

Simulation and Optimization of High Temperature Superconductor based Electromagnetic systems

by

Rajeev Das



A dissertation submitted to McGill University in partial fulfillment of the
requirements for the degree of Doctor of Philosophy

August 2014

©2014 – Rajeev Das

All rights reserved.

Abstract

The application of superconductors has always fascinated engineers because of their ability to provide very small resistive losses. Electric machines developed using superconductors provide the advantage of lower volume, and higher efficiency compared to conventional machines with similar specifications.

High temperature superconductors (HTS) display the property of superconductivity at relatively higher temperatures compared to conventional superconductors. The use of HTS in electric machines provides substantial benefits in terms of less energy spent on refrigeration. In the field of computational electromagnetics, HTS in low frequency devices pose a few unique challenges that demand different approaches from the existing techniques applied in normal materials and conventional superconducting materials. This dissertation investigates the properties of HTS and then attempts to address various issues that arise while modeling and optimizing HTS based devices.

The high degree of non-linearity in HTS presents significant difficulty for a solution to converge. This thesis models the behavior of HTS based low frequency examples, and suggests an approach that improves the convergence and reduces the computation time. In this effort, the successive substitution method and the Aitken approximation were used to develop the proposed algorithm.

Subsequently, a multiphysics model of HTS based current leads was developed. To analyze the magnetic field problem, a 2D technique was considered, and the analysis of thermal behavior used a 3D approach. This study helped in determining the approach to solve both types of field problems associated with current leads and coupling the two sets of problems.

Computational approaches to solve HTS based problems are still in their infancy and there is a shortage of data in this field. In order to generate more data, which could then be used in optimization or modeling of HTS devices, this work examined statistical

methods that could be applied in such scenarios. In this thesis, aspects such as sampling methods of search space and methods to construct surrogates were researched; the use of Latin hypercube sampling (LHS) and Generalized Regression Neural Network (GRNN) are proposed.

This dissertation ultimately considers a real world example of HTS current leads to address some of the challenges associated with them, specifically the reduction of heat leakage, the material used and the AC loss. To achieve these three goals, multi-objective optimization was considered and Differential Evolution (DE), a stochastic based technique was used to achieve the optimization goals. This facet of the thesis establishes the viability of stochastic methods to optimize systems with multiple goals to satisfy and having HTS a highly non-linear material in them, along with the possible objective functions and constraint relations.

Résumé

L'application des supraconducteurs a toujours fasciné les ingénieurs en raison de leur capacité de fournir de très petites pertes résistives. Les Machines électriques développées en utilisant les supraconducteurs offrent l'avantage de réduire le volume et une plus grande efficacité par rapport aux machines traditionnelles avec des spécifications similaires.

Les supraconducteurs à haute température (HTS) font preuve de propriété de supraconductivité à des températures plus élevées comparées aux supraconducteurs conventionnels. L'utilisation de HTS dans les machines électriques offre des avantages substantiels en termes de réduction d'énergie dépensée sur la réfrigération. Dans le domaine du calcul de propagation électromagnétique du champ, HTS en appareils de basse fréquence posent quelques défis uniques qui demandent différentes approches des techniques autres que celles existantes appliquées dans des conditions normales aux matériaux classiques. Cette thèse est consacrée à l'étude des propriétés de HTS et tente ensuite de traiter de diverses questions liées à la modélisation et l'optimisation d'appareils de HTS.

Le haut degré de non-linéarités de HTS présente une difficulté significative pour qu'une solution converge. Cette thèse modélise le comportement d'exemples HTS à basse fréquence et suggère une approche qui améliore la convergence et réduit le temps de calcul. Dans cet effort, la méthode de substitution successive et l'approximation Aitken ont été utilisées pour développer l'algorithme proposé.

Par la suite, un modèle multiphysique de conducteurs HTS a été développé. Pour analyser le problème de champ magnétique, une technique 2D a été considérée et l'analyse de conduite thermique a utilisé une approche 3D. Cette étude a aidé à déterminer de l'approche pour résoudre les deux types de problèmes de champ associés aux conducteurs et le couplage HTS.

Les approches informatiques pour résoudre des problèmes HTS sont encore dans l'enfance et il y a un manque de données dans ce domaine. Afin de générer plus de données, qui pourraient ensuite être utilisées dans l'optimisation ou la modélisation de dispositifs HTS, ce travail a examiné plusieurs méthodes statistiques qui pourraient être appliquées dans de tels scénarios. Dans cette thèse, des méthodes de sondage d'espace de recherche et les méthodes de construction de substituts ont été explorés; l'utilisation d'échantillonnage d'hypercube latin (LHS) et la Régression Généralisée le Réseau Neuronal (GRNN) est proposée.

Cette thèse considère finalement un exemple du monde réel conducteurs HTS mène pour adresser certains des défis qui leur sont associés; en particulier la réduction de perte de chaleur, les matériaux utilisés et les pertes CA. Pour atteindre ces trois objectifs, l'optimisation multi-objectif a été examinée et l'évolution différentielle (DE), une technique stochastique a été utilisée pour atteindre les objectifs d'optimisation. Cette facette de la thèse établit la viabilité des méthodes stochastiques pour optimiser les systèmes avec plusieurs objectifs à satisfaire et qui contiennent des HTS, de matériaux fortement non linéaires, avec possible fonctions objectifs et relations de contrainte.

Acknowledgements

I would like to thank my advisor Prof. David A. Lowther without whose guidance, this work would not have been possible. I would always be deeply indebted for his encouragement, support and a source of inspiration throughout my research work in the Computational Electromagnetics Laboratory at McGill University.

My gratitude to Prof. Jaime A Ramirez and Prof. Frederico G. Guimaraes from Universidade Federal de Minas Gerais (UFMG) for the numerous stimulating discussions and suggestions on the subject of optimization applied to electromagnetics during my stay in Belo Horizonte, Brazil.

On a more personal note, I thank all my colleagues and fellow graduate students at McGill University for their friendship, advice, and help; many have graduated and gone their separate ways. I would especially like to thank Dr. Prakash Paul, Mr. Adrian Ngoly, Mr. Tapabrata Mukherjee, Mr. Evgeny Kirshin, Dr. Jun Ouyang, Dr. Min Li, Miss Maryam Golshayan, Mr. Ryan Galagusz, and Miss Malika Meghjani.

The financial support of Natural Sciences and Engineering Research Council (NSERC) Canada, McGill University, and the Department of Foreign Affairs and International Trade (DFAIT) Canada is gratefully acknowledged.

Finally, I would like to express my sincere gratitude to my father, Mr. Sadhan Das, and my aunts for their loving support and encouragement. I thank my fiancée Miss Suvolaxmi whom I met during the course of my PhD for her love, support and patience.

I dedicate this work to my father and to the loving memory of my late mother Mrs Tandrima Das.

Table of Contents

List of figures.....	viii
List of tables.....	ix
1 Introduction.....	1
1.1 Motivation and Overview	1
1.2 Research contributions and objectives.....	3
1.3 Thesis outline	5
2 The High Temperature Superconductor.....	7
2.1 Introduction.....	7
2.2 Field properties	10
2.3 Thermal properties	12
2.4 Mechanical properties.....	14
2.5 Summary	15
3 Field modeling involving HTS	16
3.1 Introduction.....	16
3.2 Magnetic field modeling.....	17
3.2.1 The A-V formulation for AC loss determination.....	19
3.2.2 Algorithm to improve the convergence of a field solution	27
3.2.3 Simulations and Results	30
3.3 Thermal modeling.....	34
3.4 Coupling different field problems.....	40
3.5 Summary	42
4 Design process and Optimization involving HTS	43
4.1 Introduction.....	43
4.2 HTS based current leads	44
4.3 Sampling plan and Surrogate technique	45
4.3.1 Design of sampling plans for computer experiments	48
4.3.2 Surrogate methods	51
4.4 Multi-objective based optimization	56
4.5 Differential evolution.....	63
4.6 Simulations and Results.....	67
4.7 Summary	70
5 Conclusion and future work.....	72
5.1 Recommendation for Future Work	74
References.....	76
Appendix.....	86
Appendix A.....	86
Appendix B.....	91
Appendix C.....	94
Appendix D.....	97
Appendix E	103

List of figures

Fig 2.1. The superconducting region resides inside the three critical values and the normal state remains outside. -----	8
Fig 2.2. The behavior of type I and type II superconductors below the critical temperature as the field is varied. The critical magnetic field is temperature dependent, and the value changes as the temperature alters, at T_c the critical field is zero.-----	9
Fig 2.3. The presence of flux vortices in the mixed state of HTS materials, with currents encircling the cylindrical regions. -----	11
Fig 3.1. The screening currents on the edges of a cutaway HTS sample due to an applied field. -----	19
Fig 3.2. A triangular element, the As signify the nodes that are considered for assembling the matrices. -----	20
Fig 3.3. The Pseudo code for the Algorithm. -----	29
Fig.3.4. The magnified view of the discretization present in the HTS region, the dimensions are in meters. -----	31
Fig. 3.5. The plot showing the magnetic field distribution over the entire problem domain, dimension is in millimeters. -----	32
Fig. 3.6. The convergence plot comparing the algorithm, and the SSM. -----	32
Fig. 3.7. The experimental set up showing the $\frac{1}{4}$ section used for the field analysis (not to scale). -----	33
Fig. 3.8. The magnified view of the discretization present, dimension is in meters.-----	33
Fig. 3.9. Comparison between the experimental result [49] and the simulation result.---	34
Fig. 3.10. Heat transfer element inside a problem domain, involving surface heating and convection. -----	37
Fig. 3.11. Flow diagram of the coupled problem. -----	41
Fig. 4.1. (a) A HTS current lead geometry. (b) Schematic representation of gas cooled HTS and Copper current leads arrangement. -----	44
Fig. 4.2. (a) Full factorial sampling method using 8 points, providing 3D, side and top views. (b) Hammersley sampling method using the same number of 8 points with the 3D, side and top view plots. (c) Latin hypercube method using 8 points presenting the 3D, side and top views . -----	50
Fig 4.3. The architecture of a generalized regression neural network. -----	54
Fig 4.4. f_1 and f_2 are objective functions. (a) Convex Pareto front. (b) Non-convex Pareto front. (c) Multiple Pareto fronts denoted by bold line, line with dots and dash arising due to multi-modal condition present in either one or both objective functions. (d) Discontinuous Pareto front. (e) Non-Uniform Pareto front. (f) Deception in the Pareto front, the optimum exists from the bold curve, but the nature of search space favors the region from the dotted curve. -----	57
Fig 4.5. The overall view of the optimization approach. The flow diagram shows the interaction among various components that include the magnetic and thermal field solvers, a DE based optimizer and a parameter list. -----	63
Fig 4.6. The Differential Evolution flow diagram . The process is terminated when there is no appreciable change in the fitness of the population over successive iterations. ---	64
Fig 4.7. The non-optimized lead, with internal and outer radii of 5mm and 10mm respectively, and a length of 100mm. -----	67

Fig 4.8. The distribution of 50 data points using the LHS technique. A 3D representation to visualize the samples conveniently , in the above plots only three parameters are considered at a time.-----	68
Fig 4.9. The quarter section of an optimized HTS lead.-----	69
Fig 4.10. The temperature distribution from a section of the HTS lead. -----	70

List of tables

Table 3.1 Conductivity across the cross section of HTS sample-----	32
Table 4.1 Best solution from the Pareto set. -----	70

CHAPTER 1

1 Introduction

1.1 Motivation and Overview

Electric machines play a significant role in the modern era, whether it is their use in a dishwasher or an entity as big as a ship, life without them is unimaginable. The electrical machine design community has tried over the years, new designs and approaches to create machines that are more efficient, have a high power density, and low noise. In this endeavor the phenomenon of superconductivity provided a promising avenue, as materials exhibiting such behavior have zero resistive loss for direct current (DC) and very small losses for alternating current (AC). In the superconducting state, materials have the potential to allow the passage of high current density; this in turn provides the possibility of constructing machines, which have a lower volume, and higher efficiency compared to a conventional machine with similar specifications.

Superconductivity was discovered by Heike Kamerlingh Onnes in the year 1911 [1], [2], [3] and till the mid 80's, the known materials that showed this property were usually metals such as Mercury, or compounds of metals, kept at cryogenic temperatures. The high cost involved in maintaining extremely low temperatures, usually close to absolute zero, to sustain the property of superconductivity, has been the main reason behind the limited success of this technology in electric machines. The quest to find new materials that could act as superconductors at higher temperatures led to the discovery of High Temperature Superconductors (HTS) two decades ago.

The name HTS is self explanatory, it is a class of materials that attains the property of superconductivity at higher temperatures, compared to the conventional metal based superconductors. They could act as superconductors at temperatures as high as 135K, and are composite materials, usually compounds of copper oxide and other elements. It is due to this advantage of an operational temperature, which is achievable at a low price, that the potential use of HTS materials is being explored actively, and they have made significant contributions. More recently, HTS based machines have been used to drive the propellers of large ships and their application in space exploration is also being investigated [4], [5], [6], [7]. Still their wide scale

use remains elusive because of the cost involved in manufacturing such materials, and their usage in the design of machines is further complicated by the fact that the materials are brittle in nature. In such a scenario, computer based simulations and models could play a crucial role for the analysis and design of such machines. This would not only help in reducing cost, but also facilitate a wider spread of research activity.

As discussed earlier, the superconducting property of HTS allows a large amount of current to flow without any substantial resistance, which in turn could produce high magnetic fields. In the construction of electric machines, the use of HTS provides the following advantages:

- i. High power density - The HTS field windings achieve higher magnetic field densities than those of conventional machines, due to this there is significant reduction in the weight and size, and substantial increase in the torque density [4], [8].
- ii. High efficiency – The machines have reduced rotor ohmic, iron, and windage losses due to zero resistance. Even when the energy component spent on refrigeration is included for efficiency calculations, the gains are in the region of 2% for generators , and as high as 4% or more for motors [9], [10].
- iii. Low noise – They have lower sound emissions than conventional machines because of the absence of iron teeth .
- iv. Low synchronous reactance - HTS air-core (ironless rotor) machines provide greater stiffness during transients [4], [8].

To build computational models of such devices, it is essential that the material properties are investigated first and an accurate representation of the material behavior is prepared. The crystalline nature of HTS materials, introduces a high degree of non-linearity in their magnetic and electrical behavior (refer to Appendix B), and such nonlinearities introduce a major challenge in the modeling of HTS materials. There are a number of representations that have been used to describe the unique behavior of HTS materials, such as the Bean Critical State

(BCS), the E-J, and the Kim models, which involve high order mathematical functions. This, in turn, results in numerical instability and also the computation time is considerable. Apart from this, temperature too plays a crucial role; a slight variation above the critical value of the material could make it lose its superconductivity.

The inherent complications associated with the HTS materials, as described, above get reflected when they are used for constructing electric machines. The critical values, here being the current and the electric field, are dependent on the temperature and the magnetic field. When conditions cause the current, the magnetic field or the electric field to go beyond their critical limits, the material begins to behave as a normal conductor, which in turn may result in damage to the device. In addition to the physical complexity, computational models for such devices have stability issues due to the nonlinearity, and field simulations require considerable time. In brief it could be stated that HTS based systems present a multiphysics problem, in which the relation between the field and the thermal behavior is further compounded due to the high degree of nonlinearity. Under the conditions highlighted so far, the numerical modeling of HTS based devices provides a significant scope for contribution and challenges, and it becomes imperative to investigate how well computational techniques could be applied in this area.

1.2 Research contributions and objectives

The originality of this work lies mainly in the use of ideas present in other disciplines such as Physics, Applied Mathematics, Statistics, and Optimization in the design of electromagnetic devices and when a new material HTS (type Bi-2212) is present. This study considers Bi- 2212, an example of the first generation of HTS technology because it has widely been investigated over the years. The data and test cases involving Bi-2212 are relatively easy to find in the literature compared to the new generation HTS, such as YBCO [1], apart from this, the general behavior of conductivity and other material properties are similar in all the classes of HTS and the ideas that are suggested in this thesis could be extended to them also. It should be emphasized here that the second generation HTS materials provide both performance benefits in terms of higher operating magnetic fields and temperatures, and also reduced cost. An attempt has been made in this work to develop a simple implementation, which will help in the

generalization of the proposed concepts. The following, to the best of the author's knowledge, are the original contributions that this dissertation incorporates:

- i. Development of a methodology that uses an A-V based formulation and FEM to solve a magnetic field problem involving HTS. The proposed algorithm applies successive substitution and the Aitken method which provides improvement in the convergence of field computation. A scenario that comprises HTS and ferromagnetic material is also investigated to test the feasibility of the approach when materials having a non-linear conductivity and a non-linear permeability are present in the problem environment. Derivation and a comparative analysis of derivative free method of successive substitution and the derivative based Newton Raphson method are presented.
- ii. Establishes a method that couples, two different problems in HTS based current leads, an example of a multiphysics scenario that involves the determination of AC loss on one hand and the estimation of the thermal field on the other. A 3D model is used to solve the heat equation, and 2D approach to calculate AC loss that incorporates the effects of magnetic field and temperature on critical current density.
- iii. Proposes possible objective functions and constraints that could be applied to reduce heat leakage, the amount of HTS material used, and AC loss, these desirable attributes are common to HTS based devices. This work considers the concepts in multi-objective optimization (MOO) to address such issues and improve the performance of HTS devices.
- iv. Development of statistical based methods to deal with the problem of data scarcity that persists in the field of HTS, such approaches could provide immense benefits because the material is expensive, and also assist in device modelling and optimization. This work investigates sampling techniques that could be used, and suggests the application of the Latin Hypercube Sampling (LHS) method. Surrogate methods are reviewed, and the use of Generalized Regression Neural Networks (GRNN) is proposed to generate a crude HTS device representation, which is computationally less demanding. Differential

Evolution (DE), a stochastic based technique has been used for the first time in the field of HTS device optimization and the benefits of DE are highlighted.

This dissertation will first attempt to simulate the behavior (magnetic and thermal) of systems containing HTS material, and then try to optimize a simple HTS based device. In these efforts, this thesis explores a number of subjects as mentioned earlier and in the process will make the contributions highlighted above. The use of HTS to construct low frequency machines presents few interesting challenges that are different when compared to electromagnetic devices involving conventional materials and conventional superconductors. This work demonstrates the potential of computational electromagnetics to deal with them , and it will also describe the methods that could be applied in such scenarios.

1.3 Thesis outline

This dissertation is divided into five chapters. In this chapter, the High Temperature Superconducting (HTS) material is introduced by presenting the unique properties of HTS and the advantages such materials have over conventional superconductors. The present chapter also highlights the benefits the use of HTS provides in the construction of low frequency devices, and the difficulties such materials pose in developing computational models. The reasons behind exploring the field of computational approaches in HTS are then discussed, they form the basis for this research work, and the thesis contributions are outlined.

In the next chapter, a brief introduction has been included that highlights the progress made in the field of superconductivity, and also the major successes achieved in recent years in the field of HTS based engineering. This chapter further goes into the details of magnetic field, thermal field and structural behavior of HTS and discusses the necessary physics involved. The ideas mentioned in this chapter form the foundation for the numerical modeling of HTS material which is dealt in the following chapter.

In Chapter three, finite element method (FEM) and A-V formulation is used to determine the AC loss. To implement temporal discretization FDM was considered. A comparison between the mathematical treatment of the successive substitution method and the newton raphson technique

in such a scenario involving HTS has been provided. This work proposes an algorithm that considers successive substitution methods to construct the solvers for HTS materials, and in order to improve the convergence, the use of Aitken method has been suggested. Simulations that use the new algorithm to solve HTS problems are executed and the results are validated. Apart from this, an example that involves both HTS and ferromagnetic material, iron, is examined and the success of the algorithm is verified. This chapter further investigates the approach that could be applied to solve a HTS based multiphysics problem. This work provides a simple framework that considers 2D to solve the magnetic field problem and uses 3D to estimate the thermal behavior and then couples the two problems. The outcome of this approach is tested in the subsequent chapter.

Chapter four first describes HTS based current leads and the mechanism behind their operation. It then describes the possible actions that could be considered to improve the performance of the leads and reduce the cost of manufacturing them, by defining the necessary objective functions and constraints. In this work, multi objective optimization is investigated and the manner in which it could be applied in the present scenario is examined. This chapter also probes the sampling techniques that could be considered in the present optimization process, and identifies the Latin hypercube method as a suitable approach. In order to generate computationally less expensive models of the HTS lead, the use of Generalized Regression Neural Network (GRNN) is studied and the benefits of this approach are stressed. Subsequently, a stochastic method, specifically Differential Evolution (DE), is examined. The advantages are highlighted and then integrated into the proposed optimization framework to obtain the desired goals. Finally, this chapter presents the optimized lead geometry when the optimization framework is applied to HTS current lead problem.

The fifth and the final chapter summarizes and concludes this dissertation work. A few suggestions in regards to further possible exploration of this research area have been included. This chapter is followed by references and appendices.

CHAPTER 2

2 The High Temperature Superconductor

2.1 Introduction

The age of superconductivity in engineering was ushered in 1954, when the first superconducting magnet was designed at the University of Illinois Urbana-Champaign. The next major contribution in this field was again from the same university in 1957, when the Nobel Prize winning theory of superconductivity was published. This theory introduced the well known concept of Cooper pairs, the superconducting carriers described as two electrons with equal or opposite spin and momentum [1]. In the year 1962, Brian D Josephson from Cambridge University discovered a phenomenon unique to superconductors now known as the Josephson effect, which earned him a Nobel Prize. This effect is the tunneling of superconducting current through thin insulating layers separating two superconducting electrodes, which creates a phase difference between electrons in the two electrodes resulting in a potential difference. The Josephson effect forms the basis of the Superconducting quantum interference device popularly known as the SQUID. Another engineering marvel that came into existence during the late 60's with notable success in Germany, and Japan was levitation technology in the form of maglev vehicles, which uses the Meissner effect of superconductors [3].

The first commercial superconducting wire was manufactured in Westinghouse laboratories using niobium and titanium (NbTi) an example of a low temperature superconductor (LTS) in 1962. The first few topologies of rotating machines that were successfully constructed using NbTi in the 70's included DC homopolar and AC synchronous machines [11]. The earliest example of such a device was a 5MW generator for the US Air Force constructed by Westinghouse. The dawn of HTS technology could be attributed to Georg Bednorz and Alex K. Mueller, the duo while working at IBM Zurich Research Laboratory discovered high temperature superconductivity in 1986; the pair was awarded the Nobel Prize in Physics the next year. A joint initiative undertaken by the American Electric Power Research Institute (EPRI) and Reliance Electric Corp. led to the successful demonstrations of HTS based synchronous motors

in 1993; the two motors developed were of 2HP and 5HP capacity [12]. A conceptual design for a 100 MVA HTS based generator was provided by General Electric [13] in 1994 and Westinghouse and American Superconductor Corp. (AMSC) further contributed to HTS generator designs later. Over the years the use of HTS technology for electric machines has evolved and it has grown to an extent where the development of more powerful devices is now feasible [14].

Certain materials behave as superconductors when the temperature, magnetic field and current density are below specific critical values; Fig 2.1 provides the critical surface in 3-dimensional space involving the three. Under such a state, substances exhibit zero resistance and the Meissner effect. It should be highlighted here that in the Meissner effect the magnetic flux is always expelled and the field is zero inside the superconductor, whereas when a conductor is cooled to a very low temperature below the critical value, the field always remains fixed and it could have either zero or a nonzero value. To elaborate the second scenario further, when the temperature of a conductor drops to a value less than the critical temperature and then a magnetic field is applied, the flux does not enter the specimen and is excluded from its interior. On the other hand, if a magnetic field is applied prior to the cooling of a conductor, and subsequently the specimen is subjected to a temperature less than the critical value and then the external magnetic field is removed, the flux inside the sample remains unchanged.

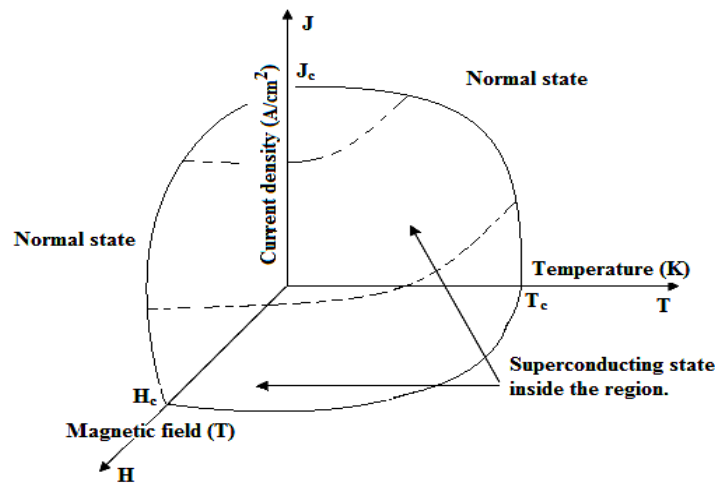


Fig 2.1. The superconducting region resides inside the three critical values and the normal state remains outside.

Superconductors are classified into two categories: i) type I, also known as soft superconductors, have a lower critical field and critical current density values, due to which they have limited practical application; and ii) type II, also referred to as hard superconductors. The distinguishing characteristic of such materials when compared to type I is that the latter always excludes an applied field and behaves as a perfect diamagnetic when the field is less than the critical magnetic field H_c . The flux penetration occurs only when the field exceeds H_c and the specimen starts exhibiting the properties of a normal conductor. In type II superconductors the applied field is completely excluded as observed in type I till a lower critical magnetic field H_{c1} , beyond this the flux begins to penetrate as the field is increased till it reaches an upper critical magnetic field H_{c2} , when the applied field is greater than H_{c2} the sample acts as a normal conductor [3], [15]. Type II superconductors exhibit a mixed state. In such a state, there is a partial flux penetration and presence of normal regions in the material, however superconductivity is not lost, but the specimens no longer display the Meissner effect, the following figure further illustrates these phenomena.

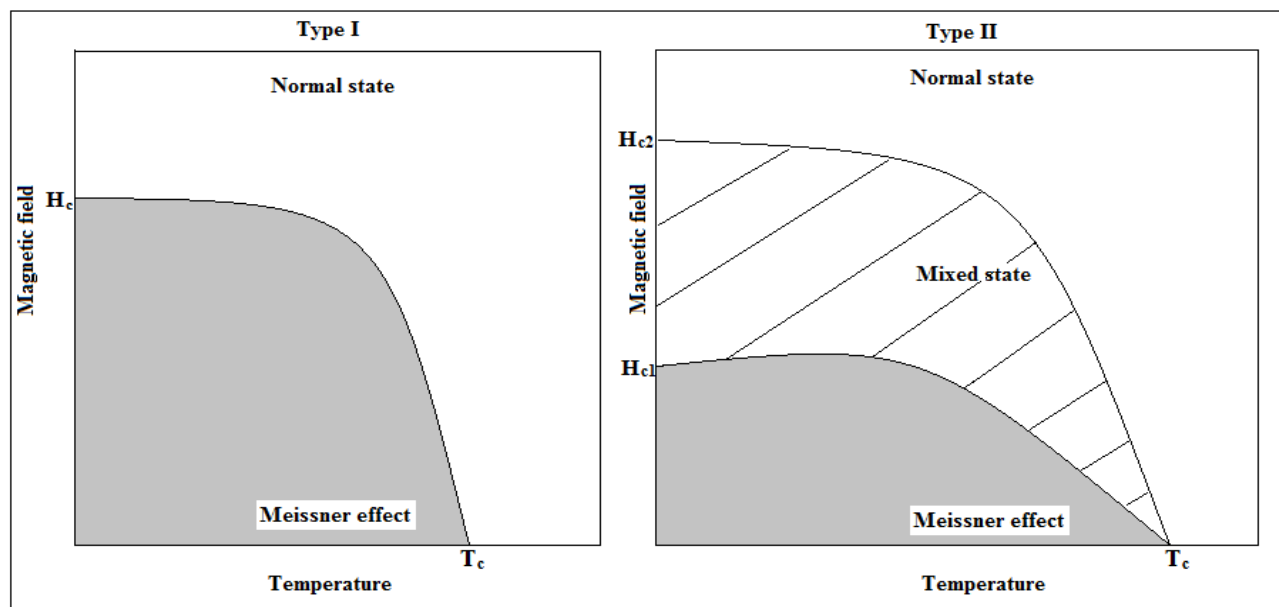


Fig 2. 2. The behavior of type I and type II superconductors below the critical temperature as the field is varied. The critical magnetic field is temperature dependent, and the value changes as the temperature alters, at T_c the critical field is zero.

There are few exceptions to the following classification; pure metals tend to behave as type I superconductors at temperatures close to the critical value and in certain instances may require the application of high pressure to transform the metals to superconductors. Alloys, certain

oxides and even few metals such as Niobium, Vanadium and Technetium behave as type II superconductors; the HTS also fall into this category. In order to model and design superconducting machines, knowledge of both the thermophysical and electromechanical properties of the material are necessary. The following sections will further provide an insight into the magnetic field behavior and the thermal behavior of HTS, and the mechanical properties to a lesser extent.

2.2 Field properties

In order to appreciate completely the behavior of type II superconductors, or in the present scenario, HTS, classical physics has a limited scope and aspects of quantum mechanics should be considered. This section discusses the subject at a fundamental level, and the treatment will help in understanding the general behavior of such superconductors. In order to retain simplicity, the inclusion of every detail involving quantum mechanics that explains such behaviors precisely has been avoided, and it is also beyond the scope of this dissertation.

A magnetic field penetrates into a superconductor to a very small extent and the field falls off exponentially over a mean distance known as the *penetration depth* (λ_L), typical values are less than 0.05 μm . λ_L is estimated using electron mass and electron charge. At a macroscopic level, such values could be neglected and it is safe to assume complete exclusion of magnetic field, but on a scale of λ_L such an approach is not applicable. In the development of an understanding of superconductivity, Ginsburg and Landau introduced the concept of *coherence length*, usually denoted by ξ [3], [15], which is a measure of the distance a pair of electrons should have that ensures interaction between the two. The coexistence of superconductivity and magnetism is dependent on the relationship between λ_L and ξ . In order to categorize superconductors, a parameter known as the Ginsburg–Landau ratio and defined as $\kappa = \frac{\lambda_L}{\xi}$ is used [3], [15]. In type I superconductors such as pure metals, ξ exceeds 0.3 μm and $\kappa < 1$; whereas in type II specimens, ξ is much smaller because the mean free path, a measure of the average distance between collisions experienced by electrons, is reduced and $\kappa > 1$. To be precise the switch between type I and type II occurs at $\kappa \approx 0.707$. An interesting fact to note here is that by altering the electron mean free path the superconducting properties of a material could be changed; such an approach

has been exploited in engineering applications by introducing lattice defects. To understand the mixed state prevalent in HTS, the thermodynamic principle involving minimization of the free energy at equilibrium is considered. Surface energy is present whenever there exist two phases of a material, in the present scenario the normal and the superconducting state. When ξ is small, as in HTS, the surface energy is negative, and the formation of coexisting normal and superconducting regions is favored because the total free energy is reduced. A negative surface energy favors many borders between the two regions to attain an equilibrium configuration. Therefore, in HTS when the applied field exceeds the first thermodynamic critical field H_{c1} small regions of normal state in the shape of flux tubes as shown below known as *fluxons* are formed. The excess field lines are localized in such pockets of the normal sections with induced circulating currents that preserve the superconducting regions outside such cylindrical cores. The fluxons arrange themselves into a regular pattern known as the *Abrikosov lattice* [16], since the fluxons are maintained by circular currents they are also known as flux vortices. The number of such vortices depends on the amount of flux that passes through the material. The theoretical foundation for such quantized flux lines, beginning at a field H_{c1} and the complete penetration at a much higher field H_{c2} was laid by Abrikosov in 1957. In 1962, Bean introduced the first macroscopic model based upon experimental results, known as the Bean's Critical State model (for more refer to Appendix B), and this model also assisted in predicting the hysteresis behavior of such materials [17].

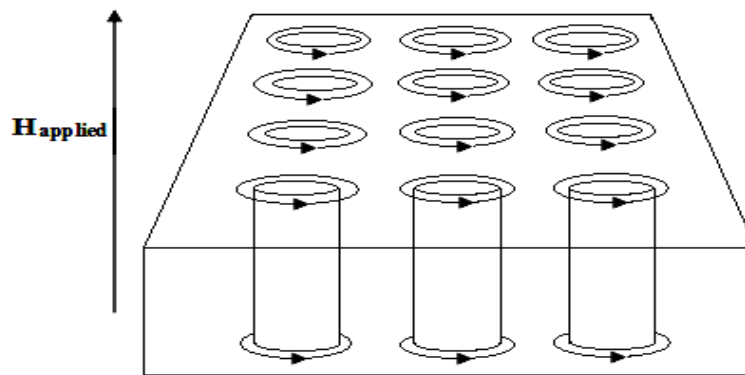


Fig 2.3. The presence of flux vortices in the mixed state of HTS materials, with currents encircling the cylindrical regions.

Whenever a transport current with a density of J flows through a type II superconducting material, it exerts a Lorentz force (F_L) given by $F_L = J \times \phi_0$, where ϕ_0 is the flux quantum. The

Lorentz force tends to make the flux vortices move sideways and such a movement is referred to as *flux flow*. Flux vortices moving with a velocity v_L generate a macroscopic electric field $E = B \times v_L$ giving rise to energy dissipation, where $B = n \phi_0$ with n the density of vortices. It is evident, if the vortices moved then there would be resistance to the flow of current; however the sideways motions are stopped by grain boundaries and impurities. This impediment to the motion of vortices is known as *flux pinning*. As long as F_L is less than the pinning force no dissipation occurs. In HTS, some amount of flux motion persists, which is attributed to thermal activation, this motion is slower in nature and is referred to as *flux creep* [18]. HTS based machines are designed to operate in the flux creep regime, and they avoid the flux flow state as the latter results in heat generation.

2.3 Thermal properties

The generation of heat in HTS can result in a significant temperature increase, which could create localized regions inside the superconductor that are above the critical temperature and display resistance. Such normal sections become an additional source of heat and an uncontrolled growth of normal regions could lead to quench in the worst case. Apart from the scenario of quench, heat generation in HTS based devices burdens refrigeration systems and could affect efficiency. Therefore a study of the thermal behavior of HTS becomes necessary, and properties such as specific heat and thermal conductivity are discussed in this section. The present discussion will not provide a mathematical treatment of the subject, and it will only focus on the fundamental points useful to this topic.

Thermal conductivity is a transport parameter that provides an estimate of the efficiency a material has in carrying heat. Specific heat is a static parameter that presents the information about the internal energy of a material. The temperature of a material is altered whenever heat is taken away or added to it, and when the thermal energy is supplied to a substance, the energy is absorbed either by the crystal lattice known as phonons, or conduction electrons, or both. Specific heat (c) is the partial derivative of internal energy U with respect to temperature, with the volume or the pressure maintained at a constant value (in solids the effects of the two are

indistinguishable), $c = \frac{\partial U}{\partial T}$ having the units J/mol-k. The specific heat of superconductors in the normal state and also of metals at low temperatures is given by

$$c = c_e + c_{lat} = \gamma T + \beta T^3 \quad (2.1)$$

where c_e is the electronic contribution, c_{lat} represents the lattice or phonon contribution, and γ is the Sommerfeld constant. γ provides an experimental estimate of the density of states at the Fermi level [15], [19], [20], [21], and coefficient β is obtained by considering the experimental data, any standard solid state physics text will provide the method to determine these values. To define the relation (2.1), the temperature is usually assumed as $T < 0.1\theta_D$, where θ_D is the Debye temperature [19], the latter is the highest temperature that can be achieved due to a single normal vibration of a crystal. A linear dependency of c_e on T ensures that at extremely low temperatures, c_e has a higher value in comparison to the falling c_{lat} . At room temperatures, the opposite happens, c_e is insignificant when compared to c_{lat} , and equation (2.1) provides a good estimate of specific heat for such materials. Some of the characteristics that are unique to superconductors are:

- i. At $T = T_c$ there is a sharp jump or discontinuity in c .
- ii. The specific heat value increases in the superconducting state near T_c .
- iii. At the lowest temperatures, the specific heat varies exponentially with respect to temperature, the behavior as described by the power law defined in (2.1) is not displayed.

All the above changes arise due to conduction electrons; the contribution due to phonons is negligible at T_c or below it. In HTS, the c_{lat} plays a major role in determining the specific heat and it also fluctuates in the neighborhood of T_c , an effect which is not prevalent in conventional superconductors. Such fluctuations are neglected while modeling HTS; apart from these unique behaviors, HTS do not display discontinuities in the values of c at T_c .

The thermal gradient ∇T (temperature difference between two separate points) imposed across a sample results in the rate of flow of heat Q across a unit cross-section perpendicular to the direction of heat flow, the relation between ∇T and Q is defined using the thermal conductivity k . As in the scenario of specific heat, k has two components, the contributions from phonons

k_{lat} and from conduction electrons k_e [19]. k_{lat} is present in all solids, whereas k_e varies and is dependent on the type of material, for instance in insulators k_e is zero. In conventional superconductors the contribution from k_{lat} above T_c is very small and the transportation of heat is due to k_e . As metals become impure the contribution from k_{lat} increases as is evident in the case of HTS, which are compounds of metals. The total thermal conductivity, even in the best crystals of HTS, is predominantly due to k_{lat} . When the temperature falls below T_c , k drops sharply in conventional superconductors and in certain instances the decrease in the magnitude could be four times the value present in the normal state. In HTS the behavior of k is substantially different below T_c , the value first increases and peaks at $\frac{T_c}{2}$ and then there is a rapid decrease. k is less dependent on temperature in the normal state for HTS and has low values.

2.4 Mechanical properties

This section will provide a brief overview of the structural properties of HTS, and the relevant concepts that are used to understand such aspects. Since this work does not deal with the mechanical properties, the discussion will not go into great detail. The inclusion was considered necessary in order to have a complete treatment of the subject matter, i.e., the use of HTS in the design of electromagnetic devices.

In many practical applications, it is significant for the conductors be able to sustain appreciable stress. For use in motors and generators, the HTS are first turned into wires, and are then wound in the form of coils. During the operation such coils are subjected to large rotational forces and variation in the working temperatures. Under such circumstances, some of the desirable features include a ductile and a flexible material, and the materials should also have considerable fatigue strength and the ability to provide a consistent behavior over long durations. Stress and strain are the fundamental parameters that define the structural behavior of any material. Stress is the force per unit area applied to a material, whereas strain is the resulting deformation. When the amount of stress is small, the deformation is elastic in nature, and materials return to the initial shape when the stress is reduced or removed. When the stress is too large, the change in the shape of the material is irreversible, and such a change in form is known as *plastic deformation*. If a stress vs. strain plot is considered for any material, then there exists a point in the curve known as the

yield point where there is a sharp bend. Any stress or strain beyond the yield point results in the material breaking. The stress that causes such a break is known as the *tensile strength*, and when a break occurs due to the strain, it is known as *fracture elongation* [19]. Fatigue quantifies the response of a material to time varying stress or strain, provided the stress does not exceed the tensile strength. A fracture occurs when tiny cracks begin to grow in the material, such cracks propagate and when they exceed a certain critical length, the material breaks. In metals there is always a presence of a distinct plastic state prior to the occurrence of material failure. Such a behavior does not exist in HTS, they exhibit elasticity right up to the point where they break, and it is due to this they are brittle in nature. HTS also demonstrate little ductility. Such issues in terms of material strength and limitations with HTS have been a major hurdle in using their full potential in the field of machine design. The study of the mechanical properties of HTS could further help in the development of their engineering applications; this area still lacks a detailed investigation. Apart from this, the effects of temperature and magnetic field on structural behavior have not been adequately researched.

2.5 Summary

This chapter presented the developments made in the field of superconductors over the years, and it also highlighted some of the advances achieved in low frequency electromagnetic devices involving superconductivity. A discussion on theoretical aspects associated with the magnetic field and thermal properties was included that provided a detailed treatment of the subject. Properties such as flux flow and flux creep that are peculiar to HTS were discussed. The next chapter will deal with the subject of numerically solving problems associated with HTS. A greater emphasis would be laid in using FEM to deal with such problems and how to improve their convergence. The discussions will also introduce a technique to couple magnetic field and thermal field problems in such scenarios.

CHAPTER 3

3 Field modeling involving HTS

3.1 Introduction

The previous chapter highlighted the properties of HTS materials with an emphasis on the type II superconductors. The following sections will describe the mathematical relations involved, and provide the derivations necessary to design computational models for estimating the magnetic and the thermal fields associated with HTS based devices. This will also provide an insight into the possible coupling mechanism of the magnetic and thermal field problems involved in such examples. There exists abundant resources in the scientific literature that describe various approaches for modeling conventional materials [22], [23], [24]. Such ideas could be extended to HTS, but not all methods are applicable because of the differences that exist between the two, which pose a different set of challenges. The distinguishing factors could be summarized as i) HTS materials have a non-linear conductivity and linear permeability, whereas the conventional materials have their behavior reversed; ii) properties such as the critical values of magnetic field and current density vary rapidly with temperature in HTS; iii) the non-linearity in conductivity is extreme for HTS, as demonstrated by the exponential term in the power law [25]. Apart from this, HTS materials display phenomena such as granularity, flux creep, and flux flow, which further complicate determining their behavior precisely.

There are a number of numerical techniques available for the analysis of electromagnetic field problem. These include the finite difference method (FDM) [26], method of moments (MOM) [27], finite element method (FEM), and the meshless methods (MM) [28], [29] that could be considered in a HTS scenario. Each method has its own advantages and shortcomings, for instance FDM is easy to understand, and implement, whereas, it is not suitable for a field with a rapidly changing gradient, or for problems with curved boundaries. FEM is the most developed and widely used approach in computational software packages for low frequency applications; it is well suited to problems with complicated geometries and a complex distribution of media, but the mesh generation process involved is computationally expensive. In recent years the use of FEM in 3D HTS simulations involving different formulations has been investigated by the EPEC

superconductivity group at the University of Cambridge, this work highlighted the associated benefits and shortcomings [30]. The meshing could be avoided in MOMs, but this numerical procedure in electromagnetics is still at a nascent stage, and there is considerable research work going on in this area. MOM uses the integral form and is suitable for open regions, as it could easily manage the truncation issue in such instances; the method is usually applied to scattering problems.

This work considers the domain method FEM, which first involves discretization of the whole domain by regular elements. Subsequently, the weighted residual approach or a variational principle is used to derive algebraic equations for the partial differential equations (PDE) corresponding to a specific problem. The discussions below will provide details for solving 2D magnetic field problems, and 3D thermal field problems using FEM and the weighted residual method.

3.2 Magnetic field modeling

The electromagnetic phenomena associated with HTS materials are explained by Maxwell's equations. In order to model the field behavior of a HTS device carrying AC current, a popular technique involves estimation of the AC losses. The solution to this kind of problem requires solving either the self-field loss due to the AC transport current, or the magnetization loss produced by an alternating external field [31], [32], [33]. The hysteresis component is dominant in the superconducting state, and both kinds of AC losses, i.e., the magnetization and the self-field loss could be attributed to it. In magnetization loss, the currents are induced at the surface similar to the skin effects found in normal conductors that could be determined using Faraday's law, and there is partial flux penetration. Unlike normal conductors, HTS have infinite conductivity, which results in screening currents even at zero frequency. The induced current in HTS is equivalent to the critical current density J_c , which is dependent on the magnetic field and it decreases as the latter increases, on the other hand, in normal skin effect the induced current density is proportional to the field amplitude provided the frequency is constant. When the applied field is reversed, the presence of the pinning property and altered current density prevents the flux from tracing the same path resulting in the hysteresis phenomenon. The flow of

energy in an isolated HTS sample experiencing such an effect could be obtained using the Poynting vector $\mathbf{P}$ as shown below.

$$\mathbf{P} = \mathbf{E} \times \mathbf{H} \quad (\text{W}) \quad (3.1)$$

where $\mathbf{E}$ is the electric field and $\mathbf{H}$ is the magnetic field intensity. With V as the volume enclosed by a closed surface S , and Q representing the total energy in V , the rate of energy loss is

$$-\frac{dQ}{dt} = \oint \mathbf{P} \cdot d\mathbf{S} = \int_V \nabla \cdot \mathbf{P} dV \quad (3.2)$$

the above equation is obtained using the divergence theorem. The $\nabla \cdot \mathbf{P}$ is replaced by $\mathbf{E}$ and $\mathbf{H}$ in equation (3.3). Subsequently, the vector identities and Maxwell's equations are considered (for more refer to Appendix A), resulting in the following relations for a magneto-quasistatic state

$$\nabla \cdot \mathbf{P} = \nabla \cdot (\mathbf{E} \times \mathbf{H}) \quad (3.3)$$

$$= (\nabla \times \mathbf{E}) \cdot \mathbf{H} - (\nabla \times \mathbf{H}) \cdot \mathbf{E} \quad (3.4)$$

$$= -\frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{H} - \left(\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) \cdot \mathbf{E} \quad (3.5)$$

considering integrals on both sides, and neglecting the displacement current term, the loss per cycle is given by

$$Q = \mu_0 \oint dt \int_V \mathbf{H} \cdot \frac{\partial \mathbf{M}}{\partial t} dv + \oint dt \int_V \mathbf{J} \cdot \mathbf{E} dv \quad (\text{J}) \quad (3.6)$$

where $\mathbf{M}$ is the magnetization due to the bounded currents furnished by the constitutive rule $\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M})$.

The presence of granularity in HTS results in two kinds of critical current density: the intra current $\mathbf{J}_{\text{cintra}}$ that acts within the boundary of the grains and the inter current, $\mathbf{J}_{\text{cinter}}$, which flows across the boundaries. In modeling HTS based devices, it is the $\mathbf{J}_{\text{cinter}}$ that should be considered, because this component flows through the entire bulk sample creating the screening phenomena when a field is applied. Moreover, in the design of electrical machines, finer samples are desirable where the minute granules exist more as large grain segments, and in such specimens the $\mathbf{J}_{\text{cinter}}$ plays a dominant role [33]. This work considers the inter current in all the formulations and is denoted by $\mathbf{J}_c$. In equation (3.6), the first term in the right hand side is a resultant of the

bounded current, and is not considered for modeling the field behavior; the second term is due to the free current.

In self-field loss, the passage of the AC transport current generates a magnetic field around the HTS. The field partially penetrates the superconductor and the screening currents are generated. The cyclic nature of the signal gives rise to hysteresis loss as described above. In both the losses i.e., the magnetization and the self-field loss, the dissipative energy involved is attributed to the screening currents, which are used for moving the flux lines and depinning them. Such a form of energy is converted to heat which is undesirable, and could even damage HTS based machines. The following figure depicts the screening current in the white region due to the presence of a varying field; the dark section signifies a region without any field inside.

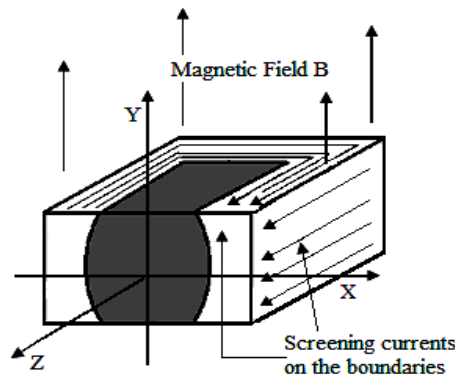


Fig 3.1. The screening currents on the edges of a cutaway HTS sample due to an applied field.

3.2.1 The A-V formulation for AC loss determination

As highlighted earlier, the ideas used for determining the behavior of conventional materials could be extended to model superconductors, and this section describes one such approach namely the A-V form. The A-V represents the magnetic vector potential and the electric scalar potential respectively, and is used to estimate the eddy current behavior or the diffusion phenomena. The other procedure to model the eddy currents is the electric vector-magnetic scalar potential method, popularly represented as the T- Ω form [34]. There are a number of factors that determine the technique to be considered for a given problem. The computation time tends to increase as the number of variables to solve rises. It is also observed that if the

derivatives are not smooth, the solution oscillates and may never converge. Apart from this, the conditionality of the associated matrices, the number of times to differentiate in order to obtain a value and the ease of writing a computer program are the other essential considerations.

When the geometrical dimension is smaller than the wavelength λ (dimension $\ll \lambda = c/f$), where c is the speed of light and f is the frequency of the applied signal, low frequency approximations are utilized. To calculate eddy currents such approximations are considered, where the displacement current is ignored and ϵ_0 is set to 0. In order to implement equations based on this idea, the magnetic vector potential $\mathbf{A}$ was used, the curl of which delivers the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$, and it provides the additional benefit that the condition $\nabla \cdot \mathbf{B} = 0$ is automatically satisfied. In addition to this, to simplify many practical problems, the scenarios are solved in 2D, which allows $\mathbf{A}$ to be scalar with only one component, and in such a case it is sufficient to solve for the single component [35], [36].

The associated diffusion equation in terms of $\mathbf{A}$ using the $\mathbf{A}$ - V form to determine the self-field loss in a HTS sample having an infinite length along the z -axis is given by

$$\frac{1}{\mu_0} \nabla \times \nabla \times \vec{\mathbf{A}}_z + \sigma_z \frac{\partial \vec{\mathbf{A}}_z}{\partial t} + \sigma_z \frac{dV}{dz} = 0 \quad (3.7)$$

where μ_0 is the permeability of free space, σ_z is the conductivity of the HTS, and V represents the scalar potential. It is assumed that the current flows along the z -axis, and the above equation is obtained using Ampere's law and the relation $\mathbf{B} = \nabla \times \mathbf{A}$ (for further details, refer to Appendix A)

with $\mathbf{E}_z = -\frac{\partial \vec{\mathbf{A}}_z}{\partial t} - \nabla V$. The 2D form of equation (3.7) is represented in the following manner:

$$\frac{1}{\mu_0} \left(\frac{\partial^2 \vec{\mathbf{A}}_z}{\partial x^2} + \frac{\partial^2 \vec{\mathbf{A}}_z}{\partial y^2} \right) + \sigma_z \frac{\partial \vec{\mathbf{A}}_z}{\partial t} + \sigma_z \frac{dV}{dz} = 0 \quad (3.8)$$

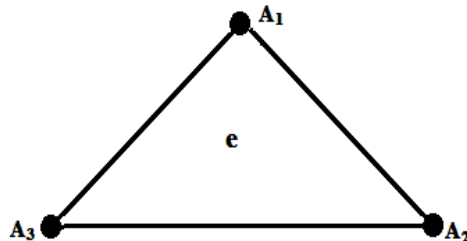


Fig 3.2. A triangular element, the A s signify the nodes that are considered for assembling the matrices.

To numerically solve the 2D electromagnetic problem using FEM, a spatial discretization involving triangular elements as shown in Fig 3.2 was used. In order to obtain the element form of equation (3.8) and the residual formulation, the Galerkin method was used (for more refer to Appendix C), which resulted in the following expression

$$[R_z]_{3 \times 1}^e = [S]_{3 \times 3}^e [A_z]_{3 \times 1}^e + \sigma_z^e [K]_{3 \times 3}^e d[A_z]_{3 \times 1}^e / dt - \sigma_z^e [Q]_{3 \times 1}^e E \quad (3.9)$$

where R_z is the residual matrix and $\mathbf{E} = -\nabla V$ was assumed to be uniform over the HTS domain, the coefficient matrices associated with each element are presented below

$$S_{ij} = \frac{1}{\mu_0} \iint \left(\frac{\partial N_i}{\partial x} \times \frac{\partial N_j}{\partial x} + \frac{\partial N_i}{\partial y} \times \frac{\partial N_j}{\partial y} \right) dx dy = \frac{1}{\mu_0} \frac{a_i a_j + b_i b_j}{4\Delta^e} \quad (3.10)$$

$$a_i = y_j - y_k, \text{ and } b_i = x_k - x_j \quad (3.11)$$

$$K_{ij} = \iint N_i N_j dx dy = \frac{\Delta^e}{6} \text{ when } (i = j), \text{ or } \frac{\Delta^e}{12} \text{ when } (i \neq j) \quad (3.12)$$

$$Q_i = \iint N_i dx dy = \frac{\Delta^e}{3} \quad (3.13)$$

N in the above equations signifies the first order shape functions, the subscripts i, j, k denote the nodes, and Δ^e is the area of the triangular element. Equation (3.11) provides the difference between the coordinate values of the nodes in an element. The temporal discretization was achieved using the FDM, and a backward time difference was applied

$$[R_z]^e = [S]^e [A_z]_{n+1}^e + \sigma_z^e \frac{1}{\Delta t} [K]^e ([A_z]_{n+1}^e - [A_z]_n^e) - \sigma_z^e [Q]^e E_{n+1} \quad (3.14)$$

in the above relation n and Δt are the time step number and time step size respectively; it is assumed $\frac{d[A_z]^e}{dt} \approx \frac{1}{\Delta t} ([A_z]_{n+1}^e - [A_z]_n^e)$ and σ_z remained constant inside each element. In self-field loss determination, the variable E_{n+1} is unknown and is calculated using the transport current as described below

$$I(t) = \iint_S \sigma_z \left(\mathbf{E} - \frac{\partial \vec{A_z}}{\partial t} \right) dS \quad (3.15)$$

where $I(t)$ is the instantaneous value of current and S denotes the area of HTS. The above equation leads to the following relation after rearranging the terms

$$\iint_S \sigma_z \mathbf{E} \, dS = \iint_S \sigma_z \frac{\partial \vec{A}_z}{\partial t} \, dS + I(t) \quad (3.16)$$

$$E \cdot \sum_{e=1}^{N_{\text{hts}}} \sigma_z^e \Delta^e = \sum_{e=1}^{N_{\text{hts}}} \iint_e \sigma_z^e \frac{dA_z^e}{dt} dx dy + I(t) \quad (3.17)$$

as performed earlier, temporal discretization using FDM and a backward time difference was applied to (3.17)

$$E_{n+1} \cdot \sum_{e=1}^{N_{\text{hts}}} \sigma_z^e \Delta^e = \sum_{e=1}^{N_{\text{hts}}} \left[\frac{\sigma_z^e [Q]^e{}^T}{\Delta t} ([A_z]_{n+1}^e - [A_z]_n^e) \right] + I_{n+1} \quad (3.18)$$

N_{hts} above is the number of finite elements in the HTS domain, the unknown variable E_{n+1} in (3.14) is substituted by

$$E_{n+1} = \frac{[Q]^e{}^T}{\Delta^e \Delta t} ([A_z]_{n+1}^e - [A_z]_n^e) + \frac{I_{n+1}}{\sigma_z^e \Delta^e N_{\text{hts}}} \quad (3.19)$$

The non-linearity in the residual formulation (3.14) arises due to the conductivity term σ_z that necessitates the use of an iterative approach, σ_z is dependent on the current density and the electric field as shown below, and was determined using the constitutive E-J power law (for more refer Appendix B)

$$\sigma_z^e = \frac{1}{\frac{E_c}{J_c} \left(\frac{E}{E_c} \right)^{1-\frac{1}{\alpha}} + \rho_0} \quad (3.20)$$

where E_c is known as the critical electric field, a test criterion, and is usually assigned a value of 1 V/cm, and E is the average value of the electric field in an element. To maintain the superconducting state, current densities should be lower than the critical current density J_c . The latter is determined from DC measurements with the electric field set to E_c , α is dependent on the pinning energy, and typically the value is in the range 10 to 20 [37]. Depending on the problem under consideration, the effects of the magnetic field $\mathbf{B}$ on J_c should be included in calculating σ_z as described in Chapter 2 and in Appendix B. There could be numerical instability at times, because the calculated conductivity values can be very large, to avoid such a scenario; a small value of resistivity $\rho_0 = 10^{-15} \Omega / \text{m}$, as suggested in [38], is introduced to prevent divide by zero

cases. The subsequent discussions provide an insight into the mathematical approach to modeling a scenario in the presence of a varying external field, the procedure involved is similar to the analysis presented so far.

To estimate the magnetization loss using the **A-V** form, the following expression is considered

$$\frac{1}{\mu_0} \nabla \times \nabla \times \vec{A}_z + \sigma_z \frac{\partial \vec{A}_z}{\partial t} - \vec{J}_z = 0 \quad (3.21)$$

the parameters μ_0 and σ_z are the permeability of free space, and the conductivity of the HTS respectively, $\vec{J}_z$ is the current density that creates the alternating field and is along the length of the component in the z-axis.

$$\frac{1}{\mu_0} \left(\frac{\partial^2 \vec{A}_z}{\partial x^2} + \frac{\partial^2 \vec{A}_z}{\partial y^2} \right) + \sigma_z \frac{\partial \vec{A}_z}{\partial t} - \vec{J}_z = 0 \quad (3.22)$$

On using spatial discretization and the Galerkin method (for further details refer to Appendix C), the residual equation in matrix form for an individual element is given by

$$[R_z]_{3 \times 1}^e = [S]_{3 \times 3}^e [A_z]_{3 \times 1}^e + \sigma_z^e [K]_{3 \times 3}^e d[A_z]^e / dt - [Q]_{3 \times 1}^e \quad (3.23)$$

a first order function was considered as the interpolation function, and the coefficient matrices presented above are derived in similar manner to that performed previously for the self-field loss scenario, the geometric constants are identical in both the instances, which allows this flexibility. The temporal discretization provides the following equation

$$[R_z]^e = [S]^e [A_z]_{n+1}^e + \sigma_z^e \frac{1}{\Delta t} [K]^e ([A_z]_{n+1}^e - [A_z]_n^e) - [Q]^e \quad (3.24)$$

In order to accomplish clarity, the structures of the stiffness and the source matrices involved in the residual equation above are shown below.

$$S = \frac{1}{4 \Delta^e \mu_0} \begin{bmatrix} a_1^2 + b_1^2 & a_1 a_2 + b_1 b_2 & a_1 a_3 + b_1 b_3 \\ \vdots & a_2^2 + b_2^2 & a_2 a_3 + b_2 b_3 \\ \text{symmetric} & \dots & a_3^2 + b_3^2 \end{bmatrix} \quad (3.25)$$

$$\sigma_z^e \frac{1}{\Delta t} K = \frac{\sigma_z^e \Delta^e}{12\Delta t} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad (3.26)$$

$$Q = \frac{\Delta^e}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad (3.27)$$

The elemental matrices are coefficients of the unknown variables, they are assembled to form the global matrix SS, the vectors associated with the known terms are moved to the right hand side, resulting in the vector U, which subsequently provides the following standard form of linear system

$$SS A = U \quad (3.28)$$

where A is the vector potential to be calculated. The discussions so far presented the treatment for 2D Cartesian coordinates using the FEM and the Galerkin method; there are other important geometrical scenarios such as 3D problems and axisymmetric cases, where the fundamental approach remains the same, except with some variations, the reference texts [35], [36] provide an insight to deal with such examples using FEM.

A simple technique to solve the equation (3.28) iteratively is the successive substitution method (SSM) also known as the *fixed-point iteration* [39], [40] that attempts to obtain the solution using linear estimates of the non-linear E-J relation. It involves an initial approximation of A for all the nodes, E is then computed for individual elements using the following differential technique, and the relations are considered according to the problem under consideration

$$E = -\frac{1}{3} \cdot \frac{1}{\Delta t} \sum_{i=1}^3 (A_{z,n+1,i} - A_{z,n,i}) + E_{n+1} \quad (3.29)$$

$$E = -\frac{1}{3} \cdot \frac{1}{\Delta t} \sum_{i=1}^3 (A_{z,n+1,i} - A_{z,n,i}) \quad (3.30)$$

the conductivity σ_z^e is evaluated using (3.20) and it is then used to calculate the contribution to the coefficient matrices. The global matrix SS and the source vector U are assembled as described earlier and the necessary boundary conditions are imposed. Equation (3.28) is solved

and the new value of A is determined, the iterative procedure is terminated if the following error criterion is satisfied, else it is repeated using the values of A from the previous iteration

$$\text{Error} = \sum_{i=0}^M \sqrt{(A_{z,i,\text{new}} - A_{z,i,\text{old}})^2} / \sum_{i=0}^M \sqrt{(A_{z,i,\text{new}})^2} \leq 10^{-5} \quad (3.31)$$

the M above signifies the total number of nodes used in the analysis. The SSM converges in most scenarios, but it has a linear convergence rate as a result of which the procedure is slow. This thesis work uses the substitution method with some modifications in the general iterative algorithm discussed above to solve problems involving HTS, the next section will present the mathematical details of the procedure implemented.

A discussion on the use of an alternative iterative approach known as the *Newton Raphson* (NR) method, which is a special case of the more general SSM has been included here to provide a complete treatment of the subject. This research work does not investigate in detail the use of NR in such scenarios, but it would highlight the procedure to implement it in HTS problems. The quadratic convergence is the main advantage NR has over the successive technique, but it has continuity constraints on functions that are difficult to satisfy at times. In addition to this, an appropriate initial guess is necessary for the solution to converge. The magnetization loss example has been used here to explain the mathematical formulation when NR is involved, considering the residual equation (3.24), the partial derivative with respect to the unknown variable A_z is

$$\frac{\partial R_i^e}{\partial A_{z,n+1,j}} = S_{ij} + \sigma_z^e \cdot \frac{1}{\Delta t} \cdot K_{ij} + \left(\frac{1}{\Delta t} \sum_{m=1}^3 K_{ik} (A_{z,n+1,m} - A_{z,n,m}) \right) \cdot \frac{\partial \sigma_z^e}{\partial A_z} \quad (3.32)$$

with i and j = 1, 2, 3, and the conductivity derivative above is determined using equation (3.20) in the following manner

$$\frac{\partial \sigma_z^e}{\partial A_{z,n+1}} = (\sigma_z^e)^2 \left[-\frac{E_c}{(J_c)^2} \frac{\partial J_c}{\partial A_{z,n+1}} \left(\left(\frac{E}{E_c} \right)^{1-\frac{1}{\alpha}} \right) + \frac{1}{J_c} \cdot \left(1 - \frac{1}{\alpha} \right) \left(\frac{E}{E_c} \right)^{-\frac{1}{\alpha}} \frac{\partial E}{\partial A_{z,n+1}} \right] \quad (3.33)$$

the first term inside the square bracket has a derivative associated with J_c , this occurs due to **B**. Such a treatment is problem dependent as the variation in J_c due to B is small below a limiting

value B_0 , and the effect on J_c could be neglected. The procedure to obtain the current density and the electric field derivatives is presented below

$$\frac{\partial J_c}{\partial A_{z,n+1}} = \frac{\partial J_c}{\partial B_x} \cdot \frac{\partial B_x}{\partial A_{z,n+1}} + \frac{\partial J_c}{\partial B_y} \cdot \frac{\partial B_y}{\partial A_{z,n+1}} \quad (3.34)$$

$$\frac{\partial J_c}{\partial B_x} \approx \frac{J_c(B_y, B_x + \Delta B_x) - J_c(B_y, B_x)}{\Delta B_x} \quad (3.35)$$

$$\frac{\partial J_c}{\partial B_y} \approx \frac{J_c(B_x, B_y + \Delta B_y) - J_c(B_x, B_y)}{\Delta B_y} \quad (3.36)$$

$$B_y = -\frac{\partial A_{z,n+1}}{\partial x} = -\frac{a_1 A_{z,n+1,1} + a_2 A_{z,n+1,2} + a_3 A_{z,n+1,3}}{2\Delta^e} = -\frac{[a]^T}{2\Delta^e} [A_z]_{n+1}^e \quad (3.37)$$

$$B_x = -\frac{\partial A_{z,n+1}}{\partial y} = -\frac{b_1 A_{z,n+1,1} + b_2 A_{z,n+1,2} + b_3 A_{z,n+1,3}}{2\Delta^e} = -\frac{[b]^T}{2\Delta^e} [A_z]_{n+1}^e \quad (3.38)$$

$$\frac{\partial B_y}{\partial A_{z,n+1}} = -\frac{[a]^T}{2\Delta^e}, \text{ and } \frac{\partial B_x}{\partial A_{z,n+1}} = -\frac{[b]^T}{2\Delta^e} \quad (3.39)$$

to determine the electric field derivative, equation (3.30) was considered

$$\frac{\partial E}{\partial A_{z,n+1}} = -\frac{1}{3} \frac{1}{\Delta t} \quad (3.40)$$

The matrix system to solve a NR based iterative technique is comprised of the Jacobian $\left[\frac{\partial R_z}{\partial A_z}\right]$ assembled by using (3.33), the unknown vector $[\delta A_z]$, and $[R]$ the residual in the right hand side

$$\left[\frac{\partial R_z}{\partial A_z}\right] [\delta A_z] = -[R] \quad (3.41)$$

R is derived from the matrix product of the global matrix (SS) and the magnetic vector potential (A_z), the latter is obtained from the previous iteration i.e., $[SS][A_z]$. The value of $[A_z]$ is updated by adding it to $[\delta A_z]$ after every iteration, and the NR gets terminated using the condition defined in (3.31). The discussions above reflect the challenges associated with the NR method formulation; reference [41] provides some good insight into the behavior of NR in a HTS environment using a commercial package. The paper does not include an in-depth mathematical

treatment and the FEM details as presented here, but it shows some success using a relaxation factor, and comes with a caveat concerning the convergence issues, and the limited ability to extend it to a wider set of problems due to the onset of instabilities.

3.2.2 Algorithm to improve the convergence of a field solution

The convergence rate criterion can be used to classify iterative methods [42], [43], in the following manner

DEFINITION: Consider a sequence $\{x_n\} \subset \mathbb{R}^N$ | an element x^* , satisfies the condition $x^* \in \mathbb{R}^N$

- i. A sequence converges quadratically, if $x_n \rightarrow x^*$ and there is $C > 0$ | $\|x_{n+1} - x^*\| \leq C \|x_n - x^*\|^2$, for all n sufficiently large.
- ii. A sequence converges superlinearly with order β , where $\beta > 1$, if $x_n \rightarrow x^*$ and there is $C > 0$ | $\|x_{n+1} - x^*\| \leq C \|x_n - x^*\|^\beta$, for all n sufficiently large.
- iii. A sequence converges linearly, if $x_n \rightarrow x^*$ | $\|x_{n+1} - x^*\| \leq K \|x_n - x^*\|$ where $K \in (0, 1)$, for all n sufficiently large.

It is a desirable characteristic to have a faster convergence rate in an iterative procedure. The discussions presented so far, would always favour the use of the NR technique or a superlinearly convergent method. This reasoning is viable in a scenario where the cost of a single iterate is the same for a quadratic or a superlinear converging sequence when compared to a linearly convergent case. However, in many practical electromagnetic problems the cost per iterate in a quadratic system is so high due to oscillations introduced by instabilities, that the use of a slower method is justified. The design of algorithms to solve non-linear electromagnetic problems involves the use of approximation at various stages to simplify the process, such as the use of spatial or time discretization and polynomials that attempt to estimate the continuous nature of a field, or the use of linear approximates to calculate a non-linear relation. The accuracy of any algorithm is dependent on the outcome of such approximations, and the error gets introduced

into the method due to discretization and rounding, the latter is dependent on the machine precision. It is one of the reasons that causes the cost per iterate to increase. The other aspect that affects the computation cost is inherent in nature, and is dependent on the sensitivity of the problem under consideration towards perturbations. Since algorithms of this nature involve matrix computation, the condition number κ as defined below provides the sensitivity measure [44], which in turn determines the stability of the process. It should be emphasized here that the stability of an algorithm does not guarantee accuracy of the solution.

$$\kappa(A) = \|A\| \cdot \|A^{-1}\| \quad (3.42)$$

where A represents a matrix and $\|\cdot\|$ is a p-norm. In the present scenario, if the NR method is considered, the Jacobian in the formulation will include the conductivity derivative; refer to equation (3.33), in an event when the field E is close to zero, or has a very small magnitude, then $\sigma_z \approx \frac{1}{10^{-15}}$ (equation (3.20)), and the conductivity derivative with the term σ_z^2 will have a value $\approx \left(\frac{1}{10^{-15}}\right)^2$, such large derivatives affect the conditionality and the latter could get as high as 10^5 to 10^6 at times. The analysis so far, presented one case when the stability deteriorates, at this juncture, it can be argued that scaling, or some form of relaxation factor could improve the outcome, but it is not always practicable and it is extremely difficult to generalize. To avoid the complexity in the formulation associated with the NR method and the possible convergence issues highlighted above, the proposed algorithm uses the fixed-point technique. The fundamental approach to solve a non-linear function $f(\cdot)$ numerically, involves finding a value x from the set of sequence X , such that the following condition is satisfied

$$f(x) = 0 \quad (3.43)$$

where $f(X) \in \mathbb{R}$, for all elements of X , and $X \in \mathbb{R}$. The value x above is known as the root, and in the problems that are being considered here, the following form of equation system is used

$$\begin{aligned} f_1(x_1, x_2, x_3, \dots, x_{n-1}) &= 0 \\ &\vdots \\ f_{n-1}(x_1, x_2, x_3, \dots, x_{n-1}) &= 0 \end{aligned} \quad (3.44)$$

The fixed-point iterations use a mapping of X into itself, i.e. $M \mid X \rightarrow X$, and a fixed point of M is a $x \in X$ that results in $Mx = x$. The term fixed-point signifies the repeated application of the same mapping function on the values, a more rigorous mathematical treatment of the approach has been provided in the Appendix D. To improve the convergence and gain computation speed, a technique introduced by Alexander Aitken known as the *Aitken method* was used [45]. According to this method, in a converging sequence $S_n \rightarrow S$, estimation could be applied using the following relation, as $n \rightarrow \infty$,

$$T_n = S_n - \frac{(\nabla S_n)^2}{\nabla^2 S_n} \quad (3.45)$$

such that T_n above converges to S faster than S_n , where $\nabla S_n = S_n - S_{n-1}$, and $\nabla^2 S_n = S_n - 2S_{n-1} + S_{n-2}$. The *Aitken* approximation above assumes that the sequence converges linearly to the true value [45], [46], [47]. The Aitken approach helps in predicting the next root without solving the entire problem, it suggests that as the approximate values of the root approach the true value, the ratios of the errors e.g., ϵ_2/ϵ_1 , ϵ_3/ϵ_2 gradually become constant, when such a condition occurs, the errors could be rearranged as $\epsilon_2^2 \approx \epsilon_1 \epsilon_3$, which yields the relation provided in (3.45) (for further details refer to Appendix D), assuming the terms in S now represent the roots for the non-linear equation, and ϵ here is the difference between the approximate root and the true solution.

```

1  FOR (t = 0, totalTime, StepSize)
2      error = 1
3      AitkenCount = 0
4      errorAtk = 1
5      WHILE (error > toleranceOne)
6          IF (error < toleranceTwo) THEN
7              AitkenCount = AitkenCount + 1
8              IF (AitkenCount > 3) THEN
9                  Calculate: Aiken Approximation ( $A_n^{\text{aitken}}$ )
10                 Calculate: Approximation Error (errorAtk)
11                 AitkenCount = 0
12                 IF (errorAtk < error) THEN
13                     Update: Value of vector potential (A) with  $A_n^{\text{aitken}}$ 
14                     Update: error with errorAtk
15                     errorAtk = 1
15                     GO TO: step 5
16                 Assemble System Matrices
17                 Calculate and Update: Value of vector potential (A)
18                 Calculate and Update: Approximation Error (error)
19             END WHILE
20 END FOR

```

Fig 3.3. The Pseudo code for the Algorithm.

The suggested algorithm has been shown in Fig 3.3; a step size in time is defined first, denoted by the term *StepSize*, in order to iterate through the entire interval, *totalTime*. The SSM is initiated in line 5 first with a predefined tolerance value of 1.0E-05 represented as *toleranceOne*. The algorithm has a second check based on the error estimates of the solution in line 6, which decides whether to apply the extrapolation. Using trial and error, the optimal values of *toleranceTwo* for the problems discussed here was determined to be in between 4.0E-03 and 9.0E-04, inclusive of the two limits. The approximation to the solution is estimated after every third iteration inside the SSM loop, and this is achieved by using the counter *AitkenCount*. Subsequently, error estimation is performed, and then it is verified whether there is an improvement in the result, as shown in the lines 10 and 12 respectively. In the event of an improvement, the approximation and the error estimation are considered, and the algorithm moves to line 5, or else both the values are discarded, and the system solves the matrix equation again as shown in the lines 16-18. The next section presents the results that were obtained using the algorithm introduced.

3.2.3 Simulations and Results

The AC loss simulation for HTS based devices requires a prior knowledge of the critical current density J_c , reference [48] provides the measured values of J_c under different external field conditions. To ascertain the magnitude of J_c , and gain an insight into its behavior, the critical current I_c flowing inside the HTS specimen is calculated. This task is achieved by considering DC or static conditions, where the electric field is set to the critical field value of 1μV/cm as defined in equation (3.20). The formulation for DC conditions involves the relations given by (3.7) and (3.15) without the time derivative terms as shown below.

$$\frac{1}{\mu_0} \nabla \times \nabla \times \vec{A}_z - \sigma_z \mathbf{E}_c = 0 \quad (3.46)$$

$$I_c(t) = \iint_S \sigma_z \mathbf{E}_c dS \quad (3.47)$$

The critical current density is subsequently obtained by dividing the critical current I_c by the cross section area S in the following manner

$$J_c = \frac{I_c}{S} \quad (3.48)$$

The discussions involving the DC scenario describe a method that delivers an initial estimate of J_c and provide a simple approach to model it. As mentioned earlier, the previous chapter and the Appendix B present the procedure to determine the behavior of J_c under AC conditions.

The proposed algorithm was verified using a HTS rod and an open core transformer arrangement that depicted the self field loss and the magnetization loss scenarios respectively. In the first experiment, a 4×4 mm square HTS rod was considered for the simulation. To apply the boundary conditions properly, the peripheries of the problem domain were set far from the rod, and the formulation presented in (3.7) was used to obtain the solution. The geometrical symmetry provided the advantage of reducing problem size, and only $\frac{1}{4}$ of the geometry was analyzed, this in turn decreased the computational efforts. The Dirichlet boundary conditions were applied on the right and the top boundary edges, and the Neumann boundary conditions were used at the bottom and the left edges. The problem domain was discretized using 9539 elements; Fig 3.4 below highlights a magnified view that shows the high density of elements used in the HTS region for better accuracy. The simulation was executed using an alternating current defined as $I = I_A \sin(\omega t)$, where I_A was set to a value of 0.7 to 0.8 times the critical current density of 1.0×10^7 A/m². The latter should be multiplied by the area, and then considered in the formulation.

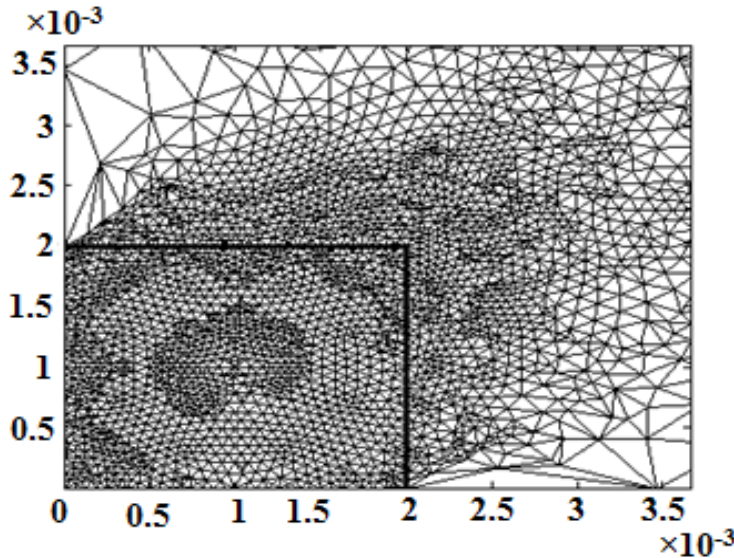


Fig 3.4. The magnified view of the discretization present in the HTS region, the dimensions are in meters.

The field distribution over the entire problem domain taken at one instant of time ($t = 0.01\text{s}$) is presented in Fig 3.5. The time step was $5.0\text{E-}04\text{ s}$, and the frequency for the AC signal was 50Hz .

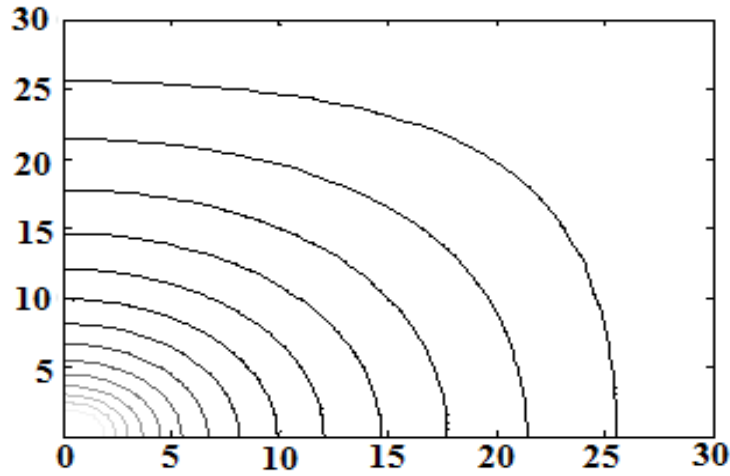


Fig 3.5. The plot showing the magnetic field distribution over the entire problem domain, dimension is in millimeters.

Table 3.1 below shows the variation in the conductivity across the cross-section of the HTS specimen, with the Y-coordinates equal to $1.15\text{E-}03\text{ m}$. This highlights the existence of non-linearity even within the confines of small spatial changes.

Table 3.1 Conductivity across the cross section of HTS sample

X-coordinate(m)	0	$5.0\text{E-}04$	$1.0\text{E-}03$	$1.50\text{E-}03$	$2.0\text{E-}03$
Conductivity(S/m)	$3.43\text{E+}13$	$3.24\text{E+}13$	$6.84\text{E+}12$	$8.68\text{E+}10$	$8.47\text{E+}10$

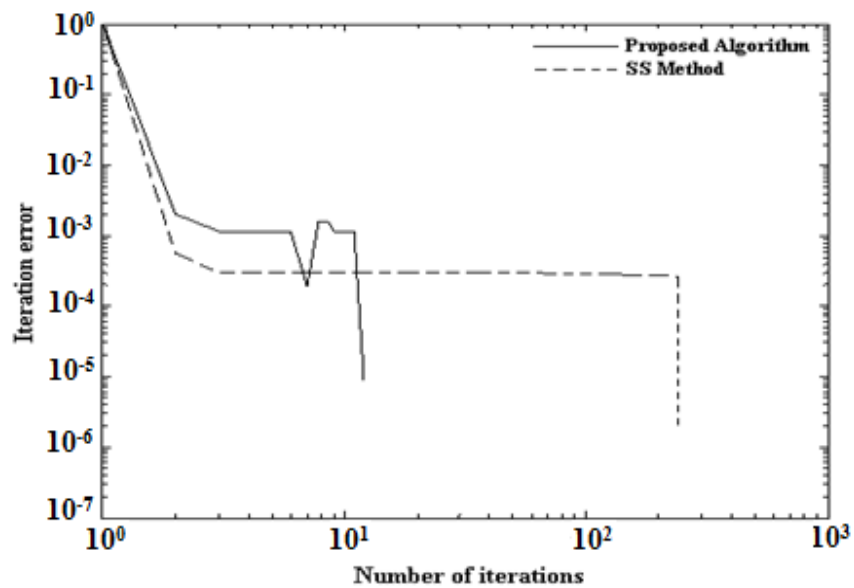


Fig 3.6. The convergence plot comparing the algorithm, and the SSM.

The above figure depicts the improvement in convergence over the SSM method. In the next endeavor, the open core transformer configuration comprising of iron, copper and HTS components as shown below was considered.

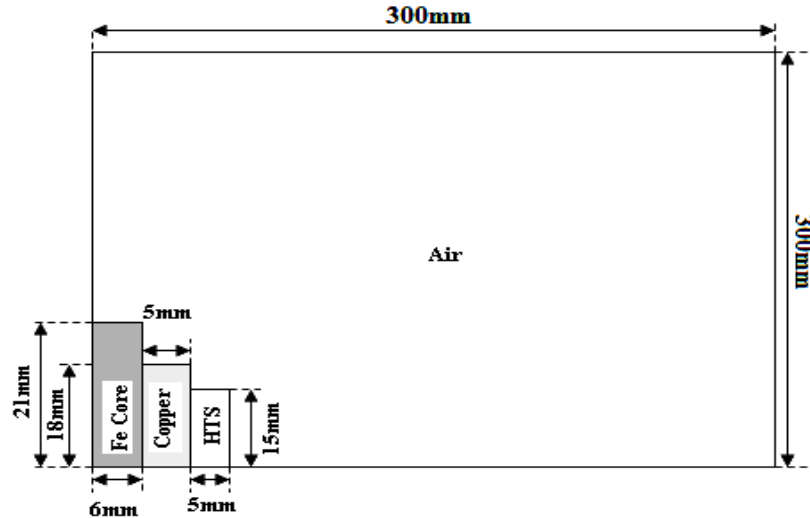


Fig 3.7. The experimental set up showing the 1/4 section used for the field analysis (not to scale).

In this experiment, the HTS carried a fixed DC current density of 4.6 A/mm^2 . To consider the effects of iron on the magnetic field computation, the formulation included the non-linear permeability of iron in the stiffness matrix. The copper component carried an AC current with an average density of 1 A/mm^2 , and a frequency of 50Hz. In order to solve the problem, the region was discretized using 11675 elements as shown in Fig 3.8. The time step size considered here was $5.0\text{E-}04 \text{ s}$. The boundary conditions were the same as those presented in the self field loss example. It is interesting to note here that the iron component had a non-linear permeability and constant conductivity, whereas the HTS unit had a constant permeability and a non-linear conductivity.

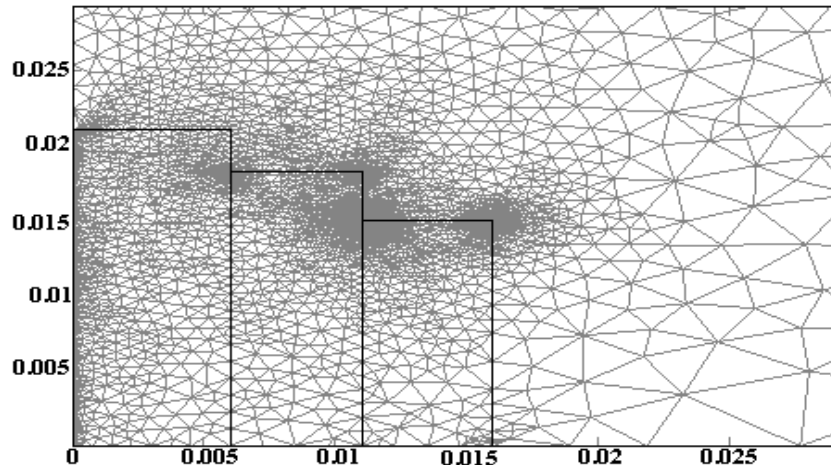


Fig 3.8. The magnified view of the discretization present, dimension is in meters.

The procedure was successful in solving the described problem, and the following figure represents the magnetic field intensity across the HTS sample compared to the experimental values.

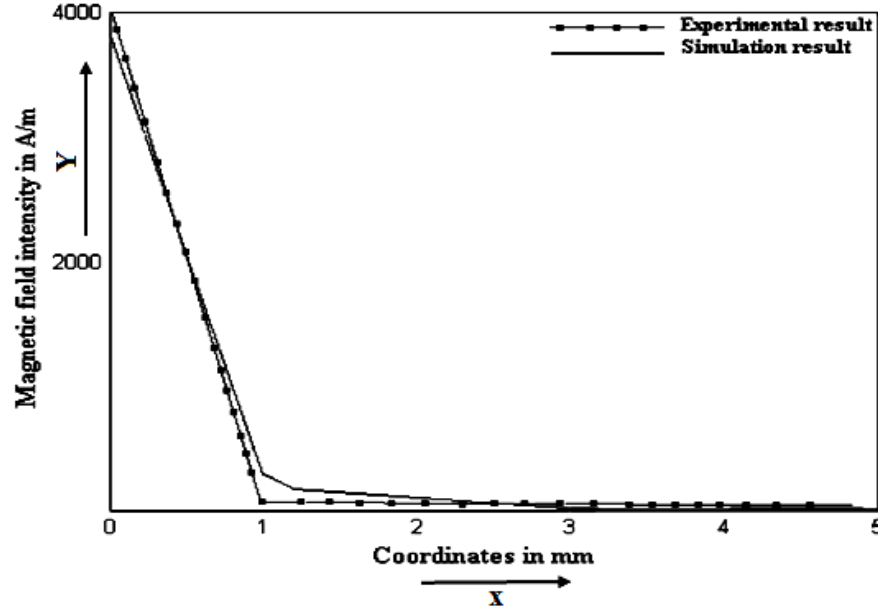


Fig 3.9. Comparison between the experimental result [49] and the simulation result.

3.3 Thermal modeling

In order to model the thermal behavior of a system, the transfer modes i.e., conduction, convection and radiation are investigated. Conduction is the transfer of thermal energy through matter because of a temperature gradient and does not involve motion of the material. In convection mode, the energy is transported by material motion. Radiation has two distinguishing characteristics when compared to conduction and convection i) no medium is required for heat transfer ii) unlike the first two methods where transfer of energy is proportional to the temperature difference between two locations, in radiation, the transfer is proportional to difference of the two temperatures each raised to the fourth power. This section discusses the use of FEM to solve heat transfer scenarios. It will be evident later in the discussions that the nature of the partial differential equation (PDE), which describes a thermal problem involving conduction is similar to the diffusion relation defined in equation (3.21). As performed earlier to model the AC losses, the approach here too will involve extending the ideas used for conventional materials to HTS. In such scenarios, the PDE that defines the governing relation

involving the conduction phenomenon is given by Fourier's law [50], [51] shown below in Cartesian coordinates

$$q_x = -k \frac{\partial T}{\partial x} \quad (3.49)$$

$$q_y = -k \frac{\partial T}{\partial y} \quad (3.50)$$

$$q_z = -k \frac{\partial T}{\partial z} \quad (3.51)$$

where q depicts the rate of heat flow per unit area, k is the thermal conductivity and depending on the problem may be a function of the temperature denoted by T above. The negative sign signifies that the thermal energy transfers from a warmer to a colder region. In vector form the above relation using the gradient is defined as

$$\mathbf{q} = -k \nabla T \quad (3.52)$$

The following thermal energy equation is derived using the first law of thermodynamics and Fourier's law

$$\nabla \cdot (k \nabla T) - \rho c \frac{\partial T}{\partial t} = -Q \quad (3.53)$$

Q in (3.53) is the thermal source, a contribution due to the heat loss arising from Joule's effect or hysteresis, or mechanical friction depending on the case investigated. ρ in the energy expression is the mass density; c denotes the specific heat and the time is represented by t . The above equation is non-linear if the specific heat or the thermal conductivity or both are dependent on temperature. In the field computation, the boundary conditions are required to ensure a well posed problem with a unique solution. The typical boundary conditions used to solve thermal problems can be classified in the following manner:

- i. Prescribed Temperature - This is an example of Dirichlet condition or boundary condition of the first kind, where the temperature is specified as a constant, or a function of one or more variables at the boundaries.

- ii. Prescribed Heat Flow - In this form, the boundary condition is the rate of heat flow across a boundary specified as a constant or a function of one or more variables at the boundaries. This case is an example of Neumann boundary condition or boundary condition of the second kind, defined as

$$-k \frac{\partial T}{\partial n} = q_s \quad (3.54)$$

n above is normal to the boundary, the significance of the remaining parameter and variables has been described earlier, refer to equations 3.49-3.51.

- iii. Convective Heat Exchange - It involves scenarios where the flow of heat is proportional to the difference between the surface temperature T_s of a body and the convective temperature or the ambient temperature T_e of an adjacent fluid across the boundary. The form is defined in the following manner

$$-k \frac{\partial T}{\partial n} = h(T_s - T_e) \quad (3.55)$$

h in (3.55) is the heat transfer coefficient.

- iv. Radiation Heat Exchange - This condition considers the rate of flow across the boundary in terms of the difference in energy emitted from the surface of a body at a temperature T_s , and the energy absorbed by it due to the incident thermal energy, emitted or reflected from other bodies at a temperature T_e present in the problem domain under investigation.

$$-k \frac{\partial T}{\partial n} = \sigma \epsilon (T_s^4 - T_e^4) \quad (3.56)$$

where the terms σ and ϵ are the Stefan-Boltzmann constant and the surface emissivity respectively.

To solve problems of such a nature numerically using FEM, the first step involves spatial discretization, as performed in the magnetic field problems presented earlier. The following discussions derive the 2D element formulation for heat conduction involving convective exchanges.

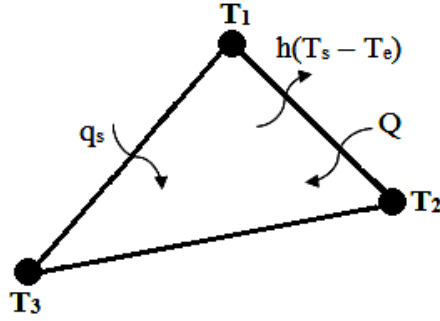


Fig 3.10. Heat transfer element inside a problem domain, involving surface heating and convection.

Consider a single element for simplicity as presented in Fig 3.10, where the boundary conditions are applied to the face. The derivation involves the use of the thermal energy equation (3.53), the convection component and the application of the Galerkin method (for more refer to Appendix C), resulting in the matrix expression shown below

$$[R_z]_{3 \times 1}^e = [C]_{3 \times 3}^e \frac{d[T]_{3 \times 1}^e}{dt} + [K]_{3 \times 3}^e [T]_{3 \times 1}^e - [G_Q]_{3 \times 1}^e + [G_q]_{3 \times 1}^e - [G_h]_{3 \times 1}^e \quad (3.57)$$

where $[R_z]_{3 \times 1}^e$ is the residual vector associated with a single element. A first order function, N , was considered as the interpolation function in the above expression, the coefficient matrices are defined in the following manner

$$[C] = \int_S \rho c [N]^T [N] t_k dA \quad (3.58)$$

C is known as the capacitance matrix, S represents the surface integral. In the present case t_k is the element thickness and dA is the differential area. N is given by

$$N_i = \frac{a_i + b_i x + c_i y}{2\Delta}$$

$$a_i = x_j y_k - x_k y_j ; b_i = y_j - y_k ; c_i = x_k - x_j , \quad i = 1, 2, 3$$

$$\text{and } \Delta \text{ is the area of triangle} \quad (3.59)$$

$$[K] = \int_S [B]^T [D] [B] t_k dA + \int_S h [N]^T [N] dA \quad (3.60)$$

the first integral in the right hand side is the conduction matrix (K_c) and the second integral is the

surface convection matrix (K_h), h in the latter is the convection heat transfer coefficient. B in (3.60) is the derivative matrix and D is the thermal conductivity matrix.

$$[B] = \frac{1}{2\Delta} \begin{bmatrix} b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

$$\frac{\partial N_i}{\partial x} = \frac{b_i}{2\Delta} ; \frac{\partial N_j}{\partial y} = \frac{c_j}{2\Delta} \quad (3.61)$$

$$[D] = \begin{bmatrix} k_x & 0 \\ 0 & k_y \end{bmatrix}, k \text{ is the thermal conductivity and in 2D there are two entries} \quad (3.62)$$

$[G_Q]$, $[G_q]$, and $[G_h]$ are the heat generation vector, the surface heating vector and the surface convection vector respectively.

$$[G_Q] = \int_S Q[N]^T tk \, dA \quad (3.63)$$

$$[G_q] = \int_S q_s[N]^T \, dA \quad (3.64)$$

$$[G_h] = \int_S hT_e[N]^T \, dA \quad (3.65)$$

In order to solve the transient thermal problem, the temporal discretization is achieved using FDM, and one such approach is the backward time difference method shown below

$$[R_z]^e = \frac{1}{\Delta t} [C]^e ([T]_{n+1}^e - [T]_n^e) + [K]^e [T]_{n+1}^e - [G_Q]^e + [G_q]^e - [G_h]^e \quad (3.66)$$

where n and Δt are the time step number and time step size respectively. The following equations provide the matrix structure of the coefficients defined above

$$\frac{1}{\Delta t} [C] = \frac{\rho c tk \Delta^e}{12\Delta t} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad (3.67)$$

$$[K_C] = \frac{K_x tk}{4\Delta^2} \begin{bmatrix} b_1^2 & b_1 b_2 & b_1 b_3 \\ b_2 b_1 & b_2^2 & b_2 b_3 \\ b_3 b_1 & b_3 b_2 & b_3^2 \end{bmatrix} + \frac{K_y tk}{4\Delta^2} \begin{bmatrix} c_1^2 & c_1 c_2 & c_1 c_3 \\ c_2 c_1 & c_2^2 & c_2 c_3 \\ c_3 c_1 & c_3 c_2 & c_3^2 \end{bmatrix} \quad (3.68)$$

$$[K_h] = \frac{h\Delta^e}{12} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad (3.69)$$

$$[G_Q] = \frac{Q_{tk}\Delta^e}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad (3.70)$$

$$[G_q] = \frac{q_s\Delta^e}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad (3.71)$$

$$[G_h] = \frac{hT_e\Delta^e}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad (3.72)$$

In a scenario where the edge of a triangle coincides with a boundary that has a heat transfer surface, an additional conductance matrix, surface heating vector, and surface convection vector are evaluated in the following manner.

$$[K_h] = \frac{h_{tk}L}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad (3.73)$$

$$[G_q] = \frac{q_s tk L}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (3.74)$$

$$[G_h] = \frac{hT_e tk L}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (3.75)$$

in the above relations, L is the length of the edge, and the product $tk L$ represents the surface of an edge, the reduction in the size of matrix and vectors occurs because only two nodes are considered. The analysis presented above did not consider the effects due to radiation, which requires use of the relation in (3.56), for more on this subject, and also on the treatment of 3D and axisymmetric scenarios refer to [50], [51]. This work used the ThermNet tool from Infolytica to model the thermal behavior and the radiation effects were not considered in the simulation. The next chapter considers a simple HTS device that will demonstrate the use of ThermNet and the thermal aspects of HTS.

The nature of the PDE to solve thermal problems given by the energy equation (3.53) is applicable to HTS cases. The general physics for low-frequency applications involving HTS and that of conventional materials is similar to an extent. These common aspects allow the extension of FEM ideas used for conventional materials to HTS based systems. The outcomes of the

energy equation are dissimilar because of the parameters involved, the specific heat c and the thermal conductivity k , which are different for both materials. At the micro level, as highlighted in the previous chapter c is comprised of two components i) specific heat due to the conduction of electrons c_e , and ii) specific heat due to the lattice vibrations c_{lat} . The term due to the flow of electrons is appreciable at low temperatures, whereas the contributions from the lattice behavior dominate at room temperatures. A dramatic change in c_e occurs at superconducting transition phase, which is not prevalent in c_{lat} . In metals and conventional superconductors, the c presents a linear behavior corresponding to a change in temperature, whereas in HTS the relationship is of higher orders, usually second order or more. Similarly, the conductivity k has two components i) conductivity due to moving electrons or free carriers k_e , and ii) k_{lat} because of the lattice vibrations. In metals and conventional superconductors, there is a substantial contribution from free carriers to thermal conductivity such that $k \approx k_e$, whereas in HTS the k_{lat} component plays a dominant role. The cumulative effects of such properties are reflected at the macro level, but an in depth analysis of these phenomena is beyond the scope of this work, for further details refer to [19].

3.4 Coupling different field problems

To model electric machines precisely, it is imperative to consider and incorporate the effects of the electromagnetic field, the thermal field, and mechanical strain, the latter a consequence of deformation. The dependencies of one such field or strain on the remaining two and vice-versa complicate the analysis procedure. This work investigates the interactions that exist between the electromagnetic and thermal fields in HTS materials; it will not study the effects due to strains. In general, such problems involving the study of different fields and their interactions are known as coupled problems. The numerical approaches to solve examples of this nature are classified into two categories: i) strongly coupled and ii) weakly coupled. In strongly coupled systems, the equations to model the effects are solved simultaneously, and are handled at the matrix level [52], [53]. This is achieved by including the coupling relations in the coefficient matrices. In weak coupling approaches, the problem is segregated into electromagnetic and thermal models, and solved in successive steps [54], [55]. Such methods provide the flexibility of using different solvers and time steps for the associated electromagnetic and thermal problems. The coupling is attained by first updating, and then transferring the dependent data to the problems defined and

subsequently solving them. The thermal field changes at a slower rate when compared to the electromagnetic field in low frequency examples, the large difference in the time scale or the time constant plays a crucial role in deciding the computational approach. The ratio between the largest and smallest time constant is known as the stiffness ratio, a parameter that determines the use of a particular coupling technique, the ratio could be as high as 10^5 to 10^7 in the present scenarios. Weakly coupled methods are preferred for problems with high stiffness ratios, involving the use of time-harmonic and time-transient approaches for the electromagnetic and thermal systems respectively.

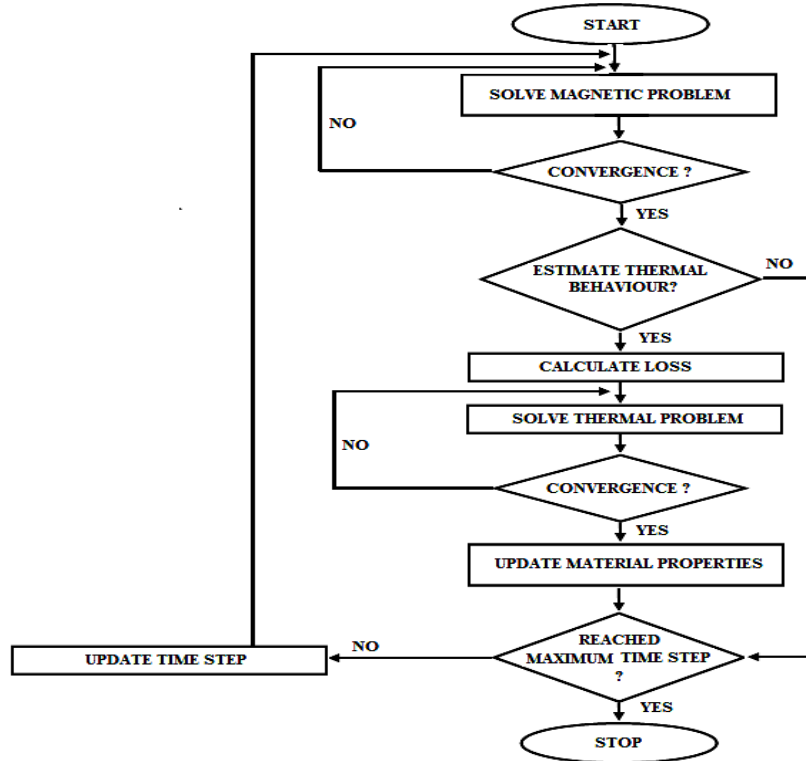


Fig 3.11. Flow diagram of the coupled problem.

The figure above highlights the flow associated with the coupling method used in this work. The first step involves electromagnetic analysis that estimates the magnetic vector potential using the relations defined in (3.7) or (3.21) depending on the scenario under investigation, and specifying the boundary conditions. A decision block was included to avoid frequent calls to the thermal solver in order to save computation time; it could be viewed as an introduction of delay to the thermal call attained by skipping few time steps. The use of such a delay function is problem dependent, and it may not be applicable in certain complex scenarios, for instance when the HTS

operates close to the quench region. If at any instance the thermal calculation is necessary, the AC loss values are estimated as discussed earlier involving numerical integration for every element. The thermal analysis is then executed using the formulation in (3.53) and applying the boundary conditions, and the new temperature distribution is determined. The values of the parameters to solve the thermal problem (equation 3.53) are available online in the National Institute of Standards and Technology (NIST) database, and in [58]. Subsequently, the time is incremented and the temperature distribution is fed back to electromagnetic analysis block, where the conductivity values are calculated and the cycle is repeated.

The ideas discussed in this section could be extended further to include structural analysis in the coupled problem described above. In order to implement such a scenario, a structural solver is required that would estimate the deformation arising due to the variation in temperature and electromagnetic force. The structural solver would calculate the deformation in a HTS based device and then feed the corresponding strain distribution to the magnetic solver. Strain affects the shape and material properties that results in the variation of magnetic field and this necessitates the use of magnetic solver. The field distribution and the heat loss are reestimated by the magnetic solver, and a similar cycle as presented in Fig 3.11 is followed subsequently.

3.5 Summary

This chapter discussed the possible numerical methods that could be used to solve electromagnetic field problems in HTS materials. The use of FEM to solve such examples was analyzed, and an algorithm based on the Aitken approach was introduced to improve their convergence. The simulations pertaining to the new technique were presented, and its effectiveness was verified. This part of the dissertation also included the FEM treatment for thermal field problems involving HTS, and an approach to couple the magnetic field and the thermal field scenarios in them was provided. The next chapter will present a procedure to optimize HTS based current leads, which involves the use of a stochastic method and the application of ideas from the discipline of multi-objective optimization (MOO).

CHAPTER 4

4 Design process and Optimization involving HTS

4.1 Introduction

This part of the dissertation will discuss and present the techniques involved in the computer based design and optimization of a simple electromagnetic device involving HTS, and in the process will highlight the challenges involved. In chapter 2, the theoretical aspects associated with the thermal and magnetic field behavior of HTS were presented. The techniques to solve numerically, the two types of field problems in HTS using FEM were examined in chapter 3. In order to fully appreciate the intricacies involved, and to address the challenges unique to the design of computational models for such materials, a HTS based current lead is considered for the study. The investigations in this chapter will provide insights to deal with the challenges unique to HTS based devices. Furthermore, the concepts from the field of multi-objective optimization are investigated, and their feasibility to obtain an optimal design for HTS leads are explored. In order to implement MOO in such scenarios, Differential Evolution (DE), a stochastic based approach is used.

The following sections will first describe HTS current leads and then underline the desired characteristics of such leads. To achieve the goal of obtaining an optimal model through simulations, a few samples from the search space are gathered first, as it is infeasible to test every point in the space for optimization. One of the sections reviews certain statistical techniques and verifies their scope in the scenarios under consideration. Subsequently the chapter examines and demonstrates the possible objective functions and constraints that could be considered to apply the multi-objective optimization method. It should be highlighted here that the study involves a multiphysics scenario which takes into account the magnetic field and thermal behavior associated with HTS current leads in operation. The discussions will provide the necessary theoretical aspects associated with the optimization process, and they will also include an insight into DE and highlight the benefits the latter carries.

4.2 HTS based current leads

The electrical connection between two components maintained at different temperatures is achieved using current leads [56], [57]. HTS based leads play a crucial role to deliver power from current carrying copper components maintained at room temperature to superconducting magnets made from NbTi that operate at liquid helium temperature (4.2 K). As highlighted earlier, HTS materials have no Joule heating and very low thermal conductivity compared to metals. Such properties help in reducing heat transfers to the cold regions, which provides a substantial benefit, as they reduce refrigeration requirements. When compared to low-temperature superconductors (LTS), HTS have a much wider transition region between superconducting and normal states; hence they provide more flexibility in terms of operation. Since HTS leads carry a high current, any variation in temperature, magnetic field or current density beyond their respective critical values could introduce resistance in the leads, and may cause burnout by Joule heating [57], [58].

A popular method for manufacturing HTS leads is the Melt Casting Process (MCP). In MCP, bulk parts in the desired shapes and sizes suitable for electrical engineering applications can be developed, and it avoids, to an extent, the issues that arise due to brittleness in HTS. The process involves the use of melt casting into different shapes and subsequently annealing them, the technique is convenient compared to conventional ceramic shaping approaches for constructing complex geometries [48]. For instance, to fabricate tubes or hollow cylinders as shown in Fig 4.1, the melt is poured into a rotating cylindrical steel mould and gets evenly distributed inside due to centrifugal force. To avoid cracks or any form of fracture in the cylindrical lead, the speed of rotation is varied during the period when the material solidifies. The solidified melt after casting

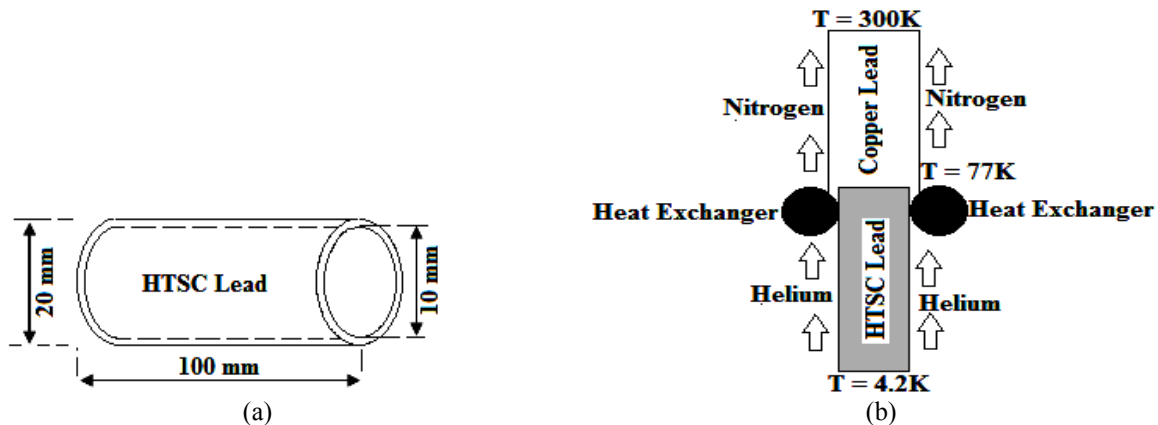


Fig 4.1. (a) A HTS current lead geometry. (b) Schematic representation of gas cooled HTS and Copper current leads arrangement.

and annealing is machined by methods such as sawing, drilling, mill cutting etc. In the system described in Fig 4.1(b), the bottom of the lead comprising the HTS section is usually immersed in liquid helium (He) and the top of the lead is kept at a higher ambient temperature. The thermal exchanger is maintained at an intermediate temperature, and it assists in reducing the total consumption of electric energy by removing heat at a higher temperature. The lead configurations are classified in terms of the cooling approach used as: conduction cooled, and vapor or gas cooled. Conduction cooled leads are simple to construct but heat leaks are higher in them, whereas vapor or gas cooled systems provide better heat leakage reduction. Cryogenics such as liquid nitrogen (N) are used at the warmer end. This work considers a conduction cooled HTS lead for the optimization process, and the method was applied only to the HTS section. To simulate the behavior of a HTS lead, it was assumed that the lower end of the lead was in close vicinity to liquid helium maintained at 4.2 K, and the upper segment was kept at 77K using liquid nitrogen.

The use of HTS materials for current leads was conceived because they provided a scope to spend less energy in refrigeration when compared to LTS based leads. LTS can operate only at temperatures close to 4K or less, whereas HTS have the potential to work at temperatures as high as 135 K, and this aspect is the main reason behind energy saving. HTS leads still operate under considerable low temperature conditions when compared to conventional materials as evident from the discussions so far, and it is always desirable under such settings to reduce the heat leakage and the AC loss. A simple method to reduce the heat leakage involves varying the cross-section along the length, as heat leakage is directly proportional to the cross-section area. The hypothesis lies in the fact that heat leakage could be minimized by reducing the cross-section area at the cold end and the regions close to it [59]. Section 4.4 will also highlight the influence of geometrical factors (cross-section area) on AC loss. HTS are expensive materials, and it is always desirable to reduce the production cost of the leads by minimizing the use of superconducting materials. The reduction of heat leakage, AC loss and the use of materials are the objectives that the optimization method attempts to achieve.

4.3 Sampling plan and Surrogate technique

Optimization methods involve the execution of computer programs that usually use a vector of

design variables as inputs and deliver a vector of responses. Over the years there have been significant advances in computational capabilities, but in certain cases, computer codes using FEM for the analysis of electromagnetic examples can consume a considerable amount of time to reach even a single solution. The optimization of a current lead in this scenario requires FEM simulation and multiphysics problem solving. Under such circumstances, it is prudent to assume that the optimization process will be a computationally intensive task. Many times engineers employ trial and error techniques, where a limited set of inputs and responses are considered to design devices. Though such an approach could reduce the computational cost, it may fail to capture a functional relationship between the inputs and the outputs. Trial and error methods without any guidance lack the ability to provide or predict those input values that generate optimal solutions. In such an event, the subject of statistics comes to the rescue, and it overcomes these difficulties by constructing an approximation of the input-output behavior which is not only efficient to execute but also provides an insight into the input-output relationship. The representation of physical systems, that captures the input-output response first mathematically and then the implementation in the form of computer codes is known as *Surrogate modeling* [56]. The following discussions will provide the mathematical relation first, representing the Surrogate approach and then highlight the challenges involved.

DEFINITION: Considering an unknown multivariate function f that denotes the performance behavior of a process or a system and defined as $f: D \rightarrow \Omega^N \mid D \subset \mathbb{R}^d$, with the function values being $f(X) = \{f(x_1), f(x_2), \dots, f(x_k)\} \subset \Omega^N \forall$ distinct sample points $X = \{x_1, x_2, \dots, x_k\} \subset D$. The Surrogate technique constructs an approximate function $\hat{f}$ from an approximation space $\mathbb{S} \mid \hat{f} : D \rightarrow \Omega^N \in \mathbb{S}$, and $\hat{f}$ closely emulates the behavior of f measured by some criterion ϵ .

A close observation of the above definition reveals that Surrogate modeling is a form of regression. The concept of regression originates from the field of statistics, where the technique is used to generate a good estimate of data sets, consisting of dependent variables (y) in response to a set of independent variables (x). It involves using some form of mapping applied to the pair of (x) and the respective outputs (y) that delivers a model with a potential to predict the dependent variables. To develop such models, and improve their accuracy further a closeness measure is introduced into the procedure through some error term such as the Euclidean distance.

The experimenter determines the values of x and it is usually assumed they are free of errors, whereas the magnitudes of y may have an error in them, hence the purpose of error measurement. In Surrogate methods the sampled points used as inputs could be viewed as the independent variables (x).

In order to implement Surrogate techniques, the first step involves the identification of design variables or inputs that have a significant impact on f . The idea is to find the shortest design variable vector that could capture most of the behavior of a system or a process, and it is accomplished by varying the values of the individual design elements or parameters within their specific range [61]. At this stage, the ranges of the various design variables are also established. The next step is selecting a number of vectors N , that are adequate to represent the design space as thoroughly as possible. N is usually small, and the number is constrained by the computational cost associated with finding the sample points. It is reasonable to assume that the use of a higher number of design variables in a problem would require more locations to measure the objective function for constructing a predictor with good accuracy. Considering a certain level of accuracy is attained by sampling a one-dimensional space in m locations, to obtain similar sample density in a d -dimensional space, m^d observations are necessary. Scenarios as described for higher dimensional spaces often cause the issue of the *curse of dimensionality*. For instance, assume in an electromagnetic problem the analysis and design process requires one hour of computation time for a single design that manipulates only one input parameter. Let there be a constraint in the computational budget of around five hours, and five simulations provide a fairly accurate predictor. In order to further refine the generated model additional variables could be considered, and as the numbers increase from the present case of one to say six parameters with the sample density remaining same, the computation requirement jumps to 5^6 executions, which is a 3000 fold increase or is equivalent to 125 days.

It could be concluded from the above discussions that evaluation of the objective function for every possible combination of all the design variables is an expensive proposition. Besides the assessments associated with an objective function, the number of design variables also contributes significantly to the number of experiments. An answer to the issues highlighted lies in seeking variables that have less or no effect on the objective function, so that such variables

could be left out of the design investigation. There is no standard set of guidelines which clearly specifies the amount of time that should be utilized in selection of variables; the application of any form of screening or sensitivity analysis is problem dependent. Some engineering intuition could play an important role, for instance, if a significant number of variables do not contribute much then the use of screening could increase the model accuracy. If on the other hand, all the variables play an active role, the application of screening will not offer any substantial benefits. The following section will provide an insight into the sampling methods, with a special emphasis on the procedure that has been considered in the present work.

4.3.1 Design of sampling plans for computer experiments

Computational models involve empirical investigation known as *computational experiment*; the procedure executes a computer program at different values of inputs to obtain the responses and analyses are performed on the results. Physical experiments and computational experiments are similar concepts, both of them have human and systematic errors, the aspect that differentiates them is the latter do not have a random error component, hence computational experiments are deterministic in nature [60]. As highlighted earlier, runs of computer codes are time consuming hence the initial process of identifying inputs that have the potential to influence results significantly. The inputs that are selected for the experiment constitute the experimental design, and the region corresponding to such inputs over which the response of the model is investigated is known as the experimental region. Many researchers suggest the use of space filling plans for computer experiment designs, when the accuracy of the predictor is of primary interest [61]. The basis for such a recommendation arises because predictors are usually interpolators, and the error in prediction for any input is dependent on the location of the design point. For instance, in designs that are not space filling, there could be scenarios where the points concentrate on the boundary of experimental design space, and this could lead to poor outcomes in regions that have sparse points. There are various approaches to ensure the design points are spread evenly throughout the region. The selection of points could be based on certain sampling algorithms such as stratified structures, or based on measures of distance between the points quantifying the extent of an even spread; another alternative method could use the measure of uniform point distribution throughout the region as the basis for selection. This work considers a space filling

plan based on sampling techniques, and the following discussions describe the methods that use a similar idea:

- i. Full factorial (FF) sampling - This procedure attempts to sample points in a uniform manner by using rectangular (square) grids. In a FF design, the number of design points is defined by the product of the number of levels (predefined values) for each factor (design variables) [62]. To elaborate further, assume there are 2 levels and m variables, then FF will generate 2^m sample points. The major drawbacks include an exponential increase in the size of experiments as the number of factors rise, which could lead to a substantial increase in the number of experiments; apart from this at times in high dimensions, the FF method, fails to ensure space filling in the design space.
- ii. Hammersley sequence sampling - In this approach, representations of a decimal number in the inverse radix form (a unique fraction between 0 and 1) are used to generate the sequence for sample selection [63]. The values of the radix (represented by n with subscripts below) are chosen as the first $(p-1)$ prime numbers, where p is the number of dimensions. To fully appreciate the concepts behind the Hammersley sequence, consider an integer n that could be represented in the following form:

$$n = n_0 + n_1 R^1 + n_2 R^2 + \dots + n_k R^k \quad (4.1)$$

where $k = [\log_R n]$, the square bracket signifies that only the integer value is considered, and R denotes a prime integer. The above definition in (4.1) could be applied to any integer. A fractional value is obtained by reversing the order of the digits about the decimal point resulting in the relation shown below

$$\varphi_R(n) = n_0 R^{-1} + n_1 R^{-2} + \dots + n_k R^{-k-1} \quad (4.2)$$

The Hammersley sequence for N points in a m -dimensional cubic space is given by

$$z_m(n) = \left(\frac{n}{N}, \varphi_{R_1}(n), \varphi_{R_2}(n), \dots, \varphi_{R_{m-1}}(n) \right) ; \quad n = 1, 2, \dots, N \quad (4.3)$$

- iii. Latin hypercube sampling (LHS) - A 2D example, in the form of Latin square is considered initially for simplicity to explain the fundamentals behind LHS. In a Latin

square scenario when p design points are required, a $p \times p$ grid is constructed first, and each cell is then populated by selecting values from the interval $1, 2, \dots, p$, the arrangement is performed in a manner that ensures an integer appears only once in each row and in each column of the grid, these integers could be viewed as sample point representations. This concept in 2D when extended to higher dimensions forms the basis for the LHS method. In LHS, the design space is split into hypercubes of equal dimensions, and such hypercubes could be viewed as boxes. The points (integers in the Latin square case) are arranged in the boxes, each box carrying only one point. To guarantee that the property of unique cells in each row and column is retained in the LHS, the strategy to fill the boxes assumes that the paths running parallel to any of the axes should always encounter a filled box only once. LHS is popular because it is simple to generate, and it could be easily tailored to provide a uniform spread even with few points.

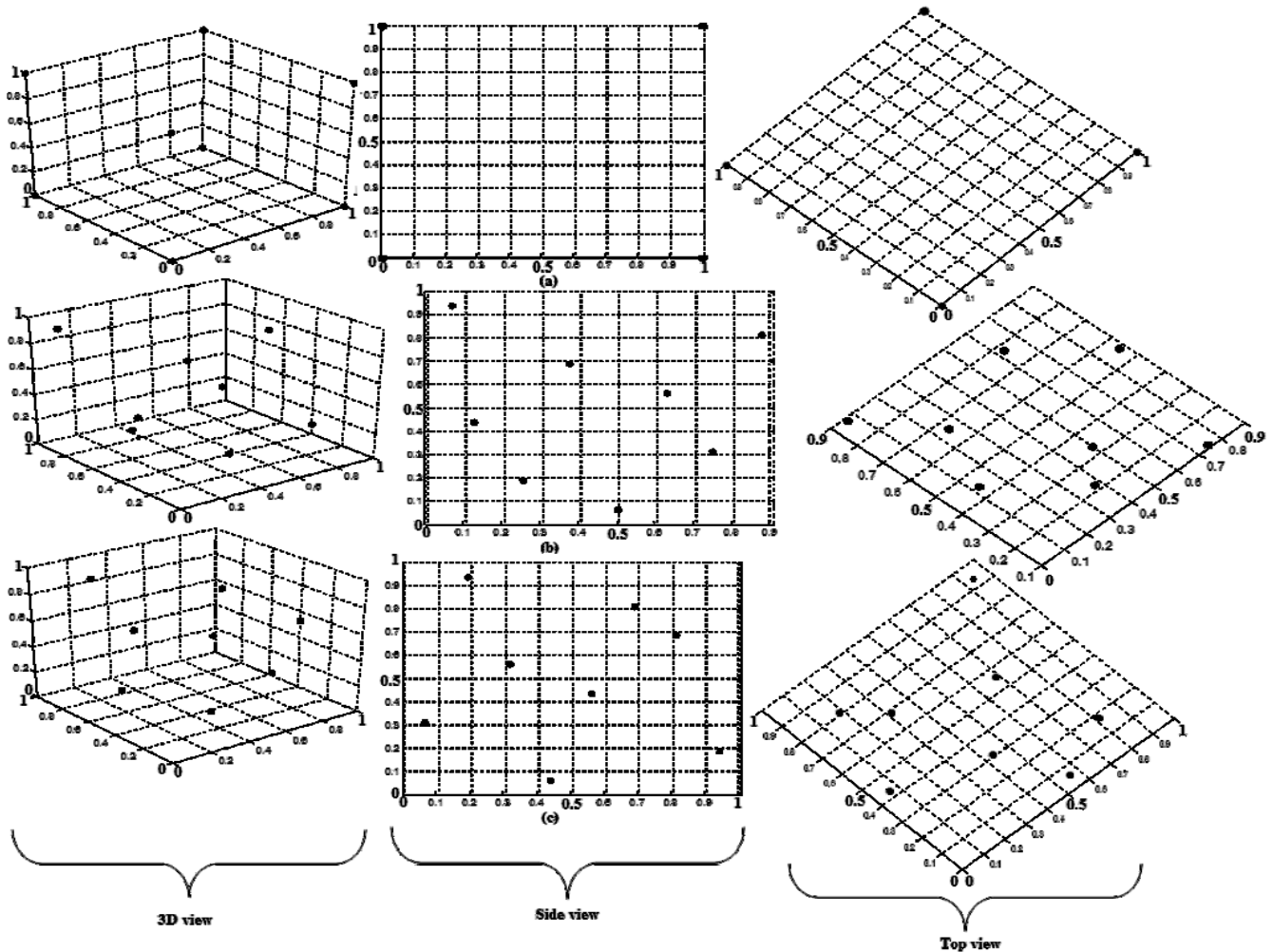


Fig 4.2. (a) Full factorial sampling method using 8 points, providing 3D, side and top views. (b) Hammersley sampling method using the same number of 8 points with the 3D, side and top view plots. (c) Latin hypercube method using 8 points presenting the 3D, side and top views .

In Fig 4.2 eight sampling points were generated in a cubic space using FF, Hammersley, and LHS procedures respectively. It is evident from the above plots that in the FF method, when the points are projected onto the axes, they will overlap which is undesirable. The remaining two methods ensure the uniformity in projections. As highlighted earlier, in the LHS scenario, all the eight points are visible from six faces of the cube, with each row and each column carrying only one point. It can be argued here, that there is no significant difference in the distribution between the results from LHS and Hammersley methods. The use of a particular procedure in a given scenario could be subjective, theoretically the Hammersley sequence is able to maintain uniformity in all dimensions, but it is complicated to implement at times. A few other sequences that are frequently used for space-filling strategies include Sobol, Halton, and Faure sequences [61], [63]. This work used the LHS procedure, and the following section describes the use of such samples to create a surrogate model of the HTS current lead.

4.3.2 Surrogate methods

After deciding on the approach of experiment design and performing the necessary simulations, the next phase in the optimization process attempts to collect information from the input-output pair to construct an approximate function $\hat{f}$ that closely replicates the actual behavior f . The concept involves searching the entire search space for a suitable $\hat{f}$, and if the approach is a mere trial and error, this could take infinite time. In many real world problems, the outcomes $\hat{f}$ from a rudimentary technique as described earlier, could be difficult to generalize, and may not be ideal in predicting responses at new points. The following discussions present the methods that attempt to estimate a meaningful $\hat{f}$ efficiently for problems under consideration. To gain a better understanding of Surrogate techniques, a simple approach that uses polynomials, as applied in the well known Response Surfaces (RS) is explained first [64], then the Generalized Regression Neural Network (GRNN), a procedure used in this study, is considered [65].

Assume a response y due to the influence of a vector $\mathbf{X}$. The independent vector comprises of terms $x_1, x_2, \dots, x_m$ where m denotes the dimension of the design vector. The relationship between the two is defined by

$$y = f(\mathbf{X}) + \epsilon \quad (4.4)$$

where ϵ represents a random error. As highlighted earlier in a real world scenario, the response $f(\mathbf{X})$ at every point in the design space is unknown, and this necessitates the use of approximation techniques. A popular method to estimate $f(\mathbf{X})$ is by using low-order polynomials as shown in the following relation.

$$\hat{y}(\mathbf{X}, \beta) = \beta_0 + \sum_{i=1}^m \beta_i x_i \quad (4.5)$$

in the above equation, β denotes the coefficients, and the order of the polynomial is one. In the case of a second order polynomial, the approximation relation is defined as

$$\hat{y}(\mathbf{X}, \beta) = \beta_0 + \sum_{i=1}^m \beta_i x_i + \sum_{i=1}^m \sum_{j=1, i \leq j}^m \beta_{ij} x_i x_j \quad (4.6)$$

the values of β in the above equations are determined by regression analysis (refer to Appendix E). The difference in values fitted by the polynomial relation and the observed data provides a measure of the error. There are a number of ways to reduce the error magnitude, one such method that is widely used in regression techniques and also considered in this work is the least square minimization method as defined below.

$$\min \sum_{i=1}^n [y^{(i)} - \hat{y}^{(i)}(\mathbf{X}, \beta)]^2 \quad (4.7)$$

where n is the number of observed data points. In the above relation, it is assumed that ϵ in (4.4) is a small constant value around each point, and the random error factor was neglected. The cross validation method is another popular approach to estimate the parameters of the polynomials [60], but the procedure could be more challenging from an implementation perspective using computer codes. Interpolations often have the issue of overfitting, which occurs when interpolates oscillate in an attempt to fit all the data points. This situation arises due to the relation (4.7), as the regression process tries to reduce the error value. In order to avoid this problem, a simple approach involves introduction of a penalty factor, when the value of β_i is large, to the above equation as shown below.

$$\min \sum_{i=1}^n [y^{(i)} - \hat{y}^{(i)}(\mathbf{X}, \beta)]^2 + \lambda \|\beta\|^2 \quad (4.8)$$

where the value of λ is determined based on the situation. There exist numerous model fitting techniques, apart from the RS method described above, which could be applied in the present

scenario. A good taxonomy exists in [60] that provides a few suggestions for model approximation approaches, and tries to relate the experiment design, choice of model, the number of dimensions (variables) and sample points. The main challenges involved while using computers for constructing surrogates of electromagnetic devices include i) the use of a few design points to sample the search space, and identify such points efficiently, ii) the entire landscape of the search space should be well covered by the sample points, a uniform distribution in the initial stages of design is beneficial, iii) improving the accuracy of surrogate models by incorporating advanced approximation functions that are adequate to capture the essential details of the design space, they should be effective in dealing not only with linear behaviors, but also the non-linear nature present in some systems, and iv) the ease of implementation. As highlighted earlier, there is a large pool of approximation techniques available for constructing surrogates, such as response surfaces, neural networks (NN), kriging, and many more [64]. At this juncture, a relevant question that arises is to identify the approach best suitable for a given problem. In the scientific literature, there are numerous works that describe the application of such methods to deal with a variety of problems, but comparative studies of these procedures are scarce. However, there exist few investigations that do compare them in engineering design problems, though many of them do not involve electromagnetics. Such examples provide some good insights into the behavior of these methods under different conditions. Some of the well known documented works include, the comparison between kriging and response surface as performed in [66], NNs and RSs have been studied in [67], and [68] provides a comparison among response surface, radial basis functions (RBF), kriging, and adaptive regression splines.

The Response Surface is a simple method, and if the scenario involves small dimensions and is devoid of high degree of non-linearities, the computational expense could be reduced. To model surfaces with significant non-linearities, higher order polynomials are recommended. However, to estimate all the coefficients in the polynomial equation having a higher order, more sample points are required because of the increase in the number of coefficients. Apart from this, instabilities may also arise. Kriging incorporates a polynomial model plus a stochastic term; the latter is a spatial correlation function usually Gaussian in nature with zero mean and non-zero covariance (refer to Appendix E), which relates any two sampled points. The polynomial

function provides a global perspective of the design space, whereas the stochastic process attempts to include local information in the kriging model. Kriging is a very flexible approach as it can provide an exact interpolation of the data or a very smooth curve depending on the type of correlation function used, and there is a wide range of functions to select from. Another attractive feature of kriging is its ability to perform a screening operation on data and determine the important factors, which assists in generating the predictor (model). The process to construct a kriging model can be computationally intensive, which is a major drawback of this technique, in addition to this, when the sample points are close to one another, the correlation matrix (refer to Appendix E) can become singular. Kriging also has the ability to deal with non-linear scenarios but it is complex when compared to neural networks for adaptation, apart from this, kriging is ideal when the number of factors (variables) is less (less than 50) [64], whereas a neural network provides the flexibility of dealing with many variables (10,000 or more) [60], [64]. Neural networks outperform response surfaces in modeling highly non-linear behavior, whereas the latter consumes less computational time. References [69], [70], [71] describe the use of such approaches when conventional materials are present in electromagnetic devices. In an ideal scenario, the methods to construct surrogates should not be computationally expensive and adequate to capture all the essential behavior of the model. The presence of HTS stretches the limits of these requirements, and the issue of limited available data further aggravates the situation. The theoretical aspects discussed so far with the associated limitations, and the ability of a specific approach to generalize easily were the main reasons behind selecting a neural network as a surrogate in the proposed optimization framework of HTS leads. The following discussions will provide some insights into the neural network technique used in this work.

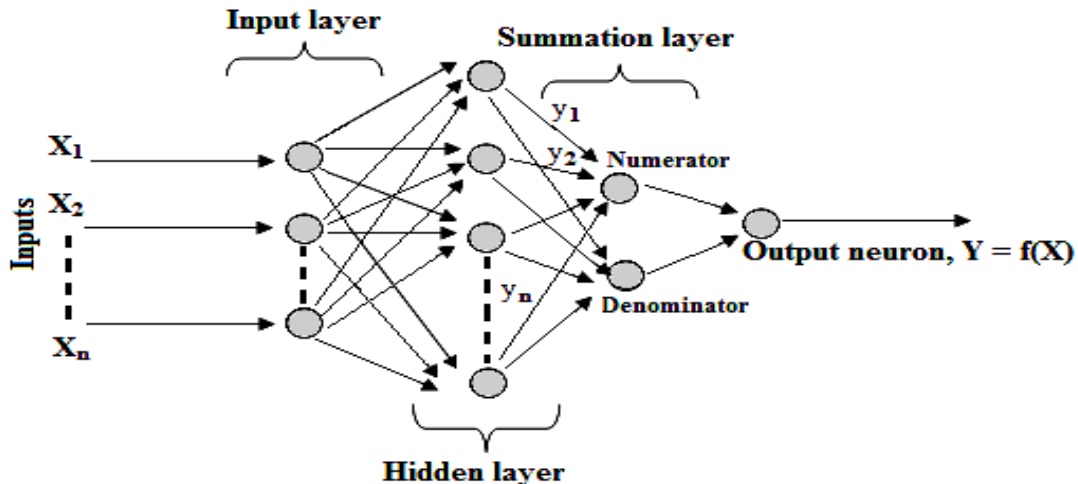


Fig 4.3. The architecture of a generalized regression neural network.

In general, a neural network can be defined as a massive network of neurons (processing units) that processes signals separately and simultaneously as a parallel distributed processing system. Such networks can learn from examples and make the knowledge available for use [72]. Neural networks, in terms of the architecture, can broadly be classified into feedforward neural networks (FNNs), recurrent neural networks (RNNs), and their combinations. The generalized regression neural network (GRNN) was proposed by Donald F. Specht, and has a FNN architecture which usually has three layers of nodes (neurons), namely the input layer, the hidden layer and the output layer node [65], [73]. The GRNN has a minor variation because there is an additional layer of summation nodes between the hidden and the output layers as shown in Fig. 4.3 . The input neurons merely provide scaled values of the input variables to the nodes of the hidden layer. There are two sets of neurons in the hidden layer, one that computes the numerator term and the other calculates the denominator of the probability density function described below. In the summation layer, the numerator terms of the hidden layer are added by multiplying them with the target value, whereas addition is only performed on the values present in the denominator. The output layer predicts the targeted value by simply dividing the accumulated numerator with the denominator obtained from the summation layer.

DEFINITION: A dataset $D = (X, Y)$ | $X = \{x_1, x_2, \dots, x_n\}$ and $Y = \{y_1, y_2, \dots, y_n\}$ generated by n sample points of the unknown function $f: X \in \mathbb{R}^d \rightarrow Y \in \mathbb{R}$. X and Y represent the set of independent random variables and the set of outcomes respectively; d denotes the dimension of the sample space. The approximation $\hat{f}$ of f provided is given by

$$\hat{f} = E[y | x] = \frac{\int_{-\infty}^{\infty} y \cdot p(x, y) dy}{\int_{-\infty}^{\infty} p(x, y) dy} \quad (4.9)$$

where x is the value of a random variable and y being the corresponding output, the notation E in the above relation is the conditional mean of y given x , and $p(x, y)$ is an unknown probability density function. A sample of observations comprising X and Y is used to estimate $p(x, y)$, and usually the consistent estimators suggested by Parzen [74] are applied and provide the following equation.

$$p(x, y) = \frac{1}{n(2\pi)^{\frac{d+1}{2}} \sigma^{d+1}} \sum_{i=1}^n \exp\left(-\frac{\|x-x_i\|^2 + (y-y_i)^2}{2\sigma^2}\right) \quad (4.10)$$

σ in the above relation is standard deviation and regulates smoothness of the function; using equations (4.9) and (4.10) renders

$$\hat{f}(x) = \frac{\sum_{i=1}^n y_i \exp\left(-\frac{\|x-x_i\|^2}{2\sigma^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{\|x-x_i\|^2}{2\sigma^2}\right)} \quad (4.11)$$

the above form (4.11) defines the neurons present in the output layer of GRNN and is obtained by integrating analytically after the two relations are decomposed in terms of x and y . Equation (4.11) in statistics is popularly known as *Nadaraya–Watson* regression estimator [72] and can be applied to numerical data. $\hat{f}(x)$ can be viewed as a weighted average of all of the observed values Y , and each observed value is weighted exponentially according to its Euclidean distance from x_i .

A probabilistic approach [75] such as GRNN generally requires fewer samples and less training time to learn from the input-output relation compared to the back-propagation (BP) learning mechanism, a popular paradigm used in neural network to implement learning, the latter utilizes weights to achieve this task [72]. The performance of BP may show improved performance in certain scenarios, but it is usually preferred in cases where accuracy is of great significance and when a large amount of input-output data is available, which justifies the additional time. The two main reasons in determining the use of GRNNs in the present scenario is their ability to approximate functions from sparse data and they can be trained rapidly. Besides these, the software simulation of a GRNN is easy to implement and use.

4.4 Multi-objective based optimization

Problems which involve simultaneous optimization of several objective functions that often are competing are best dealt by a multiple objective optimization (MOO) approach, and such scenarios are common in the field of electromagnetics. The methodology of MOO in computational electromagnetics has garnered great interest among researchers to obtain optimum designs of electromagnetic machines, under a given set of constraints. The approach has gained success in conventional machine design [76], and is still being actively researched. The MOO method involves the concept of non-dominated solutions, to elaborate further the contrary case of

a dominated solution is defined first. A solution is dominated if it is worse than another solution in at least one objective, while not being better than that solution in any other objective. In multiple objective optimization, the aim is to find not a single, but a set of, non-dominated solution points also known as the *Pareto Set* (PS) [76], [77]. There are several computational tools to achieve such a target, e.g., Game Theory, Genetic Algorithms (GA) and Evolutionary Strategies (ES) [78], etc. The presence of a large volume of published work involving procedures based on genetic and evolutionary approaches signifies the success accomplished by these methods [76]. In brief, such tools automatically discover the regularities or unique features in the problem landscape, and then exploit them in the form of decomposition of the problem through the combination of pieces of promising solutions found so far and perturbing the solutions slightly.

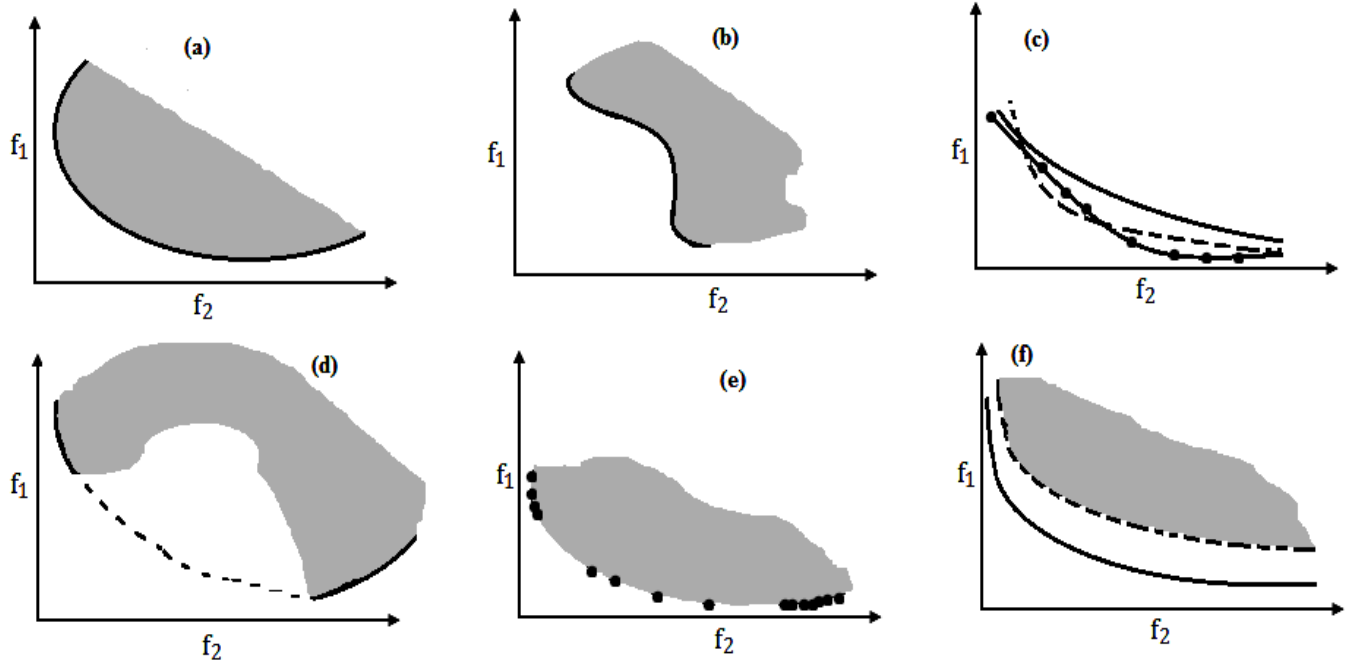


Fig 4.4. f_1 and f_2 are objective functions. (a) Convex Pareto front. (b) Non-convex Pareto front. (c) Multiple Pareto fronts denoted by bold line, line with dots and dash arising due to multi-modal condition present in either one or both objective functions. (d) Discontinuous Pareto front. (e) Non-Uniform Pareto front. (f) Deception in the Pareto front, the optimum exists from the bold curve, but the nature of search space favors the region from the dotted curve.

DEFINITION: Given a function $f: \Omega \subseteq S = \mathbb{R}^n \rightarrow \mathbb{R}, \Omega \neq \Phi$, for $\mathbf{x} \in \Omega$ the $f^* \equiv f(\mathbf{x}^*) > -\infty$ is called a global minimum iff

$$\forall \mathbf{x} \in \Omega: f(\mathbf{x}^*) \leq f(\mathbf{x}) \quad (4.12)$$

where, $\mathbf{x}^*$ is the global minimum solution(s), f is the objective function, and the set Ω is the feasible region ($\Omega \subset S$).

DEFINITION: In general, a MOO problem minimizes $F(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_k(\mathbf{x}))$ subject to $g_i(\mathbf{x}) \leq 0, i = 1, \dots, m, \mathbf{x} \in \Omega$, where $\mathbf{x}$ is an n dimensional decision variable vector ($\mathbf{x} = x_1, \dots, x_n$) [76].

DEFINITION: Considering two vectors $\mathbf{u}$ and $\mathbf{v}$, the vector $\mathbf{u} = (u_1, \dots, u_k)$ is said to Pareto dominate $\mathbf{v} = (v_1, \dots, v_k)$ (denoted by $\mathbf{u} \leq \mathbf{v}$) iff $\mathbf{u}$ is less than or equal to $\mathbf{v}$, i.e., $\forall i \in \{1, \dots, k\}, u_i \leq v_i \wedge \exists i \in \{1, \dots, k\}: u_i < v_i$.

DEFINITION: A solution $\mathbf{x} \in \Omega$ is said to be Pareto optimal with respect to Ω iff there is no $\mathbf{x}' \in \Omega$ for which $\mathbf{v} = F(\mathbf{x}') = (f_1(\mathbf{x}'), \dots, f_k(\mathbf{x}'))$ dominates $\mathbf{u} = F(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_k(\mathbf{x}))$.

DEFINITION: For a given MO problem $F(\mathbf{x})$, the Pareto optimal set ($\mathbf{P}^*$) is

$$\mathbf{P}^* = \{\mathbf{x} \in \Omega \mid \neg \exists \mathbf{x}' \in \Omega F(\mathbf{x}') \leq F(\mathbf{x})\} \quad (4.13)$$

DEFINITION: In a MO problem $F(\mathbf{x})$ and Pareto optimal set $\mathbf{P}^*$, the Pareto front ($\mathbf{PF}^*$) is given by

$$\mathbf{PF}^* = \{\mathbf{u} = F(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_k(\mathbf{x})) \mid \mathbf{x} \in \mathbf{P}^*\} \quad (4.14)$$

The above definitions provide the fundamental mathematical concepts associated with MOO. In general, functions in MOO provide a mathematical description of performance criteria that conflict, the optimization attempts to find a solution of all the objective functions having values that may be acceptable to the decision maker (DM). In most cases there exist restrictions known as *constraints* that describe the dependencies among decision variables in the example; they arise because of the characteristics of the problem under investigation, and are expressed in the form of mathematical inequalities or equalities. The number of equality constraints should be less than the number of decision variables, else the problem is overconstrained and there are no degrees of freedom left for optimizing. The difference in the number of decision variables and the number of equality constraints renders the degrees of freedom. The entity DM mentioned earlier is usually a person who is an expert in the field with considerable knowledge who introduces the

preferences related to the objectives and different solutions in a certain form. A good example includes the case when all the objective functions of a problem are combined into a single objective function, in such a scenario, the individual objective functions are assigned weights decided by the DM before incorporating them into a single function. The purpose of computational tools such as evolutionary algorithms is to retain the diversity in the population of function and parameter spaces, and to direct the search towards the optimal Pareto region. Some of the conditions that prove to be an obstacle in convergence towards the Pareto-optimal are:

- i. Multi-modality – Occurs due to the presence of more than one Pareto-optimal front, as shown in Fig 4.4 resulting in difficulties in obtaining the optimum solution.
- ii. Isolated optimum – In a situation where the search space is flat, finding an optimum point becomes difficult because useful information to guide a search method is absent. Optimization procedures use features of the search space to determine the next move, and if the optimum point is surrounded by a flat region then an exhaustive search is required.
- iii. Deception – When there is more than one local optimum solution in the search space, a condition may arise when the features of the search space favour any optimum point other than the global optimum (refer to Fig 4.4), leading to the problem of convergence towards a non-global optimum point. In scenarios where the global optimum is desirable, deception could result in a multiple objective optimization procedure missing such an optimum point or may result in considerable time to converge.
- iv. Convexity or non-convexity – In optimization algorithms, at times there is a tendency to find more intermediate solutions than to find individuals near the best solutions resulting in a bias towards some section of a Pareto-optimal region. Such an issue is found when the global Pareto-optimal front is non-convex and there exist convex local Pareto-optimal fronts.
- v. Discontinuity – This problem arises when the global Pareto-optimal front exists as a collection of discretely spaced sub- regions rather than as an entity in a continuous form.

In such a scenario, the optimization procedure has the additional task of approximating missing sections of the Pareto-optimal front.

- vi. Non-uniformity – The presence of a higher solution density in certain sections of the Pareto-optimal front compared to other regions causes the problem of non-uniformity. This uneven distribution may result in some bias towards higher density regions and also reduce the diversity in solutions.

The above issues may affect any MOO method and are problem dependent. The situations discussed until now are prevalent in the present scenario also, where the problem concerns solving electromagnetic and thermal fields and the shape of the device, there are numerous examples in the scientific literature that have discussed similar scenarios; the difference here arises because of the HTS material.

The discourses so far described the multiple objective optimization method, highlighted some of the issues associated with it, and briefly discussed the ideas to implement it using evolutionary algorithms. There are a few other methods that deviate slightly, and do not explicitly incorporate the concept of Pareto optimality in their selection structure. One such method as highlighted earlier is the *Aggregating Function*, in this procedure each objective function is first multiplied by a weighted coefficient and then all the resulting values are added as shown below.

$$\min \sum_{i=1}^n w_i f_i(x) \quad (4.15)$$

where $w_i \geq 0$ denotes the weighted coefficients. A shortcoming of this approach is its inability to generate non-convex portions of the Pareto front; this behavior is independent of any combination of weight used, and is more prevalent in linear aggregating functions. This failure could lead to a scenario where the Pareto front misses a substantial block of possibly good optimal solutions. Apart from Aggregating Functions, certain population based techniques such as Vector Evaluated Genetic Algorithm (VEGA) has a preference (bias) for the edge regions of the Pareto front, and it may miss optimal solutions that are present in the central section of the Pareto front. Such a method is applied for problems, in which a bias is desired in the selection

process, for instance, in scenarios where constraints are considered as objectives in formulating single objective optimization methods. These aspects associated with multi-objective optimization had not been investigated in this work, but there could be some benefits in terms of saving computation time. A discussion has been included here to present a thorough treatment of the subject matter.

As highlighted earlier, this work optimizes a HTS current lead by reducing heat leakage, AC loss and the amount of material used. In order to achieve these goals, objective functions that define the relationship between the specific targets such as heat leakage reduction and the parameters that could be manipulated to obtain the desired result are identified first. In the present scenario, the lead operates in a superconducting state; and the hysteresis component plays a dominant role in determining loss. To model the AC loss reduction, the self-field loss in a tubular structure was estimated using the following mathematical expression; this approach is based on the ideas first suggested by W. T. Norris [79], [80], [81]

$$Q = \frac{\mu_0 I_c^2}{\pi \delta^2} \left[(1 - \delta F) \ln(1 - \delta F) + \delta F \ln \left(1 - \frac{\delta F}{2} \right) \right] \text{ (J/m)} \quad (4.16)$$

where δ is defined as $\delta = \left(1 - \left(\frac{R_i}{R_o} \right)^2 \right)$ and F is described as $F = I_p/I_c$, the term R_i is the internal radius, the external radius is given by R_o , I_p represents the peak value of operating current, and I_c denotes the critical current. In order to reduce the heat leakage and the material used, a linear variation in the area along the length was considered, this approach is similar to the ideas presented in [59]. A higher order function could also be used here to generate even surfaces, but the first order form was employed due to its simplicity, and is defined as

$$A(x) = \left(1 + \left(\frac{A_H}{A_C} - 1 \right) \frac{x}{L} \right) A_C \quad (4.17)$$

in the above relation, $A(x)$ is the area at a distance x from the reference point, in the present scenario, reference point could either be the cold end or the warm edge, A_H represents the cross-section at the warm end, A_C denotes the area at the cold end, and L is the total length of the lead. The volume integral provides estimates for the amount of the material used. The constraint

relation applicable in this scenario is introduced in the form of the operating current density $\mathbf{J}$, this component should be less than or equal to the critical current density $\mathbf{J}_c$ ($\mathbf{J} \leq \mathbf{J}_c$), the latter in turn is largely affected by the temperature and to a lesser extent by the magnetic field. The dependency of temperature on the critical current density is defined using the following linear relationship

$$\mathbf{J}_c(T) = \mathbf{J}_{c0} \frac{1 - \frac{T}{T_c}}{1 - \frac{T_0}{T_c}} \quad (4.18)$$

where T_c is the critical temperature having the value of 89K; T_0 is 77K, the temperature of liquid Nitrogen, and $\mathbf{J}_{c0}$ denotes the current density of the example under consideration [82], [83]. As mentioned earlier, the magnetic field also influences $\mathbf{J}_c$, and this aspect is incorporated using the Kim relation (refer to Appendix B) as shown below

$$\mathbf{J}_c(\mathbf{B}) = \frac{\mathbf{J}_{c0}}{1 + \frac{\|\mathbf{B}\|}{\mathbf{B}_0}} \quad (4.19)$$

$\mathbf{J}_{c0}$ in equation (4.19) is the critical current density when the field is zero, experimental data is used to determine the value of $\mathbf{B}_0$ and it is usually assigned a value in the range of 20-30 mT, the region in which $\mathbf{J}_c$ falls to half the value of $\mathbf{J}_{c0}$. The variation in $\mathbf{J}_c$ due to $\mathbf{B}$ is small below $\mathbf{B}_0$, as evident from the available data [48], and depending on the problem its effect on $\mathbf{J}_c$ could be neglected.

In Fig 4.5 below, a comprehensive representation of the optimization procedure applied in this study has been presented. The internal radius at four different points along the length of the current lead were considered as the parameters for the optimization method. Additional parameters could be considered in this investigation, for instance a larger set of internal radii comprising more points along the length, or the inclusion of a set of outer radius points as additional parameters along with the set of internal radii, but it was avoided to reduce the computation cost. The process involves initialization of the parameters first; subsequently calls to the magnetic field and thermal solvers are made. It should be pointed out here that to avoid the repeated calls to computationally expensive solvers, the surrogate model was considered at this juncture to generate solutions, the response surface obtained by using GRNN provided a look-up

table for reference. After a solution is obtained, based on the results it is determined whether to apply any further optimization or exit the procedure. A Differential Evolution (DE) based optimizer has been employed in this study, which then updates the parameters and the process is repeated by a call to the solvers again.

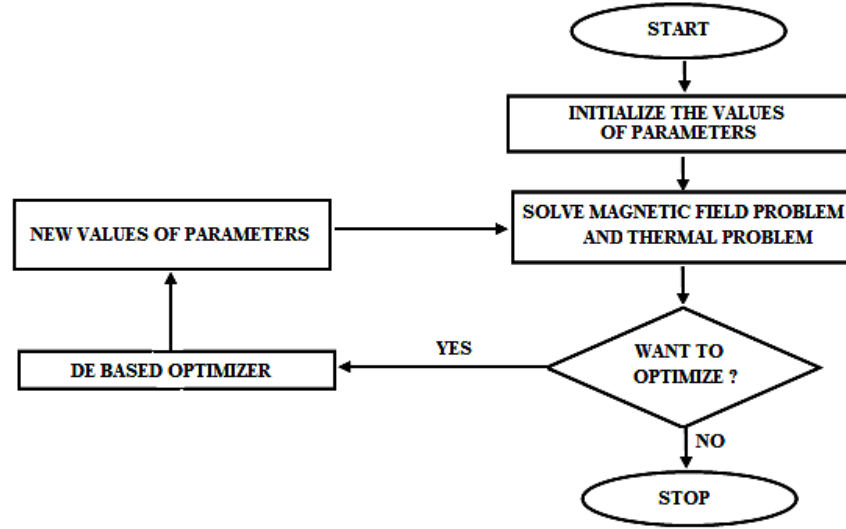


Fig 4.5. The overall view of the optimization approach. The flow diagram shows the interaction among various components that include the magnetic and thermal field solvers, a DE based optimizer and a parameter list.

The next section will provide a detailed discussion on DE concerning the theoretical aspects, the benefits it provides and the reason behind its usage. Subsequently, in section (4.6), the approach to implement the ideas presented so far will be discussed along with the results.

4.5 Differential evolution

The problem defined above can be optimized using either Deterministic approaches such as the Penalty function method, the Lagrangian method and Sequential quadratic programming or techniques that consider a different concept known as the *Stochastic methods*. The two procedures have their own advantages and drawbacks, and at times pose a challenge in determining the benefits one has over the other for the scenario under consideration. A deterministic based algorithm requires computation of the gradient and, in certain scenarios, the Hessian of the involved variables. In its favor, the proponents of Deterministic strategies argue that if a function is differentiable and convex, the convergence towards the minimum solution is

faster compared to the Stochastic mechanisms. The expression “faster” signifies that the procedure requires fewer function evaluations to reach the optimum solution. Apart from this, the mathematical treatment in Deterministic methods is more rigorous and hence it could be argued that they are more direct and replicable. In many practical engineering problems, there exists little or no prior information about the overall system behavior; in such conditions, it is very difficult to predict the response of an objective function. Stochastic methods are well equipped to deal with problems of this nature because they are zeroth order techniques, which utilize the function values only, and thereby avoid the requirement of continuity of the objective function. Apart from this, unlike deterministic algorithms, stochastic approaches are not intrinsically single objective and they can overcome local minima; the treatment of constraints is also simple and involves replacement of infeasible solutions by new feasible ones without the use of penalty functions which are prevalent in some deterministic techniques. One of the major issues with Stochastic procedures is the high number of function evaluations associated with them, but with the advent of powerful computers this problem has been addressed to an extent, and the methods have found extensive use.

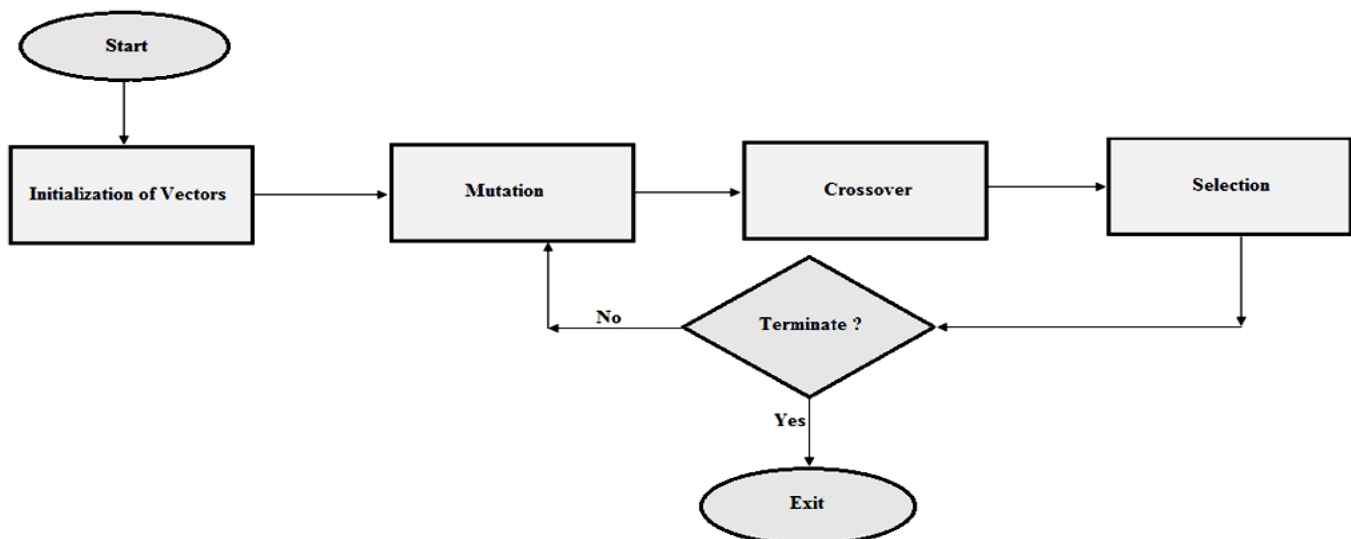


Fig 4.6. The Differential Evolution flow diagram . The process is terminated when there is no appreciable change in the fitness of the population over successive iterations.

Differential Evolution (Fig.4.6) is a simple to implement evolutionary algorithm (EA) introduced by Kenneth Price and Rainer Storn [84]. This procedure is an outcome of their research to implement simulated annealing to the Chebyshev polynomial fitting problem, a well known function used to test the performance of optimization algorithms. In general all nature-inspired

evolutionary algorithms such as Evolutionary Strategies, Genetic Algorithm , and DE involve the concepts of recombination, mutation and survival of the fittest. However, there exist a few distinctions, for instance, in ES, the parameters are encoded using floating-point numbers and they are manipulated with arithmetic operators, due to this, ES is more suitable as a continuous function optimizer. In case of GAs the parameters are often encoded as bit strings and logical operators are used to manipulate them. Genetic Algorithms are better suited for combinatorial optimization because of the encoding approach, but there are examples where GAs have been modified to operate on floating-point number and they function more as ES-type algorithms. Some form of probability distribution function is used in most of the procedures highlighted so far to determine the next generation, but in DE the offspring are generated by perturbing the present generation members with the scaled differences of randomly selected and distinct population members. In order to elaborate further, the four fundamental steps associated with DE are discussed below

- i. Population initialization – The parameters used in the DE algorithm are represented in a vector, an approach, similar to the ones used in other forms of evolutionary algorithm. The lower and upper bounds for each parameter are specified first, this is performed because parameters are a measure of physical quantities that have limited range, which is an outcome of the restrictions that occur naturally or due to design specifications. After specifying the bounds, a random number generator is used to assign values to each parameter within the bounds of the defined range as shown below

$$X_{j,i,0} = \text{rand}_j(0,1)(b_{j,U} - b_{j,L}) + b_{j,L} \quad (4.20)$$

where j represents the new random value for each parameter of the vector, $(i,0)$ indicates the generation, $\text{rand}_j(0,1)$ is a function that returns a uniformly distributed random number from within the range $[0,1)$, the subscripts U and L are the upper and lower bounds respectively, and b denotes the bound value.

- ii. Mutation – In the world of life sciences the term “mutation” implies a sudden change in the characteristics of a chromosome. This aspect is incorporated in the DE computing

paradigm by creating a perturbation with a random element. After the population of size N_p is initialized, DE mutates and recombines the population to produce a population of trial vectors. In order to create vectors for the mutant population ($\vec{P}_{v,G}$), three distinct parameter vectors $\vec{X}_{r_i^v}$ are sampled randomly from the present generation denoted by G . Subsequently, the difference of any two of these three vectors is scaled by a scalar number F and added to the third as presented in the following relation

$$\vec{P}_{v,G} = \vec{X}_{r_1^v,G} + F * (\vec{X}_{r_2^v,G} - \vec{X}_{r_3^v,G}) \quad (4.21)$$

typical values of the scalar F is in the interval of $[0.5, 1]$, the subscript v denotes distinct base vectors, and the indices r are selected from the range $[1, N_p]$, and are mutually different.

- iii. Crossover – The procedure of mutation is complemented by the technique of crossover, also known as *discrete recombination*, the purpose of this method is to gain diversity in the population and create trial vectors $\vec{U}_{j,v,G}$. A user defined criterion known as *crossover rate* (Cr) that typically lies in the range of $[0,1]$, is used to determine the source that contributes to the creation of a trial vector as described below

$$\vec{U}_{j,v,G} = \begin{cases} \vec{P}_{j,v,G} & \text{if } (\text{rand}_j(0,1) \leq Cr \text{ or } j = j_{\text{rand}}) \\ \vec{X}_{j,v,G} & \text{otherwise} \end{cases} \quad (4.22)$$

in the above relations Cr is compared to the output of a random number generator and if the value of the latter is less or equal, then the mutant vector is considered, else the parent vector also known as the *target vector* is selected.

- iv. Selection – This method is used to maintain the population size constant, over subsequent generations. Selection helps in determining whether the target or the trial vector is considered as a member for the next generation ($G+1$). This is achieved by comparing the values of the objective function $f()$ obtained by considering the trial vector and the

target vector. If the yield of the trial is less than the target, the former replaces the corresponding target vector in the next generation; otherwise the target is retained in the population as represented in the following mathematical relation.

$$\vec{X}_{j,v,G+1} = \begin{cases} \vec{U}_{j,v,G} & \text{if } (f(\vec{U}_{j,v,G}) \leq f(\vec{X}_{j,v,G})) \\ \vec{X}_{j,v,G} & \text{otherwise} \end{cases} \quad (4.23)$$

From the discussions so far it could be summarized that DE is less mathematically complicated than the deterministic methods, and the former includes randomness in the search space; the latter aspect is beneficial in obtaining the global minimum and performing a more thorough investigation of the design space. Apart from this, a notable property of Differential Evolution is that the parameters Cr and F highlighted above require a lesser amount of fine tuning than is necessary in many other EAs.

4.6 Simulations and Results

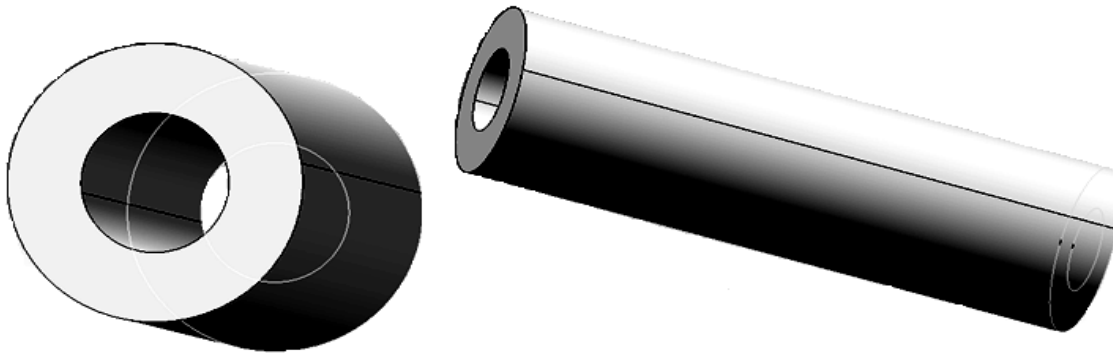


Fig 4.7. The non-optimized lead, with internal and outer radii of 5mm and 10mm respectively, and a length of 100mm.

A set of experimental data from a non-optimized HTS lead as provided in [56] (shown in the above figure 4.7) was considered for the initial parameters and conditions, in order to evaluate the ideas discussed and presented so far. The outer radius was fixed to a constant magnitude of 10mm, the initial internal radius along the entire length was set to 5mm. The lead model was specified a length of 100mm; the temperature at the warm end ($x=100\text{mm}$) was set to 77K and the cold end ($x=0$) at 4.2K. The critical current of the sample at 77K was 1680 A, and the operating current was set to 0.7-0.9 times the initial critical value with a frequency of 50Hz. In the optimization procedure, all subsequent samples carried the same boundary conditions, outer

radius and length. The variation in the internal radius was enforced along the length at four different locations ($x = 0\text{mm}$, 50mm , 75mm and 100mm), and acted as the four parameters for the optimization method. To generate more sample points in the search space the evenly distributing and space filling scheme of the Latin hypercube as described earlier in this chapter was used. This design approach was used to generate 50 sample points, and with a view to approximate the response surface and estimate its behavior, an interpolation method using the Generalized Regression Neural Network was implemented by applying the *newgrnn* MATLAB routine. The parameters F and Cr associated with Differential Evolution (refer to section 4.5) were assigned a value of 0.8 and 0.7 respectively.

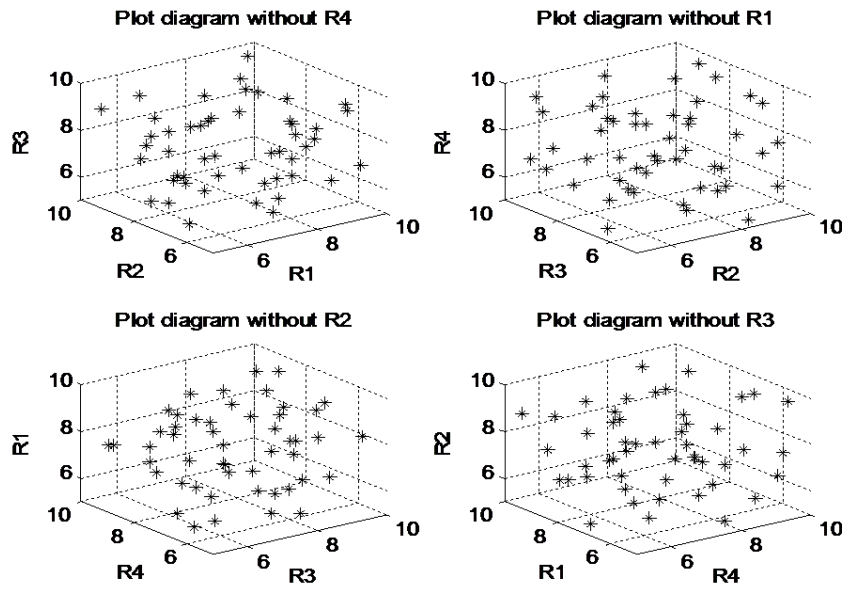


Fig 4.8. The distribution of 50 data points using the LHS technique. A 3D representation to visualize the samples conveniently, in the above plots only three parameters are considered at a time.

The following discussions provide results obtained by applying the ideas presented in this chapter. The sample points for the four parameters as shown in Fig. 4.8 were generated using latin hypercube sampling and were in the range of $5\text{mm} - 9.8\text{mm}$. Limits for the inner radii were specified to investigate the viable search space. For instance, if they reduce beyond 5mm , it will cause an increase in the cross section area and also the volume, which is not what this approach attempts to achieve and the reference [56] used for the initial estimates does not provide any data for leads having an internal radius less than 5mm . The upper limit was set to 9.8mm because if an inner radius is equal to the outer radius, then it will cause a divide by zero error (refer to equation 4.16).

Using the set of 50 sample points each carrying four parameters the developed magnetic field solver and the ThermNet solver were used to perform the steady state and transient field analysis, respectively, and the corresponding values of cross-section area, AC loss and volume were stored. Only $\frac{1}{4}$ of the geometry was analyzed because of the symmetry involved, thereby reducing the computational cost. Subsequently, to develop the Generalized Regression Neural Network based surrogate model the MATLAB routine newgrnn (for more refer to the MathWorks Documentation Center available online) was used and in order to train it, the set of 50 parameters acted as the input and the respective values of cross-section area, AC loss and volume were used as the output. Thereafter the optimization process using the stochastic method as shown in Fig. 4.5 was carried out. There were 1248 function calls for the Pareto set, and following the estimation of solution set, the next task involved determining a feasible solution and in this process, the role of the decision maker as highlighted earlier comes into play. There exist numerous methods to implement a selection procedure of this nature [85], and in this work the following criteria were used for this stage.

- i. Consider the solutions that have the lowest 15% values for AC loss.
- ii. Lead volumes greater than $1.77\text{E}+04 \text{ mm}^3$ (approximately 75% of the non-optimized lead) were excluded from the selection list.
- iii. Among the remaining solutions, the one that had the smallest cross-section area at the cold end was chosen.

The above criteria are alterable and are dependent on the approach a decision maker considers.

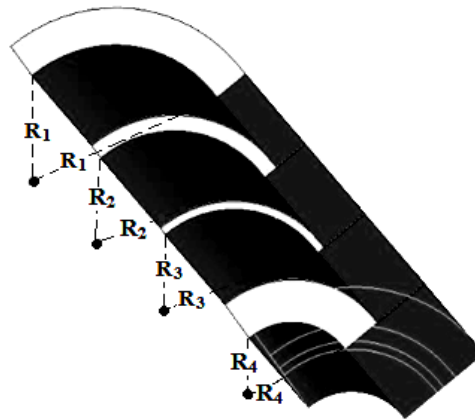


Fig 4.9. The quarter section of an optimized HTS lead.

In Fig.4.9 above the structure of an optimized HTS current lead is presented and the following table 4.1 provides the optimum values obtained through the approach presented in this work, as evident the extent of variation in the radius had less impact on the AC loss. The figure in 4.10 below describes the temperature distribution along the length of the lead.

Table 4.1 Best solution from the Pareto set.

	R_1 mm	R_2 mm	R_3 mm	R_4 mm	Loss J/m	Volume mm ³
Non - optimized	5	5	5	5	0.121	2.36E+04
Optimized	9.45	8.37	7.8	5	0.119	1.22E+04
Cross section mm² (optimized)	33.60	94.07	123.02	235.62		

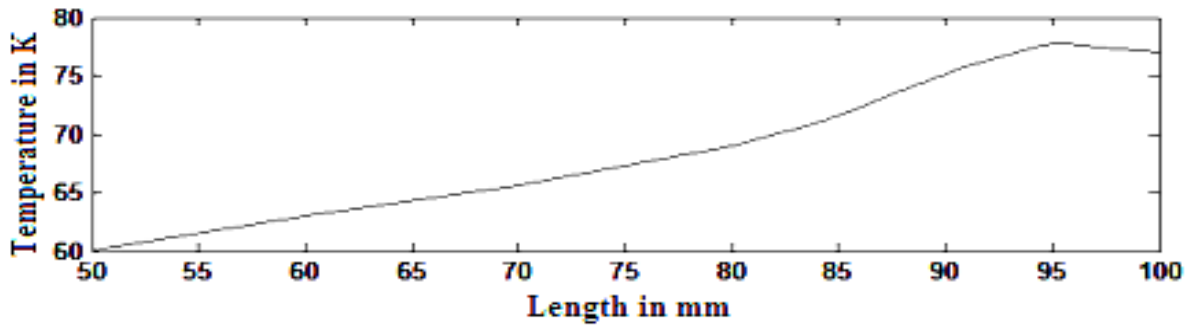


Fig 4.10. The temperature distribution from a section of the HTS lead.

4.7 Summary

The HTS based current lead was used as an example to investigate the techniques that could be applied to optimize a simple electromagnetic device constructed from HTS, this study also includes the influence of the thermal behavior of HTS in such a process. This part of the dissertation describes HTS current leads, the conditions in which they operate and highlights the scope for improving the performance and cost saving. To achieve the necessary goals, a discussion on sampling techniques, an area often neglected was presented and the use of LHS was suggested. A GRNN based surrogate technique was recommended to generate an inexpensive computational model which considered both the magnetic field and thermal field effects.

Subsequently, the fundamentals of multi-objective optimization were analyzed and the concepts involving the Pareto optimal front were used. Differential Evolution, a stochastic based optimization method, involving MOO-DE approach was applied to achieve the goals, and the benefits of such an approach were emphasized. In the field of stochastic algorithms there exist the well known “No Free Lunch Theorem”, which states that for an algorithm, any enhancement in the performance over one class of problems is offset by performance over another class. Similar to the issue in the area of applied electromagnetics, in HTS based devices too, it still remains an open question as to which algorithm is best suited for a given class of problem. This thesis does not attempt to compare DE and other evolutionary based strategies in the scenarios involving HTS. The main reason behind its preference in the present case is the ease of implementation provided by Differential Evolution. This chapter concludes by presenting the simulation results based on the ideas discussed.

CHAPTER 5

5 Conclusion and future work

Electric machines are complex devices that are difficult to design. Many times engineers use knowledge and experience to design and construct motors, generators and actuators, but it is a known fact that such an approach involves considerable time. In such a scenario, the use of computers and techniques adapted to draw the advantages from computational methods provide immense benefits in the form of time savings, better solutions and reduced wastage in the design process. This project investigated a scenario when a new material, in the present case HTS, is used to construct electromagnetic devices. The use of new materials in any system introduces some unique characteristics and changes in the overall behavior of the system, the same occurs when HTS is used in the design of electric machines. This thesis first highlighted the properties of HTS, and then mentioned the benefits and drawbacks associated with its use. From a computational perspective, HTS poses its own set of challenges and this work demonstrated the manner in which the area of computational electromagnetics addresses them. The flow of this research provided an interesting insight into how different problems associated with HTS based low frequency devices are dealt from the microscopic level to the macroscopic scale in the entire design cycle.

The research goals were stated in chapter one and could be summarized as:

- i. Define an approach to deal with the convergence issues that exist in solving magnetic field problem in HTS, and to find a scope for improvement in convergence.
- ii. Find a scheme to solve the multiphysics problem associated with HTS based current leads, that attempts to determine the solutions to two different problems, particularly magnetic and thermal field problems using 2D and 3D models respectively, and then couple the two.
- iii. Design a technique to deal with the issues common to HTS based devices specifically the reduction of heat leakage, AC loss and the material used. In this effort, investigate the

MOO method and its use in this scenario, and identify the possible objective functions and constraints.

- iv. Probe statistical based methods to determine the manner in which they could be applied to generate data of HTS devices, and address the problem of data scarcity. In addition, assist the optimization process by investigating and using computationally less expensive surrogate methods, and suggest a feasible method. Finally, recommend a method that is simple to implement and efficiently executes the MOO problem defined earlier for HTS leads.

In order to achieve the goals highlighted above, numerical methods to solve 2D magnetic field problems involving HTS were discussed in chapter three. The A-V form was considered to solve the diffusion equation involved. FEM was used for the spatial discretization and the temporal discretization was achieved using FDM, a backward time difference method was utilized to compute the time steps. An algorithm based on the Aitken method was proposed to improve the convergence in such problems. The approach was verified in a self field scenario that involved a HTS sample carrying transport current. An external field simulation was implemented using a case that consisted of a HTS sample and a ferromagnetic iron specimen, the results from the recommended method were verified. Afterwards, the numerical formulation to solve the heat equation were discussed, FEM was used again, but now it involved 3D treatment of the problem. To solve the thermal problem, the ThermNet package from Infolytica was used. Subsequently, the techniques to couple thermal and magnetic field problems were discussed and an approach based on weak coupling was recommended.

In chapter four, the properties and the working of HTS current leads were discussed. To improve their performance and reduce costs, an optimization approach based on MOO was recommended. The possible objective functions and constraints in the MOO process were defined. This thesis further explored various statistical based methods with aim to provide good sampling data points for the optimization process and an efficient implementation, the study suggested the use of the LHS procedure. This work, then examined possible surrogate methods that could be used to develop a less computationally expensive model of a HTS lead, the use of GRNN was proposed and it was implemented by using the *newgrnn* MATLAB routine. DE, a stochastic method was

considered to achieve the optimization goals and the reasons behind its use were highlighted. Finally, simulation results from the proposed optimization framework were presented.

5.1 Recommendation for Future Work

There are many areas of research relating to this dissertation that were not explored due to time constraints. Some of the major thrust areas that hold considerable potential for contributions in the future are:

- i. The application of the finite element method (FEM) to solve magnetic field problems involving high temperature superconductors (HTS) requires the use of a high density of elements. As stressed in this work on numerous occasions, HTS materials are highly nonlinear and this requires the use of high element densities to determine their field behavior accurately. The non-linearity causes slow convergence and thereby increases the computation cost, the latter aspect is aggravated further due to the computationally expensive process of mesh generation. In order to avoid this problem the use of a meshless methods (MM) in the calculation of the magnetic field could be investigated as this would skip the expensive mesh generation process.
- ii. In order to develop accurate computational models of complex systems such as motors and generators, it is imperative to investigate three-dimensional magnetic field calculations involving HTS. The conventional approaches are based on magnetic field estimation for scenarios having a small air gap, HTS based machines have larger air gaps and the existing methods may lack accuracy. Moreover, large flux densities at the end parts of coils further necessitate the use of 3D methods.
- iii. Development of a unified computational model that not only has the ability to replicate the behavior of a HTS based system in the superconducting state, but also has the capability to reproduce near quench state behavior. In this regard the flux movement phenomenon should be included in the formulation.

- iv. The multiphysics scenario presented in this work involved only the study of the magnetic field and the thermal behavior. This could further be extended to include the structural behavior of systems that are made of HTS, one of the major hurdles that lies in this aspect is the lack of data concerning the mechanical properties of HTS.
- v. Evolutionary algorithms require extensive exploration of the search space to deliver a near optimum solution. Individuals of a population-based method move towards improvement through a randomness controlled by a set of possible guidelines. For instance, in the present scenario, DE explores new regions of the search space by combining and mutating repeatedly using promising solutions. Such manipulations of the building blocks (partial solutions of a problem) either break the building blocks frequently or do not mix them effectively. The performance of the algorithm degrades further when the building blocks are spread across a large problem space. An alternative approach that could make such algorithms faster is by generating new solutions through the extraction of information from the entire set of promising solutions. The method can then devise mechanisms to learn the structure of the problem on fly and use such information to guide the exploration of the search space through proper mixing of building blocks. In this regard, probabilistic modeling of solutions such as Bayesian procedures hold a lot of promise. The optimization techniques for HTS based devices which are computationally intensive could substantially benefit from such an idea.

References

- [1] Peter J. Lee, ed., *Engineering Superconductivity*. New York, Wiley-Interscience, 2001.
- [2] Swarn S. Kalsi, *Applications of High Temperature Superconductors to Electric Power Equipment*. John Wiley and Sons Inc, 2011.
- [3] Thomas P. Sheahen, *Introduction to High- Temperature Superconductivity*. Boston, MA, USA: Kluwer Academic Publishers, 2002.
- [4] Swarn S. Kalsi, Konrad Weeber, H. Takesue, Clive Lewis, Heinz-Werner Neumueller, and Richard D. Blaugher, “Development Status of Rotating Machines Employing Superconducting Field Windings,” *Proceedings of the IEEE*, vol. 92, No. 10, pp.1688–1704, October, 2004.
- [5] Donald U. Gubser, “Superconducting motors and generators for naval applications,” *Physica C*, pp.1192–1195, 2003.
- [6] W Nick, M Frank, P Kummeth, J J Rabbers, M Wilke and K Schleicher, “Development and construction of an HTS rotor for ship propulsion application,” *Supercond. Sci. Technol.*, vol.234, pp.1–9, 2010.
- [7] Philippe J.Masson, and Cesar A. Luongo, “High Power Density Superconducting Motor for All-Electric Aircraft Propulsion,” *IEEE Trans. Appl. Supercond.*, vol. 15, no. 2, pp.2226–2229, June, 2005.
- [8] J Frauenhofer, J Grundmann, G Klaus and W Nick, “Basic concepts, status, opportunities, and challenges of electrical machines utilizing high temperature superconducting (HTS) windings,” *Supercond. Sci. Technol.*, vol.97, pp.1–10, 2008.

- [9] Gunar Klaus, Markus Wilke, Joachim Fraunhofer, Wolfgang Nick, Heinz-Werner Neumüller, “Design Challenges and Benefits of HTS Synchronous Machines,” *IEEE Power Engineering Society General Meeting*, pp.1–8, June, 2007.
- [10] Naoki Maki, Mitsuru Izumi, Masayoshi Numano, Kiyoshi Aizawa, Kagao Okumura, Katsunori Iwata, “Design Study of High-Temperature Superconducting Motors for Ship Propulsion Systems,” *Proceeding of International Conference on Electrical Machines and Systems*, Seoul, Korea, pp.1523–1527, October, 2007.
- [11] James L. Kirtley and Frederick J. Edeskuty, “Application of Superconductors to Motors, Generators, and Transmission Lines,” *Proceedings of the IEEE*, vol. 77, No. 8, pp.1143–1154, August, 1989.
- [12] C. H. Joshi, C. B. Prum, R. F. Schiferl and D. I. Driscoll, “Demonstration of Two Synchronous Motors Using High Temperature Superconducting Field Coils” *IEEE Trans. Magn.*, vol. 5, no. 2, pp.968–971, June, 1995.
- [13] Paul N. Barnes, Michael D. Sumption, and Gregory L. Rhoads, “Review of high power density superconducting generators: Present state and prospects for incorporating YBCO windings,” *Elsevier, Cryogenics*, vol. 45, pp.670–686, 2005.
- [14] K. F. Goddard, B. Lukasik, and J. K. Sykulski, “Alternative Designs of High-Temperature Superconducting Synchronous Generators,” *IEEE Trans. Appl. Supercond.*, vol. 19, no. 6, pp.3805–3811, December 2009.
- [15] David A. Cardwell, David S. Ginley, *Handbook of Superconducting Materials Volume 1: Superconductivity, Materials and Processes*. CRC Press, 2002.
- [16] A.A. Abrikosov, “On the magnetic properties of superconductors of the second group,” *Sov. Phys. JETP*, vol.5, pp.1174–1182, 1957.

- [17] Bean C P, “Magnetization of high-field superconductors,” *Rev. Mod. Phys.*, vol.36, pp.31–39, 1964.
- [18] P. W. Anderson, “Theory of flux creep in hard superconductors,” *Phys. Rev. Lett.*, vol.9, pp.309–311, 1962.
- [19] Charles Poole, Horacio Farach, and Richard Creswick, *Handbook of Superconductivity*. Elsevier Academic Press, 1999.
- [20] S.J. Collocott, R. Driver and C. Andrikidis, “Specific heat of the ceramic superconductor $\text{Bi}_2\text{Sr}_2\text{CuO}_6$ from 0.4 to 20 K” *Physica C*, vol.173, pp.117-124,1991.
- [21] J. Robert Schrieffer, and James S. Brooks. *Handbook of High-Temperature Superconductivity. Theory and Experiment*. Springer, 2007.
- [22] Anders Bondeson, Thomas Rylander, and Par Ingelstrom, *Computational Electromagnetics*. Springer, 2005.
- [23] A. M. Campbell, “An introduction to numerical methods in superconductors,” *J. Supercond. Novel Magn.*, pp.27-33, 2011.
- [24] Pei-bai Zhou, *Numerical Analysis of Electromagnetic Fields*. Springer-Verlag, 1993.
- [25] J Rhyner, “Magnetic properties and AC-losses of superconductors with power law current–voltage characteristic,” *Physica C*, vol. 212, pp.292-300, 1993.
- [26] J.K.Sykulski, R.L.Stoll and A.E.Mahdi, “Modelling HTc Superconductors for AC Power Loss Estimation,” *IEEE Trans. Magn.*, vol. 33, no. 2, pp.1568–1571, March 1997.
- [27] G Barnes, M McCulloch, and D Dew-Hughes, “Computer modeling of type II superconductors in applications,” *Supercond. Sci. Technol.*, vol 12, pp 518-522, 1999.

- [28] Rajeev Das and David A. Lowther, “Electromagnetic Device Analysis Using a Meshless Approach Coupled to a Kohonen Network,” *IEEE Trans. Magn.*, vol. 46, no. 8, pp 3381-3384, August, 2010.
- [29] Rajeev Das and David A. Lowther, “Magnetic Field Computation Using Compact Support Radial Basis,” *IEEE Trans. Magn.*, vol. 47, no. 5, pp 1362-1365, May, 2011.
- [30] M. Zhang and T. A. Coombs, “3D modelling of high- T_c superconductors by finite element software,” *Supercond. Sci. Technol.*, vol. 25, no. 1, pp. 015009-1–015009-7, January 2012.
- [31] Markus Zahn, *Electromagnetic Field Theory: A Problem Solving Approach*. Massachusetts Institute of Technology: MIT Course Ware.
- [32] D. Ruiz-Alonso, T.A. Coombs and A.M. Campbell, “Computer modeling of high-temperature superconductors using an A–V formulation,” *Supercond. Sci. Technol.*, vol.17, pp.305–310, 2004.
- [33] W.J. Carr. Jr, “Program to Study AC Loss in High- T_c Superconductors,” *Technical Report*, WJC Research and Development Jefferson Heights Pittsburgh, and United States Air Force (USAF).
- [34] Mark D. Ainslie, Weijia Yuan, Zhiyong Hong, Ruilin Pei, Tim J. Flack, and Tim A. Coombs, “ Modeling and Electrical Measurement of Transport AC Loss in HTS-Based Superconducting Coils for Electric Machines,” *IEEE Trans. Appl. Supercond.*, vol. 21, no. 3, pp.3265–3268, June 2011.
- [35] J. P. A. Bastos and N. Sadowski, *Electromagnetic Modeling by Finite Element Methods*. Boca Raton, FL, USA: CRC Press, 2003.
- [36] Gérard Meunier, *The Finite Element Method for Electromagnetic Modeling*. John Wiley and Sons Inc, 2008.

- [37] S.Elschner, F. Breuer, A.Wolf,M. Noe, L. Cowey, and J.Bock, “Characterization of BSCCO 2212 bulk material for resistive current limiters,” *IEEE Trans. Appl. Supercond.*, vol. 11, no. 1, pp.2507–2510, March, 2001.
- [38] N. Nibbio, S. Stavrev, and B. Dutoit, “Finite element method simulation of AC loss in HTS tapes with B-dependent E-J power law,” *IEEE Trans. Appl. Supercond.*, vol. 11, no. 1, pp.2631–2634, March, 2001.
- [39] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*. Society for Industrial and Applied Mathematics (SIAM). Philadelphia, PA, USA, 2003.
- [40] Michael. T. Heath, *Scientific Computing—An Introductory Survey*. New York, USA: McGraw-Hill, 2002.
- [41] Emmanuel Vinot, Gérard Meunier, and Pascal Tixador, “Different Formulations to Model Superconductors,” *IEEE Trans. Magn.*, vol. 36, no. 4, pp.1226–1229, July, 2000.
- [42] Jorge Nocedal, and Stephen J. Wright, *Numerical Optimization*. Springer, 2006.
- [43] Nicholas J. Higham, *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics (SIAM). Philadelphia, PA, USA, 2002.
- [44] David S. Watkins, *Fundamentals of Matrix Computations*. John Wiley and Sons Inc, 2002.
- [45] A. C. Aitken, “On Bernoulli’s numerical solution of algebraic equations,” *Proc. Roy. Soc. Edinburgh*, vol. 46, pp. 289–305, 1926.
- [46] Christopher J. Zarowski, *An Introduction to Numerical Analysis for Electrical and Computer Engineers*. John Wiley and Sons Inc, 2004.

- [47] Rajeev Das and David A. Lowther, “Acceleration of Field Computation Involving HTS,” *IEEE Trans. Magn.*, vol. 49, no. 5, pp. 1785–1788, May, 2013.
- [48] J. Bock, H. Bestgen, S. Elscbner and E. Preisler, “Large Shaped Parts of Melt Cast BSCCO for Applications in Electrical Engineering,” *IEEE Trans. Appl. Supercond.*, vol. 3, no. 1, pp. 1659–1662, March, 1993.
- [49] V. Sokolovsky, V. Meerovich, and V. Dikovsky, “Penetration of AC magnetic field into bulk high-temperature superconductors: Experiment and simulation,” *J. Appl. Phys.*, vol. 95, no. 11, pp. 6693–6695, June, 2004.
- [50] Kenneth H. Huebner, Donald L. Dewhirst, Douglas E. Smith, and Ted G Byrom. *The Finite Element Methods for Engineers*. John Wiley and Sons Inc, 2001.
- [51] Roland W. Lewis, Perumal Nithiarasu Kankanhalli, and N. Seetharamu. *Fundamentals of the Finite Element Method for Heat and Fluid Flow*. John Wiley and Sons Inc, 2004.
- [52] Johan Driesen, Ronnie Belmans, and K. Hameyer, “Methodologies for Coupled Transient Electromagnetic-Thermal Finite Element Modeling of Electrical Energy Transducers,” *IEEE Trans. Ind. Appl.*, vol. 38, no. 5, pp. 1244-1250, September, 2002.
- [53] K. Hameyer and R. Belmans, *Numerical Modelling and Design of Electrical Machines and Devices*. WIT Press, 1999.
- [54] K. Hameyer, Johan Driesen, Herbert De Gersem and Ronnie Belmans, “The Classification of Coupled Field Problems,” *IEEE Trans. Magn.*, vol. 35, no. 3, pp. 1618-1621, May, 1999.
- [55] J.P. A. Bastos, N. Sadowski, and R. Carlson, “A Modeling Approach of a Coupled problem between Electrical current and its Thermal effects,” *IEEE Trans. Magn.*, vol. 26, no. 2, pp. 536-539, March, 1990.

- [56] Jian Li, Mingzhe Rong, Yi Wu, and Jianyong Lou, "Electromagnetic-Thermal Characteristics of MCP BSCCO-2212 Tube During Quenching Process in Self-Field," *IEEE Trans. Magn.*, vol. 42, no. 4, pp. 931-934, April, 2006.
- [57] John R. Hull, "High-Temperature Superconducting Current Leads," *IEEE Trans. Appl. Supercond.*, vol. 3, no. 1, pp. 869-875, March, 1993.
- [58] S. Y. Seol, Y. S. Cha, R. C. Niemann, and J. R. Hull, "Prediction of Burnout of a Conduction-Cooled BSCCO Current Lead," *IEEE Trans. Appl. Supercond.*, vol. 7, no. 2, pp. 696-699, June, 1997.
- [59] S. Y. Seol, J. R. Hull, and M.C. Chyu, "Optimization Of High-Temperature Superconductor Current Leads," *IEEE Trans. Appl. Supercond.*, vol. 5, no. 2, pp. 785-788, June, 1995.
- [60] Alexander Forrester, Andras Sobester, and Andy Keane, *Engineering Design via Surrogate Modelling*. John Wiley & Sons, 2008.
- [61] Thomas J. Santner, Brian J. Williams, and William I. Notz, *The Design and Analysis of Computer Experiments*. Springer-Verlag, 2003.
- [62] Max D Morris, "Factorial Sampling Plans for Preliminary Computational Experiments," *Technometrics*, vol.33, no.2, pp.161-177, May, 1991.
- [63] Urmila M. Diwekar and Jayant R. Kalagnanam, "Efficient Sampling Technique for Optimization under Uncertainty," *American Institute of Chemical Engineering Journal.*, vol. 43, no.2. pp. 440-447, 1997.
- [64] T. W. Simpson, J. D. Peplinski, P. N. Koch and J. K. Allen, "Metamodels for Computer-based Engineering Design: Survey and recommendations," *Springer-Verlag.*, vol.17, no.2, pp. 129-150, July, 2001.

- [65] D. F. Specht, "A general regression neural network," *IEEE Trans. Neural Netw.*, vol. 2, no. 6, pp. 568–576, November, 1991.
- [66] R. Jin, W. Chen, and T. Simpson, "Comparative studies of metamodeling techniques under multiple modeling criteria," *AIAA, Tech. Rep.* 2000-4801, pp. 1–13, 2000.
- [67] Varadarajan, S., Chen, W. and Pelka, C., "The Robust Concept Exploration Method with Enhanced Model Approximation Capabilities," *Engineering Optimization*, vol. 32, no. 3, pp. 309-334, 2000.
- [68] Giunta, A. and Watson, L. T., "A Comparison of Approximation Modeling Techniques: Polynomial Versus Interpolating Models," *Technical report, AIAA 98-4758*, vol. 1, pp. 392-404, 1998.
- [69] Zahid Malik and Kashif Rashid, "Comparison of Optimization by Response Surface Methodology with Neurofuzzy Methods," *IEEE Trans. Magn.*, vol. 36, no. 1, pp. 241-257, January, 2000.
- [70] Ujjwal Sinha, "A Design and Optimization Assistant for Induction Motors and Generators," Ph.D. dissertation, Massachusetts Institute of Technology, June, 1998.
- [71] Glenn Hawe, "Kriging Methods for Constrained Multi-Objective Electromagnetic Design Optimization," Ph.D. dissertation, University of Southampton, February, 2008.
- [72] S. Haykin, *Neural Networks; A Comprehensive Foundation*. Englewood Cliffs, NJ: Prentice-Hall, 1999.
- [73] John Y. Goulermas, Xiao-Jun Zeng, Panos Liatsis, and Jason F. Ralph, "Generalized Regression Neural Networks With Multiple-Bandwidth Sharing and Hybrid Optimization," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol.37, No.6, pp.1434-1445, December, 2007.

- [74] E. Parzen, “On estimation of a probability density function and mode,” *Ann. Math. Statist.*, vol. 33, pp. 1065-1076, September, 1962.
- [75] D. F. Specht, “Generation of Polynomial Discriminant Functions for Pattern Recognition,” *IEEE Trans. Electronic Computers*, vol.16, no. 3, pp. 308-319, June, 1967.
- [76] Paolo Di Barba, *Multiobjective Shape Design in Electricity and Magnetism*. Lecture Notes in Electrical Engineering, Springer, 2010.
- [77] Jurgen Branke, Kalyanmoy Deb, Kaisa Miettinen, Roman Słowiński, *Multiobjective Optimization- Interactive and Evolutionary Approaches*. Springer-Verlag, 2008.
- [78] Carlos A. Coello Coello, Gary B. Lamont, David A. Van Veldhuizen, *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer-Verlag, 2007.
- [79] W. T. Norris, “Calculation of hysteresis losses in hard superconductors carrying ac: isolated conductors and edges of thin sheets,” *J. Phys. D: Appl. Phys.* 3, pp. 489-507, 1970.
- [80] F.Gomory and L.Gherardi, “Transport AC losses in round superconducting wire consisting of two concentric shells with different critical current density,” *Physica C*, vol. 280, pp. 151–157, 1997.
- [81] Rajeev Das, Fernando B Oliveira, Frederico G Guimaraes, and David A. Lowther, “The Optimal Design of HTS Devices,” *IEEE Trans. Magn.*, vol. 50, no. 2, pp. 249– 252, February, 2014.
- [82] K. Berger, J. Lévêque, D. Netter, B. Douine, and A. Rezzoug, “Influence of Temperature and/or Field Dependences of the E J Power Law on Trapped Magnetic Field in Bulk YBaCuO,” *IEEE Trans. Appl. Supercond.*, vol. 17, no. 2, pp. 3028– 3031 , June, 2007.

- [83] Y. B. Kim, C. F. Hempstead, and A.R Strnad, “Magnetization and Critical Supercurrents,” *Phys. Rev.*, vol.129, pp.528–535, 1963.
- [84] Kenneth V. Price, Rainer M. Storn and Jouni A. Lampinen, *Differential Evolution A Practical Approach to Global Optimization*. Springer, 2005.
- [85] Frederico G. Guimarães, Felipe Campelo, Rodney R. Saldanha, Hajime Igarashi, Ricardo H. C. Takahashi, and Jaime A. Ramírez, “A Multiobjective Proposal for the TEAM Benchmark Problem 22,” *IEEE Trans. Magn.*, vol. 42, no. 4, pp.1471–1474, April, 2006..
- [86] Naoyuki Amemiya, Shun-ichi Murasawa, Nobuya Banno, Kengo Miyamoto, “Numerical modelings of superconducting wires for AC loss calculations,” *Elsevier, Physica C*, vol. 310, pp.16–29, 1998.
- [87] T.A. Coombs, A.M. Campbell, A. Murphy and M. Emmens, “A fast algorithm for calculating the critical state in superconductors,” *Compel*, vol.20, pp.240-252, 2001.
- [88] Y.S. Cha, “Magnetic diffusion in high- T_c superconductors,” *Elsevier, Physica C*, vol.330, pp.1–8, March, 2000.
- [89] David Ruiz-Alonso, T. A. Coombs, and A. M. Campbell, “Numerical Analysis of High-Temperature Superconductors with the Critical-State Model,” *IEEE Trans. Appl. Supercond*, vol. 14, no. 4, pp.2053–2063, December, 2004.
- [90] A Gurevich, H. kupfer, “Time scales of the flux creep in superconductors,” *Phys. Rev.*, vol. 48, pp.6477–6487, September, 1993.
- [91] Ernst Helmut Brandt, “Superconductor disks and cylinders in an axial magnetic field. I. Flux penetration and magnetization curves,” *Phys. Rev.*, vol.58, pp.6506–6522, September, 1998.

Appendix

Appendix A

A discussion on the fundamental equations involved in electromagnetics has been presented here. The Maxwell's Equations and its various forms are described briefly to highlight their significance in different kinds of electromagnetic problems. Both sets of equations i.e., the integral and the differential forms have been included.

Faraday's Law – It couples the electric and magnetic fields, and forms the working principles of many electromagnetic devices such as transformers, generators etc. In brief, it could be stated that a changing magnetic field produces an induced voltage and is defined by the following equations.

$$\nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t} \quad (1)$$

$$\oint \mathbf{E} \cdot d\mathbf{l} = - \frac{d\phi}{dt} \quad (2)$$

where $\frac{\partial \mathbf{B}}{\partial t}$ and $\frac{d\phi}{dt}$ are the change in the magnetic field and magnetic flux respectively with respect to time and $\mathbf{E}$ is the electric field generated. The integral in equation (2) is a closed contour integral over the path l .

Ampere's Law - It describes the relationship between the magnetic field intensity and the current that generates it. The two are related in the following manner.

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \quad (3)$$

$$\oint \mathbf{H} \cdot d\mathbf{l} = \int \left(\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) \cdot d\mathbf{s} \quad (4)$$

$\mathbf{H}$ in the above two equations is the magnetic field intensity, $\mathbf{J}$ is the current density, and $\mathbf{D}$ is the displacement current density. When there is no source $\mathbf{J}$ will not exist. In equation (4) the left hand side is a closed contour integral and the right hand side is an integral over the closed

surface s . The term $\mathbf{D}$ was introduced by Maxwell to guarantee charge conservation, and allows an interpretation that is converse of the Faraday's Law, implying a changing electric field can produce a magnetic field.

Gauss's Law (for E) – It states that the electric flux through a closed surface is proportional to the enclosed charge, or in other words

$$\nabla \cdot \mathbf{D} = \rho_v \quad (5)$$

$$\oint \mathbf{D} \cdot d\mathbf{s} = Q \quad (6)$$

the integration in equation (6) is an integral over the closed surface s , ρ_v is the charge density distributed over the volume, Q is the total charge enclosed by surface s . It should be emphasized here that any charge outside the surface does not contribute to the flux through the surface. In source free region space, ρ_v and Q are assigned zero.

Gauss's Law (for B) – It defines that the total magnetic flux through a closed surface is zero, and this relation is represented as

$$\nabla \cdot \mathbf{B} = 0 \quad (7)$$

$$\oint \mathbf{B} \cdot d\mathbf{s} = 0 \quad (8)$$

similar to equation (6), the equation in (8) is also a closed surface integral over s . This law signifies that there are no monopoles in magnetism.

The above four laws constitute the Maxwell's equations, but to apply them at the macroscopic levels there are three constitutive relations required, which take into account the properties of the medium, the first connects the magnetic field $\mathbf{B}$ and the magnetic field intensity $\mathbf{H}$, the second relates the displacement current density $\mathbf{D}$ with the electric field $\mathbf{E}$, and the remaining provides the relationship between the current density $\mathbf{J}$ and the electric field $\mathbf{E}$ as shown in equations (9), (10) and (11) respectively.

$$\mathbf{B} = \mu \mathbf{H} \quad (9)$$

$$\mathbf{D} = \epsilon \mathbf{E} \quad (10)$$

$$\mathbf{J} = \sigma \mathbf{E} \quad (11)$$

μ in (9) represents the permeability of the medium, ϵ in (10) is the permittivity and σ in (11) is the conductivity of the medium.

To appreciate the above discussions so far and also to help in understanding the chapters presented, the following paragraphs digress slightly to provide certain aspects of vector calculus. The operators involved in such a mathematical approach, their importance and the manner they work on a scalar function U and a vector function $\mathbf{A}$. In many areas of engineering as in the case of electromagnetics, the problems are defined through partial differential equations (PDEs), without deviating further; the emphasis here will be to highlight the differential techniques. There are three types of operators involved, namely, the **gradient** (∇) which works on a scalar function, the **divergence** ($\nabla \cdot$), and the **curl** ($\nabla \times$) the latter two relate to vector functions.

Gradient – It provides the maximum spatial rate of change of a scalar function both in terms of the magnitude and direction.

$$\nabla U = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy + \frac{\partial U}{\partial z} dz \quad (12)$$

The gradient points in the direction of maximum change in scalar function as mentioned above, and it provides a vector function. It is always normal to a constant value surface.

Divergence – It is defined as the net flow of the flux of vector $\mathbf{A}$ out of a small volume through the surface enclosing the volume, or in other words it is a scalar field that describes the strength of local sources or sinks. The divergence in cartesian coordinates is given by

$$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \quad (13)$$

Curl - It renders a vector field that describes the local rate of rotation of the vector $\mathbf{A}$ over a small area and acts in a direction normal to the area. It is represented as

$$\nabla \times \mathbf{A} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{i} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{j} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{k} \quad (14)$$

The magnitude of the curl provides the maximum rotation of the vector per unit area at a point.

Till now the physical interpretations along with the formulation involved in the differential form of vector calculus were stressed, subsequent details will throw more light into the possible combinations of the operators and the vector identities. Such relations help in the study of electromagnetics and in understanding of the scenarios involved with them.

$$\nabla \cdot (\nabla U) \quad (15)$$

$$\nabla \times (\nabla U) \quad (16)$$

$$\nabla(\nabla \cdot \mathbf{A}) \quad (17)$$

$$\nabla \cdot (\nabla \times \mathbf{A}) \quad (18)$$

$$\nabla \times (\nabla \times \mathbf{A}) \quad (19)$$

Considering additional functions a scalar V and a vector $\mathbf{B}$, some of the identities that are useful are as follows

$$\nabla(UV) = U(\nabla V) + V(\nabla U) \quad (20)$$

$$\nabla \cdot (U\mathbf{A}) = U(\nabla \cdot \mathbf{A}) + (\nabla U) \cdot \mathbf{A} \quad (21)$$

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = -\mathbf{A} \cdot (\nabla \times \mathbf{B}) + (\nabla \times \mathbf{A}) \cdot \mathbf{B} \quad (22)$$

$$\nabla \times (U\mathbf{A}) = U(\nabla \times \mathbf{A}) + (\nabla U) \times \mathbf{A} \quad (23)$$

As mentioned earlier, the relations above are useful in terms of formulating governing equations in electromagnetics, for instance if equations (12), (13) and (15) are considered, then on substitution and further simplification, the following *Laplace operator* is obtained i.e.,

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (24)$$

In electromagnetics the Laplace operator is applied to define the widely used Poisson's and Laplace equations.

In chapter 3, the formulations for Eddy currents were described. The two main approaches, i.e., the A-V formulation and the T- Ω method are widely used to handle problems of such nature. Extending the discussions further here to have completeness of the matter, under static conditions, the electric scalar potential V and the magnetic scalar potential Ω could be used in place of the intensity components namely, the electric field intensity $\mathbf{E}$ and the magnetic field intensity $\mathbf{H}$. This is possible because when the time derivative component in equation (1) and the right hand side of the equation (3) are zero then $\mathbf{E}$ becomes curl free, and so does the curl of $\mathbf{H}$ respectively, providing the following relations:

$$\mathbf{E} = -\nabla V \quad (25)$$

$$\mathbf{H} = -\nabla \Omega \quad (26)$$

equations (25) and (26) are useful in combination with the vector functions such as the magnetic vector potential and the electric vector potential, in addressing general electromagnetic field problems.

$$\mathbf{J} = \begin{cases} \mathbf{J}_c, & \text{for partial or complete magnetic flux penetration} \\ 0, & \text{for no magnetic flux penetration} \end{cases} \quad (1)$$

If the superconducting material exhibits high exponential value, i.e., the n term in the power law discussed in the chapters earlier and the following sub section, which signifies very strong pinning, then the BCS model is valid. Usually in practical applications, the pinning force is not that strong to satisfy the condition present in the BCS.

E-J model – The phenomena of flux creep as shown in Fig 1 is caused by thermal activations [90]. The BCS model does not account for this behavior of superconductors and hence fail in predicting the response of the material for AC losses. To overcome this limitation J Rhyner [25] proposed the E-J power law, and is defined as

$$\frac{\mathbf{E}}{\mathbf{E}_c} = \left(\frac{\mathbf{J}}{\mathbf{J}_c} \right)^n \quad (2)$$

where $\mathbf{J}$ is the current density, $\mathbf{J}_c$ is the critical current density, $\mathbf{E}$ and $\mathbf{E}_c$ are the electric field and the critical electric field at $\mathbf{J}_c$ respectively and n is the exponent term. The value of $\mathbf{E}_c$ is usually considered as $\mathbf{E}_c = 1 \times 10^{-4}$ V/m. The parameter n is dependent on the pinning strength and has the limiting case of $1 < n < \infty$, when n is equal to 1, it corresponds to a purely resistive material satisfying the Ohm's law, and for very large values such as $n \geq 50$ [91], the power law is close to the BCS.

Kim model – As discussed earlier, the BCS model states that wherever there is flux there will be current and the magnitude will be equal to the constant current density $\mathbf{J}_c$. To incorporate the flux creep phenomena and correctly predict the behavior of a superconductor, the power law was introduced. Both these procedures lack the ability to include the effects of the magnetic field $\mathbf{B}$. To overcome this shortcoming, Kim extended the BCS model to take into account the magnetic field dependence [83]. This technique also known as the Kim model or the magnetic field dependence model provides the following relationship between the critical current density and the magnetic field.

$$\mathbf{J}(\mathbf{B}, T) = \frac{\mathbf{J}_{c0}(T)}{1 + \frac{\|\mathbf{B}\|}{B_0}} \quad (3)$$

$\mathbf{J}_{c0}(T)$ is the critical current density at zero applied field and at a given temperature, B_0 is a constant that is determined from experimental data, and is the value of the field when $\mathbf{J}_c$ reduces to half the value of $\mathbf{J}_{c0}(T)$. As evident from Fig 1 the critical current density tends to decrease with the increase in magnetic field.

Anderson's model – It is also known as the Anderson–Kim model, and relates the critical current density to the pinning strength and the electric field [18] in the following manner

$$\mathbf{J}_c = \mathbf{J}_{c0} \left[1 - \frac{kT}{U_0} \ln \left(\frac{Bd\omega}{E_{min}} \right) \right] \quad (4)$$

where k is the Boltzmann constant, U_0 an energy parameter known as the pinning strength, having dependencies on magnetic field and temperature, $\mathbf{B}$ is the induced magnetic field, E_{min} is the minimum electric field that defines $\mathbf{J}_c$, d is the hopping distance of the flux quanta, and ω is the frequency of a flux hopping occurrence. The present discussion does not delve further into the details of this relation and the material behavior at the microscopic levels, for more refer [3].

Appendix C

This section discusses the ideas behind the FEM formulation when the method of weighted residuals (MWR) is used; it will also present the well known Galerkin method which is a special case of MWR. The residual method is used to approximate solutions to linear and non-linear PDEs, and helps in formulating the element equations in FEM. Consider the following differential governing equation

$$A(\varphi) = f \quad (1)$$

equation (1) is applicable to a domain Ω bounded by the surface Γ , and f represents the known values. The boundary conditions are specified on Γ . In such a scenario to apply the MWR, the solution for φ is approximated using $\hat{\varphi}$. The approximate function $\hat{\varphi}$ is made to satisfy the differential equation (1) along with the boundary conditions. This is achieved by specifying $\hat{\varphi}$ in terms of the functional dependence on all but one independent variable, the latter is left unspecified. In order to realize this, a linear combination of basis functions is chosen from a linearly independent set in the following manner

$$\hat{\varphi} = \sum_{i=1}^m \phi_i C_i \quad (2)$$

where ϕ_i is the assumed function, and C_i is the basis function of the independent variable. Usually C_i are constants or functions of time for steady-state or unsteady problems respectively. The summation limit m represents the number of C_i unknowns. When $\hat{\varphi}$ is substituted in equation (1), the resultant carries an error also known as the *residual error* (R), i.e.,

$$A(\hat{\varphi}) - f \neq 0 \quad (3)$$

Or,

$$A(\hat{\varphi}) - f = R \quad (4)$$

The concept behind MWR is to force the error in equation (4) to zero in some average sense. This is accomplished by considering a weighted average of the residual error carrying the condition that the weighted average vanishes over Ω . It could be represented as

$$\int_{\Omega} [A(\phi) - f] W_i d\Omega = \int_{\Omega} R W_i d\Omega = 0, \quad i = 1, 2, 3, \dots, m \quad (5)$$

where W_i is the *weighting function* also known as the *test function*. The result is a set of m algebraic equations for the unknown constants C_i . The nature of error distribution in the relation (5) is dependent on the weighting function considered, and based on the latter MWR could further be classified as:

- i. Collocation method.
- ii. Least square method.
- iii. Galerkin method.
- iv. Method of moments.

The present work uses the Galerkin method; and hence the following paragraphs would restrict themselves in highlighting the various aspects of this approach, for more on the remaining techniques refer [24].

The Galerkin technique is popular because it provides greater generality, apart from this there are certain areas in machine analysis where the variational expression cannot be applied. In this technique, the weighting function (W_i) has the same form as the finite element shape function. In terms of FEM, the approximate function as described in (2) could be represented as

$$\phi = \sum_{i=1}^n \phi_i N_i \quad (6)$$

where n now represents the total number of nodes in the meshes, ϕ_i is the function value for node i , and N_i is the shape function associated with the node i . To elaborate further, consider the equations (4) and (5) again, as mentioned earlier, the shape function is now the weighting function, hence the residual error for node i is given by

$$R = \int_{\Omega} [A(\phi) - f] N_i d\Omega = 0 \quad (7)$$

Assuming there is l number of nodes with unknown values, as there could be few nodes among n that exist on the boundary with known values, the representation in (7) will furnish a set of l equations. The local residual error in each element could be written as

$$R^e = \int_{\Omega^e} [A(\phi^e) - f] N_i^e d\Omega^e = 0 \quad (i = 1, 2, 3) \quad (8)$$

In the above relation triangular elements made from three nodes were considered, hence i has numbers from one to three. The approximate function for each element is

$$\phi^e = \sum_{i=1}^3 \phi_i^e N_i^e \quad (9)$$

Appendix D

A detailed mathematical description and the theoretical aspects behind the SSM and the Aitken method are discussed in this section. It will provide an in-depth insight, and thereby assist in appreciating their use in solving the problems discussed in this work.

DEFINITION: A space S is a metric space if the following conditions are satisfied. Considering two elements $x, y \mid x, y \in S$. Let d represent the distance between x and y defined as

$$d = d(x, y) \quad (1)$$

the satisfaction of all the following properties ensures S is a metric space.

$$d(x, y) = d(y, x) \quad (2)$$

$$d(x, y) \geq 0 \quad (3)$$

$$d(x, y) = 0, \text{ iff } y = x \quad (4)$$

$$d(x, y) \leq d(x, y) + d(x, z) \mid z \in S \quad (5)$$

DEFINITION: Assuming a sequence $\{x_n\} \in \mathbb{R}^N$, $\{x_n\}$ converges to a point x^* that could be written as $\lim_{m \rightarrow \infty} x_m = x^*$, if for any $\epsilon > 0$, there is an index Z , where $Z \in I \mid$

$$\|x_m - x^*\| \leq \epsilon \quad \forall m \geq Z \quad (6)$$

DEFINITION: A sequence is a Cauchy sequence, if for any $\epsilon > 0$ there is an index $Z \mid$

$$\|x_k - x_l\| \leq \epsilon \quad \forall k \geq Z \text{ and } l \geq Z, k \neq l \quad (7)$$

A sequence converges, iff it is a Cauchy sequence.

DEFINITION: Lipschitz continuity is used to measure change in function values with respect to change in the independent variable. A function f of the nature $f: D \rightarrow \mathbb{R}^N$, if x_1 and $x_2 \in D$, then

variation in the input $|x_1 - x_2|$ will have a corresponding change of $\|f(x_1) - f(x_2)\|$ in the output. f has Lipschitz continuity, if there is a constant $L > 0$ known as the Lipschitz constant

$$\|f(x_1) - f(x_2)\| \leq L \|x_1 - x_2\| \quad \forall x_1, x_2 \in D \quad (8)$$

The value of L is dependent on the function f .

DEFINITION: A mapping $M : X \rightarrow X$ is known as contraction mapping on X , if there is $\alpha \in \mathbb{R}$ $| 0 < \alpha < 1$, and $\forall x, y \in X$

$$d(Mx, My) \leq \alpha d(x, y) \quad (9)$$

When M is applied to the points x and y , the points get closer if M is contracting in nature, the function $d()$ provides the distance between points inside a metric space, α is a Lipschitz constant.

To prove the convergence property of SSM, a closed subset $\Omega \subset \mathbb{R}^N$ is considered, and let P be a contraction mapping on Ω with Lipschitz constant $0 < \alpha < 1$ $| P(x) \in \Omega \quad \forall x \in \Omega$, here Ω is a metric space. Considering an iteration of the following nature that depicts the SSM behavior,

$$x_{n+1} = P(x_n) \quad (10)$$

the proof involves the existence of Cauchy sequence that converges linearly in Ω with a factor α , and the presence of an unique fixed point x^* of P , where $x^* \in \Omega$. The nature of the sequence when P is repeatedly applied on x_0 is given by

$$x_1 = P x_0, x_2 = P x_1 = P^2 x_0, \dots x_n = P^n x_0 \quad (11)$$

Using the above equation and the contraction mapping definition the following relation is obtained

$$\begin{aligned} d(x_{i+1}, x_i) &= d(Px_i, Px_{i-1}) \\ &\leq \alpha d(x_i, x_{i-1}) \\ &= \alpha d(x_{i-1}, x_{i-2}) \end{aligned} \quad (12)$$

$$\begin{aligned}
&\leq \alpha^2 d(x_{i-1}, x_{i-2}) \quad , \text{ after applying the mapping } P \\
&\vdots \\
&\leq \alpha^i d(x_1, x_0)
\end{aligned}$$

from (12), it could be concluded that the sequence $\{x_n\}$ is bounded $\forall i \geq 1$, therefore

$$\|x_n - x_0\| = \left\| \sum_{i=0}^{n-1} x_{i+1} - x_i \right\| \quad (13)$$

$$\leq \sum_{i=0}^{n-1} \|x_{i+1} - x_i\| \quad , \text{ from the triangular inequality} \quad (14)$$

$$\leq \|x_1 - x_0\| \sum_{i=0}^{n-1} \alpha^i \quad , \text{ as performed in (12)} \quad (15)$$

$$= \|x_1 - x_0\| \frac{(1-\alpha^n)}{(1-\alpha)} \quad , \text{ after using geometric series sum} \quad (16)$$

$$\leq \|x_1 - x_0\| \frac{1}{(1-\alpha)} \quad , \text{ since } 0 < \alpha < 1, \text{ and } 1 - \alpha^n < 1 \quad (17)$$

now assuming $\forall n, k \geq 0$, and considering the summation starting at 1 instead of 0 will provide

$$\|x_{n+k} - x_n\| = \|P(x_{n+k-1}) - P(x_{n-1})\| \quad (18)$$

$$\leq \alpha \|x_{n+k-1} - x_{n-1}\|$$

$$\leq \alpha \|P(x_{n+k-2}) - P(x_{n-2})\|$$

$$\leq \alpha^2 \|x_{n+k-2} - x_{n-2}\|$$

$$\vdots$$

$$\leq \alpha^n \|x_k - x_0\|$$

$$\leq \|x_1 - x_0\| \frac{\alpha^n}{(1-\alpha)} \quad , \text{ after using triangular inequality and the geometric series sum} \quad (19)$$

In the above equation (19), the value in the right hand side of the inequality could be made small by considering n very large due to the condition $0 < \alpha < 1$ and the value of $\|x_1 - x_0\|$ being fixed. Hence $\{x_n\}$ is a Cauchy sequence, $\Rightarrow x_n \rightarrow x^* \in \Omega$. To prove the uniqueness, the triangular inequality and the convergence property described above are applied as shown below

$$d(Px^*, x^*) \leq d(x^*, x_k^*) + d(x_k^*, Px^*) \quad (20)$$

$$\leq d(x^*, x_k^*) + \alpha d(x_{k-1}^*, x^*) \quad (21)$$

if $k \rightarrow \infty$, and $x_n \rightarrow x^*$ then $d(x_k^*, Px^*) = 0 \Rightarrow Px^* = x^*$ from (4) as used in the definition of metric spaces, hence x^* is a fixed point as $Px^* = x^*$.

Assume the existence of another fixed point y , i.e., $Py = y$ then $d(x^*, y) = d(Px^*, Py) \leq \alpha d(x^*, y) \Rightarrow x^* = y$, from the convergence relation presented in (18) - (19), and the metric space property defined in equation (4), consequently x^* is unique.

The following paragraphs would derive the formulation used in Chapter 3 for the Aitken method.

DEFINITION: Assuming a sequence $\{x_n\} \in \mathbb{R}^N$, the first backward difference between the elements is given by

$$\nabla x_n = x_n - x_{n-1} \quad (22)$$

the subsequent backward differences are obtained by applying the above relation recursively, hence for l^{th} step, the difference is

$$\nabla^l x_n = \nabla(\nabla^{l-1} x_n) \quad (23)$$

the second and the third differences are provided below

$$\nabla^2 x_n = \nabla(\nabla x_n) = \nabla x_n - \nabla x_{n-1} \quad (24)$$

$$= x_n - 2x_{n-1} + x_{n-2} \quad (25)$$

$$\nabla^3 x_n = \nabla(\nabla^2 x_n) = \nabla^2 x_n - \nabla^2 x_{n-1} \quad (26)$$

$$= \nabla(\nabla x_n) - \nabla(\nabla x_{n-1}) \quad (27)$$

$$= \nabla x_n - \nabla x_{n-1} - \nabla x_{n-1} + \nabla x_{n-2} \quad (28)$$

$$= x_n - 3x_{n-1} + 3x_{n-2} - x_{n-3} \quad (29)$$

for higher order differences, the above approach is repeated.

As mentioned earlier the SSM has a linear convergence, and was defined in chapter 3 (refer section 3.2.2) with the converging point x^* as

$$\|x_{n+1} - x^*\| \leq K \|x_n - x^*\| \text{ where } K \in (0, 1), \text{ for all } n \text{ sufficiently large} \quad (30)$$

in order to simplify the equation form later and derive the Aitken relation, the expression (30) is re-written in the following manner with x_t denoting the converging point, and n having a very large value.

$$\|x_n - x_t\| \leq K \|x_{n-1} - x_t\| \quad (31)$$

$$\frac{x_n - x_t}{x_{n-1} - x_t} \approx K \quad (32)$$

$$x_n - x_t = K (x_{n-1} - x_t) \quad (33)$$

$$x_n = K x_{n-1} + (1 - K) x_t \quad (34)$$

introducing the term x_{n-1} on both sides

$$x_n - x_{n-1} = K x_{n-1} + (1 - K)x_t - x_{n-1} \quad (35)$$

$$x_n - x_{n-1} = (x_t - x_{n-1})(1 - K) \quad (36)$$

$$x_t = x_{n-1} + \frac{(x_n - x_{n-1})}{(1-K)}, \text{ after rearranging the terms in (36)} \quad (37)$$

equation (22) is used to substitute the difference term that exists in the numerator, and the above relation delivers

$$x_t = x_{n-1} + \frac{\nabla x_n}{(1-K)} \quad (38)$$

on similar lines as that of the relation presented in (31), the same sequence could be written in the following manner for all n after a certain point, when the latter is very large

$$\frac{x_n - x_{n-1}}{x_{n-1} - x_{n-2}} \approx K \quad (39)$$

or

$$K = \frac{x_n - x_{n-1}}{x_{n-1} - x_{n-2}} \quad (40)$$

using the backward difference relation defined earlier, the following expression is obtained

$$K = \frac{\nabla x_n}{\nabla x_{n-1}} \quad (41)$$

the Aitken relation is given by substituting the above equation in (38) and using the relation (24), i.e.,

$$x_t = x_{n-1} + \frac{\nabla x_n \nabla x_{n-1}}{(\nabla x_{n-1} - \nabla x_n)} \quad (42)$$

$$x_{\text{Aitken}} = x_{n-1} - \frac{\nabla x_n \nabla x_{n-1}}{\nabla^2 x_n} \quad (43)$$

Appendix E

Covariance, correlation and regression are statistical methods that form the basis for many surrogate and optimization techniques discussed earlier in chapter 4; this section provides mathematical notes that describe these approaches. In regression, the existing data is used to define a mathematical relation for prediction of a dependent variable vector or response variable Y based on the value of independent variable vector also known as *predictor variable* X . It is used to interpolate between the existing data. The association between two variables (X, Y) is measured using correlation, and the procedure quantifies the strength of such a relationship, it cannot be used for prediction. The evaluations are performed on only the existing data, and do not have the aspect of interpolation as prevalent in regression. Covariance measures the correlation between sets of random variables, and correlation could be defined as the ratio of covariance to the standard deviations of two variables (X, Y) .

For simplicity, assume the dimension of the independent variable X to be one, and the model to determine the response as linear, given by

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (1)$$

where ϵ is the error due to any individual value y that does not follow the regression linear model, β_0 and β_1 are constants known as *model regression coefficients*. Another assumption considered in the above scenario is that the range of observations stays within the acceptable limits for the model to be accurate. The representation in terms of the elements is

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i ; \quad i = 1, 2, \dots, n \quad (2)$$

n in the above equation denotes the number of observations. The linearity in the above model is because of the linear relation that exists when the parameters are considered, for instance,

$$y = \beta_0 x_1 + \beta_1 x_2^2 + \beta_2 x_1 x_2 + \epsilon \quad (3)$$

equation (3) is linear with respect to x_1 , whereas the following definition is not

$$y = \beta_0 + \beta_1 x_1^a x_2^b + \epsilon \quad (4)$$

in (3) and (4), ϵ represents the error associated with one response. Assuming there are n observations made, the corresponding response and predictor vectors will be $Y = \{y_1, y_2, \dots, y_n\}$ and $X = \{x_1, x_2, \dots, x_n\}$ respectively, considering the model defined in (2), the unknown regression coefficients β_0 and β_1 are determined in the following manner,

$$E = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \quad (5)$$

differentiation of the above equation with respect to β_0 and β_1 , and equating them to zero renders

$$\frac{\partial E}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0 \quad (6)$$

$$\frac{\partial E}{\partial \beta_1} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i = 0 \quad (7)$$

The least square estimates of β_0 and β_1 are given by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (8)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (9)$$

where $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ and $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$, ultimately the fitted value of Y or the regression model is defined as

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X \quad (10)$$

The above approach can be extended to matrix form as shown below, with n observations and the X, Y pair.

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (11)$$

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad (12)$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \quad (13)$$

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} \quad (14)$$

$$\epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad (15)$$

The matrix representation for the regression model is

$$Y = X\beta + \epsilon \quad (16)$$

In the present scenario the error relation is given by

$$\epsilon^T \epsilon = (Y - X\beta)^T (Y - X\beta) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \quad (17)$$

the least square minimization as performed in (6) and (7) for $\epsilon^T \epsilon$ results in the following matrix form of $\hat{\beta}$

$$X^T X \hat{\beta} = X^T Y \quad (18)$$

or,

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (19)$$

in relation(19), it is assumed that $X^T X$ exist.

The covariance between the predictor and response pair (X, Y) is defined as

$$\text{Cov}(X, Y) = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{n-1} \quad (20)$$

where $\bar{x}$ and $\bar{y}$ are defined earlier in the discussion. When $\text{Cov}(X, Y) > 0$ the relationship between X and Y is positive, whereas when $\text{Cov}(X, Y) < 0$ the pair has a negative association, the plots presented below would throw more light on this behavior.

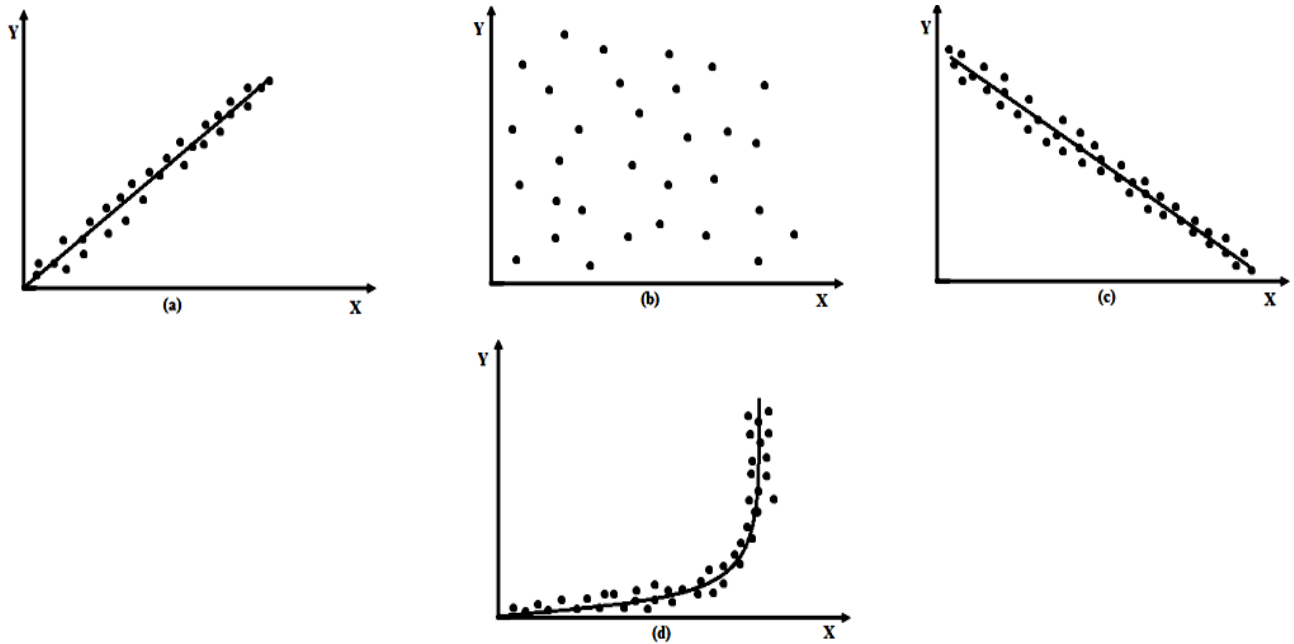


Fig 1 a) A positive covariance relationship, (b) no covariance, (c) a negative covariance relation, and (d) a non-linear covariance.

As highlighted earlier, there exists a relation between covariance and correlation, and this aspect has been used to define the latter as shown below.

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y} \quad (21)$$

where $\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$ and $\sigma_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}}$ are the standard deviations. On simplifying equation (21) renders

$$\text{Cor}(X, Y) = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (22)$$