designs for the CMA; (e), (f) flat and (g), (h) SARS designs for the BMA. Axis dimensions are in cm. Map values are displayed in parts per million, measured relative to B_0 for each design. The dashed circles represent the boundary of a 40 cm DSV.

on the specification introduced in Chapter 5 that successful passive shimming requires a value of less than 500 ppm in the volume of interest, these results indicate that the BMA would not be suitable for medical imaging within both 30 and 40 cm DSVs with either flat or SARS pole piece designs.

Inclusion of the yoke bore necessitates a removal of magnetic material from within the yoke and pole structures. Due to the loss of magnetic material that contributes to the net magnetic field at isocenter, a slight decrease in B_0 of 0.006 and 0.015 T was calculated for the flat and SARS pole piece designs, respectively. The extent of the corresponding fringe fields remained unchanged for the flat pole piece, but was slightly reduced from $(x_{5G}, y_{5G}, z_{5G}) = (2.4, 2.4, 2.8)$ to $(2.4, 2.4, 2.7)$ m for the SARS design.

6.3.2 Nonlinear Optimization

6.3.2.1 Design Comparison

An optimal DARS pole piece design was obtained after 186 iterations of the optimization scheme described in § 6.2.4, while optimized AGP and NAGP designs were obtained after 672 and 1022 iterations, respectively. The resulting designs are visually compared in Figure 6.7 alongside the flat and SARS designs. Magnetic field properties corresponding to the DARS, AGP, and NAGP designs are compared in Table 6.3.

The greatest level of field uniformity was obtained with the NAGP design, which resulted in the lowest values for the objective function f and magnetic field inhomogeneity. Specifically, ΔB values of 77 and 921 ppm were obtained for 30 and 40 cm DSVs, respectively. This is compared to values of 180 and 1127 ppm obtained for the AGP design, and 279 and 2116 ppm for the DARS design, respectively. Thus, with regards to the inhomogeneity over a 40 cm DSV, the NAGP pole piece provided a relatively marginal

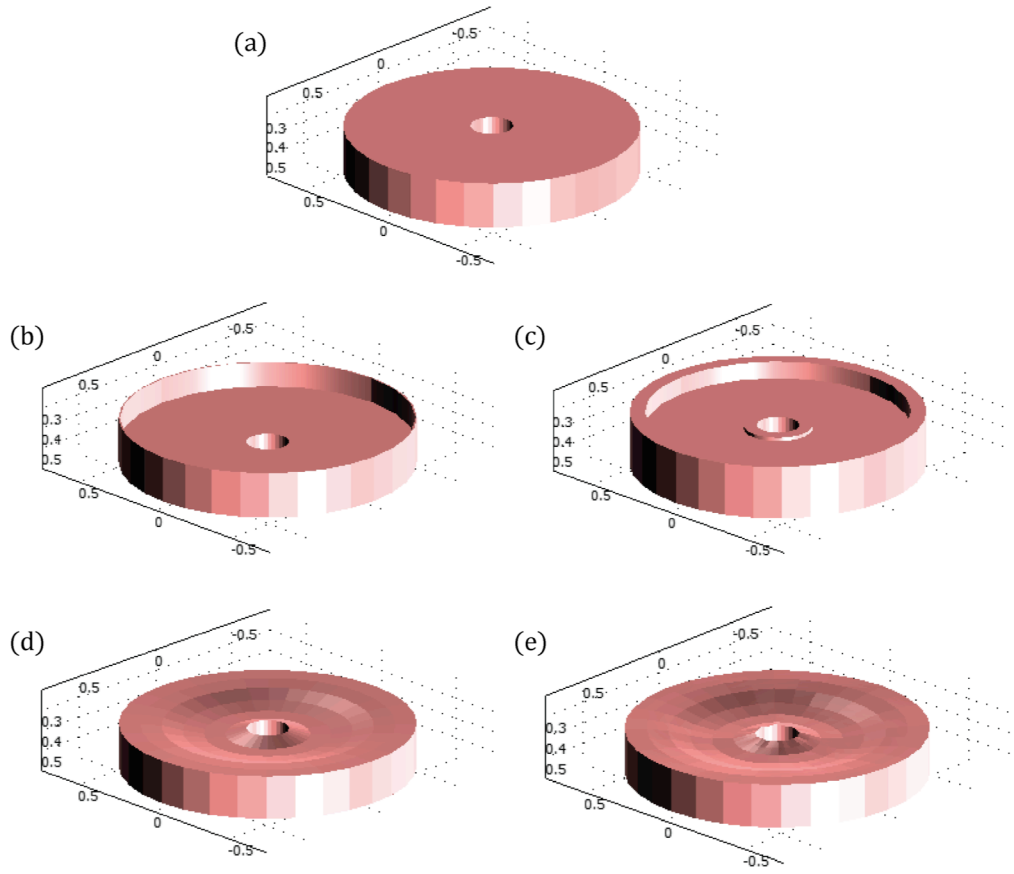


FIGURE 6.7: 3D visualizations of the (a) flat, (b) SARS, (c) DARS, (d) AGP, and (e) NAGP optimized pole pieces. Axis dimensions are displayed in meters.

improvement over the AGP design. Furthermore, as can be seen in Figure 6.7, the surface geometries of these designs are visually similar. These results contrast those obtained in the preceding chapter, for which the surface topology of the NAGP design involved more pronounced non-axisymmetric features, and resulted in a greater reduction of ΔB within a 40 cm DSV. The limited advantage of the NAGP design here is further visualized in the magnetic field maps in the xy and xz planes at isocenter provided in Figure 6.8. Although this design is effective at reducing the non-axisymmetric field variation near isocenter, the overall improvement is

TABLE 6.3: Summary of optimization results and comparison of magnetic field properties for the bored magnet designs with optimized pole pieces.

Quantity	DARS Pole Piece	AGP Pole Piece	NAGP Pole Piece
Iterations	186	672	1019
B_0 (T)	0.171	0.179	0.177
f (10^{-11} T^2·m^3)	1.685	0.719	0.353
ΔB, 30 cm DSV (ppm)	279	180	77
ΔB, 40 cm DSV (ppm)	2116	1127	921
x_{5G} (m)	2.4	2.4	2.4
y_{5G} (m)	2.4	2.4	2.4
z_{5G} (m)	2.8	2.8	2.7

extenuated over larger volumes due to the rapid drop off in the field magnitude as the yoke bore is approached along the z direction.

The DARS, AGP, and NAGP pole piece designs all yield inhomogeneity levels within a 30 cm DSV that are less than the 500 ppm limit specified in order to permit successful passive shimming. In particular, the simple and effective DARS design achieves an order of magnitude improvement in f as compared to the SARS design, while reducing the inhomogeneity by greater than a factor of four over a 30 cm DSV. Although this design does not perform as well as the AGP and NAGP designs, the optimization requires significantly less computation time, and the resulting design would be simpler to construct.

The field strengths resulting from the DARS, AGP, and NAGP designs were relatively similar, with values in the range of 0.171−0.179 T. Furthermore,

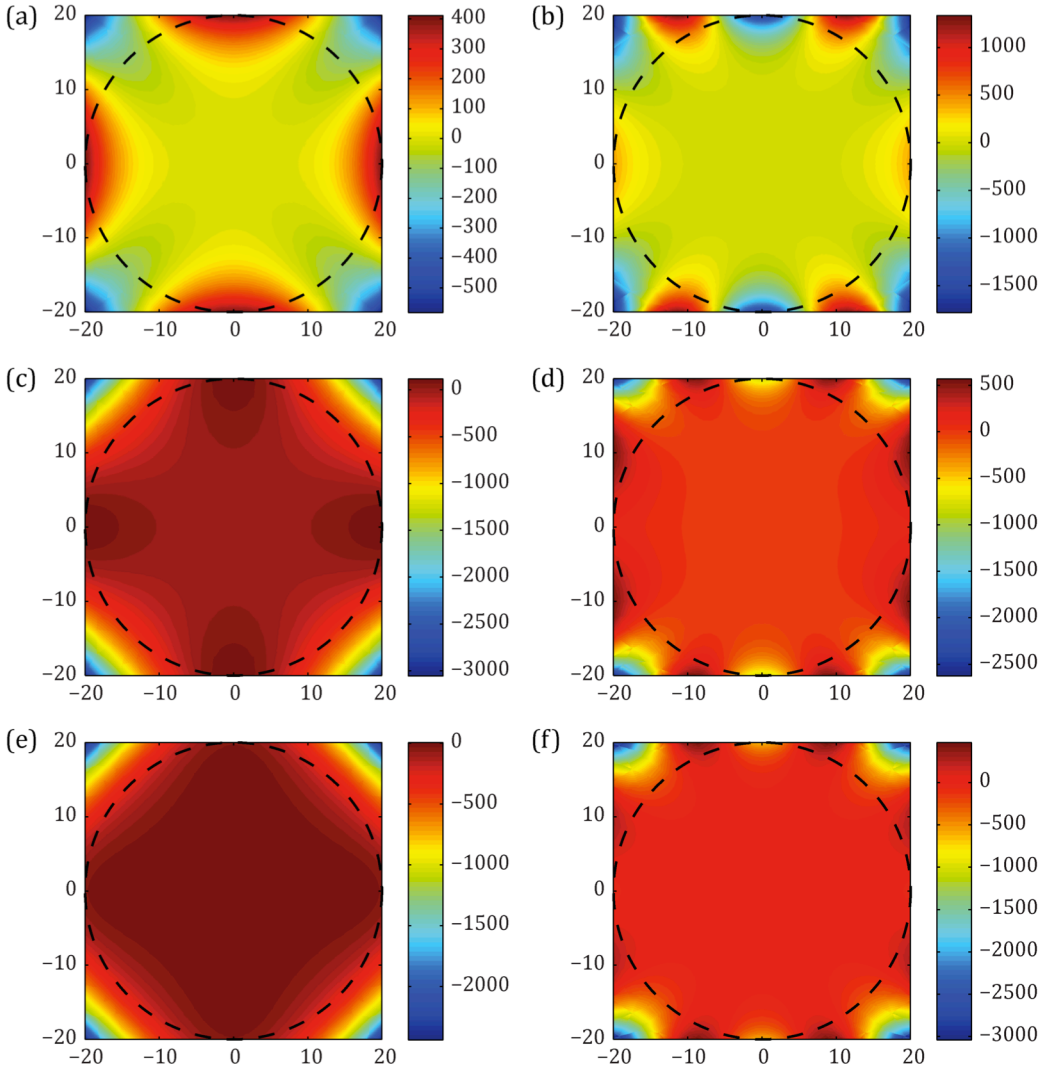


FIGURE 6.8: Magnetic field maps in the xy (first column) and xz (second column) planes at isocenter corresponding to the: (a), (b) DARS, (c), (d) AGP, and (e), (f) NAGP designs obtained for the BMA. Axis dimensions are in cm. Map values are displayed in parts per million, measured relative to B_0 for each design. The dashed circles represent the boundary of a 40 cm DSV.

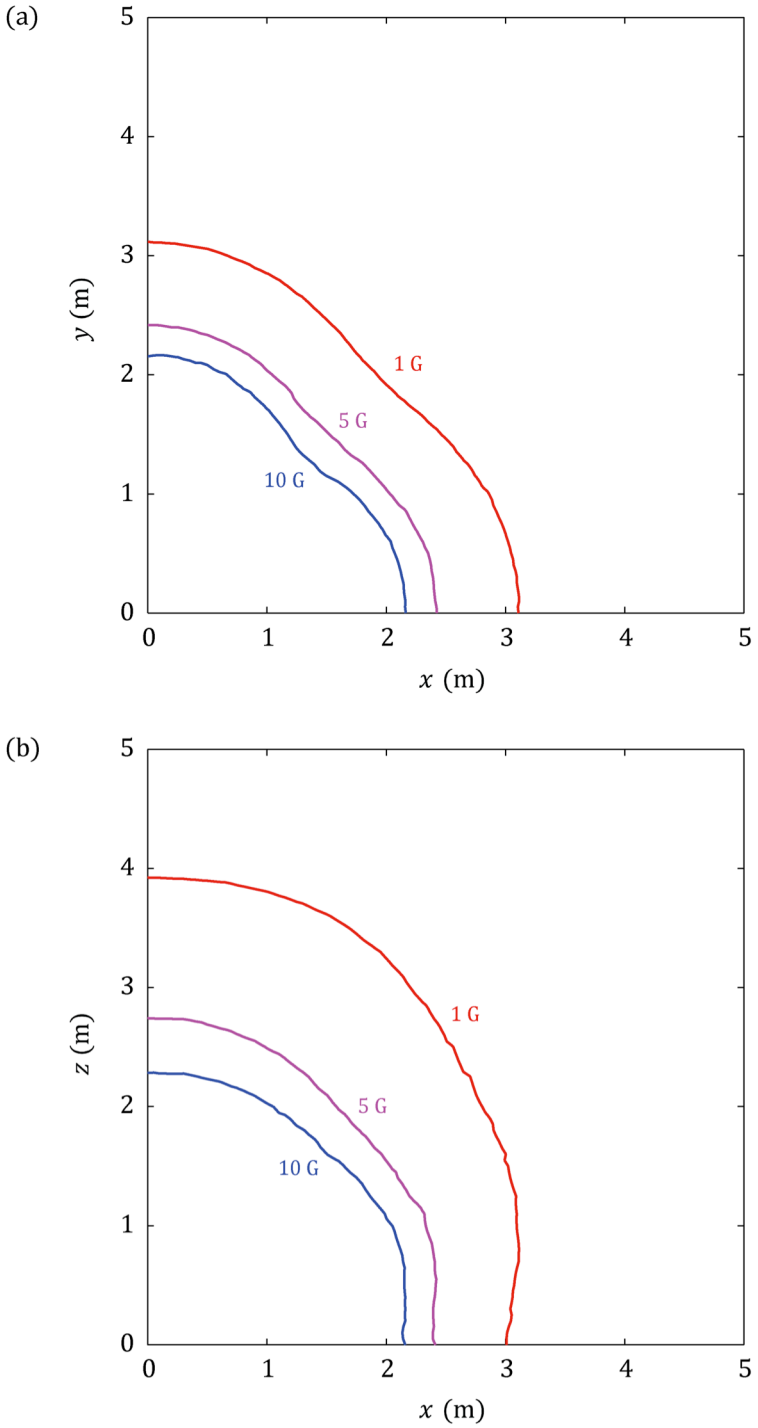


FIGURE 6.9: Plots of the 1, 5, and 10 G magnetic field contours projected onto the (a) xy and (b) xz planes for the BMA with the NAGP pole piece design obtained via the SD method.

little change was observed in the strength of the fringe magnetic fields, as seen in the values provided in Table 6.3. The corresponding 1, 5, and 10 G magnetic field contours have been plotted in Figure 6.9.

6.3.2.2 Algorithm Performance

The evolution of the objective function f and gradient norm $|\nabla f|$ is plotted in Figure 6.10. These graphs correspond to the optimization of the AGP and NAGP pole piece designs, using the SD version of the optimization scheme developed in Chapter 5. From Figure 6.10(a), it can be seen that AGP designs that surpasses the SARS and DARS designs are obtained in just 17 and 65 iterations, respectively. After this short segment of rapid improvement, the evolution of f is dominated by a long plateau. This decline in performance was characteristic of the optimization algorithm during previous use with the CMA, for which this behaviour was attributed to the first-order nature of the SD method employed in calculating the search directions. Nevertheless, the optimization algorithm yielded a reduction in the objective function of greater than three orders of magnitude relative to the initial design state describing a simple flat pole piece surface. As seen in Figure 6.10(b), a similar overall reduction in the gradient norm $|\nabla f|$ was achieved, indicating that the optimization algorithm was successful in approaching a stationary point.

The evolution of the gradient norm $|\nabla f|$ illustrated in Figure 6.10(b) is notably non-smooth. In particular, the sharp variations evident in this graph are seemingly more random and of greater magnitude than those previously observed during optimization of the CMA. If a similar bound δf on the absolute error of f is assumed (due to the similarity of the CMA and BMA FEM models), then the larger value of $|\nabla f| \approx 2 \times 10^{-9}$ T^2·m^2 near the termination of the NAGP optimization suggests a smaller estimate for the noise-level perturbation interval[17] of $\delta \check{p} \approx 0.001$ mm (see § 5.3.2.2). Thus,

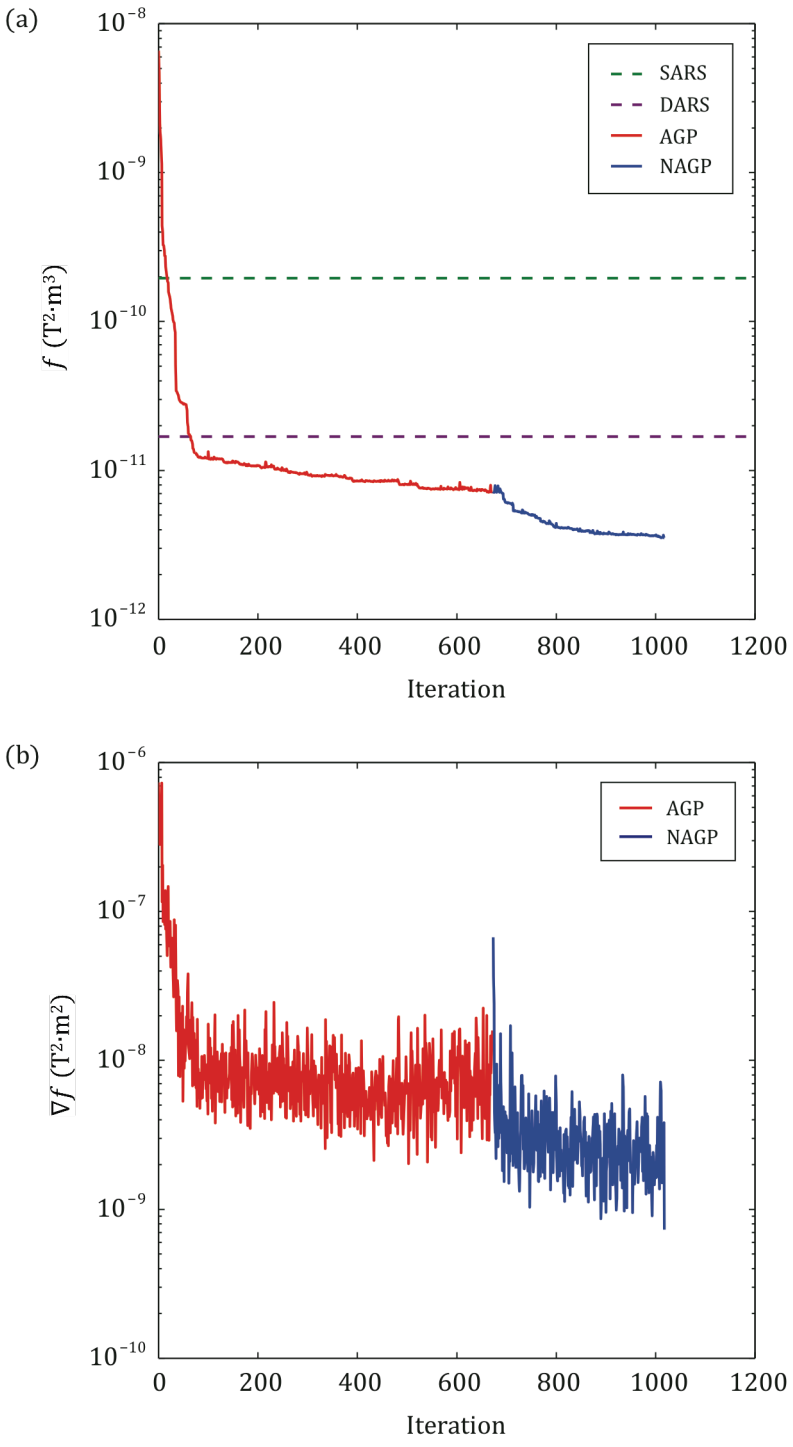


FIGURE 6.10: Evolution of the (a) objective function f and (b) gradient ∇f during optimization via the SD method.

the perturbation interval $\delta p = 0.1$ mm used here in approximating ∇f is of sufficient magnitude to produce a meaningful change in f that is greater than the noise-level. That said, too large of a perturbation can also be harmful, due to the truncation error encountered when neglecting higher-order terms in the linear approximation for $\partial f / \partial p_i$, upon which the finite-difference approximation is founded. Therefore, improved performance of the optimization algorithm might be possible through a more thorough investigation into the optimal choice of δp.

6.3.3 Yoke Bore Magnetic Field Analysis

Linear profiles of the field components B_z and $B_\perp$, oriented in the z direction and positioned at $x = 0, 10, 20, \text{and } 50$ mm, have been provided in Figure 6.1, for which the inset plots correspond to the region where magnetically sensitive equipment or devices may be located.

The various B_z profiles obtained are relatively similar in their shape and magnitude. A common trait of these profiles is a drop from the field strength $B_0 = 0.177$ T at $z = 0$, to a large negative value of approximately -0.300 T within the bore and adjacent to the permanent magnet pole. This is followed by an increase back to a small positive peak of approximately 47 G, at a position approximately 18 cm exterior to the yoke. The region of negative B_z in these profiles indicates that the magnetic field reverses direction within the yoke bore. Indeed, this behaviour can be observed in the magnetic field line plot generated by COMSOL MULTIPHYSICS in Figure 6.12.

Although the $B_\perp$ profiles retain a similar shape at various positions, the peak magnitude of this component changes dramatically. The transverse field component is nulled at $x = 0$ due to symmetry about the central bore axis, and then increases in magnitude with increasing off-axis position as the inner bore surface is approached. This effect is further illustrated in Figure 6.13,

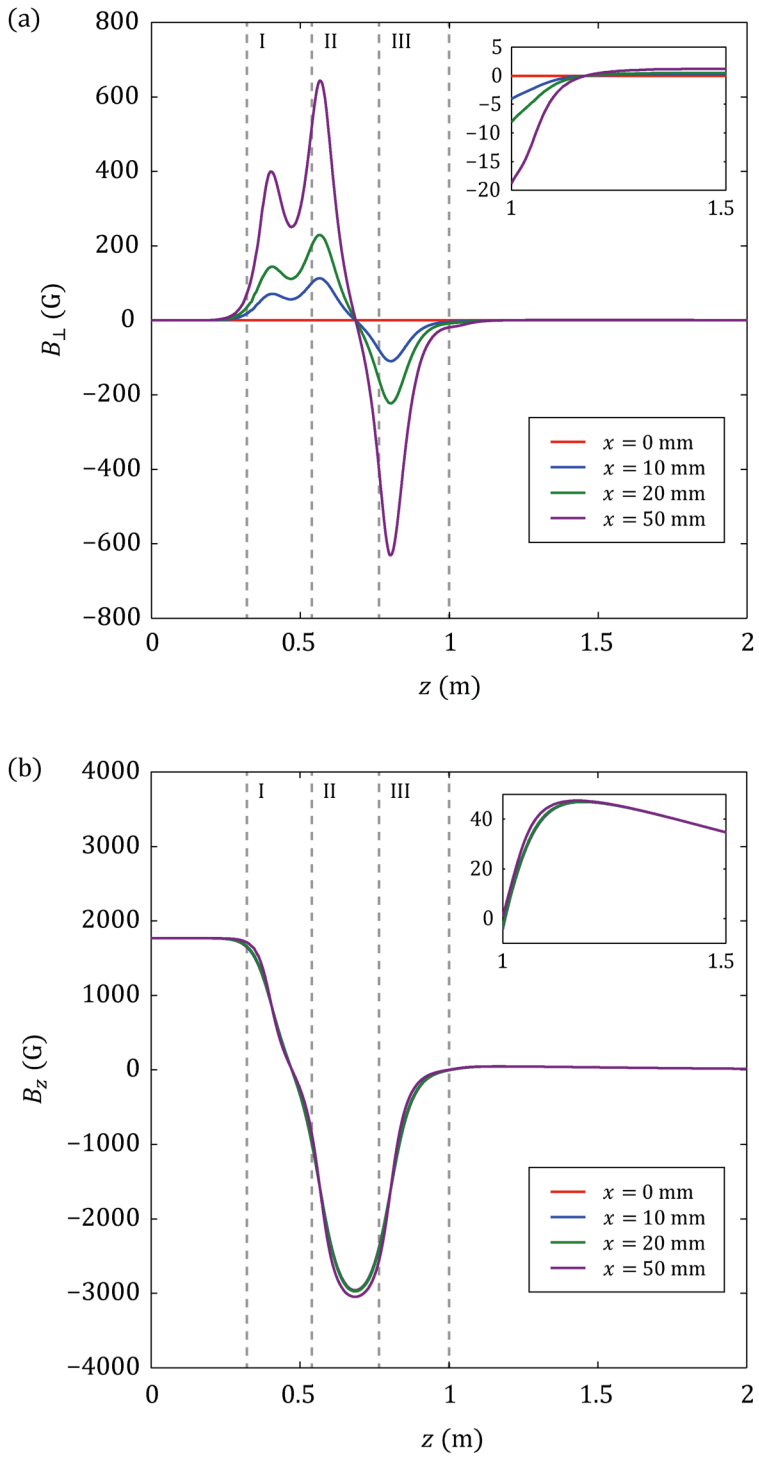
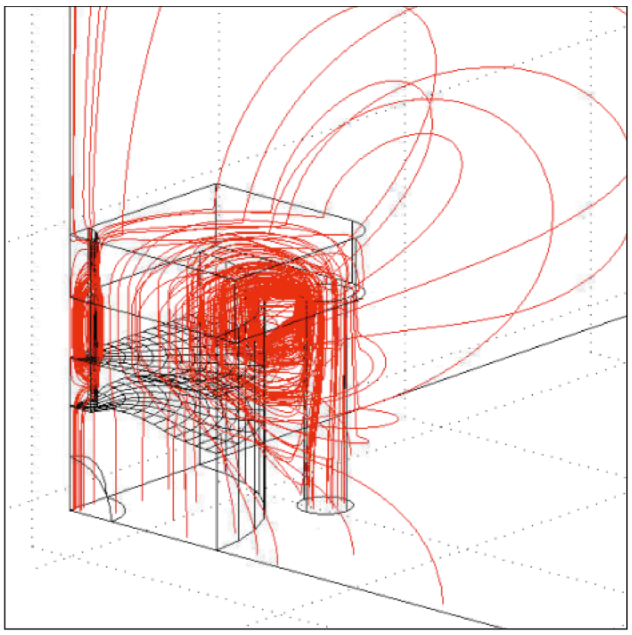


FIGURE 6.11: (a) $B_\perp$ and (b) B_z field profiles from the xz plane, taken at x = 0, 10, 20, and 50 mm. The inset plots correspond to the region along the z direction in which sensitive auxiliary equipment or devices may be located. The dotted lines indicate the axial extent of the pole piece (I), permanent magnet pole (II), and yoke (III).

(a)



(b)

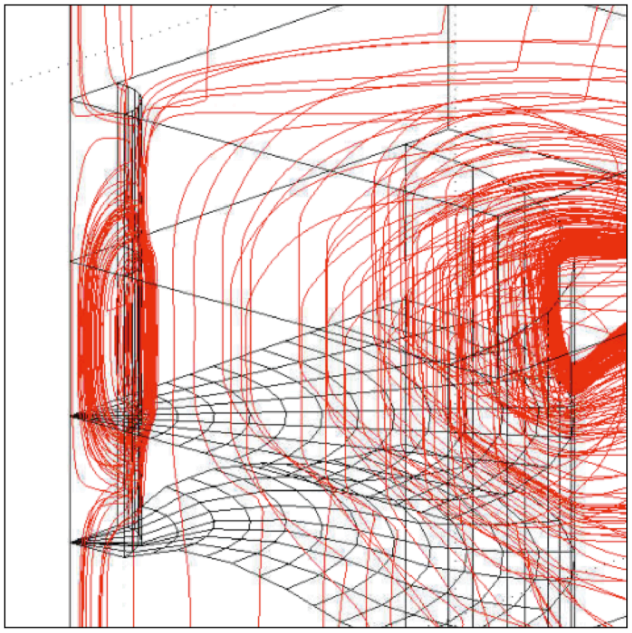


FIGURE 6.12: (a) Section of a 3D magnetic field streamline plot illustrating the reversal of the magnetic field direction within the bore of the magnet assembly with the NAGP optimized pole piece. (b) A zoomed view of the streamline plot detailing the field reversal.

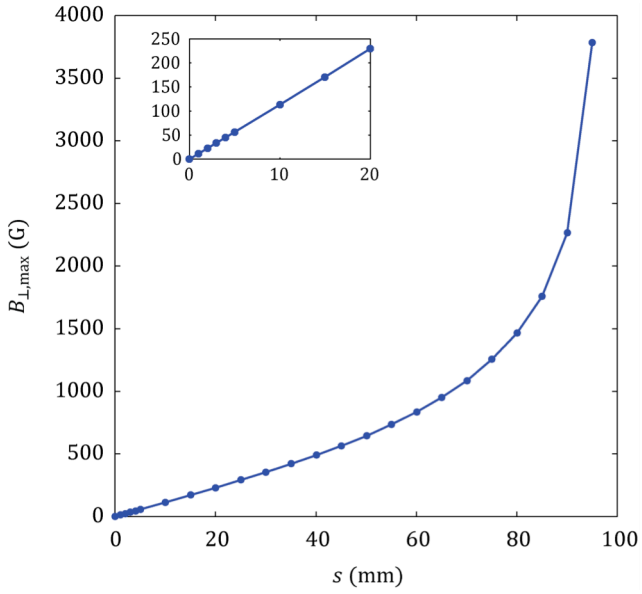


FIGURE 6.13: The peak magnitude of $B_\perp$ within the bore volume, as a function of distance s from the z axis.

which shows the peak magnitude of the transverse component $B_{\perp,\max}$ within the bore volume as a function of the off-axis distance s. Within the cylindrical region identified by $s < 40$ mm, $B_{\perp,\max}$ increases linearly. However, exterior to this region, the rate of change of $B_{\perp,\max}$ begins to increase rapidly as the magnetic field lines necessarily bend to form the closed loops directly evident in Figure 6.12 and indicated by the polarity change in the $B_\perp$ profiles.

In summary, the magnetic field within the bore is complex and highly non-uniform, as shown in Figure 6.11 and discussed above. In particular, both the axial and transverse components reverse direction in this region, with the total field magnitude peaking around 0.300 T. In comparison, the region exterior to the bore exhibits a significantly weaker field (< 50 G) with less variation, and therefore may lend itself as a better candidate for the location

of a magnetically sensitive device (such as a linac). Now armed with a complete three-dimensional vector description of the magnetic field produced by the BMA, the electron dynamics within the linac waveguide can be simulated in order to precisely evaluate its performance in proximity to this magnet system.

6.3.4 Design Sensitivity

The sensitivity analysis introduced in Chapter 5 was performed on the optimal NAGP pole piece design obtained in this work. For each population of 100 perturbed designs obtained, the statistical parameters P_{50}, P_{80}, and P_{90} were calculated for 10, 20, 30, and 40 cm DSVs at isocenter. These results are plotted as a function of the maximum variation amplitudes $\sigma_{\max}$ in Figures 6.14 and 6.15.

The field uniformity corresponding to the NAGP pole piece design is relatively stable for perturbation amplitudes of $\sigma_{\max} \leq 0.5$ mm. This is indicated by the result that over 90% of the perturbed optimal designs corresponding to $\sigma_{\max} = 0.5$ mm provided inhomogeneity values better than 250 ppm over a 30 cm DSV. Furthermore, despite exhibiting greater field inhomogeneity than the CMA over 30 and 40 cm DSVs, similar values for P_{50}, P_{80}, and P_{90} were obtained for the BMA over 10 and 20 cm DSVs. Specifically, the NAGP design for the CMA resulted in P_{90} values, corresponding to $\sigma_{\max} = 0.5$ mm, of 17 and 75 ppm over 10 and 20 cm DSVs, respectively, compared to values of 22 and 91 ppm for the BMA.

6.4 CONCLUSION

The design and optimization of a novel bored biplanar permanent magnet assembly has been presented. A unique feature of this magnet assembly is the presence of a 20 cm diameter hole that was vacated from the yoke and

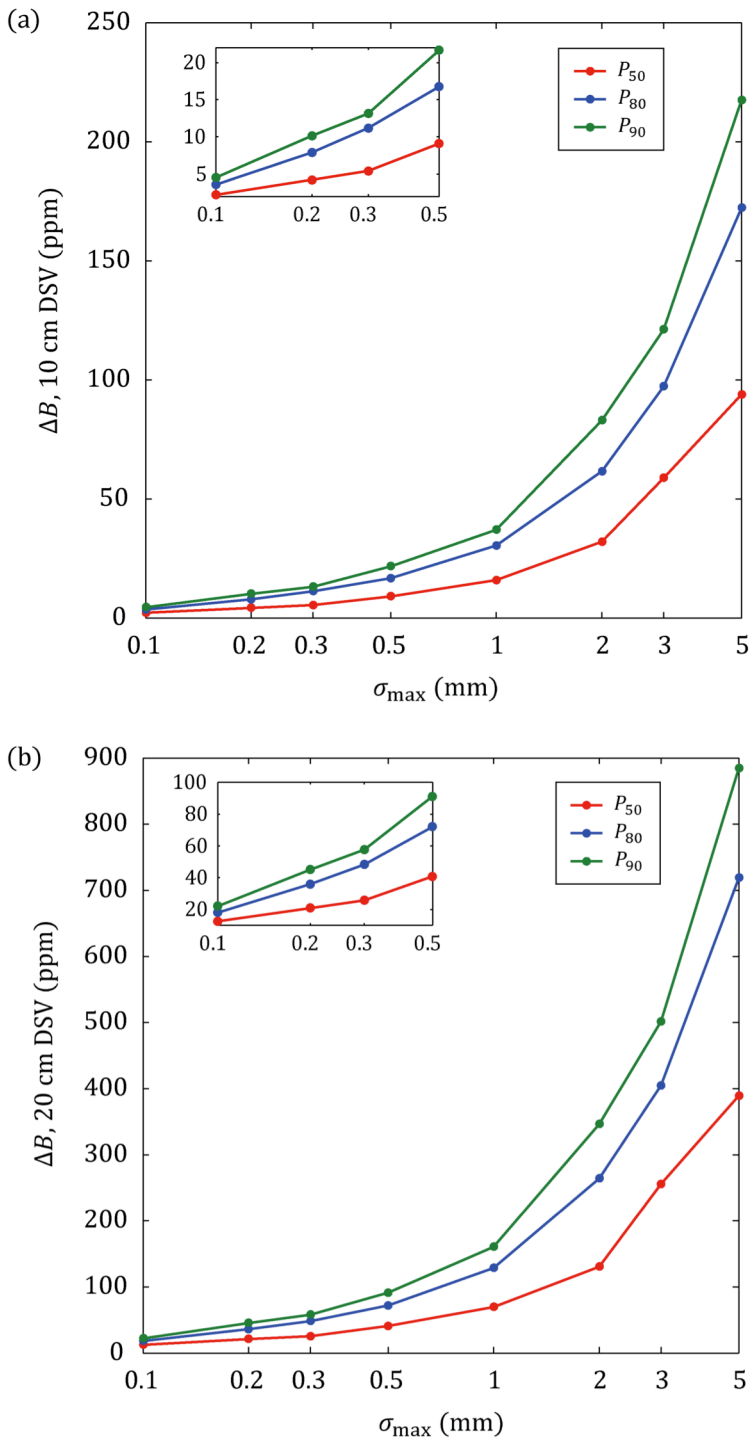


FIGURE 6.14: Plots of the statistical parameters P_{50}, P_{80}, and P_{90} as a function of the maximum variation amplitude σ_{max} for (a) 10 cm and (b) 20 cm DSVs.

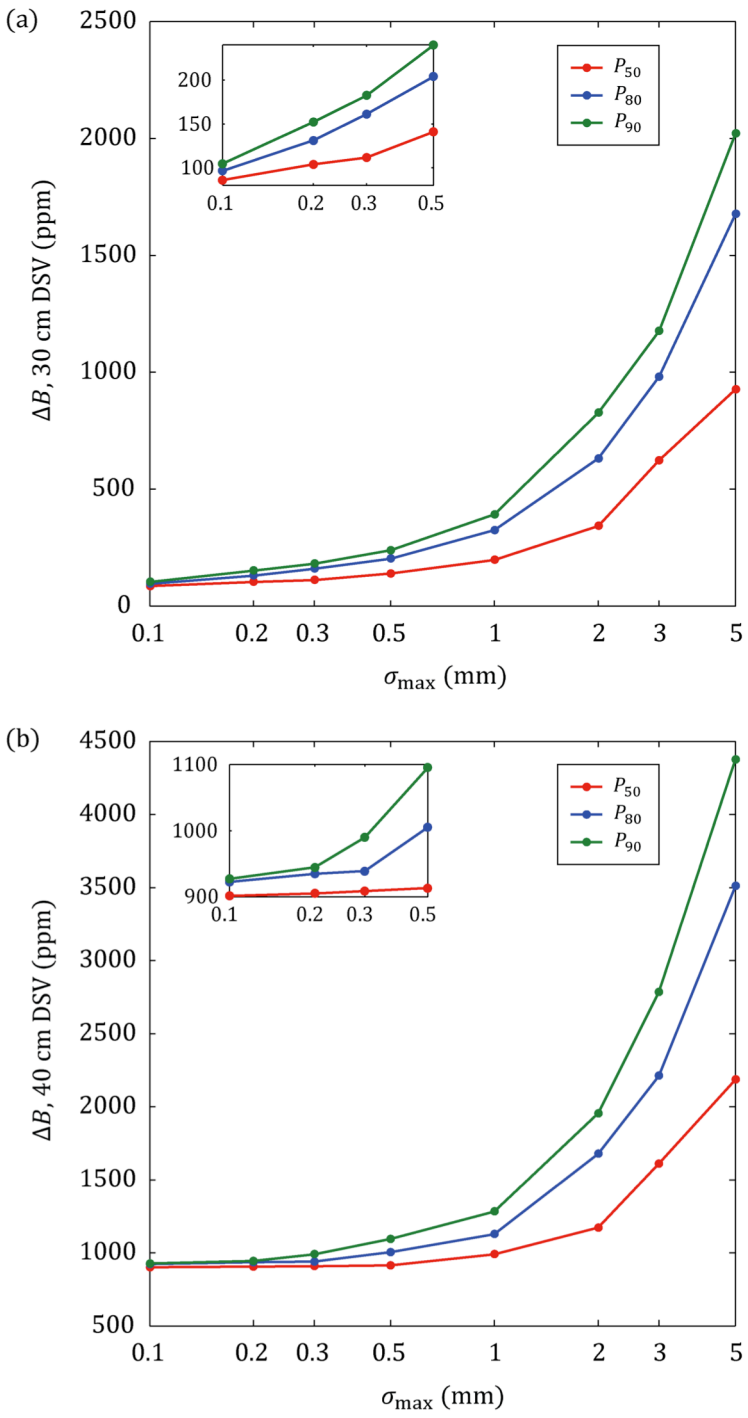


FIGURE 6.15: Plots of the statistical parameters P_{50}, P_{80}, and P_{90} as a function of the maximum variation amplitude σ_{max} for (a) 30 cm and (b) 40 cm DSVs.

pole structures to permit the inclusion of additional devices within or near the magnet structure that may benefit from being oriented along the direction of the main magnetic field. Through the use of magnetic field calculations with the FEM, it was demonstrated that the presence of the large bore degrades the magnetic field uniformity in the imaging volume. Therefore, optimized DARS, AGP, and NAGP pole pieces were required in order to permit a level of homogeneity suitable for successful passive shimming within a 30 cm DSV. Although the NAGP design resulted in the greatest field uniformity, the newly introduced DARS pole piece was notably effective, despite the simplicity of the design surface. An examination of the axial and transverse field components along the yoke bore revealed that the magnetic field reverses direction within the vacated region of the magnet assembly. Consequently, a generally non-zero transverse field component was observed, which is nulled on the central bore axis due to symmetry, but increases in peak magnitude with off-axis distance. Furthermore, the magnetic field magnitude peaks at approximately 3000 and 46 G within and exterior to the yoke bore, respectively, indicating that the bore interior may not be a suitable location for magnetically sensitive devices. Lastly, a sensitivity analysis was performed with the NAGP pole piece design, which indicated that the uniformity over a 30 cm DSV was relatively stable for perturbations with a maximum amplitude that was less than 0.5 mm. Therefore, this value represents a reasonable upper limit for the specification of manufacturing tolerances on the design surface.

6.5 REFERENCES

[1] A. Podol'skii, "Development of permanent magnet assembly for MRI devices," *IEEE Trans. Magn.*, **34**(1), 248−252 (1998).

[2] A. Trequattrini, G. Coscia, and S. Pittaluga, "Double-cavity open permanent magnet for dedicated MRI," *IEEE Trans. Appl. Supercond.* **10**(1), 756−758 (2000).

[3] A. Podol'skii, "Permanent-magnet assemblies for magnetic resonance imaging devices for various purposes," *IEEE Trans. Magn.* **38**(3), 1549−1552 (2002).

[4] Y. Yao, Y. Fang, C. S. Koh, and G. Ni, "A new design method for completely open architecture permanent magnet for MRI," *IEEE Trans. Magn.* **41**(5), 1504−1507 (2005).

[5] A. Trequattrini, S. Besio, S. Pittaluga, V. Punzo, and L. Satragno, "A novel 0.25 T dedicated MRI apparatus," *IEEE Trans. Appl. Supercond.* **16**(2), 1505−1508 (2006).

[6] Y. Zhang, D. Xie, B. Bai, H. S. Yoon, and C. S. Koh, "New development of monohedral magnet for MRI using the combination of genetic algorithm and FEM-NESM," *IEEE Trans. Appl. Supercond.* **44**(6), 1266−1269 (2008).

[7] B. G. Fallone, M. Carlone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, K. Wachowicz, and C. Kirkby, "Development of a linac-MRI system for real-time ART," *Med. Phys.* **34**(6), 2547−2547 (2007).

[8] B. G. Fallone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, and S. Vidakovic, "First MR images obtained during megavoltage photon irradiation from a prototype integrated linac-MR system," *Med. Phys.* **36**(6), 2084−2088 (2009).

[9] B. G. Fallone, "Real-time MR-guided radiotherapy: integration of a low-field MR system," *Med. Phys.* **36**(6), 2774−2775 (2009).

[10] T. Tadic and B. G. Fallone, "Three-dimensional non-axisymmetric pole piece shape optimization for biplanar permanent magnet MRI systems," *IEEE Trans. Magn.* **47**(1), 231–238 (2011).

[11] N. B. S. Gloria, M. C. L. Areiza, I. V. J. Miranda, and J. M. A. Rebello, "Development of a magnetic sensor for detection and sizing of internal pipeline corrosion defects," *NDT&E Int.* **42**(8), 669−677 (2009).

[12] M. J. Tannenbaum, A. K. Ghosh, K. E. Robins, and W. B. Sampson, "Magnetic properties of the iron laminations for CBA magnets," *IEEE Trans. Nucl. Sci.* **30**(4), 3472−3474 (1983).

[13] R. Skomski and J. M. D. Coey, *Permanent Magnetism* (Institute of Physics Publishing, Bristol, United Kingdom, 1999).

[14] COMSOL MULTIPHYSICS, ver. 3.4 (Comsol AB, Stockholm, Sweden, 2007).

[15] MATLAB, ver. 7.11 (The MathWorks, Natick, MA, 2010).

[16] A. Cauchy, "Méthodes générales pour la résolution des systèmes d'équations simultanées," *C. R. Acad. Sci.* **25**, 536−538 (1847).

[17] P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, "Computing forward-difference intervals for numerical optimization," *SIAM J. Sci. Stat. Comput.* 4(2), 310−321 (1983).

CHAPTER 7
OPTIMIZATION OF SUPERCONDUCTING MRI MAGNET SYSTEMS WITH MAGNETIC MATERIALS

A version of this chapter has been accepted for publication. T. Tadic and B. G. Fallone, "Design and optimization of superconducting MRI magnet systems with magnetic materials," *IEEE Trans. Appl. Supercond.* (2012).

7.1 INTRODUCTION

The optimal design of superconducting magnets for magnetic resonance imaging (MRI) has been the subject of considerable research. Many approaches have been reported which focus on the optimization of coil configurations to yield compact shielded magnets exhibiting a high degree of uniformity and confined fringe fields.[1–11] Despite yielding significantly different coil arrangements, the majority of the magnet designs resulting from these methods achieve similar performance.[12] Consequently, there has been a recent thrust to further develop techniques that also minimize the conductor volume.[13–25] Since the cost of a superconducting magnet system is strongly related to the amount of conductor used, this has become a growing factor of importance in the increasingly competitive market for MRI magnets.

The majority of optimization algorithms for superconducting magnets only consider designs which either have a cylindrical geometry,[1, 25–35] or do not contain any magnetic materials,[2–11, 13–24] thereby permitting a simplification of the required magnetic field analysis. This was largely justified in the past by the fact that most commercial full-body magnets satisfied these assumptions.[12] Field calculations for coil-only systems can be performed

rapidly with formulas derived from the Biot-Savart law (see § 3.1.1), or by employing well known spherical harmonic expansions.[35] When considering magnetized materials with nonlinear susceptibilities, the magnetic field analysis becomes substantially more complicated. Equivalent magnetization current methods and other direct techniques utilizing analytical formulas for the magnetic fields due to magnetized rings have been proposed,[29–32] although these schemes are strictly limited to cylindrical geometries. The finite element method (FEM) has also been applied in problems involving magnetic materials due to its high accuracy and ability to handle complicated geometries. However, to reduce the number of mesh elements and the associated computation time, the majority of optimization methods employing the FEM assume simplified cylindrical geometries, permitting the use of smaller two-dimensional (2D) models.[26–28, 33, 34]

Corresponding to the current trend towards improving patient access and decreasing patient claustrophobia, there is a growing demand for compact open biplanar magnet systems.[12, 36, 47] These magnets typically incorporate non-cylindrical yoke structures, for which accurate analysis of the complete three-dimensional (3D) magnetic fields and associated inhomogeneities are required during the design process. We are therefore motivated to develop new optimization methods that can be applied to the design of these systems.

In the present work, an iterative method is disclosed for the optimal design of homogenous superconducting magnet assemblies with general non-axisymmetric magnetic yoke structures. A linear programming (LP) sub-problem is solved at each iteration of this method, in order to obtain a minimum volume coil configuration while constraining the magnetic field inhomogeneity within an arbitrary target volume. The FEM is then used to calculate the complete 3D magnetic field produced by the entire magnet system, thereby taking into account the presence of any magnetic materials. This method exhibits good performance and typically requires a small

number of iterations and FEM analyses to converge. Consequently, it is feasible to employ large 3D FEM models that may require significant computation time due to high mesh densities or complicated geometries.

The effectiveness of the proposed method was demonstrated with the optimal design of an open and compact 0.5 T biplanar superconducting magnet system with a four-column yoke structure. A special feature of this magnet design is a coil assembly constructed from MgB_2 high-temperature superconducting (HTS) material, intended for use in a conduction-cooled cryogen-free environment. Due to the absence of a massive cryostat vessel and safety ventilation system, rotation of this magnet structure is feasible. Hence, this magnet design is suitable for integration with a medical linear accelerator (linac) in a perpendicular configuration.[38–40]

7.2 METHODS

7.2.1 Nonlinear Optimization

7.2.1.1 Algorithm Overview

The ultimate goal of the method presented in this work is to obtain minimum volume coil configurations subject to magnetic field constraints for superconducting magnet systems involving non-cylindrical ferromagnetic yoke structures. This was achieved through the iterative optimization scheme summarized in the flow diagram of Figure 7.1. The steps encountered during the optimization procedure are outlined below.

The design process begins with a complete specification of the magnetic yoke structure and pole pieces, for which the geometry and properties of the materials involved are held constant for all subsequent steps. The feasible coil domain is then specified, which defines the permissible region to be

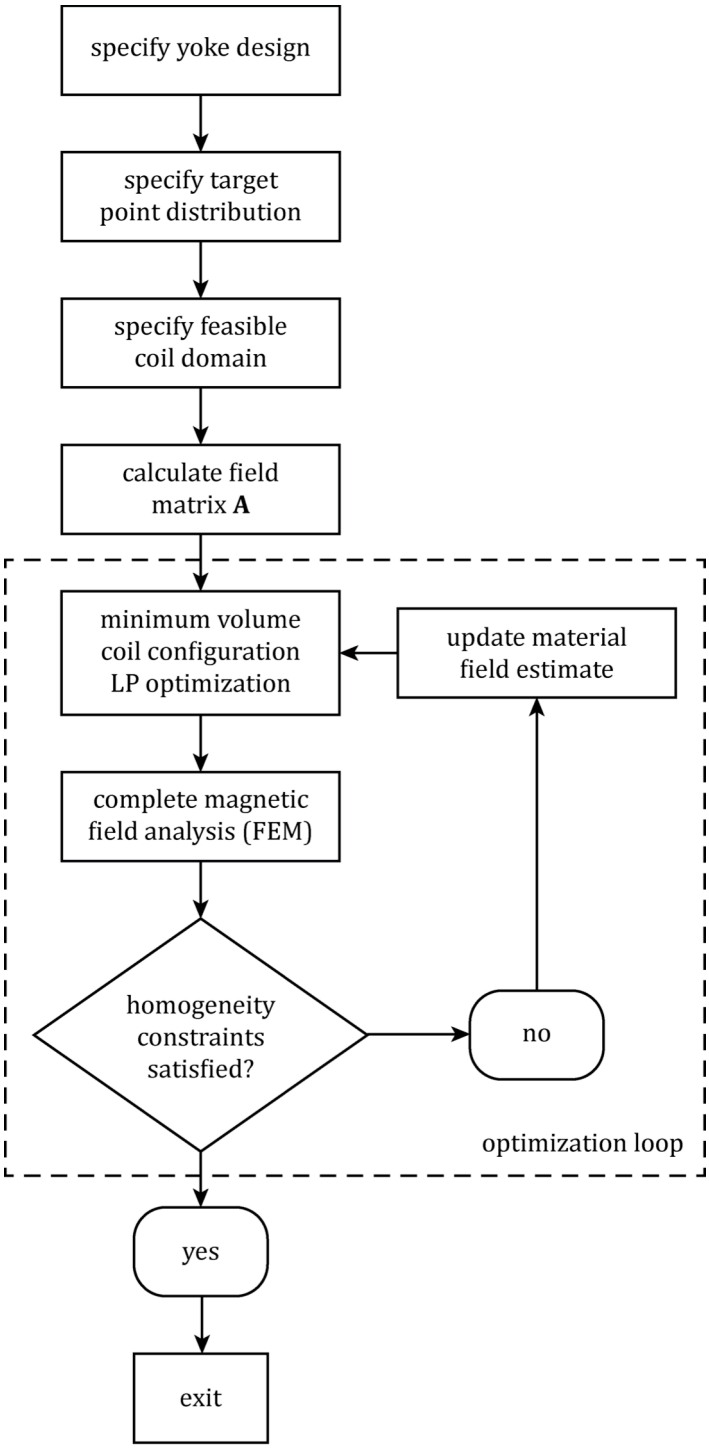


FIGURE 7.1: Iterative optimization flow diagram.

occupied by the coil configuration. The precise role of the feasible coil domain in this process is detailed in the following section. A distribution of N_t target points are specified next, at which locations the axial magnetic field uniformity is to be constrained. As such, the target points are typically chosen on the surface of a large diameter spherical volume (DSV) that is designated for imaging.

Progression through the remaining steps of the optimization process is best understood through an analysis of the constraint expressions, as it is from these constraints that the optimal coil configuration is developed. By denoting the location of the ith target point as $\boldsymbol{r}_i$, the preliminary field uniformity constraints are given as

$$|\Delta B_z(\boldsymbol{r}_i)| \leq \kappa_1', \quad \text{for } i = 1, \dots, N_t, \tag{7.1}$$

where the parts per million (ppm) field variation $\Delta B_z(\boldsymbol{r})$ at the point $\boldsymbol{r}$ is defined by

$$\Delta B_z(\boldsymbol{r}) = \frac{B_z(\boldsymbol{r}) - B_0}{B_0} \cdot 10^6. \tag{7.2}$$

In these expressions, B_z refers to the axial component of the magnetic field, B_0 is the desired field strength at isocenter, and κ_1' is the relative field error tolerance specified in units of ppm.

The radial component of the magnetic field can be considered negligible when the axial component is homogenous,[14] and therefore it is not included in the constraint formulation above (see § 2.3.1). Furthermore, it is assumed with this design method that the yoke structure itself provides adequate reduction of the fringe magnetic fields produced by the magnet assembly, as

typically found with modern passively shielded biplanar MRI systems.[12] Consequently, active shielding constraints are not employed when optimizing the coil configuration.

By introducing the notation $b_{zi} = B_z(\boldsymbol{r}_i)$, the field uniformity constraints in Eq. (7.1) may be rewritten as

$$|b_{zi} - B_0| \leq \kappa_1 B_0, \quad \text{for } i = 1, \dots, N_t, \tag{7.3}$$

where $\kappa_1 = \kappa_1' \cdot 10^{-6}$. For convenience, b_{zi} can be separated into the two components:

$$b_{zi} = b_{si} + b_{mi}, \tag{7.4}$$

where the coil field b_{si} is the contribution strictly due to the free current configuration, and the material field b_{mi} is the contribution due to the nonlinear magnetization induced in the magnetic materials. Thus, it is important to note that although b_{si} can be calculated with knowledge of only the current configuration, the calculation of b_{mi} is more complicated as it requires knowledge of the entire system configuration including the distribution of both current and magnetic material distributions. A target coil field b_{ti} is then defined as

$$b_{ti} = B_0 - b_{mi}, \tag{7.5}$$

and the constraints in Eq. (7.3) are rewritten to give

$$|b_{si} - b_{ti}| \leq \kappa_1 B_0, \quad \text{for } i = 1, \dots, N_t. \tag{7.6}$$

Prior to iterating through the optimization loop, it is assumed that a coil configuration is not initially defined. Thus, in the absence of any magnetic source, the magnetic field is zero at all points:

$$b_{si}^{(0)} = b_{mi}^{(0)} = b_{zi}^{(0)} = 0, \tag{7.7}$$

where the superscript denotes the iteration number. Therefore, since the constraints in Eq. (7.6) are not initially satisfied, the iterative optimization loop is engaged.

During the first step of each iteration, an LP method adapted from the literature[14] is used to determine a minimum volume coil arrangement such that Eq. (7.6) is satisfied. The details of this procedure are discussed in the following section. Since the true material field is unknown at this stage, the target coil field in Eq. (7.6) is calculated by using the material field from the preceding iteration as an estimate.[27, 28] In other words, the LP method is terminated at the kth iteration when a coil configuration is obtained that satisfies

$$\left| b_{si}^{(k)} - b_{ti}^{(k-1)} \right| \leq \kappa_1 B_0, \quad \text{for } i = 1, \ldots, N_t, \tag{7.8}$$

where the target coil field is given by

$$b_{ti}^{(k-1)} = B_0 - b_{mi}^{(k-1)}. \tag{7.9}$$

A FEM model that combines the new coil configuration with the specified yoke design is then generated and the complete axial field at each of the target points can be calculated. Since the coil field is already known from the LP step, the updated material field at the kth iteration is now easily determined:

$$b_{mi}^{(k)} = b_{zi}^{(k)} - b_{si}^{(k)}, \tag{7.10}$$

where $b_{zi}^{(k)}$ has been extracted directly from the result of the kth FEM analysis.

The true material field typically differs from the estimate used in Eqs. (7.8) and (7.9), since a change in the current configuration necessitates a change in the magnetization of the magnetic materials. Therefore, the constraint in Eq. (7.6) is not trivially satisfied.

For example, the initial coil configuration calculated during the first iteration ($k = 1$) is always obtained in the absence of magnetic materials. This is because prior to entering the optimization loop ($k = 0$), the material field is assumed identically zero in accordance with Eq. (7.7). Hence, the magnetic materials are not accounted for in the calculation of the constraints in Eq. (7.8). However, after an initial coil configuration is established, the magnetic materials will become magnetized, and the true material field will no longer be negligible. If the true material field due to the initial coil configuration is then determined (with the FEM), it is almost certain that the constraint in (7.3) will no longer be satisfied.

Successful evolution through the optimization loop relies on the assumption that the difference in the material field between successive iterations tends to progressively decrease:

$$\left|\boldsymbol{b}_{\boldsymbol{m}}^{(k)} - \boldsymbol{b}_{\boldsymbol{m}}^{(k-1)}\right| \leq \left|\boldsymbol{b}_{\boldsymbol{m}}^{(k-1)} - \boldsymbol{b}_{\boldsymbol{m}}^{(k-2)}\right|, \text{ for } k \geq 2, \tag{7.11}$$

where $\boldsymbol{b}_{\boldsymbol{m}}$ is the material field vector with the N_t elements b_{mi}. Thus, by iterating the LP and FEM steps described above, the material field estimate

$\boldsymbol{b}_{\boldsymbol{m}}^{(k-1)}$ at the kth iteration is assumed to eventually provide a close approximation to the true material field $\boldsymbol{b}_{\boldsymbol{m}}^{(k)}$. Or equivalently, it is assumed that $b_{ti}^{(k-1)} \approx b_{ti}^{(k)}$, for sufficiently large k.

In practice, it is beneficial to relax the termination criteria in Eq. (7.6) so that an excessive number of iterations are avoided. Therefore, the optimization loop is terminated when

$$\left|b_{si}^{(k)} - b_{ti}^{(k)}\right| \leq \kappa_1 \kappa_2 B_0, \quad \text{for } i = 1, \dots, N_t. \tag{7.12}$$

where $\kappa_2 \geq 1$ is the specified convergence factor, and hence

$$\left|b_{ti}^{(k)} - b_{ti}^{(k-1)}\right| \leq \kappa_1 (1 + \kappa_2) B_0, \quad \text{for } i = 1, \dots, N_t. \tag{7.13}$$

In terms of the quantities used to define the preliminary constraints in Eq. (7.1), these constraints can be stated as

$$|\Delta B_z(\boldsymbol{r}_i)| \leq \kappa_1' \kappa_2, \quad \text{for } i = 1, \dots, N_t. \tag{7.14}$$

According to Eq. (7.12), the expected final peak-to-peak field variation ΔB_z at the target points is bounded above by $2\kappa_1 \kappa_2 B_0$ (in absolute units), for which many combinations of the parameters κ_1 and κ_2 will yield the same value. However, the particular choice for each of these parameters has additional consequences. First, the parameter κ_1 plays the critical role of constraining the field variation in Eq. (7.8) during the LP step. As will be discussed in the following section, the resulting number of coils in the configuration determined at this step is strongly related to the extent that the field uniformity is constrained. Thus, choosing too strict of a value for κ_1 can lead to an unnecessarily complicated final optimal coil configuration. Second, the

parameter κ_2 plays the practical role of limiting the total number of iterations in the optimization loop. Thus, if this parameter is chosen is too close to unity, then the speed of the optimization algorithm will suffer, as an unnecessarily large number of computationally expensive FEM simulations will be performed.

The optimization algorithm described above was implemented in the MATLAB[24] scripting environment. The commercially available 3D FEM software package[42] OPERA-3D was used to calculate the complete magnetic fields produced by the entire magnet assembly. In order to integrate the FEM analysis into the optimization script, a Component Object Model (COM) interface was established between the MATLAB and OPERA-3D application processes. In this way, OPERA-3D could be controlled from input commands executed within the MATLAB script. This permitted the automatic and iterative generation and analysis of the FEM models derived from the optimization variables. Details of the FEM modeling procedure are provided in § 7.2.3.

7.2.1.2 Linear Programming Formulation

At each iteration of the proposed method, the problem of determining a minimum volume coil configuration that satisfies the field uniformity constraint in Eq. (7.8) is encountered. The details of the LP method used to achieve this are disclosed below.

The specified feasible coil domain introduced in the preceding section is first segmented with a dense grid that defines an array of n candidate coils, each of which is located at the center of a grid cell (Figure 7.2). These grid cells do not represent physical cross-sections of the candidate coils, but rather only serve to define their spacing and location. The $N_t \times n$ axial magnetic field matrix **A** can then be introduced, for which the element A_{ij} is the field per

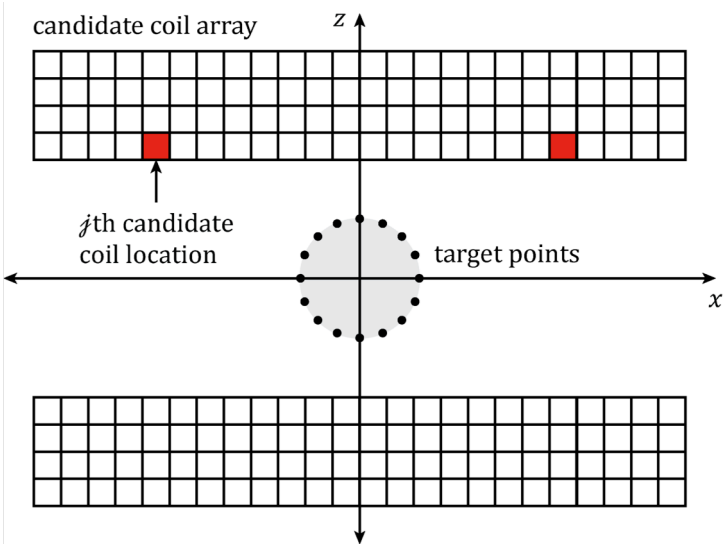


FIGURE 7.2: Illustrated section of a feasible coil domain and target point distribution. The feasible coil domain is segmented by an array of candidate coils defined on a grid, shown here coaxial with the z axis. The grid cell corresponding to the location of the jth candidate coil is highlighted red. The target points at which locations the field inhomogeneity is constrained are typically distributed on the surface of a large DSV designated for imaging.

unit current at the ith target point due to the jth candidate coil. This permits the matrix expression:

$$\mathbf{A}\boldsymbol{p} = \boldsymbol{b}_s, \tag{7.15}$$

where $\boldsymbol{b}_s$ is the vector of axial coil magnetic fields at the N_t target points, and $\boldsymbol{p}$ is the vector of currents, such that the element p_j (for $j = 1, \dots, n$) is the current passing through the jth candidate coil.

The total conductor volume $\mathcal{V}_c$ in the array of n circular coils is given by

$$\mathcal{V}_c = \sum_{j=1}^{n} 2\pi s_j a_j, \tag{7.16}$$

where s_j and a_j are the radius and cross-section for the jth coil, respectively. Assuming a constant current density J in all of the candidate coils, the individual currents satisfy

$$|p_j| = J a_j. \tag{7.17}$$

Hence, the total conductor volume can be rewritten as

$$\mathcal{V}_c = \frac{2\pi}{J} \sum_{j=1}^{n} s_j |p_j|. \tag{7.18}$$

By combining Eqs. (7.8), (7.15), and (7.18), the coil volume minimization problem at the kth iteration may be expressed as

$$\text{minimize } f(\boldsymbol{p}) = \sum_{j=1}^{n} s_j |p_j|, \tag{7.19a}$$

$$\text{subject to } -\kappa_1 B_0 \mathbf{1}_{N_t \times 1} \leq \mathbf{A}\boldsymbol{p} - \boldsymbol{b}_{\boldsymbol{t}}^{(k-1)} \leq \kappa_1 B_0 \mathbf{1}_{N_t \times 1}, \tag{7.19b}$$

where $\mathbf{1}_{u \times v}$ denotes a $u \times v$ matrix with all entries equal to 1. In this form, these equations constitute an L_1-norm minimization problem. Through the use of an algebraic transformation developed below, the solution to Eq. (7.19) can be obtained via an equivalent LP problem, for which robust numerical methods are readily available (see § 4.3.2).

First, the auxiliary vector $\boldsymbol{t}$ is introduced, consisting of n elements denoted t_j (for $j = 1, \ldots, n$). In particular, these elements are required to satisfy the constraint

$$0 \leq |p_j| \leq t_j, \quad \text{for } j = 1, \ldots, n. \tag{7.20}$$

Next, the augmented constraint matrix $\overline{\mathbf{A}}$, constraint vector $\overline{\boldsymbol{b}}$, cost vector $\overline{\boldsymbol{\varsigma}}$, and design vector $\overline{\boldsymbol{p}}$ are defined as

$$\overline{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathbf{0}_{N_t \times n} \\ -\mathbf{A} & \mathbf{0}_{N_t \times n} \\ -\mathbf{I}_n & -\mathbf{I}_n \\ \mathbf{I}_n & -\mathbf{I}_n \end{bmatrix}, \tag{7.21a}$$

$$\overline{\boldsymbol{b}} = \begin{bmatrix} \boldsymbol{b}_{\boldsymbol{t}}^{(k-1)} + \kappa_1 B_0 \mathbf{1}_{N_t \times 1} \\ -\boldsymbol{b}_{\boldsymbol{t}}^{(k-1)} + \kappa_1 B_0 \mathbf{1}_{N_t \times 1} \\ \mathbf{0}_{N_t \times 1} \\ \mathbf{0}_{N_t \times 1} \end{bmatrix} = \begin{bmatrix} B_0(\kappa_1 + 1)\mathbf{1}_{N_t \times 1} - \boldsymbol{b}_{\boldsymbol{m}}^{(k-1)} \\ B_0(\kappa_1 - 1)\mathbf{1}_{N_t \times 1} + \boldsymbol{b}_{\boldsymbol{m}}^{(k-1)} \\ \mathbf{0}_{N_t \times 1} \\ \mathbf{0}_{N_t \times 1} \end{bmatrix}, \tag{7.21b}$$

$$\overline{\boldsymbol{\varsigma}} = \begin{bmatrix} \mathbf{0}_{n \times 1} \\ \boldsymbol{s} \end{bmatrix}, \tag{7.21c}$$

$$\overline{\boldsymbol{p}} = \begin{bmatrix} \boldsymbol{p} \\ \boldsymbol{t} \end{bmatrix}. \tag{7.21d}$$

In these expressions, $\mathbf{0}_{u \times v}$ denotes a $u \times v$ matrix with all entries equal to 0, and $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. An augmented LP problem comprised of the quantities above can then be stated:

$$\text{minimize } \bar{f}(\overline{\boldsymbol{p}}) = \overline{\boldsymbol{\varsigma}}^T \overline{\boldsymbol{p}}, \tag{7.22a}$$

$$\text{subject to } \overline{\mathbf{A}}\overline{\boldsymbol{p}} \leq \overline{\boldsymbol{b}}. \tag{7.22b}$$

Since the only constraints on the elements of $\boldsymbol{t}$ are given by Eq. (7.20), the solution $\overline{\boldsymbol{p}}^{*T} = [\,\boldsymbol{p}^{*T} \quad \boldsymbol{t}^{*T}\,]$ to Eq. (7.22) implies the equality

$$\left|p_j^*\right| = t_j^*, \quad \text{for } j = 1, \ldots, n. \tag{7.23}$$

Hence, $\boldsymbol{p}^*$ must minimize f, in accordance with

$$f(\boldsymbol{p}^*) = \sum_{j=1}^{n} s_j \left|p_j^*\right| = \sum_{j=1}^{n} s_j t_j^* = \bar{f}(\overline{\boldsymbol{p}}^*). \tag{7.24}$$

Since the first $2N_t$ rows of Eq. (7.22b) trivially satisfy the constraint in Eq. (7.19b), it can be concluded that $\boldsymbol{p}^*$ solves the L_1-norm minimization problem. Therefore, the equivalency of the problems in Eqs. (4.21) and (7.19) has been established.

A beneficial feature of L_1-norm minimization problems is that the resulting solution is sparse, in that the minimum number of non-zero elements of $\boldsymbol{p}$ required to satisfy the constraints are obtained. In the formulation presented here, this means that the solution to Eq. (7.22) will correspond to the minimum number of candidate coils (with non-zero currents) that are needed to produce a magnetic field with the desired level of uniformity. Furthermore, since Eq. (7.22) constitutes an LP problem, the solution obtained is guaranteed to be globally optimal (see § 4.3.2). In the current implementation, Eq. (7.22) was solved with the built-in function `linprog` from MATLAB, which employs an interior-point predictor-corrector algorithm based on the LIPSOL method.[43]

The solution to Eq. (7.22) ultimately represents the optimal coil configuration that corresponds to the particular discretization specified for the feasible coil domain. Thus, improved accuracy in the solution can always

obtained by specifying a denser coil grid. However, increasing the grid resolution is matched with an increase in the number of unique degrees of freedom n in the LP problem, which is directly related to the computational effort. Moreover, the field matrix **A** is inherently ill-conditioned. Increasing the number of candidate coils further spoils this property, thereby hampering the ability of the LP method to arrive at a true solution to Eq. (7.22).

In an effort to achieve a balance between solution accuracy and computational efficiency, the LP step of the optimization loop was split into three intermediate stages characterized by successively finer grid resolutions. During the first intermediate stage, a relatively coarse coil grid spacing of $1-3$ cm is typically defined. When performing the magnetic field calculations required in assembling **A** at this stage, each of the candidate coils are approximated by an ideal current loop with zero cross-sectional area. This permits the use of simple axial field formulae involving complete elliptic integrals of the first and second kind,[44] which can be computed rapidly using the `ellipke` routine from MATLAB (see § 3.2.1). For the second and third intermediate stages, the grid cells corresponding to the coils with non-zero currents from the preceding stage are identified. Only these grid cells are subsequently divided into finer coil sub-grids, with a coil spacing that is reduced by the integer grid reduction factor γ_r. When assembling the field matrix **A** during the second and third stages, the candidate coils are modeled with finite rectangular cross-sections[45] (see § 3.2.2). The required cross-sectional areas (and dimensions) of the coils are determined from the optimal currents from the preceding stage, together with an aspect ratio γ_a specified prior to optimization. In this way, by choosing a typical value of $\gamma_r = 3$, a final grid resolution of approximately $1-3$ mm can be obtained.

7.2.2 Magnet Assembly Specifications

The method described in this work was applied to the optimal design of an open and compact 0.5 T biplanar superconducting magnet assembly for potential use in a linac-MRI system with a perpendicular configuration.[38–40] Accordingly, the outer dimensions of the four-column yoke structure illustrated in Figure 7.3 were chosen to permit rotation of the magnet assembly within a typical radiotherapy vault of 3.50 m height. This results in a practical vertical clearance of 90 cm, allowing sufficient room for additional peripheral equipment.[36, 47] The remaining dimensions of this assembly were selected such that the total weight of magnetic material was within 10% of the 20 tonne yoke structure used in a commercially available 0.5 T biplanar system with an acceptable fringe field.[46] In addition, the dimensions of the yoke and pole structures were chosen to provide adequate space for the coil configuration and associated patient gap, while also providing an approximate balance between field strength enhancement at isocenter, passive shielding of fringe fields, and mechanical stability of the rotating magnet assembly.

Similar to previous chapters, the yoke and column structures of the magnet assembly are composed of AISI 1020 plain carbon steel and the pole plates are composed of a special Armco magnetic steel. The nonlinear magnetization curves[22, 23] for these materials were provided in Figure 5.2.

The design proposed in this work is intended to employ a cryogen-free coil configuration constructed from MgB_2 HTS material. In the absence of any liquid coolant, the requirement of a bulky cryostat vessel comprised of a safety ventilation system can be eliminated, thereby making mechanical rotation of the entire scanner feasible. A practical coil operating temperature of 12 K could be realized through use of a conduction-cooled split cryostat design similar to that detailed in the literature.[46, 49, 50] Within a vacuum

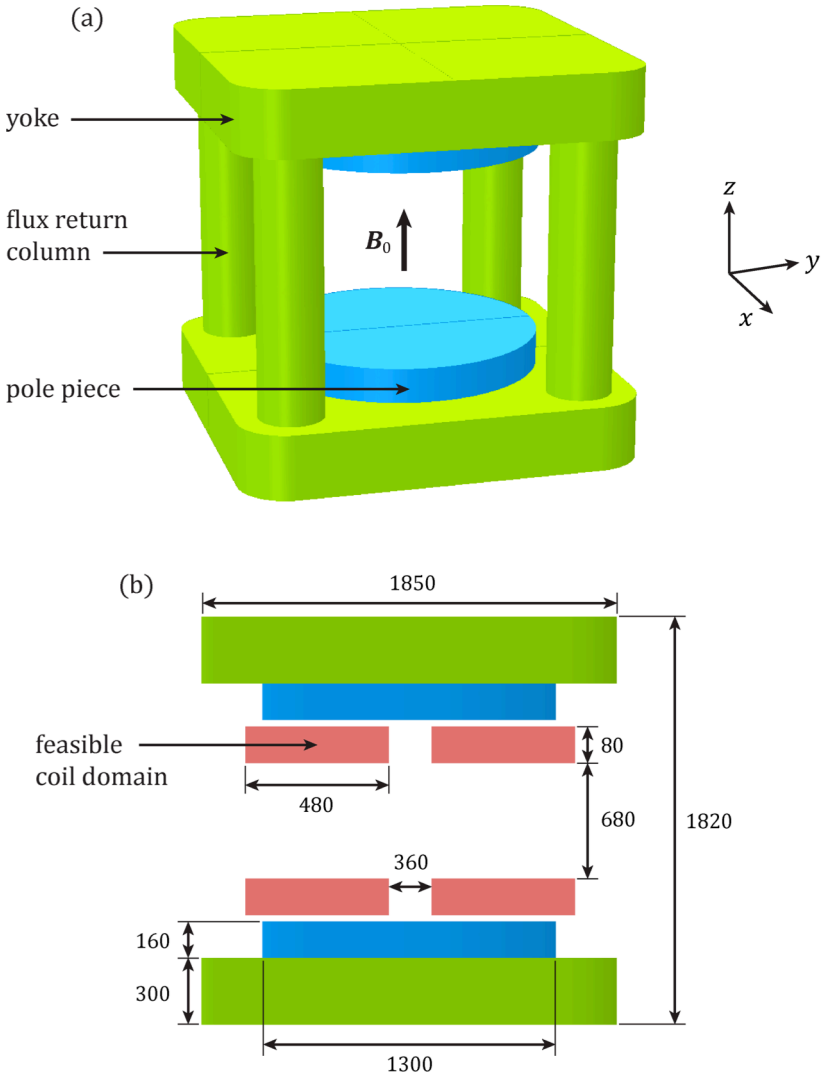


FIGURE 7.3: (a) A 3D schematic of the four-column magnetic yoke structure with the relative orientation of the Cartesian coordinate axes shown. (b) Cross-sectional view with dimensions in millimeters.

chamber encompassing each of the magnet poles, a thermal screen surrounding the HTS coil arrangement could be thermally coupled to the first stage of a Gifford-McMahon (GM) cryocooler subassembly. The second stage of this cryocooler could then be coupled to a set of thermally conductive rings and sheets embedded within, and bonded adjacent to, the coil forms

and windings. By employing superconducting joints recently developed for MgB_2 tapes,[51] a persistent mode of operation could be permitted. It is expected that a cryostat design of this type, including the associated support structures, would only occupy an additional $4-6$ cm of space surrounding the coils, resulting in a nominal pole separation of at least 60 cm. As previously noted in Chapter 5, the presence of pancake gradient coils would further reduce the patient gap by approximately $6-10$ cm.

A realistic target for the engineering critical current density[24] for MgB_2 coils is 28 kA/cm^2 at 12 K and 4 T. Hence, a working current density of $J = 21$ kA/cm^2 was specified at approximately 75% of the critical value.

The feasible coil domain was defined in the region surrounding the pole plates such that adequate space was supplied for the cooling and mounting system described above, while providing a patient gap of approximately 60 cm. The location of the feasible coil domain is illustrated in Figure 7.3(b). An initial array of $n = 24 \times 4 = 96$ circular candidate coils with a grid spacing of 2 cm was specified for each magnet pole. This grid spacing was further reduced to 2.2 mm during the LP refinement process, through specification of the grid reduction parameter $\gamma_r = 3$ and coil aspect ratio $\gamma_a = 1.2$ (see § 7.2.1.2). Complete solution of the LP step at each iteration requires less than two minutes of computation time on a 3.0 GHz Intel Xeon quad-core PC workstation with 16 GB of RAM.

Since the coil configuration in this particular magnet design is completely axisymmetric, only axisymmetric field inhomogeneities can be constrained. Accordingly, a distribution of 34 target points was defined along a single polar arc, extending from $\theta = 0$ to $\theta = \pi/2$, on the surface of a 40 cm DSV at isocenter, as illustrated in Figure 7.4.

The target field strength $B_0 = 0.5$ T was specified for the optimization, in addition to a relative field error $\kappa_1 = 5\ \mu$T, and a convergence factor $\kappa_2 = 2$. This resulted in an expected peak-to-peak axial field variation of at most 40 ppm at the target points.

7.2.3 The Finite Element Method

7.2.3.1 Problem Formulation

The details of the FEM implementation employed by OPERA-3D have been provided in § 3.3.6, for which a brief outline is provided here. The FEM problem formulation begins with Maxwell's equation for the divergence of the magnetic field vector $\boldsymbol{B}$:

$$\nabla \cdot \boldsymbol{B} = 0. \tag{7.25}$$

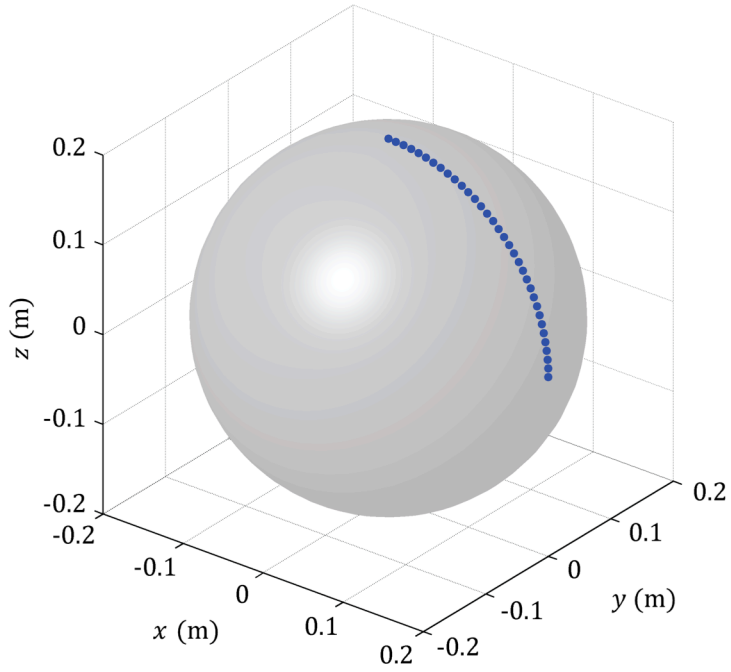


FIGURE 7.4: Target point distribution on the surface of a 40 cm DSV at isocenter.

In a region with magnetic permeability μ and zero remanent magnetization, $\boldsymbol{B}$ is related to the auxiliary field $\boldsymbol{H}$ through the constitutive relationship

$$\boldsymbol{B} = \mu \boldsymbol{H}. \tag{7.26}$$

For convenience, the spatial domain of the problem can be divided into a collection of smaller regions, each belonging to one of two types. In regions of the first type, free currents are prohibited. Hence, $\boldsymbol{H}$ can be expressed in terms of the scalar magnetic potential Φ_m:

$$\boldsymbol{H} = -\nabla \Phi_m. \tag{7.27}$$

In regions of the second type, free currents are permitted and μ is assumed to be spatially uniform. Thus, $\boldsymbol{H}$ can be separated into the two components

$$\boldsymbol{H} = \boldsymbol{H_s} + \boldsymbol{H_m}, \tag{7.28}$$

where $\boldsymbol{H_s}$ is the solenoidal field due to the free current distribution only, and $\boldsymbol{H_m}$ is the remaining irrotational component due to any magnetic materials. In this way, $\boldsymbol{H_s}$ is given by the Biot-Savart integral in Eq. (3.4), and $\boldsymbol{H}_m$ can be expressed in terms of the scalar magnetic potential Φ_m as before:

$$\boldsymbol{H_m} = -\nabla \Phi_m. \tag{7.29}$$

By combining Eqs. (7.25) and (7.26) with either of the expressions for $\boldsymbol{H}$ above, the governing second-order partial differential equation (PDE) can be obtained:

$$-\nabla \cdot (\mu \nabla \Phi_m) = 0. \tag{7.30}$$

The magnetostatics solver TOSCA in OPERA-3D can be used to numerically solve Eq. (7.30) for Φ_m at discrete points within the model geometry, subject to appropriate boundary conditions. In doing so, the L_2-norm of the total FEM residual is reduced to a value below a specified tolerance of 10^{-8}.

7.2.3.2 Model Description

The FEM model constructed in the present work consists of the biplanar magnet assembly positioned at the center of a large cubic global model volume (GMV). As noted in Chapter 5, the GMV is required in order to simulate the far field behaviour of the magnetic field within the vacated region surrounding the magnet structure. However, due to limited computational resources, the GMV must be truncated at a finite distance to limit the number of degrees of freedom in the FEM analysis. The B_0 strength of the magnet considered in the present work is roughly 2.5 times that of the permanent magnet assemblies previously considered. Since the stray magnetic fields are expected to approximately scale with B_0, the GMV was appropriately truncated at a larger distance of 15 m. Through a procedure comprising progressive enlargement of the GMV, it was estimated that this truncation distance resulted in an acceptable error in the magnetic field solution of at most 2 μT (approximately 4 ppm) within a 40 cm DSV at isocenter. Furthermore, through application of the same boundary conditions described in the preceding chapters regarding symmetries in the magnet structure, the FEM model was reduced to just one eighth of the complete geometry.

The simplified FEM model was partitioned into approximately 1.017×10^6 isoparametric tetrahedral quadratic Lagrange elements (see § 3.3.2). This mesh was automatically generated with a Delaunay triangulation algorithm, in accordance with a manual selection of the maximum element size within each of the model components. This resulted in a mesh-independent

solution, as indicated by a variation in the magnetic field solution within a 40 cm DSV of no more than 0.5 μT (approximately 1 ppm) as the overall nodal density in the model was doubled. In particular, a maximum mesh dimension of 4 mm was specified on the surface of a 40 cm DSV, in order to obtain an estimated error bound of 0.2 μT (approximately 0.4 ppm) on this surface. An illustration of the simplified FEM model geometry and finite element discretization is provided in Figure 7.5. Calculation of the FEM solution required approximately 50 minutes on a 3.0 GHz Intel Xeon quad-core PC workstation with 16 GB of RAM.

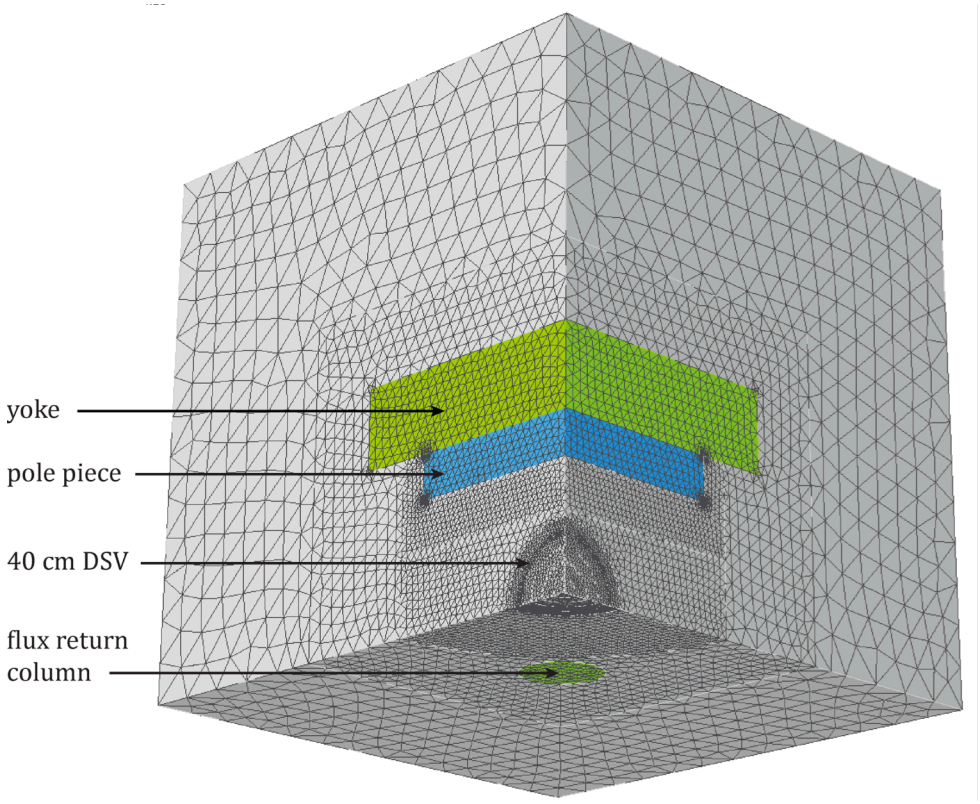


FIGURE 7.5: The simplified FEM model geometry and mesh. Conductors are not shown. The entire modeling domain that extends 15 m from isocenter is only partially shown (in grey). Due to symmetry in the magnet geometries and corresponding magnetic fields, the FEM is only applied within a fraction of the complete model geometry in an effort to reduce the computational expense.

As evident in the discussion above, a substantially more accurate field simulation was obtained here relative to previous chapters. Specifically, this was to achieve the smoothest possible field variation over the region occupied by the target points. The pole piece optimization method introduced in Chapter 5 was based on minimizing the integrated field inhomogeneity defined in Eq. (5.10). As such, local variations or oscillations in the field solution arising from the FEM discretization were not expected to have a significant impact, due to the smoothing effect of the integral operator. In contrast, the coil configuration obtained during the LP step in the scheme presented here is directly influenced by the specific material field values at each of the target points. Therefore, reducing the discretization error through an appropriately fine mesh was critical.

7.3 RESULTS AND DISCUSSION

7.3.1 Nonlinear Optimization

The optimal magnet design illustrated in Figure 7.6 was obtained in 11 iterations with the method disclosed in § 7.2.1. The total computation time required to arrive at a solution was approximately 9.9 hours. The evolution of the optimization algorithm is plotted in Figure 7.7, which illustrates a rapid reduction of the axial field inhomogeneity ΔB_z at the target points.

The optimal parameters describing the resulting coil arrangement are given in Table 7.1. The total conductor volume in the six coil pairs is $\mathcal{V}_c = 1.791 \times 10^4$ cm^3. A central magnetic field strength of B_0 =0.500 T was achieved, and a peak-to-peak axial field inhomogeneity of $\Delta B_z = 26.1$ ppm was obtained at the target points, thereby satisfying the field uniformity constraints defined in Eq. (7.12). Lastly, the peak magnetic field calculated at the coils is 3.428 T, which is well below the 4 T limit for the MgB_2 coils cited in § 7.2.2.

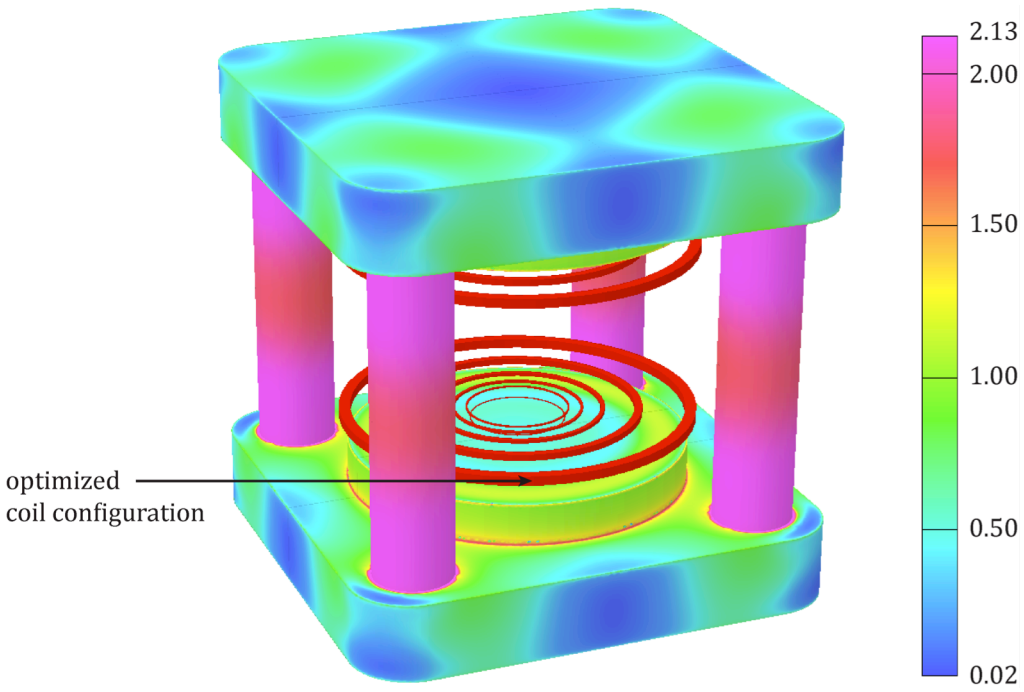


FIGURE 7.6: The complete magnet geometry and optimal coil configuration. The magnetic field solution as calculated via the FEM is displayed on the surface of the yoke structure. Values are displayed in units of T.

A total axial field inhomogeneity ΔB_z of 235 ppm was obtained within the entire 40 cm DSV, which was inherently limited by the non-axisymmetric nature of the four-column yoke structure. This can be visualized in the residual magnetic field pattern illustrated in Figure 7.8. A total field inhomogeneity ΔB of 235 ppm was also calculated, thereby verifying the assumption that the transverse field components within this volume are negligible (see § 2.3.1). The peak-to-peak inhomogeneity ΔB_z is plotted for various spherical volumes of interest in Figure 7.9. As expected, the field uniformity degrades as the volume size increases, due to the increasing proximity of the flux return columns. Nevertheless, based on the criterion introduced in Chapter 5 that successful passive-shimming within a given region of interest (ROI) requires a level of inhomogeneity no worse than

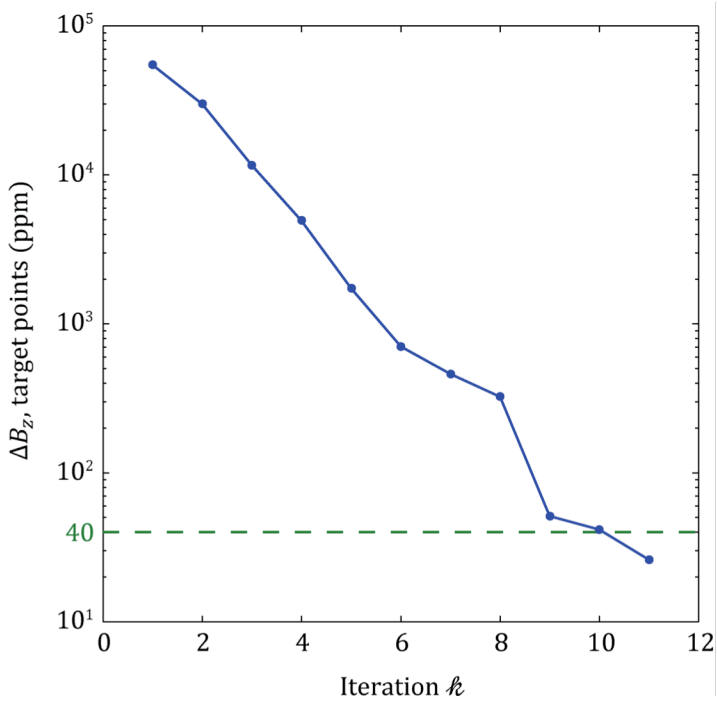


FIGURE 7.7: The evolution of the axial magnetic field inhomogeneity ΔB_z at the target points during optimization of the superconducting magnet assembly. The optimal design is achieved in 11 iterations. The dotted green line indicates the 40 ppm value corresponding to the termination criteria.

500 ppm, the results suggest that this magnet design yields a 40 cm DSV suitable for medical imaging.

As seen in Table 7.1, all but one of the coil pairs in the optimized configuration are located at the lowest attainable axial position of 341.1 mm, with the remaining coil pair possessing the smallest permissible radius. Evidently, the greatest impact per unit current, and thus per unit volume of conductor, is possible when the coils are situated on the boundaries of the feasible coil domain that are closest to the target surface. Indeed, this is the general trend observed with cylindrical magnet systems in the absence of magnetic materials,[12] for which the main coils are typically arranged nearest

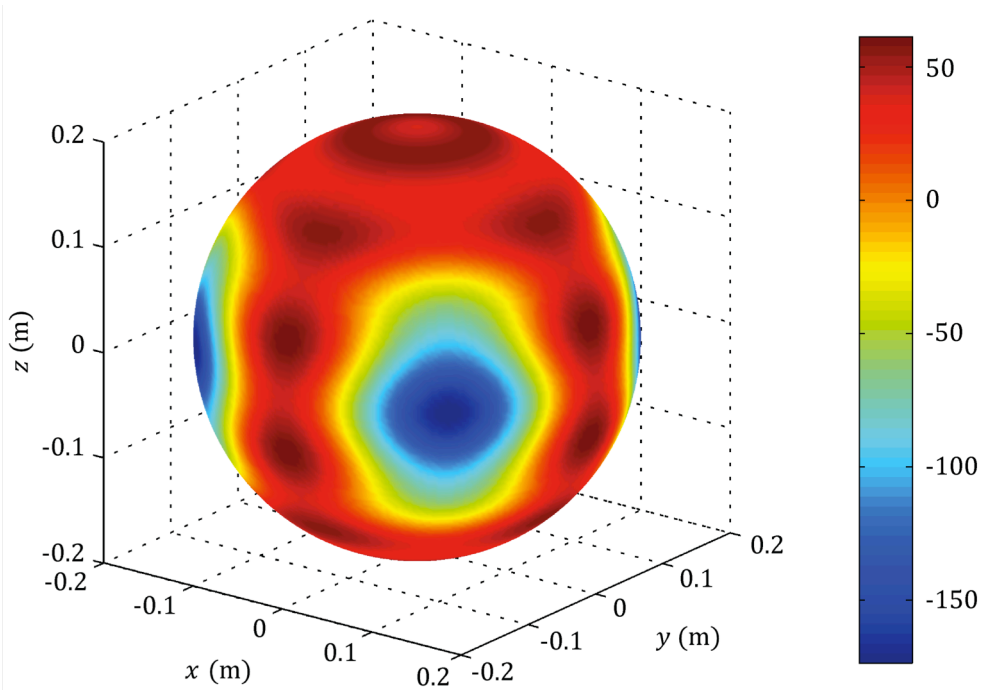


FIGURE 7.8: The residual axial magnetic field on the surface of a 40 cm DSV at isocenter. Field values are displayed in parts per million, measured relative to B_0. The limiting nature of the non-axisymmetric four column yoke structure is reflected in the magnetic field pattern shown.

the inner bore (with the exception of the active-shielding coil). This configuration is also attractive from a practical point of view, since it may simplify the design of the cryostat and support structures. Namely, it may be possible for several coils to share a common annular cooling plate, thereby leading to a reduction in the number and complexity of thermal joints and support rods within the vacuum vessel, which otherwise might be infeasible if each coil possessed a distinct location.

For completeness, the fringe magnetic fields of the optimized magnet are illustrated in Figure 7.10. The maximal extent of the 5 G field contours were calculated as $(x_{5G}, y_{5G}, z_{5G}) = (4.6, 4.6, 5.8)$ m from isocenter.

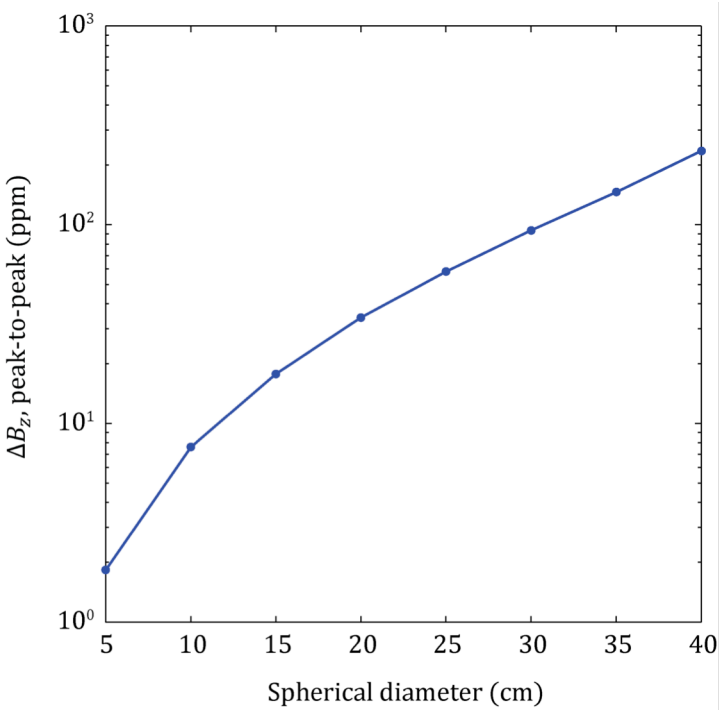


FIGURE 7.9: The residual peak-to-peak axial magnetic field inhomogeneity ΔB_z (ppm) calculated over various spherical regions of interest.

TABLE 7.1: Dimensions of the optimized coil configuration. Due to symmetry, only the dimensions for the upper magnet pole are provided.

Coil	Radial position (mm)	Axial position (mm)	Radial width (mm)	Axial width (mm)	Current (kA·turns)
1	181.1	381.1	3.7	3.1	-2.416
2	194.4	341.1	6.7	5.6	7.909
3	245.6	341.1	10.8	9.0	-20.416
4	324.9	341.1	16.4	13.6	46.888
5	461.1	341.1	29.5	24.6	-152.608
6	658.9	341.1	42.7	35.6	319.034

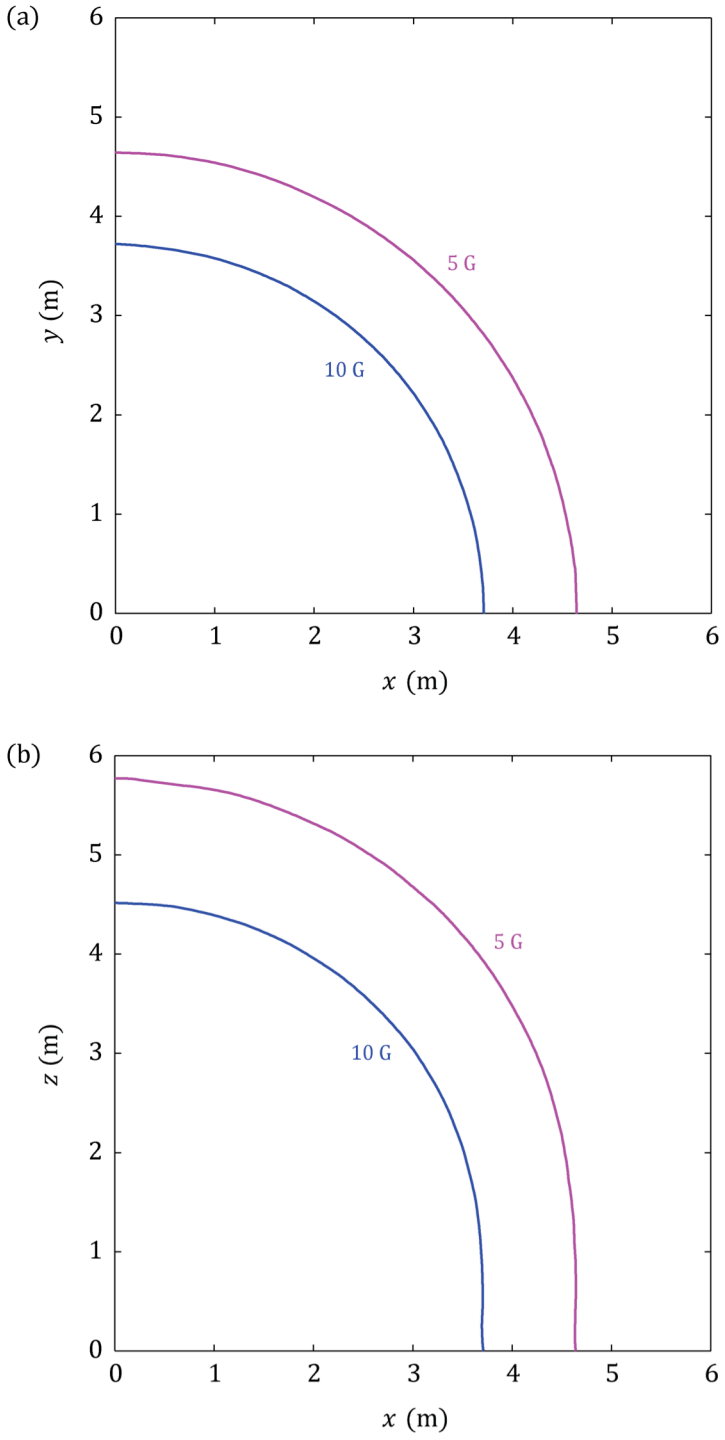


FIGURE 7.10: Plots of the 1, 5, and 10 G magnetic field contours projected onto the (a) xy and (b) xz planes for the optimized magnet. Axes dimensions are in meters.

7.4 CONCLUSION

An iterative method has been presented for the optimal design of homogenous superconducting MRI systems containing magnetic materials. In addition to minimizing the conductor volume, a key feature of this method is the ability to handle magnet designs involving complicated non-cylindrical magnetic yoke geometries, such as those found in modern passively shielded biplanar MRI scanners employing HTS coil configurations. The accurate design of these systems requires the use of computationally expensive numerical techniques for the magnetic field analysis, the burden of which is alleviated by the rapid convergence achieved with the proposed method. The effectiveness of this technique was demonstrated with the optimal design of an open and compact 0.5 T biplanar superconducting magnet system with a four-column yoke structure. This system was designed to employ a conduction-cooled cryogen-free coil configuration constructed from MgB_2 HTS material, such that rotation of the cryostat would be feasible. As such, this magnet system would be appropriate for use in an integrated linac-MRI system with a perpendicular configuration, capable of performing advanced real-time adaptive radiotherapy. An optimized solution consisting of six coil pairs was obtained in only 11 iterations with the proposed method, which resulted in a residual axial magnetic field inhomogeneity of 235.0 ppm within a 40 cm DSV.

7.5 REFERENCES

[1] A. Ishiyama, M. Hondoh, N. Ishida, and T. Onuki, "Optimal design of MRI magnets with magnetic shielding," *IEEE Trans. Magn.* **25**(2), 1885−1888 (1989).

[2] A. Ishiyama, T. Yokoi, S. Takamori, and T. Onuki, "An optimal design technique for MRI superconducting magnets using mathematical programming method," *IEEE Trans. Magn.* **24**(2), 922−925 (1988).

[3] M. Fujita, "The coil design of the superconducting MRI magnet," *IEEE Trans. Magn.* **24**(6), 2907−2909 (1988).

[4] S. Pissanetzky, "Structured coils for NMR applications," *IEEE Trans. Magn.* **28**(4), 1961−1968 (1992).

[5] M. R. Thompson, R. W. Brown, and V. C. Srivastava, "An inverse approach to the design of MRI main magnets," *IEEE Trans. Magn.* **30**(1), 108−112 (1994).

[6] S. Crozier and D. M. Doddrell, "Compact MRI magnet design by stochastic optimization," *J. Magn. Reson.* **127**(2), 233−237 (1997).

[7] P. N. Morgan, S. M. Conolly, and A. Macovski, "Resistive homogeneous MRI magnet design by matrix subset selection," *Magn. Reson. Med.* **41**(6), 1221−1229 (1999).

[8] S. Crozier, H. Zhao, and D. M. Doddrell, "Current density mapping approach for design of clinical magnetic resonance imaging magnets," *Conc. Magn. Reson.* **15**(3), 208−215 (2002).

[9] N. R. Shaw and R. E. Ansorge, "Genetic algorithms for MRI magnet design," *IEEE Trans. Appl. Supercond.* **12**(1), 733−736 (2002).

[10] G. Sinha, R. Sundararaman, and G. Singh, "Design concepts of optimized MRI magnet," *IEEE Trans. Magn.* **44**(10), 2351−2360 (2008).

[11] M. Xu, M. Ogle, X. Huang, K. Amm, and E. T. Laskaris, "Iterative EM design of an MRI magnet using HTS materials," *IEEE Trans. Appl. Supercond.* **17**(2), 2192−2195 (2007).

[12] Y. Lvovsky and P. Jarvis, "Superconducting systems for MRI- present solutions and new trends," *IEEE Trans. Appl. Supercond.* **15**(2), 1317−1325 (2005).

[13] A. K. Kalafala, "A design approach for actively shielded magnetic resonance imaging magnets," *IEEE Trans. Magn.* **26**(3), no. 3, 1181−1188 (1990).

[14] H. Xu, S. M. Conolly, G. C. Scott, and A. Macovski, "Homogeneous magnet design using linear programming," *IEEE Trans. Magn.* **36**(2), 476−483 (2000).

[15] V. Cavaliere, A. Formisano, R. Martone, G. Masullo, A. Matrone, M. Primizia, "Design of split coil magnets for magnetic resonance imaging," *IEEE Trans. Appl. Supercond.*, **10**(1), 759−762 (2000).

[16] V. Cavaliere, A. Formisano, R. Martone, M. Primizia, "A genetic algorithm approach to the design of split coil magnets for MRI," *IEEE Trans. Appl. Supercond.*, **10**(1), 1376−1379 (2000).

[17] J. H. Jensen, "Minimum-volume coil arrangements for generation of uniform magnetic fields," *IEEE Trans. Magn.*, **38(**6), 3579−3588 (2002).

[18] S. Besio, S. Pittaluga, V. Punzo, A. Trequattrini, "Elliptical coils for dedicated MRI magnets," *IEEE Trans. Appl. Supercond.*, **18**(2), 908−911 (2008).

[19] Q. Wang, G. Xu, Y. Dai, B. Zhao, L. Yan, K. Kim, "Design of open high magnetic field MRI superconducting magnet with continuous current and genetic

algorithm method," *IEEE Trans. Appl. Supercond.*, **19**(3), 2289−2292 (2009).

[20] Q. M. Tieng, V. Vegh, I. M. Brereton, "Globally optimal superconducting magnets part I: minimum stored energy (MSE) current density map," *J. Magn. Reson.*, **196**(1), 1−6 (2009).

[21] Q. M. Tieng, V. Vegh, I. M. Brereton, "Globally optimal superconducting magnets part II: symmetric MSE coil arrangement," *J. Magn. Reson.*, **196**(1), 7−11 (2009).

[22] Q. M. Tieng, V. Vegh, I. M. Brereton, "Minimum stored energy high-field MRI superconducting magnets," *IEEE Trans. Appl. Supercond.*, **19**(4), 3645−3652 (2009).

[23] C. Wang, Q. Wang, Q. Zhang, "Multiple layer superconducting magnet design for magnetic resonance imaging," *IEEE Trans. Appl. Supercond.*, **20**(3), 706−709 (2010).

[24] S. Pittaluga, S. Besio, V. Punzo, and A. Trequattrini, "Racetrack coils for dedicated MRI magnets," *IEEE Trans. Appl. Supercond.* **20**(3), 786−789 (2010).

[25] M. Kitamura, S. Kakukawa, K. Mori, and T. Tominaka, "An optimal design technique for coil configurations in iron-shielded MRI magnets," *IEEE Trans. Magn.* **30**(4), 2352−2355 (1994).

[26] H. Siebold, H. Huebner, L. Soeldner, and T. Reichert, "Performance and results of a computer program for optimizing magnets with iron," *IEEE Trans. Magn.* **24**(1), 419−422 (1988).

[27] M. D. Ogle and J. D'Angelo, "Design optimization method for a ferromagnetically self-shield MR magnet," *IEEE Trans. Magn.* **27**(2), 1689−1691 (1991).

[28] V. E. Eregin, V. S. Kashikhin, A. A. Lipatnikov, Y. P. Severgin, V. P. Shangin, and N. A. Shatil, "Formation of homogenous field in a shielded superconducting solenoid for an MRI," *IEEE Trans. Magn.* **28**(1), 675−677 (1992).

[29] S. Noguchi and A. Ishiyama, "Optimal design method for MRI superconducting magnets with ferromagnetic shield," *IEEE Trans. Magn.* **33**(2), 1904−1907 (1997).

[30] H. Zhao and S. Crozier, "Rapid field calculations for the effect of ferromagnetic material in MRI magnet design," *Meas. Sci. Technol.* **13**(2), 198−205 (2002).

[31] H. Zhao and S. Crozier, "A design method for superconducting MRI magnets with ferromagnetic material," *Meas. Sci. Technol.* **13**(12), 2047−2052 (2002).

[32] Y. Muruta, M. Abe, R. Ando, and T. Nakayama, "A novel design method of shapes of ferromagnetic materials for the superconducting MRI magnets," *IEEE Trans. Appl. Supercond.*, **19**(3), 2293−2296 (2009).

[33] X. Tang, D. Zu, T. Wang, and B. Han, "An optimizing design method for a compact iron shielded superconducting magnet for use in MRI," *Supercond. Sci. Technol.*, **23**(8), 1−6 (2010).

[34] C. Wu, J. Guo, C. Chen, G. Yan, and C. Li, "Optimal design and test of main magnet in superconducting MRI," *IEEE Trans. Appl. Supercond.* **20**(3), 1810−1813 (2010).

[35] M. W. Garrett, "Axially symmetric systems for generating and measuring magnetic fields," *J. Appl. Phys.* **22**(9), 1091−1107 (1951).

[36] T. Tadic and B. G. Fallone, "Design and optimization of a novel bored biplanar permanent magnet assembly for hybrid MRI systems," *IEEE Trans. Magn.* **46**(12), 4052−4058 (2010).

[37] T. Tadic and B. G. Fallone, "Three-dimensional non-axisymmetric pole piece shape optimization for biplanar permanent magnet MRI systems," *IEEE Trans. Magn.* **47**(1), 231–238 (2011).

[38] B. G. Fallone, M. Carlone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, K. Wachowicz, and C. Kirkby, "Development of a linac-MRI system for real-time ART," *Med. Phys.* **34**(6), 2547−2547 (2007).

[39] B. G. Fallone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, and S. Vidakovic, "First MR images obtained during megavoltage photon irradiation from a prototype integrated linac-MR system," *Med. Phys.* **36**(6), 2084−2088 (2009).

[40] B. G. Fallone, "Real-time MR-guided radiotherapy: integration of a low-field MR system," *Med. Phys.* **36**(6), 2774−2775 (2009).

[41] MATLAB, ver. 7.11 (The MathWorks, Natick, 2010).

[42] OPERA-3D, ver. 13.0 (Vector Fields, Oxford, 2009).

[43] Y. Zhang, "Solving large-scale linear programs by interior-point methods under the MATLAB environment," *Optim. Method. Softw.* **10**(1), 1−31 (1998).

[44] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists,* 6th ed. (Elsevier, Burlington, 2005).

[45] L. K. Forbes, S. Crozier, and D. M. Doddrell, "Rapid computation of static fields produced by thick circular solenoids," *IEEE Trans. Magn.* **33**(5), 4405−4410 (1997).

[46] M. Razeti, S. Angius, L. Bertora, D. Damiani, R. Marabotto, M. Modica, D. Nardelli, M. Perrella, and M. Tassisto, "Construction and operation of cryogen-free MgB2 magnets for open MRI systems," *IEEE Trans. Appl. Supercond.* **18**(2), 882−886 (2008).

[47] N. B. S. Gloria, M. C. L. Areiza, I. V. J. Miranda, and J. M. A. Rebello, "Development of a magnetic sensor for detection and sizing of internal pipeline corrosion defects," *NDT&E Int.* **42**(8), 669−677 (2009).

[48] M. J. Tannenbaum, A. K. Ghosh, K. E. Robins, and W. B. Sampson, "Magnetic properties of the iron laminations for CBA magnets," *IEEE Trans. Nucl. Sci.* **NS-30**(4), 3472−3474 (1983).

[49] D. Damiani, A. Laurenti, R. Marabotto, and M. P. Segre, "Coil with superconductive windings cooled without cryogenic fluids," U.S. Patent Application, US 2009/0315655 A1 (Dec. 2009).

[50] E. T. Laskaris, R. Ackermann, B. Dorri, D. Gross, K. Herd, and C. Minas, "A cryogen-free open superconducting magnet for interventional MRI applications," *IEEE Trans. Appl. Supercond.* **5**(2), 163−168 (1995).

[51] D. Nardelli, R. Marabotto, and A. Laurenti, "Superconducting coil having a granular superconducting junction," U.S. Patent Application, US 2009/0264295 A1 (Oct. 2009).

[52] X. Jiang, G. Shen, Y. Lai, and J. Tian, "Development of an open 0.3 T NdFeB MRI magnet," *IEEE Trans. Appl. Supercond.* **14**(2), 1621−1623 (2004).

[53] J. S. Ryu, Y. Yao, and C. S. Koh, "3-D optimal shape design of pole piece in permanent magnet MRI using parameterized nonlinear design sensitivity analysis," *IEEE Trans. Magn.* **42**(4), 1351−1354 (2006).

[54] J. St. Aubin, S. Steciw, B. G. Fallone, "Effect of transverse magnetic fields on a simulated in-line 6 MV linac," *Phys. Med. Biol.*, **55**(16), 4861−4869 (2010).

[55] J. St. Aubin, S. Steciw, B. G. Fallone, "Magnetic decoupling of the linac in a low field biplanar linac-MR system," *Med. Phys.*, **37**(9), 4755−4761 (2010).

CHAPTER 8
DESIGN AND OPTIMIZATION OF BIPLANAR SUPERCONDUCTING MAGNETS FOR AN INTEGRATED LINAC-MRI SYSTEM

A version of this chapter has been submitted for publication. T. Tadic and B. G. Fallone, "Design and optimization of biplanar superconducting magnets for an integrated linac-MRI system," (2012).

8.1 INTRODUCTION

The first prototype integrated linear accelerator (linac) and magnetic resonance imaging (MRI) system designed and built by our group involved the use of a low-field 0.2 T biplanar permanent magnet assembly.[38–40] The successful acquisition of magnetic resonance (MR) images during megavoltage irradiation with this system established the feasibility of the linac-MRI concept and demonstrated that the major technical issues regarding mutual interference of the two devices have been overcome. A permanent magnet imager was preferred for the first prototype due to the relative ease of integration offered by this type of magnet. A low field strength translated into relatively weak fringe fields, permitting expedited installation of the magnet system and a simplification of the required linac shielding. Furthermore, the absence of a cryostat and a relative lack of electrical wiring lend this system well to mechanical rotation.

In accordance with the desire to achieve improved image quality and greater acquisition speeds, we are now also considering the use of magnet systems with higher field strengths. Specifically, we are interested in cryogen-free

high-temperature superconducting (HTS) magnets with field strengths of approximately 0.5 T. As noted in the previous chapter, a particularly attractive feature of these systems is that they do not require massive cryostats filled with liquid coolant. Hence, the need for a dedicated safety ventilation line is alleviated, making mechanical rotation of these scanners practical. However, due to the limited current densities that can be achieved without quenching, conduction-cooled MRI systems employing HTS materials typically consist of large steel yoke structures similar to their permanent magnet counterparts.[9–11] These yoke structures act to increase the strength of the main magnetic field located between the magnet poles, while passively shielding the stray magnetic fields these systems produce.

The majority of cylindrical superconducting magnets for MRI have similar geometries. Since the introduction of active shielding methods, these geometries typically only vary in terms of their overall dimensions, such as their total length and inner/outer radii.[9] In contrast, there is a notable variation amongst the many possible yoke geometries used in biplanar magnet assemblies, such as the well-known four-column,[12] H-box,[13] pill-box,[14] and C-shaped[15] yoke designs. Due to the lack of previously available design methods, little work has been published in the past to optimize and evaluate multiple biplanar magnet systems with a unified approach. In particular, the design of these systems for use in radiotherapy guidance has yet to be thoroughly explored.

In this work, several theoretical designs for 0.5 T biplanar HTS magnets have been presented and evaluated for potential use in a rotating linac-MRI system with a parallel configuration. Each magnet assembly is compact in size, with a large hole longitudinally bored through the entire yoke structure to permit an unobstructed beam path for the linac. Four-column, H-box, pill-box, and C-shaped yoke geometries were considered, in addition to a novel modified four-column design. Utilizing the FEM-based optimization method

developed in the foregoing chapter,[11] a minimum volume coil configuration was determined for each design, subject to constraints on the axisymmetric magnetic field uniformity in the imaging volume. The residual non-axisymmetric field variations were then calculated and compared in terms of a spherical harmonic decomposition of the three-dimensional (3D) magnetic field patterns. In addition, the magnetic fringe field extent and passive shielding performance of each design was compared, as well as the total amount of superconducting material required to obtain the desired field strength at isocenter.

8.2 METHODS

8.2.1 Magnet Geometry Specifications

Five magnet geometries were investigated in this work: a standard four-column design with cylindrical posts (S4); an H-box design with two open sides (HB); a closed cylindrical pill-box design (PB); a C-shaped design (CS); and a modified four-column design with non-cylindrical posts (M4). Illustrations of these geometries are provided in Figure 8.1. The main magnetic field is oriented along the axial direction, as designated by the z axis, and the patient axis about which the magnet rotates is identified with the x axis (Figures 1.2 and 8.1). Each magnet assembly consists of a yoke structure constructed from AISI 1020 steel, a pole plate composed of Armco magnetic steel, and a unique optimized MgB_2 HTS coil configuration located within a region referred to herein as the feasible coil domain.

For each yoke geometry, we have considered both $l = 20$ and 30 cm diameter cylindrical beam holes longitudinally bored through the entire yoke structures, yielding a total of ten different magnet designs. For consistency, all magnet designs corresponding to a particular yoke bore diameter share

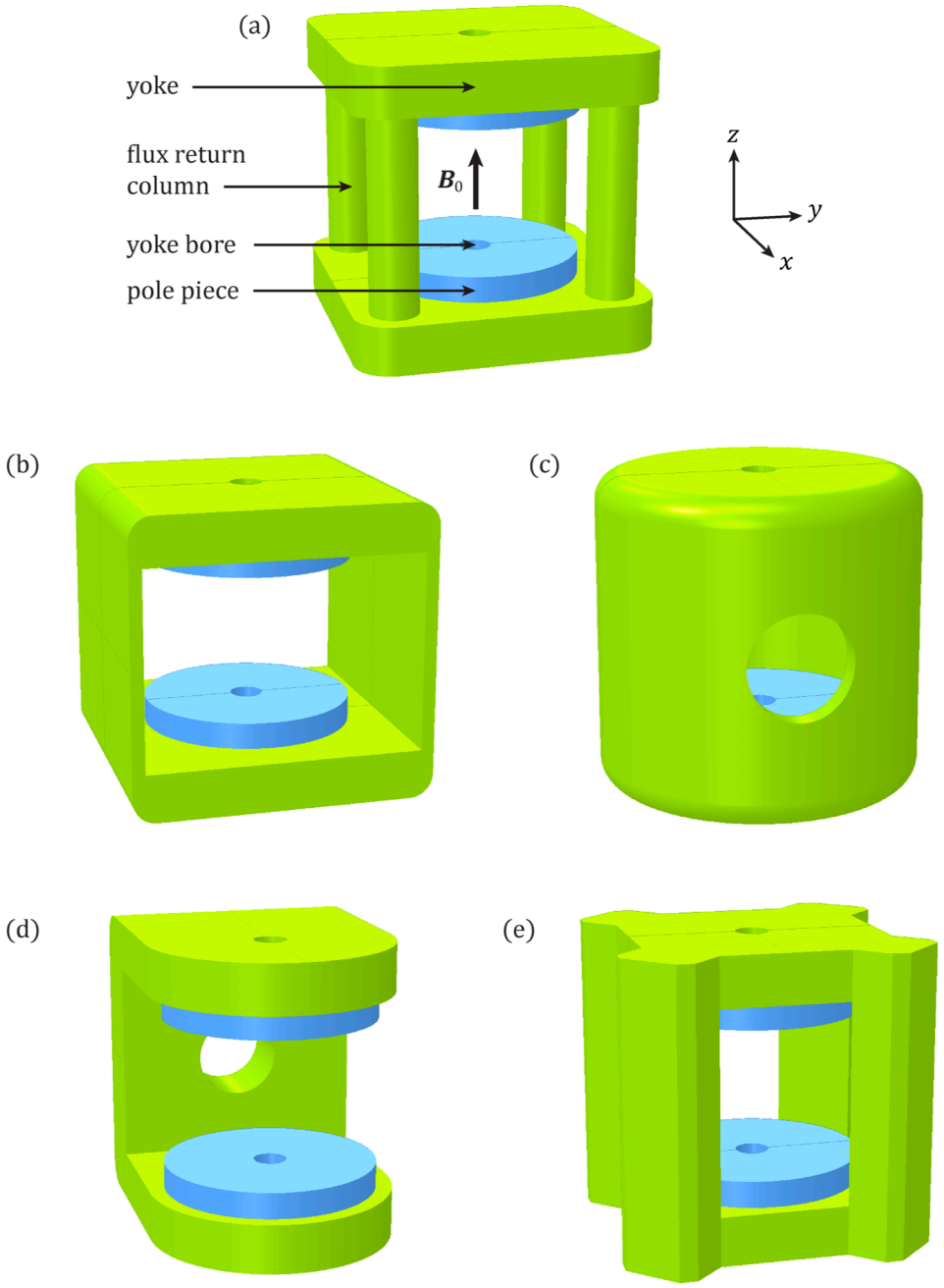


FIGURE 8.1: Three-dimensional illustrations of the magnet geometries investigated in this work: (a) a standard four-column design with cylindrical posts (S4); (b) an H-box design with two open sides (HB); (c) a closed cylindrical pill-box design (PB); (d) a C-shaped design (CS); and (e) a modified four-column design (M4).

an identical: pole plate geometry, feasible coil domain, axial yoke thickness (in the z direction), and volume of steel $\mathcal{V}_y$ comprising the yoke structure.

In order to achieve a practical vertical clearance of at least 90 cm in a radiation therapy vault of 3.50 m height, a maximum diameter of rotation of 2.60 m was specified. This is to provide adequate room for the gantry structure upon which the rotating magnet is supported, as well as to accommodate for any additional diagnostic or therapeutic devices and equipment (such as the linac and other associated structures). For a yoke extending 1.82 m in the axial direction, this constraint corresponds to a maximum lateral extent of 1.86 m. Each of the S4, HB, PB, and M4 designs has either a circular or square footprint in the xy plane. Consequently, these designs are assigned a lateral extent of 1.85 m. In contrast, the geometry and footprint of the CS design is naturally asymmetric. Hence, the lateral extent of 1.50 m assigned to this design is primarily governed by the diameter of the pole and coil structures it supports. The remaining dimensions of each magnet geometry were selected such that the total weight of magnetic material was within 10% of the 20 tonne yoke structure used in a commercially available conduction-cooled 0.5 T biplanar magnet system with an acceptable fringe field.[46] Similar to the approach in the foregoing chapter, these dimensions were chosen to provide adequate space for the coil configuration and associated patient gap, while also providing an approximate balance between field strength enhancement, passive magnetic shielding, and mechanical stability. A selection of the dimensions common to all magnet designs are shown in Figure 8.2.

The feasible coil domain for all designs extends between 34 and 42 cm from isocenter in the axial direction, such that adequate space is available for a cooling and mounting system while providing a patient gap of at least 60 cm. The inner radii of the feasible coil domains are 16 and 22 cm for the 20 and

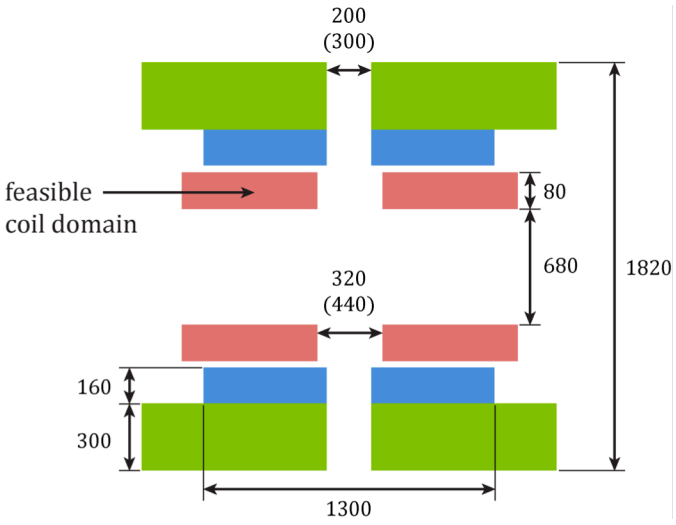


FIGURE 8.2: A typical magnet cross-section in the *yz* plane with dimensions given in millimeters. For dimensions with multiple values, the numbers with and without parenthesis correspond to the 20 and 30 cm diameter yoke bore variants, respectively. All other dimensions are common to all magnet designs investigated.

30 cm yoke bore variants, respectively, with all designs sharing an outer radius of 66 cm. For a representative linac target to isocenter distance of 135 cm, maximum radiation field sizes of 30×30 cm^2 and 45×45 cm^2 at isocenter would be possible for the 20 and 30 cm diameter yoke bore variants, respectively.

8.2.1.1 Coil Optimization Procedure

For each of the magnet designs investigated in this work, an iterative optimization scheme was applied to determine a minimum volume coil configuration, subject to magnetic field uniformity constraints over a spherical region of interest. The full details of this method have been provided in the preceding chapter,[11] for which the important aspects of our implementation are reviewed here. A flow diagram summarizing the optimization process is shown in Figure 8.3.

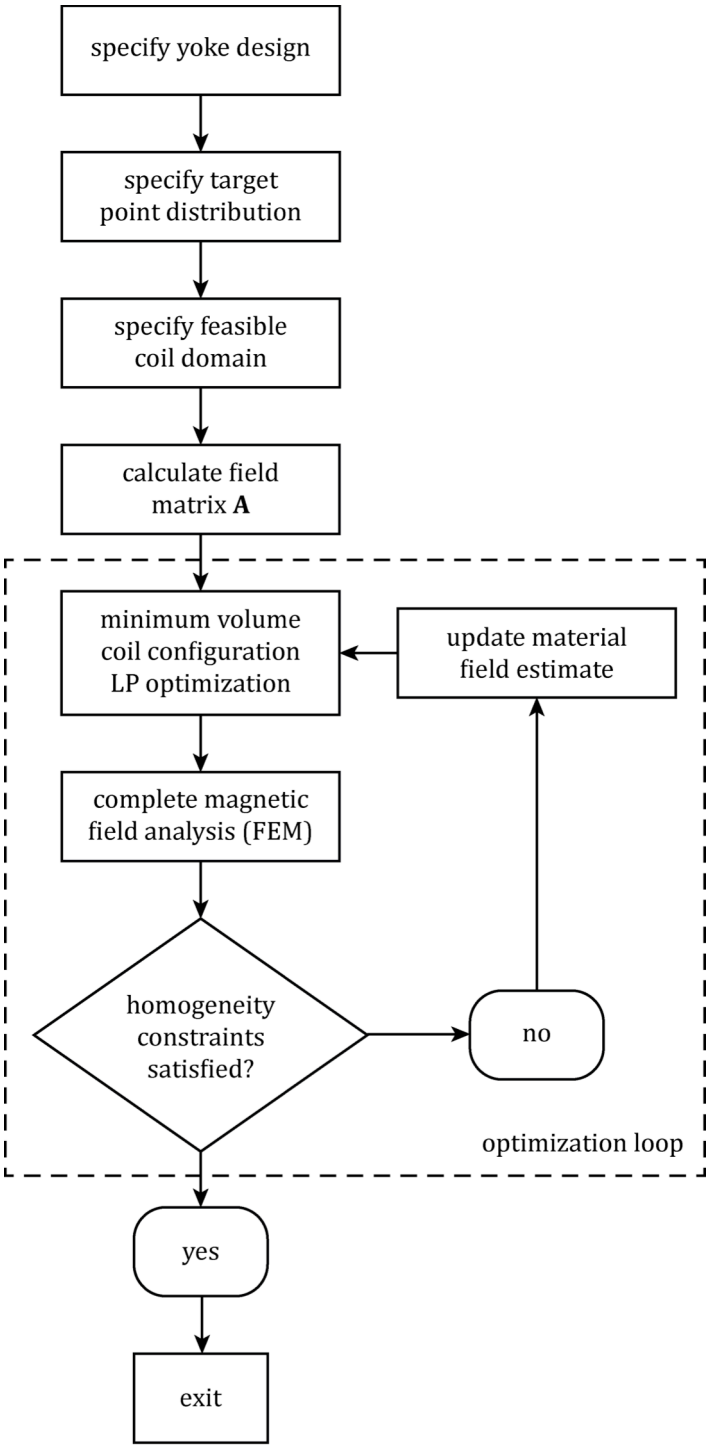


FIGURE 8.3: Iterative optimization flow diagram.

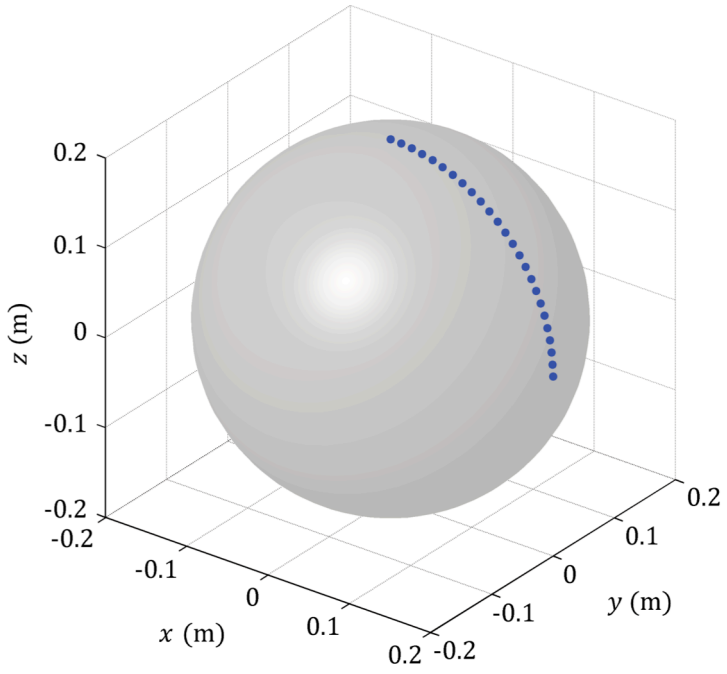


FIGURE 8.4: Target point distribution on the surface of a 40 cm DSV at isocenter.

An initial array of n circular candidate coils was established by segmenting the feasible coil domain by means of a dense grid with 2 cm spacing. This resulted in an initial array of 200 and 176 coils for the 20 and 30 cm yoke bore variants, respectively. The initial grid spacing of 2 cm was further reduced to 2.2 mm during the linear programming (LP) refinement process, through specification of the grid reduction parameter $\gamma_r = 3$ and coil aspect ratio $\gamma_a = 1.2$ (see § 7.2.1.2).

A distribution of $N_t = 25$ target points was specified along a single polar arc extending from $\theta = 0$ to $\theta = \pi/2$, on the surface of a 40 cm diameter spherical volume (DSV) at isocenter (Figure 8.4). By denoting the location of the ith target point as $\boldsymbol{r}_i$, the uniformity of the axial magnetic field at the target points was constrained according to (see § 7.2.1.1):

$$|\Delta B_z(\boldsymbol{r}_i)| \leq \kappa_1' \kappa_2, \quad \text{for } i = 1, \ldots, N_t, \tag{8.1}$$

where the axial field variation $\Delta B_z(\boldsymbol{r})$ at the point $\boldsymbol{r}$ is defined by

$$\Delta B_z(\boldsymbol{r}) = \frac{B_z(\boldsymbol{r}) - B_0}{B_0} \cdot 10^{-6}. \tag{8.2}$$

In these expressions, B_z refers to the axial component of the magnetic field, B_0 is the desired field strength at isocenter, κ_1' is the relative field error tolerance in units of ppm, and $\kappa_2 \geq 1$ is the specified convergence factor. In the present work, the values $B_0 = 0.5$ T, $\kappa_1' = 10$ ppm, and $\kappa_2 = 2$ were specified.

For a common current density J in each of the coils, the total conductor volume $\mathcal{V}_c$ is given by:

$$\mathcal{V}_c = \frac{2\pi}{J} \sum_{j=1}^{n} s_j |p_j|. \tag{8.3}$$

where s_j and p_j are the radius and current for the jth coil, respectively. Ultimately, the goal of the optimization scheme is then to minimize $\mathcal{V}_c$ subject to the axial field uniformity constraints in (8.1).

Assuming an operating temperature of 12 K, a realistic target for the engineering critical current density of the MgB_2 coils[24] is 28 kA/cm^2 at 4 T, from which a working value of $J = 21$ kA/cm^2 was specified.

The optimization algorithm was implemented in MATLAB[24] on a 3.0 GHz quad-core PC workstation with 16 GB of RAM. The minimization of Eq. (8.3) at

each iteration amounts to a linear programming (LP) that was solved with the `linprog` routine from the OPTIMIZATION TOOLBOX in MATLAB. Complete solution of the LP step and associated grid refinement process required less than two minutes of computation time. Simulation of the complete magnetic fields at each iteration was performed with the commercially available 3D FEM software package[42] OPERA-3D. The updated parameters defining the coil configuration were passed to OPERA-3D using a component object model (COM) interface that is established within the MATLAB script. In this way, iteratively updating and solving the complete FEM model is an automated step. A description of the FEM modeling procedure is given below.

8.2.2 The Finite Element Method

OPERA-3D employs the nonlinear magnetostatics FEM solver TOSCA[19] to calculate a nodal solution for the scalar magnetic potential within the model geometry. The 3D magnetic field can then be inferred from this solution. The complete details of the FEM formulation employed by this solver are provided in § 3.3.6.

8.2.2.1 Model Description

For each design investigated in this work, a FEM model was generated with the magnet assembly placed at the center of a large cubic global model volume (GMV). The GMV was truncated at a distance of 15 m from isocenter according to the discussion in the preceding chapter (see § 7.2.3.2). This results in an estimated error in the magnetic field of at most 2 μT (approximately 4 ppm) within a 40 cm DSV at isocenter. Symmetry in the magnet assemblies and associated magnetic fields was then exploited to simplify the model geometries and reduce the complexity and memory requirements of the FEM simulations. Specifically, one eighth of each magnet structure was modeled for the S4, HB, CB, and M4 designs, and one quarter of

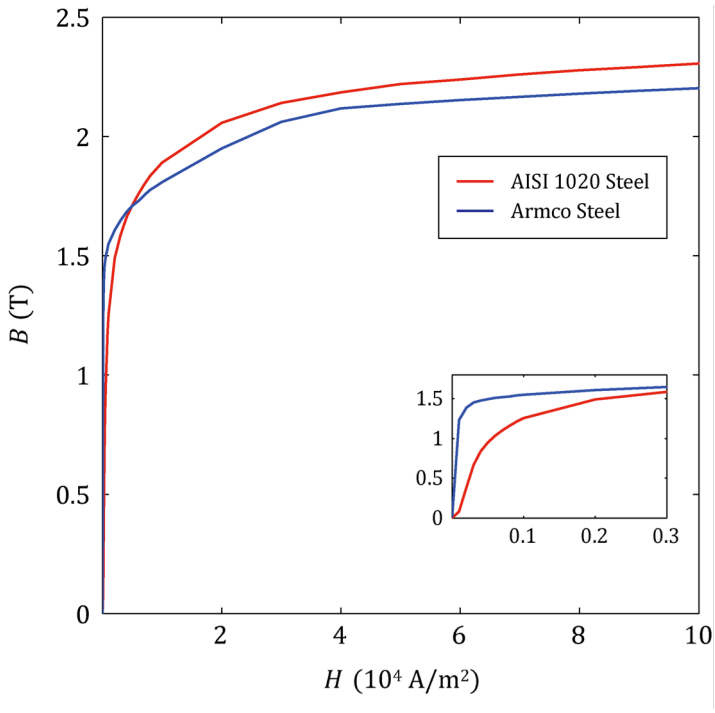


FIGURE 8.5: Nonlinear magnetization curves for AISI 1020 plain carbon steel and Armco magnetic steel.

the structure for the CS design. The required boundary conditions to achieve this reduction were discussed previously in § 5.2.2.2. In order to take into account the nonlinear properties of the magnetic materials employed, the magnetization curves[22, 23] shown in Figure 8.5 were assigned to the appropriate components.

Each simplified FEM model was partitioned with a three-dimensional mesh consisting of isoparametric tetrahedral quadratic Lagrange elements (see § 3.3.2). The manual selection of mesh parameters was identical to that specified in Chapter 7. This resulted in an estimated error of at most 0.2 and 0.5 μT (approximately 0.4 and 1 ppm) on the boundary and interior of a 40 cm DSV, respectively. The number of mesh elements and the computation times for each model are listed in Tables 8.1 and 8.2 in the results (see

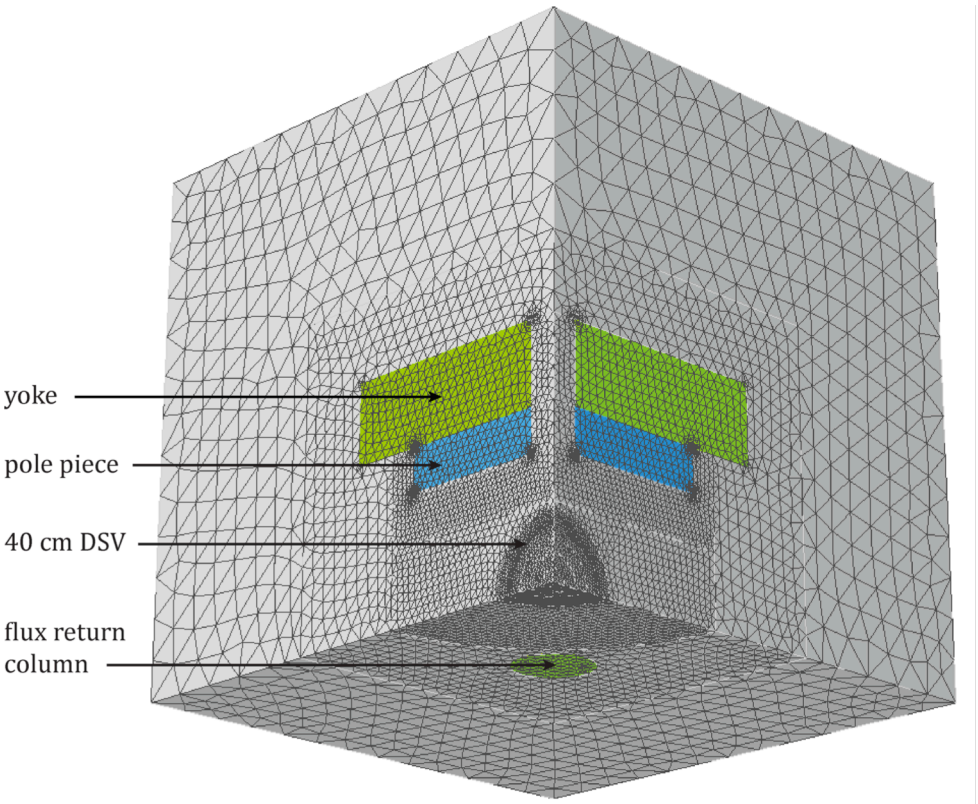


FIGURE 8.6: The simplified FEM model geometry and mesh for the S4 design with a 20 cm diameter yoke bore. Conductors are not shown. The entire modeling domain that extends 15 m from isocenter is only partially shown (in grey). Due to symmetry in the magnet geometries and corresponding magnetic fields, the FEM is only applied within a fraction of the complete model geometry in an effort to reduce the computation expense.

§ 3.3.2) The simplified model geometry and corresponding FEM mesh is illustrated for the S4 design in Figure 8.6.

8.2.3 Magnetic Field Analysis

In order to evaluate and compare the residual axial field variations produced by the magnet designs investigated, a spherical harmonic decomposition of the 3D magnetic field patterns was performed. This analysis elucidates the nature of these field distributions and how they relate to the geometrical

configurations of their corresponding yoke structures. In addition, this analysis is also helpful in determining any passive shimming advantages potentially offered by the magnet designs investigated.

In a region absent of currents and magnetic materials, the axial magnetic field component satisfies Laplace's equation

$$\nabla^2 B_z = 0. \tag{8.4}$$

Accordingly, B_z may be expressed as[22] (see § 3.1.3)

$$B_z(r,\theta,\phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} r^{\ell} Y_{\ell}^{m}(\theta,\phi), \tag{8.5}$$

where (r,θ,ϕ) denote the usual spherical coordinates and $a_{\ell m}$ are the expansion coefficients. The spherical harmonic function $Y_{\ell}^{m}(\theta,\phi)$ is defined according to

$$Y_{\ell}^{m}(\theta,\phi) = N_{\ell}^{m} P_{\ell}^{m}(\cos\theta) e^{im\phi}, \tag{8.6}$$

where $P_{\ell}^{m}(\cos\theta)$ denotes the associated Legendre polynomial of order ℓ and degree m, and the normalization factor N_{ℓ}^{m} is defined as

$$N_{\ell}^{m} = \sqrt{\frac{(2\ell+1)}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}. \tag{8.7}$$

Evidently, the orthonormal basis $\{Y_{\ell}^{m}\}$ is composed of complex functions. Thus, it is instead preferred to express B_z in terms of the basis of real spherical harmonics $\{y_{\ell}^{m}\}$, such that

$$B_z(r,\theta,\phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} c_{\ell m} r^{\ell} y_{\ell}^{m}(\theta,\phi), \tag{8.8}$$

with $y_{\ell}^{m}(\theta,\phi)$ defined according to

$$y_{\ell}^{m} = \begin{cases} \sqrt{2}\,\mathrm{Im}(Y_{\ell}^{m}) = \dfrac{1}{i\sqrt{2}}(Y_{\ell}^{-m} - (-1)^m Y_{\ell}^{m}), & \text{for } m < 0 \\ Y_0^0, & \text{for } m = 0 \\ \sqrt{2}\,\mathrm{Re}(Y_{\ell}^{m}) = \dfrac{1}{\sqrt{2}}(Y_{\ell}^{m} + (-1)^m Y_{\ell}^{-m}), & \text{for } m > 0. \end{cases} \tag{8.9}$$

Due to the orthonormality of the basis $\{y_{\ell}^{m}\}$, the expansion coefficients $c_{\ell m}$ can be calculated using the expression

$$c_{\ell m} = \frac{1}{R^{\ell}} \int_0^{2\pi} \int_0^{\pi} B_z(R,\theta,\phi) y_{\ell}^{m}(\theta,\phi) \sin\theta \, d\theta \, d\phi. \tag{8.10}$$

For this analysis, $B_z(R,\theta,\phi)$ was obtained from the FEM simulations described in the previous section and sampled on the surface of a 40 cm DSV at isocenter according to a 131st-order Lebedev quadrature scheme[22] that was used to numerically evaluate the integral in Eq. (8.10).

Although the coefficients $c_{\ell m}$ complete the expansion defined in Eq. (8.8), their values alone are not telling of the relative impact that the various harmonics have on the magnetic field variation in a particular region of interest. This is partly due to the fact that the maximum value of $|y_{\ell}^{m}(\theta,\phi)|$ generally differs from unity. Furthermore, Eq. (8.8) clearly shows that the impact of a given harmonic is proportional to r^{ℓ}. Hence, in order to compare

the relative contribution of each harmonic to the total field variation over a 40 cm DSV, the peak harmonic amplitudes $C_{\ell m}$ were calculated with

$$C_{\ell m} = c_{\ell m} R^{\ell} \max_{\theta,\phi}\{|y_{\ell}^{m}(\theta,\phi)|\}. \tag{8.11}$$

In this way, the impact of a given harmonic can easily be expressed in familiar units of ppm, simply by dividing $C_{\ell m}$ by the corresponding field strength B_0.

8.3 RESULTS AND DISCUSSION

Results of the coil optimization procedure are presented in Tables 8.1 and 8.2. Convergence to a minimum volume coil configuration, satisfying the field uniformity constraint in Eq. (8.1), required no more than 24 iterations for any magnet design. Each coil configuration is symmetric and consists of seven coils per magnet pole, with alternating current directions in adjacent coils. The optimized coil arrangement for each of the 20 cm yoke bore variants are visualized in Figure 8.7, along with the FEM solution for the magnetic field within the yoke and pole structures. The residual axial magnetic field inhomogeneities ΔB_z for these designs are visualized over the surface of a 40 cm DSV at isocenter in Figure 8.8, and the corresponding peak-to-peak values are shown in in Tables 8.1 and 8.2. To illustrate the dependence of the field variation magnitude on the imaging volume size, a plot of the peak-to-peak inhomogeneity as a function of spherical volume diameter is provided in Figure 8.10.

To aid in the evaluation of the residual magnetic field patterns, the coefficients corresponding to the real spherical harmonic decomposition in Eq. (8.10) were calculated up to a maximum order of $\ell = 20$ for each magnet design. These coefficients were then scaled by the appropriate factors in

TABLE 8.1: FEM model properties and optimization results for the 20 cm diameter yoke bore variants of the magnet designs investigated. The minimized conductor volume $\mathcal{V}_c$, final magnetic field inhomogeneities ΔB_z and ΔB, peak conductor field B_{peak}, and fringe field extent are given.

Quantity	S4	HB	PB	CS	M4
$\mathcal{V}_y$ (m^3)	2.767	2.767	2.759	2.767	2.759
Mesh elements (10^6)	1.081	1.128	1.100	2.197	1.143
Simulation time (min)	53.1	45.9	55.6	137.6	46.1
Optimization iterations	11	14	15	12	15
B_0 (T)	0.5000	0.4998	0.5000	0.4997	0.4998
ΔB_z, target points (ppm)	31	32	31	35	36
ΔB_z, 40 cm DSV (ppm)	251	1063	802	3613	399
ΔB, 40 cm DSV (ppm)	251	1063	802	3613	399
$\mathcal{V}_c$ (10^4 cm^3)	1.810	1.770	1.852	1.660	1.708
B_{peak} (T)	3.406	3.379	3.505	3.280	3.351
$x_{5\text{G}}$ (m)	4.6	4.5	1.9	2.6	1.8
$y_{5\text{G}}$ (m)	4.6	4.5	1.9	2.5	1.8
$z_{5\text{G}}$ (m)	5.8	5.6	2.2	3.4	2.3

Eq. (8.11) in order to determine the peak amplitudes of the individual harmonics over a 40 cm DSV. The scaled coefficients with $\ell < 8$ that were calculated to have a magnitude of at least 0.1 mT are provided in Tables 8.3 and 8.4.

A typical requirement for the installation of clinical MRI systems is the confinement of magnetic fields with strengths greater than 5 G to within the boundaries of the imaging suite. Accordingly, the maximal extent $(x_{5\text{G}}, y_{5\text{G}}, z_{5\text{G}})$ of the 5 G magnetic field contours (measured relative to

TABLE 8.2: FEM model properties and optimization results for the 30 cm diameter yoke bore variants of the magnet designs investigated. The minimized conductor volume $\mathcal{V}_c$, final magnetic field inhomogeneities ΔB_z and ΔB, peak conductor field B_{peak}, and fringe field extent are given.

Quantity	S4	HB	PB	CS	M4
$\mathcal{V}_y$ (m^3)	2.731	2.731	2.723	2.731	2.723
Mesh elements (10^6)	1.109	1.158	1.137	2.243	1.176
Simulation time (min)	49.6	53.3	52.8	154.0	48.2
Optimization iterations	18	8	24	7	20
B_0 (T)	0.5000	0.4999	0.4999	0.4996	0.5000
ΔB_z, target points (ppm)	26	26	38	26	29
ΔB_z, 40 cm DSV (ppm)	241	1078	805	3556	401
ΔB, 40 cm DSV (ppm)	241	1078	805	3556	401
$\mathcal{V}_c$ (10^4 cm^3)	2.036	1.989	2.104	1.837	1.923
B_{peak} (T)	3.444	3.453	3.550	3.371	3.434
x_{5G} (m)	4.7	4.6	1.9	2.6	1.8
y_{5G} (m)	4.7	4.6	1.9	2.6	1.8
z_{5G} (m)	5.8	5.7	2.3	3.5	2.3

isocenter) are also provided in Tables 8.1 and 8.2. Projections of these contours are visualized in Figure 8.9

For a given yoke bore diameter, there was less than a 6.9 % variation in the peak conductor fields B_{peak} amongst the various magnet designs investigated, as shown in Tables 8.1 and 8.2. A maximum value of 3.550 T was observed for the PB design with a 30 cm yoke bore diameter, which is well below the 4 T limit specified for the MgB_2 material used. This design also required a conductor volume of 2.104×10^4 cm^3, which was greater than that used in any of the other assemblies. In comparison, the CS design

resulted in the lowest peak field and required conductor volume, with values of 3.280 T and 1.660×10^4 cm^3, respectively.

The presence of a larger yoke bore clearly necessitates a reduction in the volume of the magnetized yoke and pole structures that contribute to the net magnetic field at isocenter. Consequently, the 30 cm yoke bore designs lead to an increase of at least 10.7% in the minimum conductor volume required to satisfy the optimization homogeneity constraints, in addition to an increase of at least 1.1% in the peak conductor fields.

As shown in Figure 8.10, the four-column S4 and M4 magnet designs were found to exhibit superior magnetic field uniformity over the full range of spherical volumes examined. In particular, the S4 design achieved the lowest peak-to-peak ΔB_z values of 251 and 241 ppm over a 40 cm DSV for both 20 and 30 cm yoke bore variants, respectively. Similarly, the M4 design achieves comparable values of 399 and 401 ppm for the 20 and 30 cm yoke bore variants, well outperforming the HB, CS, and PB designs. In comparison to the M4 design, the improved uniformity obtained with the S4 design is likely due to the use of smaller yoke columns, which have a reduced influence on the magnetic field pattern near isocenter.

The S4 and M4 magnet designs also result in the fewest number of terms contributing to the spherical harmonic decomposition of the residual axial field variation, as shown in Tables 8.3 and 8.4. In addition to the even zonal harmonics (ℓ even, $m = 0$), only those of the fourth degree ($m = 4$) with either $\ell = 4$ or 6 have a substantial amplitude ($> 1\ \mu$T). Indeed, this is the expected result due to the prominent four-fold symmetry about the *z* axis possessed by both of these yoke designs. Coupled with the relatively low magnitude of the noted sparse harmonics, these results suggest that the S4 and M4 magnet designs are likely to pose the least challenge for post-

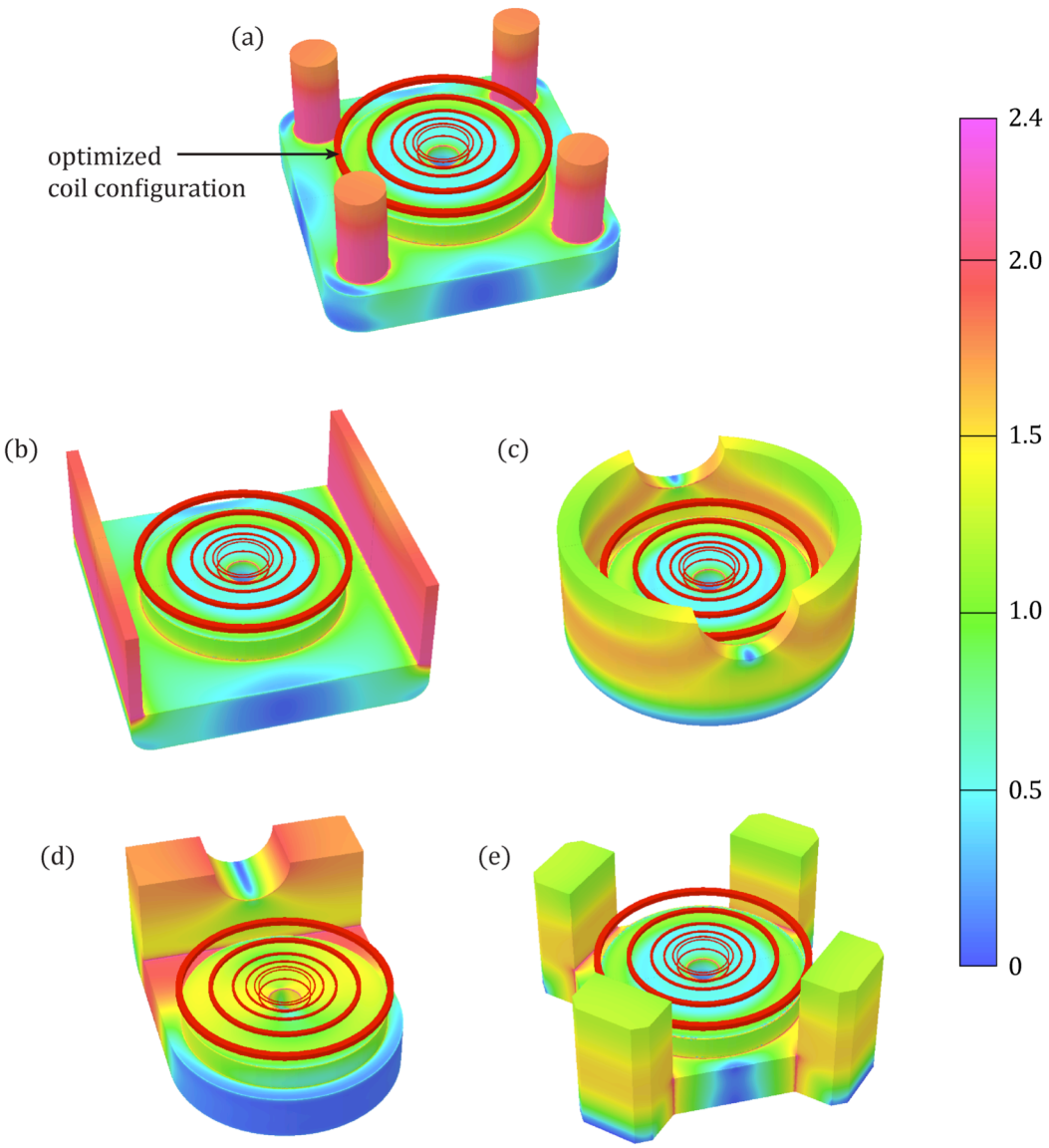


FIGURE 8.7: Half sections of the magnet geometry and corresponding optimal coil configuration for the (a) S4, (b) HB, (c) PB, (d) CS, and (e) M4 designs with a 20 cm diameter yoke bore. The magnetic field solutions are displayed on the surface of the yoke and pole structures. Values are given in units of T according to the color scale shown on the right. The material in regions for which the induced flux is greater than 1.8 T is magnetically saturated. The 30 cm diameter yoke bore variants exhibit similar field patterns.

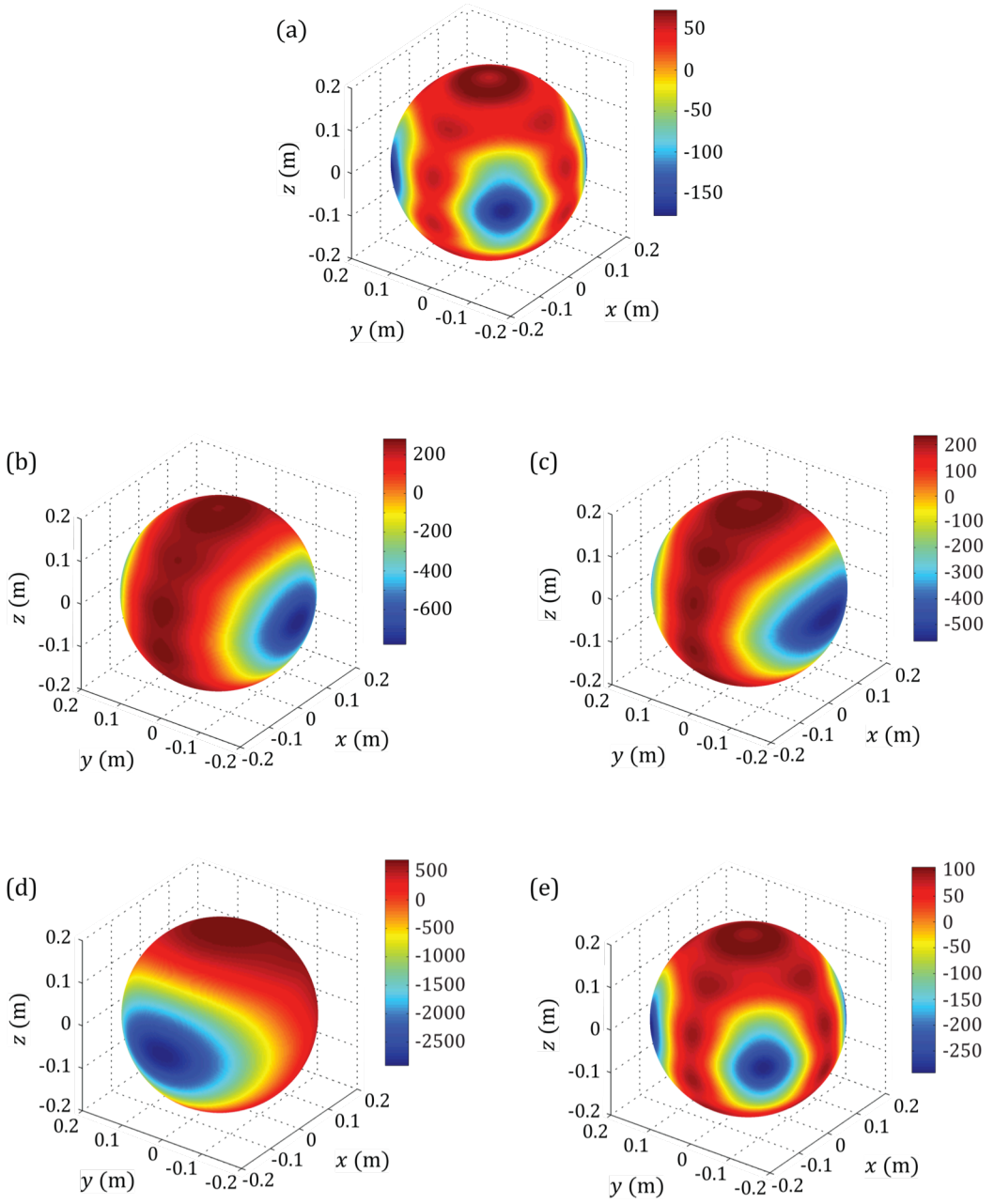


FIGURE 8.8: The residual axial magnetic field variation ΔB_z (ppm) plotted on the surface of a 40 cm DSV at isocenter for the: (a) S4; (b) HB; (c) PB; (d) CS; and (e) M4 magnet designs with a 20 cm diameter yoke bore. The limiting nature of the non-axisymmetric yoke structures is reflected in the magnetic field patterns shown. The 30 cm diameter yoke bore variants exhibit similar distributions.

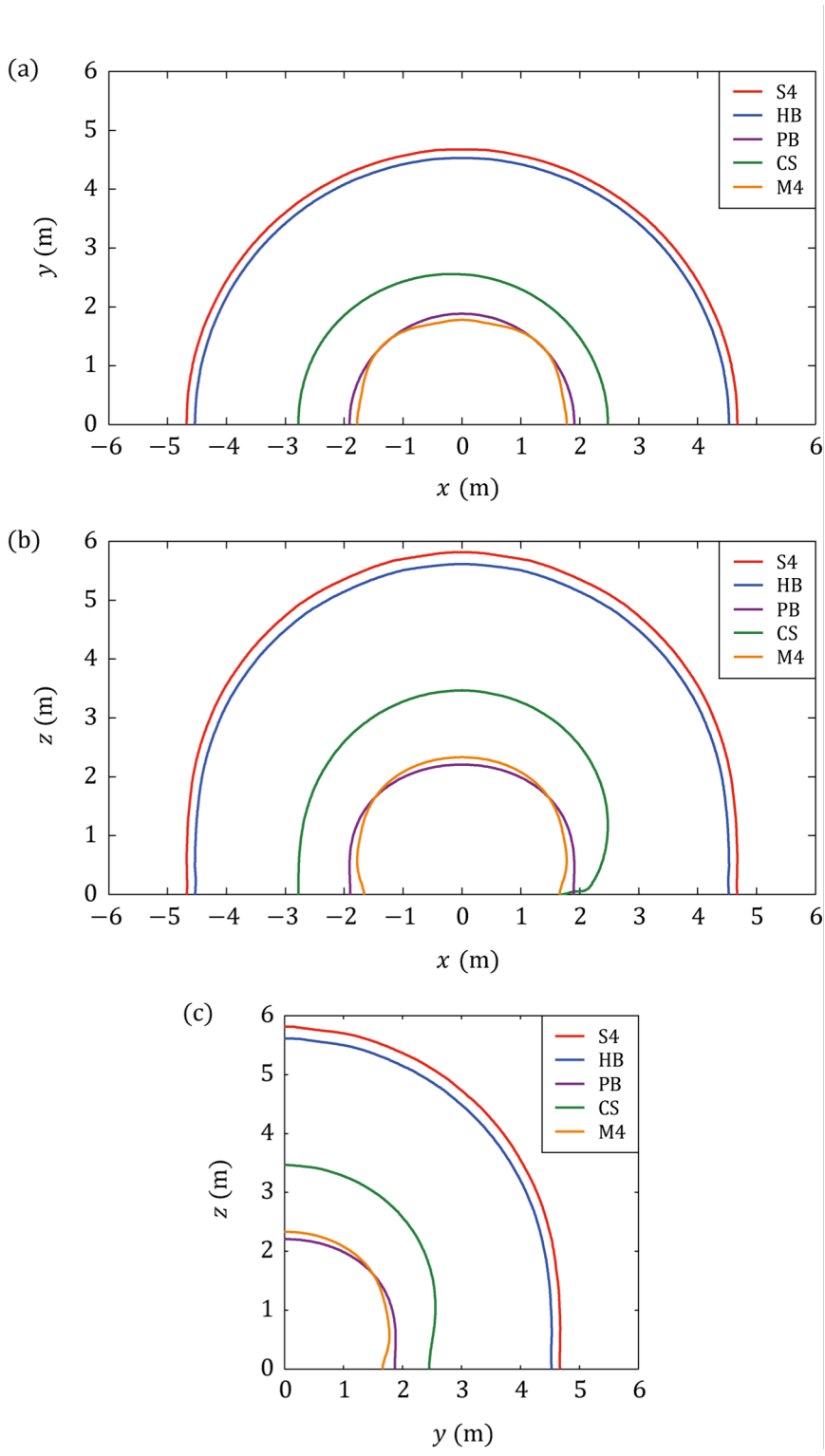


FIGURE 8.9: Plots of the 5 G magnetic field contours projected onto the (a) xy, (b) xz, and (c) yz planes for 20 cm diameter yoke bore variants of the optimized magnet designs. The M4 and PB designs provide the greatest shielding of the fringe magnetic fields.

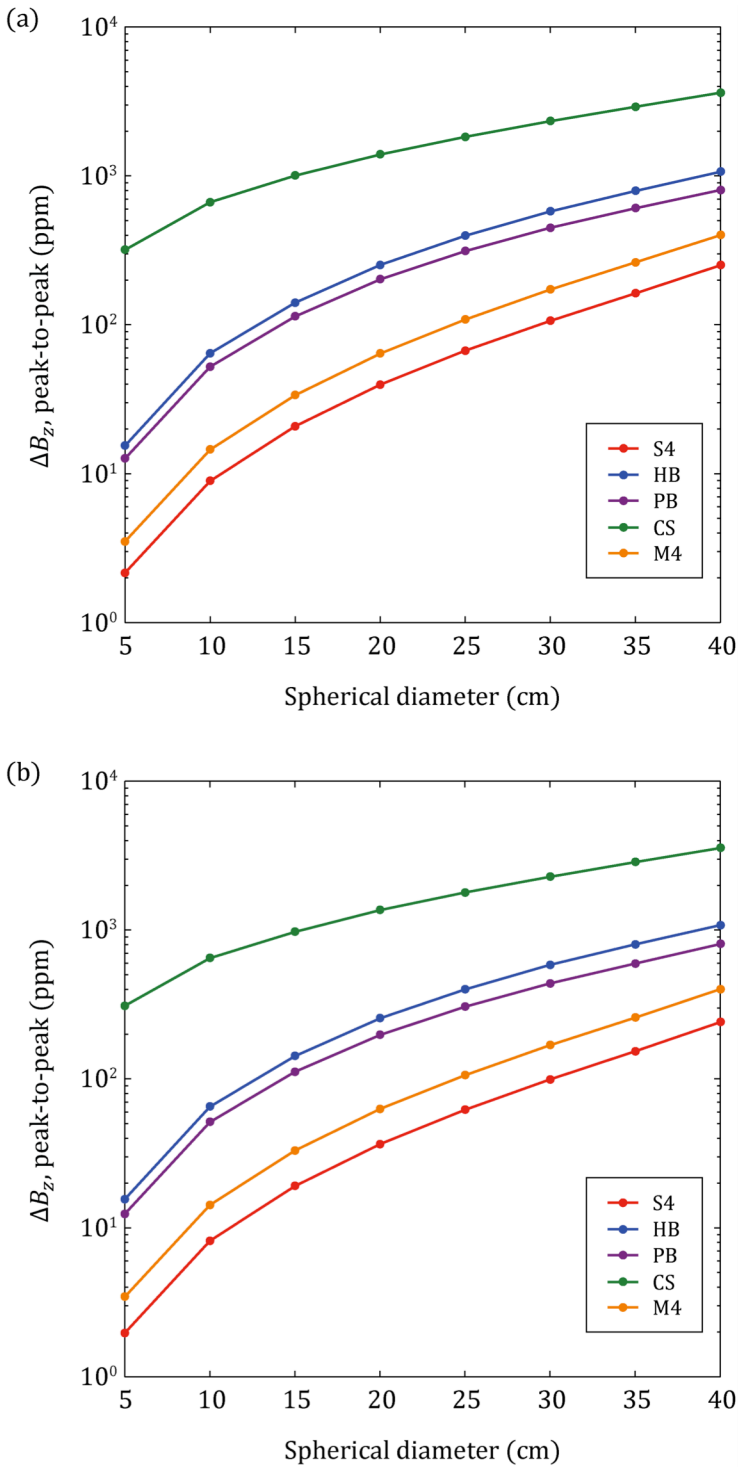


FIGURE 8.10: The residual peak-to-peak axial magnetic field inhomogeneity ΔB_z (ppm) calculated over various spherical regions of interest for the (a) 20 and (b) 30 cm diameter yoke bore variants. The four-column S4 and M4 designs exhibit superior uniformity over the complete range of spherical diameters.

TABLE 8.3: Scaled peak amplitudes (μT) for the real spherical harmonic decomposition of the residual axial magnetic field, computed for the 20 cm yoke bore diameter variants of the magnet designs investigated. The values given correspond to a 40 cm DSV at isocenter. Only harmonics of order $0 < \ell < 8$ with a magnitude of at least 0.1 μT are shown.

		Peak harmonic amplitude $C_{\ell m}$ (μT)				
Order ℓ	Degree m	S4	HB	PB	CS	M4
1	1	0.0	0.0	0.0	-633.0	0.0
2	0	45.3	182.3	164.4	436.5	73.4
2	2	0.1	222.5	160.3	-247.2	0.0
3	1	0.0	0.0	0.0	240.3	0.0
3	3	0.0	0.0	0.0	-63.8	0.0
4	0	-14.9	-45.8	-57.1	-87.8	-29.0
4	2	0.1	-46.0	-36.4	51.7	0.0
4	4	48.9	-13.0	19.2	-13.1	83.1
5	1	0.0	0.0	0.0	-20.7	0.0
5	3	0.0	0.0	0.0	9.1	0.0
5	5	0.0	0.0	0.0	-2.2	0.0
6	0	5.3	4.1	9.3	1.1	7.5
6	2	0.0	3.2	3.5	-3.1	0.0
6	4	-5.8	1.6	-2.7	1.4	-9.6
6	6	0.0	0.5	1.5	-0.3	0.0
7	1	0.0	0.0	0.0	0.9	0.0
7	3	0.0	0.0	0.0	-0.5	0.0
7	5	0.0	0.0	0.0	0.2	0.0

manufacturing passive shimming. The closed HB and PB designs have two-fold rotational symmetry, leading to a large contribution from even order harmonics with $m = 2$, in addition to a decreasing contribution from those with degrees that are multiples thereof. In contrast, the CS design lacks rotational symmetry about the *z* axis, leading to a large number of mixed harmonics. In particular, the presence of the single large flux return column

TABLE 8.4: Scaled peak amplitudes (μT) for the real spherical harmonic decomposition of the residual axial magnetic field, computed for the 30 cm yoke bore diameter variants of the magnet designs investigated. The values given correspond to a 40 cm DSV at isocenter. Only harmonics of order $0 < \ell < 8$ with a magnitude of at least 0.1 μT are shown.

		Peak harmonic amplitude $C_{\ell m}$ (μT)				
Order ℓ	Degree m	S4	HB	PB	CS	M4
1	1	0.0	0.0	0.0	-615.5	0.0
2	0	41.5	186.1	158.9	430.1	72.6
2	2	0.0	224.7	163.0	-247.8	0.0
3	1	0.0	0.0	0.0	246.5	0.0
3	3	0.0	0.0	0.0	-64.9	0.0
4	0	-16.9	-52.2	-61.9	-94.8	-37.0
4	2	0.0	-46.5	-37.1	53.3	0.0
4	4	49.6	-13.1	19.6	-13.3	84.5
5	1	0.0	0.0	0.0	-21.2	0.0
5	3	0.0	0.0	0.0	9.2	0.0
5	5	0.0	0.0	0.0	-2.2	0.0
6	0	4.7	4.8	8.4	-1.2	7.1
6	2	0.0	3.2	3.6	-3.0	0.0
6	4	-5.9	1.6	-2.7	1.4	-9.7
6	6	0.0	0.5	1.5	-0.3	0.0
7	1	0.0	0.0	0.0	0.7	0.0
7	3	0.0	0.0	0.0	-0.4	0.0
7	5	0.0	0.0	0.0	0.2	0.0

leads to a dominant contribution from low-order harmonics and those with $m = 1$. Of course, this is with exception to the harmonics with $m < 0$ or $(\ell - m)$ odd, which are necessarily zero for all of the designs due to symmetries across the xz and xy planes, respectively.

As displayed in Tables 8.1 and 8.2, the M4 and PB designs provide the greatest shielding of magnetic fringe fields. In comparison to the other magnet assemblies, these designs employ the greatest amount of magnetic material in the flux return components of their yoke structures. Consequently, magnetic saturation of these components is avoided and improved fringe field confinement is achieved. As can be seen in Figure 8.7, the magnetic field strength within the M4 and PB yoke structures is generally limited to 1.8 T. According to Figure 8.5, the comprising steel material rapidly begins to saturate above this value. In contrast, the flux return components of the S4 and HB designs are considerably saturated, and the resultant fringe fields are necessarily expansive.

It is evident that the geometric distribution of yoke material plays an important role in determining the shielding capability of a given magnet design. In particular, by reallocating the yoke material from regions with low levels of induced flux (< 0.5 T) to regions of saturation (> 1.8 T), the material is more effectively distributed along the flux return path, leading to a dramatic improvement in fringe field confinement. In addition to explaining the superior shielding of the M4 design, this suggests that the shielding performance of the CS magnet could be improved yet. By introducing a yoke design possessing a graded axial thickness (or lateral width) that increases towards the single flux return column, the saturation effect that is observed in Figure 8.7(d) could be corrected without recourse to incorporating additional material and unnecessarily increasing both the magnet cost and weight. Indeed, this is the approach observed with a similar commercially available C-shaped design that has been reported in the literature.[46]

As evident in the discussion above, each magnet design investigated in this work exhibits some extent of trade-off between the various performance measures considered. Hence, the suitability of a particular design depends

on the requirements of the specific application at hand. The primary design goal for the magnet in a rotating linac-MRI system is maximum field uniformity in the imaging region so as to minimize geometric distortions that may complicate the treatment guidance process. Consequently, the open S4 and M4 designs immediately stand out as prime candidates, due to the low magnitude and sparse harmonic content of their magnetic field inhomogeneity patterns. That said, special consideration of the fringe field extent and passive shielding performance is also required. The magnitude of potential field distortions caused by the relative motion of the rotating magnet assembly with respect to nearby ferromagnetic objects (such as an elevator or passive room shielding) is likely to depend on the extent of the stray magnetic field. Thus, the modified four-column design M4 is deemed the most suitable candidate. Although the amount of superconducting wire represents a limited fraction of the total linac-MRI system cost, the M4 design also represents an economical choice.

It is important to note that the pole plate designs considered in this work strictly possess flat opposing surfaces. It is well known that optimal contouring of these surfaces can enhance the magnetic field uniformity in the imaging volume,[24, 25] potentially reducing the conductor volume required for an optimal coil configuration. As was previously noted in the discussion above, it is also understood that the modification and optimization of the basic yoke geometries considered may result in improved uniformity and passive shielding. Nevertheless, the trends observed in this work are still expected to apply. Moreover, the results obtained here have provided a strong practical foundation and useful insight towards the appropriate selection of a candidate geometry for further design optimization.

8.4 CONCLUSION

A unified approach has been presented for the optimal design of several biplanar superconducting magnet assemblies for use in a rotating linac-MRI system with a parallel configuration. In particular, the magnet designs considered are compact in size and possess a large hole bored through the yoke assembly so as to permit an unobstructed path for the treatment beam. Minimum volume coil configurations comprised of seven coil pairs were obtained for each of the designs. Based on the optimization results, it was demonstrated that the geometrical configuration of the yoke structure had a pronounced effect on several performance characteristics. For instance, periodicity and symmetry in the flux return structures exhibited a clear impact on the spherical harmonic decomposition of the residual field distributions in the imaging volume. Furthermore, the selective allocation of yoke material to regions of high induced fields, and along the flux return path, was seen to dramatically improve the passive shielding effectiveness of the yoke design. In comparison to the other geometries investigated, the modified four-column design exhibited a relatively high level of field uniformity in addition to a sparse harmonic decomposition of the associated field pattern. These attributes are particularly desirable since obtaining a homogenous field and minimizing geometric distortions are expected to be of central importance for accurate target tracking with a linac-MRI system. In contrast, the asymmetric nature of the C-shaped design resulted in a large number of mixed harmonic terms. The modified four-column and pill-box geometries exhibited superior passive shielding of the fringe magnetic fields, which may alleviate the risk of detrimental field perturbations in the imaging volume due to magnetic interactions with nearby ferromagnetic objects upon system rotation. Taking these factors into consideration, the modified four-column geometry has been selected as an excellent candidate for further design optimization and incorporation into the linac-MRI design pursued by our group.

8.5 REFERENCES

[1] B. G. Fallone, M. Carlone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, K. Wachowicz, and C. Kirkby, "Development of a linac-MRI system for real-time ART," *Med. Phys.* **34**(6), 2547−2547 (2007).

[2] B. G. Fallone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, and S. Vidakovic, "First MR images obtained during megavoltage photon irradiation from a prototype integrated linac-MR system," *Med. Phys.* **36**(6), 2084−2088 (2009).

[3] B. G. Fallone, "Real-time MR-guided radiotherapy: integration of a low-field MR system," *Med. Phys.* **36**(6), 2774−2775 (2009).

[4] J. Lagendijk, B. Raaymakers, U. van der Heide, J. Overweg, K. Brown, C. Bakker, A. Raaijmakers, M. Vulpen, J. Welleweerd, and I. Jurgenliemk-Schulz, "In room magnetic resonance imaging guided radiotherapy (MRIgRT)," *Med. Phys.* **32**(6), 2067 (2005).

[5] J. J. W. Lagendijk, B. W. Raaymakers, A. J. E. Raaijmakers, J. Overweg, K. J. Brown, E. M. Kerkhof, R. W. van der Put, B. Hardemark, M. van Vulpen, U. A. van der Heide, "MRI/linac integration," *Radiother. Oncol.* **86**(1), 25−29 (2008).

[6] B. W. Raaymakers, J. J. W. Lagendijk, J. Overweg, J. G. M. Kok, A. J. E. Raaijmakers, E. M. Kerkhof, R. W. van der Put, I. Meijsing, S. P. M. Crijns, F. Benedosso, M. van Vulpen, C. H. W. de Graaff, J. Allen, and K. J. Brown, "Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept," *Phys. Med. Biol.* **54**(12), N229−N237 (2009).

[7] J. F. Dempsey, D. Benoit, J. R. Fitzsimmons, A. Haghighat, J. G. Li, D. A. Low, S. Mutic, J. R. Palta, H. E. Romeijn, and G. E. Sjoden, "A device for realtime 3D image-guided IMRT," *Int. J. Radiat. Oncol.* **63**(2), S202, (2005).

[8] C. Kirkby, T. Stanescu, S. Rathee, M. Carlone, B. Murray, and B. G. Fallone, "Patient dosimetry for hybrid MRI-radiotherapy systems," *Med. Phys.* **35**(3), 1019−1027 (2008).

[9] Y. Lvovsky and P. Jarvis, "Superconducting systems for MRI- present solutions and new trends," *IEEE Trans. Appl. Supercond.* **15**(2), 1317−1325 (2005).

[10] M. Razeti, S. Angius, L. Bertora, D. Damiani, R. Marabotto, M. Modica, D. Nardelli, M. Perrella, and M. Tassisto, "Construction and operation of cryogen-free MgB2 magnets for open MRI systems," *IEEE Trans. Appl. Supercond.* **18**(2), 882−886 (2008).

[11] T. Tadic and B. G. Fallone, "Design and optimization of superconducting MRI magnet systems with magnetic materials," *IEEE Trans. Appl. Supercond.* (accepted for publication, 2012).

[12] T. Miyamoto, H. Sakurai, H. Takabayashi, and M. Aoki, "A development of a permanent magnet assembly for MRI devices using Nd-Fe-B material," *IEEE Trans. Magn.* **25**(5), 3907−3909 (1989).

[13] J. H. Battocletti and T. A. Knox, "A permanent magnet for whole-body NMR imaging," IEEE Trans. Magn. **21**(5), 1874−1876 (1985).

[14] T. Miyamoto, H. Hayashi, H. Sakurai, H. Takabayashi, and Y. Ohnishi, "Magnetic field generating device for NMR-CT," U.S. Patent No. 4,679,022 (1987).

[15] I. R. Young, "Nuclear magnetic resonance imaging apparatus," U.S. Patent No. 4,534,358 (1985).

[16] S. Pittaluga, S. Besio, V. Punzo, and A. Trequattrini, "Racetrack coils for dedicated MRI magnets," *IEEE Trans. Appl. Supercond.* **20**(3), 786−789 (2010).

[17] MATLAB, ver. 7.11 (The MathWorks, Natick, 2010).

[18] OPERA-3D, ver. 13.0 (Vector Fields, Oxford, 2009).

[19] J. Simkin and C. W. Trowbridge, "Three-dimensional nonlinear electromagnetic field computations, using scalar potentials," *IEE Proc. B.* **127**(6) 368–374 (1980).

[20] N. B. S. Gloria, M. C. L. Areiza, I. V. J. Miranda, and J. M. A. Rebello, "Development of a magnetic sensor for detection and sizing of internal pipeline corrosion defects," *NDT&E Int.* **42**(8), 669−677 (2009).

[21] M. J. Tannenbaum, A. K. Ghosh, K. E. Robins, and W. B. Sampson, "Magnetic properties of the iron laminations for CBA magnets," *IEEE Trans. Nucl. Sci.* **NS-30**(4), 3472−3474 (1983).

[22] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).

[23] V. I. Lebedev and D. N. Laikov, "Quadrature formula for the sphere of 131-th algebraic order of accuracy," Dokl. Akad. Nauk **366**(6), 741−745 (1999).

[24] T. Tadic and B. G. Fallone, "Design and optimization of a novel bored biplanar permanent magnet assembly for hybrid magnetic resonance imaging systems," *IEEE Trans. Magn.* **46**(12), 4052−4058 (2010).

[25] T. Tadic and B. G. Fallone, "Three-dimensional non-axisymmetric pole piece shape optimization for biplanar permanent magnet MRI systems," *IEEE Trans. Magn.*, **47** (1), 231–238 (2011).

CHAPTER 9
CONCLUSIONS

A compelling approach to the advancement of image-guided radiotherapy (IGRT) involves the novel integration of magnetic resonance imaging (MRI) and external beam radiotherapy (EBRT). Through the rapid acquisition of cinematic magnetic resonance (MR) images bearing exquisite soft tissue contrast, these hybrid devices will permit real-time target localization and online treatment guidance.

Our group at the Cross Cancer Institute (CCI) is currently pursuing the development of an integrated medical linear accelerator (linac) and biplanar MRI system capable of performing advanced real-time adaptive radiotherapy (ART^2). The first prototype of this system has been successfully constructed and tested, for which the linac was mounted on the open end of a biplanar permanent magnet in a perpendicular configuration. In the embodiment currently pursued for the second prototype, the linac is mounted coaxial with the magnet pole of a biplanar superconducting system in the dosimetrically advantageous parallel configuration.

The rotating biplanar linac-MRI concept has introduced two unique challenges with regards to the design of the main magnet assembly. In particular, rotation of this system requires that the magnet dimensions be constrained, while an unobstructed beam path in a parallel configuration requires a large hole to be bored from the magnetic yoke structure. The work presented in this thesis investigated the consequences of these challenges and was specifically aimed at developing novel optimization methods that could be used for the design of both permanent and superconducting magnet systems that addressed them.

In Chapter 5, the finite element method (FEM) was used to simulate the magnetic fields generated by a full-body 0.2 T four-column biplanar permanent magnet assembly. It was demonstrated that reducing the lateral dimensions of this magnet system led to an increase in the magnetic field inhomogeneity near isocenter. This resulted in greater restrictions on the size of the volume over which successful passive shimming is possible, based on the assumed maximum tolerable inhomogeneity of 500 ppm. Specifically, it was shown that flat and single annular ring-shimmed (SARS) pole piece designs were unable to yield suitable levels of field uniformity within a 40 cm diameter spherical volume (DSV). In an attempt to compensate for the effects of lateral size reduction, a nonlinear iterative optimization method based on the FEM was developed and implemented. This method was demonstrated to arrive at optimal axisymmetric and non-axisymmetric (AGP and NAGP) pole piece designs that minimized the magnetic field variation within a spherical region of interest (ROI). The NAGP design resulted in superior levels of field uniformity, due to the ability to correct for the predominantly non-axisymmetric influence of the magnetic yoke structure. As such, the NAGP pole piece was the only design that permitted passive shimming within a 40 cm DSV.

In Chapter 6, the compact magnet assembly investigated in Chapter 5 was modified so as to introduce a large 20 cm diameter hole longitudinally bored through the entire yoke structure. This design would therefore permit an unobstructed beam path for a linac-MRI system with a parallel configuration. It was demonstrated through field calculations with the FEM that the presence of the bore further degraded the magnetic field inhomogeneity of this magnet assembly. Specifically, flat and SARS pole piece designs were shown to be incapable of yielding uniformity levels that permitted passive shimming within either 30 or 40 cm DSVs. Accordingly, the nonlinear optimization scheme developed in Chapter 5 was adapted to arrive at optimized double annular ring-shimmed (DARS), AGP, and NAGP pole piece

designs that permitted passive shimming within a 30 cm DSV. The NAGP design provided superior field uniformity, however, the advantage of this design was less apparent within larger volumes due to the overwhelming inhomogeneity introduced by the yoke bore. Through an examination of the fields within and proximate the magnet assembly, it was also determined that the magnetic field within the bore region reverses direction, peaks in magnitude, and comprises a transverse component that increases with off-axis distance. Consequently, it was judged that the bore interior may not be a suitable location for magnetically sensitive devices.

In Chapter 7, the focus of this thesis shifted towards the design of superconducting magnet systems with magnetic materials. A robust iterative optimization scheme based on the FEM was presented for the design of homogeneous yoked superconducting magnets with minimum volume coil arrangements. In particular, this method is appropriate for the optimization of conduction-cooled cryogen-free magnet systems employing high-temperature superconducting (HTS) coils. Magnets of this type are becoming of increasing interest due to the ability to achieve medium to high field strengths without the need of liquid cryogens or a safety ventilation line. As such, these magnets are especially well suited for use in a rotating biplanar linac-MRI system. The effectiveness of the proposed optimization method was demonstrated with the design of a compact 0.5 T superconducting magnet system with a four-column yoke structure. A solution comprising six coil pairs was rapidly obtained in 11 iterations, resulting in a level of field uniformity that would permit successful passive shimming within a 40 cm DSV.

In Chapter 8, a uniform approach was applied to the optimal design and subsequent evaluation of several full-body 0.5 T superconducting biplanar magnet systems. Standard four-column, H-box, pill-box, C-shaped, and modified four-column designs were considered, with 20 and 30 cm diameter

holes bored from the yoke structures to permit an unobstructed beam path for a linac-MRI with a parallel configuration. The optimization method developed in Chapter 7 was used to obtain minimum volume coil configurations consisting of seven coil pairs for each design. Using the results of the optimization and FEM simulations, it was verified that the yoke geometries had a notable impact on the residual magnetic field patterns and passive shielding performance of the various designs. The greatest field uniformity was observed for the standard and modified four-column designs, for which the associated spherical harmonic decomposition had the fewest number of terms. In contrast, the C-shaped design resulted in the greatest inhomogeneity and a large number of mixed harmonics. The modified four-column and pill-box designs demonstrated superior passive shielding, due to the effective allocation of magnetic material along the flux return path established with these yoke geometries. Magnetic saturation was observed with the other geometries, indicating the potential for improved shielding performance upon redesign. In any case, the relatively low residual inhomogeneity, sparse harmonic content, and superior passive shielding observed with the modified four-column design suggests that this assembly is an excellent candidate for further design and optimization.

It can be concluded based on the research presented in this thesis that the magnet design issues that challenge the integration of a medical linear accelerator and biplanar MRI system can be overcome. Specifically, using the numerical optimization methods developed in this work, the design of compact bored permanent and superconducting magnet systems is feasible. The three-dimensional FEM is an integral tool in this design process, for which flexible and powerful commercially available software packages are readily available for integration with custom optimization scripts.

There is great potential for future extensions to the present work that may further contribute to the development of a rotating biplanar linac-MRI

system. For instance, the proposed optimization methods decidedly neglect constraints on magnetic shielding, as the yoke design does not enter the optimization loop. The magnetic material comprising the yoke structure is highly nonlinear and hence incorporating this aspect into the design process is computationally expensive. A logical next step would then involve determining an efficient method for approximating the magnetic field-related impact of small changes to the geometrical features of the yoke, from the viewpoint of the fringe magnetic fields. Given such a technique, the complete optimization process could then either simultaneously or sequentially address the field shielding and homogeneity goals, as illustrated in Figures 9.1 and 9.2. This functionality would become increasingly valuable

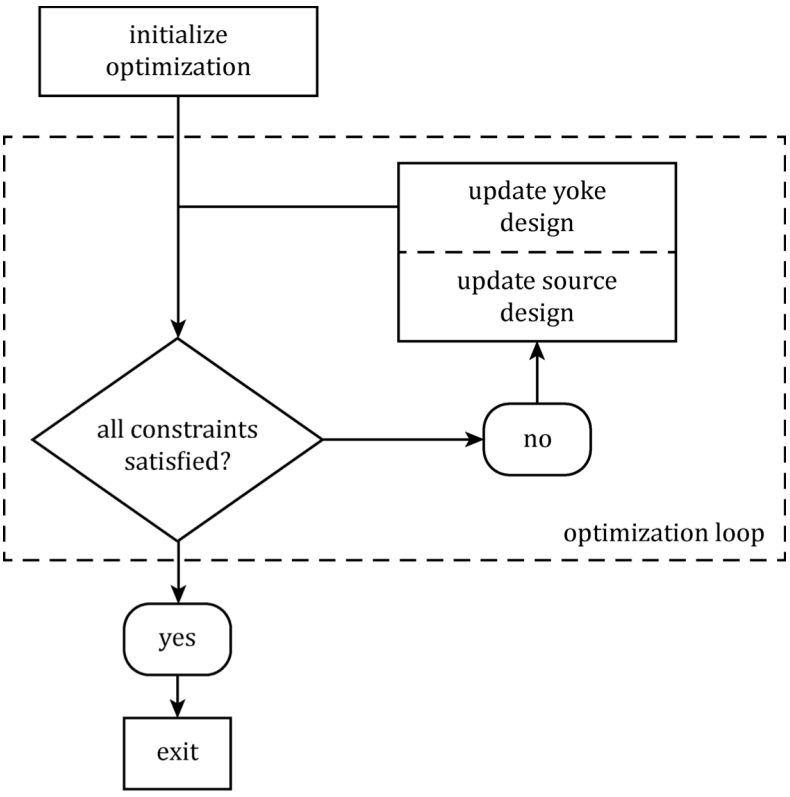


FIGURE 9.1: Iterative optimization flow diagram for the simultaneous treatment of passive shielding and field homogeneity constraints.

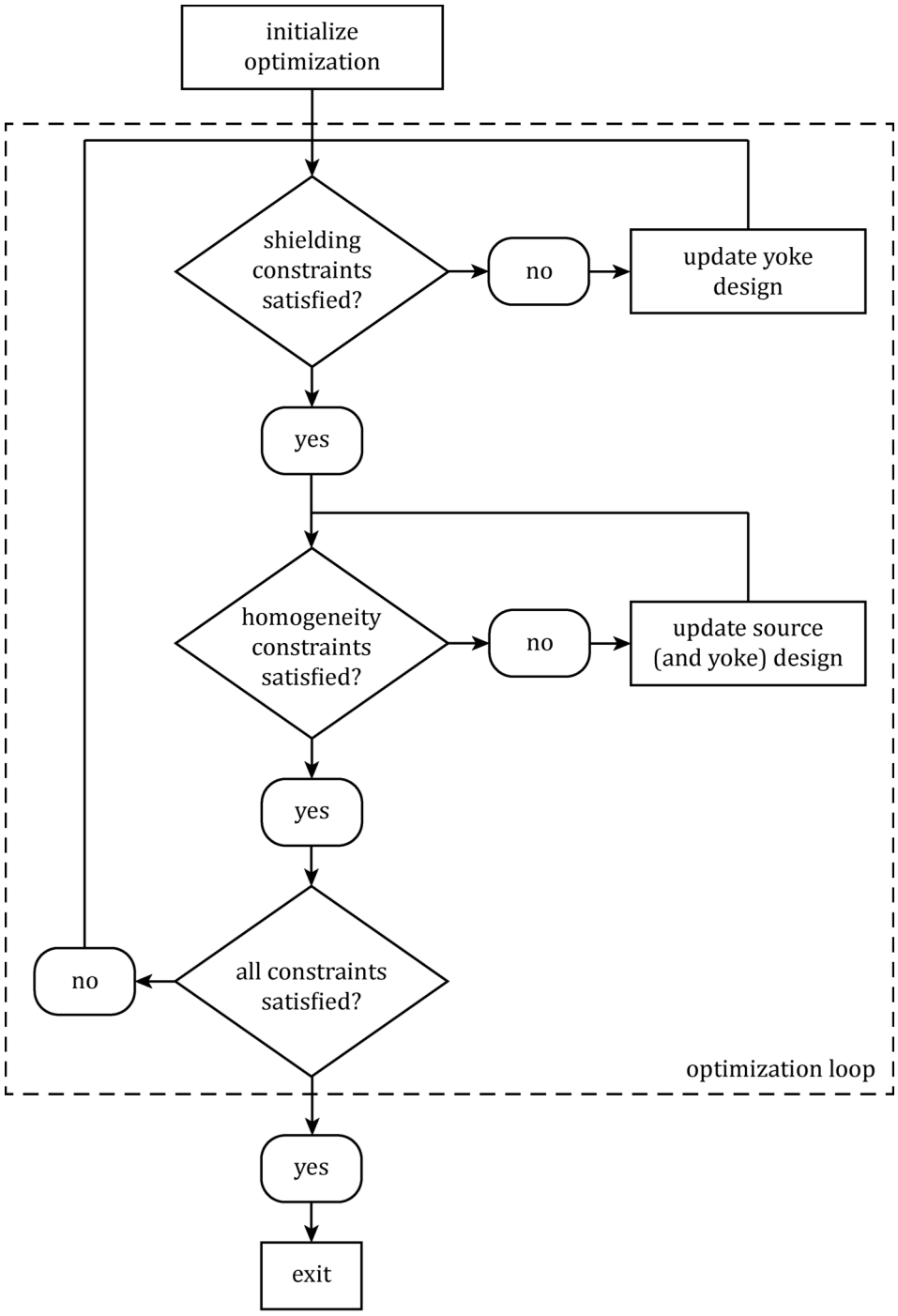


FIGURE 9.2: Iterative optimization flow diagram for the sequential treatment of passive shielding and field homogeneity constraints.

as additional magnetic structures are incorporated into the linac-MRI design, such as passive shielding components for the linac and other magnetically sensitive devices, in addition to the large steel gantry upon which the linac-MRI system is supported and rotates. These auxiliary components will undoubtedly alter the characteristics of a given magnet design and thus determining their optimal arrangement would be beneficial. Ultimately, this would permit the greatest performance of the linac-MRI system in its entirety.