Wire Antennas Optimizations on Various Platforms using Radial Basis Functions and Evolutionary Algorithms

by

Shifu Zhao

B.Engineering in Microelectronic Engineering, Second Class Honours Division A
Griffith University, 2008

Thesis submitted for the degree of

Doctor of Philosophy

in

Electrical and Electronic Engineering
University of Adelaide

2012

© 2012

Shifu Zhao

All Rights Reserved

Typeset in LaTeX 2ε

Shifu Zhao

Contents

Abstract

High Frequency (HF) and Very High Frequency (VHF) electromagnetic waves have been used as the means of long-distance communication for decades. Nevertheless, in the design of wire antennas for HF and VHF devices, size reduction is one of the critical issues due to wavelengths of in ranges from 1 to 100 meters. It is well known that inductive and capacitive loadings can effectively change the current distribution along an antenna, reducing the self-resonant frequency, and hence the antenna size. Various types of inductive and capacitive loadings can be implemented on the wire antennas using ideal lumped components or realistic winding structures, such as zig-zag and helix shapes. Nevertheless, the physical limits of electrically small antenna can greatly constrain the dimensions, and the design of optimally varying windings will significantly increase the complexity in the modeling and simulation process. Furthermore, size reduction can also introduce significant degradation in both efficiency and bandwidth, and thus, obtaining a design with balanced performance becomes a challenging task, which is addressed in this thesis.

The work presented in this thesis contributes to the research by proposing and applying a generic methodology to the optimal design of size-reduced HF and VHF wire antennas. The electromagnetic simulator, NEC-2 (Numerical Electromagnetic Codes), based on the method of moments, is used to provide fast and accurate numerical estimation of the performance for the antennas. To drive the electromagnetic simulator, an evolutionary optimizer is developed using both genetic algorithm (GA) and particle swarm algorithm (PSA) for multi-objective optimization (MOO). The combination of these tools, i.e. electromagnetic simulator and optimizers, is applied to address the trade-offs of the small antenna design as well as to achieve faster convergence efficiently to the global optimal region. The in-house developed tool is named MATNEC, and couples antenna geometry modeling, electromagnetic simulation, and evolutionary optimization into an automated program. Several strategies have been used to

reduce the simulation and optimization complexity with, in particular the application of radial basis function expansions to compactly describe the antenna structure. This effectively converts the optimization process from optimizing the antenna configuration directly to optimizing the parameters of mathematical expansion, thus achieving a significant complexity reduction.

In the application of the proposed technique in this thesis, three types of inductive loadings are successively introduced into the design of optimized wire antennas, producing a marked increase in performance in all cases. Firstly, as preliminary study, lumped inductive loadings along a monopole are used to effectively verify the optimization methodology and the antenna shortening theory. Secondly, a non-uniform zig-zag winding structure is considered to effectively verify the roles of optimized distributed inductive loadings formed by the antenna wire itself and also allowing for experimental validation of the findings. Thirdly, non-uniform helical antenna structures are also considered and verified experimentally. The optimal designs were verified both in bandwidth and in efficiency using a "Wheeler Cap" approach. The optimized results provide useful guidelines for the design of wire antennas for both HF and VHF communications.

The thesis also provides an investigation of the robustness of the optimized design in non-ideal environments. Optimized devices are integrated on various platforms or with near-by objects, and the re-optimization is carried out including the non-ideal environment. The weak impact from non-ideal environments and the similar results from re-optimization effectively demonstrate the strong functionality and robustness of the proposed design and optimization strategy for real-world applications. Mutual interaction between multiple antennas is also investigated, and the result illustrates the weak interference of the optimized antennas when used in an array environment.

Statement of Originality

This work contains no material that has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published written by another person, except where due reference has been made in the text.

I give consent to this copy of the thesis, when deposited in the University Library, being available for loan, photocopying and dissemination through the library digital thesis collection.

Signed

Date

Acknowledgments

First and foremost, I want to thank my supervisors A.Prof Chris Coleman and A.Prof Christophe Fumeaux sincerely, for their guidance and encouragement throughout my candidature. It is a great honor to be their Ph.D student, and a privilege to undertake such a challenging but also rewarding research. They have taught me, both consciously and unconsciously, how research in electrical and electronic engineering is done. I appreciate all their contributions of time, ideas and funding to make my Ph.D experience productive and simulating. Their passion and enthusiasm for research was contagious motivational for me, especially during the hard time of the Ph.D pursuit.

I want to acknowledge my colleagues and academic members in the antenna research group for their heuristic and inspiring discussion. I would also like to acknowledge the staff of the Schools of Electrical and Electronic Engineering at the University of Adelaide for their technical support and administrative assistance.

Finally, I would like to express my deeply-felt thanks to my family for their unconditional love and support. Their understanding and encouragement have always inspired me throughout the whole journey. This research could never have been completed without them.

Conventions

Typesetting

This thesis it typeset using the LaTeXsoftware. WinEdt build 5.5 was used as an effective interface to LaTeX.

Referencing

The referencing and citation style adopted in this thesis are based on the Institute of Electrical and Electronics Engineering (IEEE) Transaction style.

Units

The units used in this thesis are based on the international system of units (SI units).

Prefixes

In this thesis, the commonly used numerical prefixes to the SI units are ”p” (pico, 10^{-12}), ”n” (nano, 10^{-9}), ”μ” (micro, 10^{-6}), ”m”,(milli,10^{-3}), ”k”(kilo, 10^{3}), ”M”(mega, 10^{6}), and ”G”(giga, 10^{9}).

Spelling

Australian English spelling is used in this thesis.

Author Publications

Journal

[1] S. Zhao, C. Fumeaux, and C. J. Coleman, "Miniaturised HF and VHF antennas based on optimised non-uniform helical structures," *IET Microwaves, Antennas and Propagation*, vol. 6, pp. 603–610, 2012.

Conference

[1] S. Zhao, C. Fumeaux, and C. Coleman, "Optimal positions of loading for a shortened resonant monopole using genetic algorithm," in *International Conference on Electromagnetics in Advanced Applications, ICEAA 2010*, Sydney, Australia, September 2010, pp. 705–708.

[2] S. Zhao, C. Fumeaux, and C. Coleman, "Optimal loading configurations of a shortened resonant wire antenna using evolutionary optimizers," in *Twelfth Australian Symposium on Antennas, ASA 2011*, Sydney, Australia, February 2011.

[3] S. Zhao, C. Fumeaux, and C. Coleman, "Optimal helical antenna with continuously varying radius using evolutionary optimizers," in *IEEE International Symposium on Antennas and Propagation, AP-S 2011*, Washington, USA, July 2011, pp. 757–760.

[4] S. Zhao, C. Fumeaux, and C. Coleman, "Evolutionary optimization of zig-zag antennas using gaussian and multiquadric radial basis functions," in *Asia-Pacific Microwave Conference, APMC 2011*, Melbourne, Australia, December 2011, pp. 1594–1597.

[5] S. Zhao, C. Fumeaux, and C. Coleman, "Optimized helical monopole antennas for portable VHF communication devices," in *IEEE International Symposium on Antennas and Propagation, AP-S 2012*, Illinois, USA, July 2012.

Abbreviations

AMP Antenna Modeling Program

EFIE Electric Field Integral Equation

ESA Electrically Small Antenna

FBW Fractional Matched VSWR Bandwidth

FDTD Finite Difference Time Domain

GA Genetic Algorithm

G-RBF Gaussian Radial Basis Function

HF High Frequency

MFIE Magnetic Field Integral Equation

MoM Method of Moments

MQ-RBF Multiquadric Radial Basis Function

NEC Numerical Electromagnetic Codes

PEC Perfect Electric Conductor

PIFA Planar Inverted F Antenna

PSA Particle Swarm Algorithm

PSO Particle Swarm Optimization

RBF Radial Basis Function

SAL Small Antenna Limit

VSWR Voltage Standing Wave Ratio

VHF Very High Frequency

List of Figures

List of Tables

Chapter 1

Introduction

THIS chapter presents brief information on HF and VHF communications, electrically small antennas, and multi-objective optimizations. Size reduction strategies for electrically small antennas are introduced, being the primary concern in the investigation. Antenna design methodology and process in this thesis includes geometry modeling, electromagnetic simulation and multi-objective optimization. This chapter also provides an overview of the organization of this thesis and its contributions.

1.1 Introduction and Motivation

With the advance of science and technology, wireless communications have been developed widely and extensively to fulfill the requirements for multiple purposes. Antennas constitute an important part of wireless systems and their study is the focus of this work. Among the many properties of antennas, size, efficiency and bandwidth are particularly important and critical. At low frequencies (long wavelengths), size is a priority consideration in the design process since traditional designs are of the order of a quarter to a half of a wavelength. In the past, HF and VHF were the means of long-distance communication. However, even though they have been partially replaced by satellite or terrestrial communications, they still show a vast amount of potential and interest due to the advantages of their relatively simple systems and low cost. Nevertheless, the size of antennas, which is inversely proportional to the frequency, has always been one of the major drawbacks for HF and VHF applications. Furthermore, modern technology and applications are consistently demanding a growing data rate be transmitted, which means an increased bandwidth requirement for antennas. This, however, brings out another disadvantages of HF and VHF communications, that for many practical applications, the antenna needs to be much smaller than a wavelength. Antennas with a maximum physical dimension relatively small with respect to their operating wavelength can be categorized as electrically small antennas (ESA). Although, ESAs have been widely used for HF and VHF bands, they exhibit considerable limitations in terms of both bandwidth and efficiency. To ease these limitations, size reduction strategies have been investigated extensively in order to achieve satisfactory antenna performance in terms of efficiency and bandwidth. This thesis will put its focus into methods to reduce the dimension of HF and VHF antennas whilst exhibiting good efficiency and bandwidth.

1.1.1 HF and VHF Communications

HF and VHF, nominally the 3 - 30 MHz and 30 - 300 MHz regions in the frequency spectrum, have been investigated and used for many years for long-distance radio

communication in both military and civilian applications. Even though they were partially replaced in the 1960s to 1970s by satellite and cable communications [1], they still show a vast amount of potential due to advantages in terms of both simplicity and cost. They have been typically used in international broadcasting, national broadcasting in tropical regions, aeronautical and maritime mobile communication and over the horizon radar [2,3].

HF communication is mostly by means of ”sky-wave propagation”, a mode that exploits propagation of signal waves up to worldwide distances by using the reflection of waves between the ionosphere and earth surface for over-the-horizon communication [1,2]. The ionosphere is divided into mainly three regions or layers designated D, E, and F respectively from the earth upwards. The reflective coefficients are different for each layer and vary with respect to time of the day, season of the year, and sunspot activity. The E and F regions behave mostly as radio wave reflectors, and allow long range communication between terrestrial terminals. The lowest (the D region), however, acts mainly as an absorber of energy which induces signal attenuation in the HF range [1]. VHF communications can achieve over then horizon propagation through mechanisms such as tropospheric scatter [2].

Traditional HF and VHF antennas have dimensions that are proportional to their wavelengths (a quarter to a half for example). HF and VHF bands with wavelengths ranging from 1 - 300 m, will consequently need a fairly large size of antenna. Large antenna size inevitably increases the cost and causes inconvenience in operation, which is the primary disadvantage for HF and VHF communications. Small antennas can be used, but these introduce the above mentioned problems of bandwidth and efficiency.

Due to the nature of long-distance communications, large transmission powers are required to overcome transmission losses. This requires the antennas to perform at a very high efficiency in order to lower the power consumptions for HF and VHF systems. In addition, HF and VHF communications are significantly affected by a variety of additional losses due to the complicated nature of the propagation environment.

Consequently, efficiency is a major consideration when designing HF and VHF antennas.

1.1.2 Electrically Small Antennas

Electrically small antennas are defined as having a maximum physical dimension which is very small with respect to the operating wavelength. Generally their radiation resistance is very low, the input reactance is quite high, the bandwidth is fairly narrow, and the efficiency is low. These properties can be concluded from the various studies of fundamental limits of small antennas [4–7].

The size of a small antenna is defined in terms of the radian sphere, the smallest sphere that contains the antenna. A small antenna is normally defined as one for which the radius is smaller than radianlength, defined as $\lambda/2\pi$. For a dipole this will mean an overall length less than $\lambda/2\pi$, and for a monopole an overall length less than $\lambda/4\pi$, as shown in Figure 1.1.

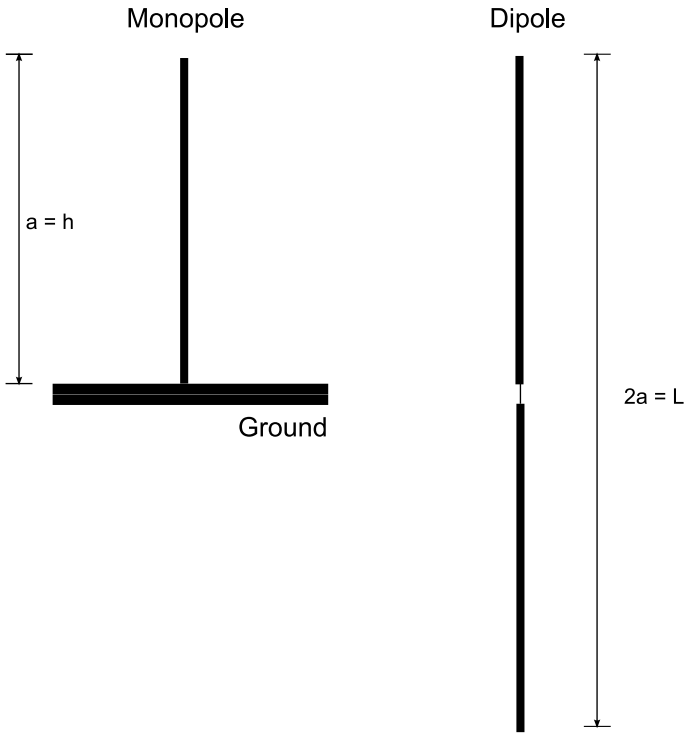

Figure 1.1. Illustrations of a monopole and a dipole.

The Q of a small antenna is defined by the ratio of reactive energy to accepted power [8–10] at the frequency ω_0:

$$Q(\omega_0) = \frac{\omega_0 W}{P} \tag{1.1}$$

where W is internal energy and P is total power accepted by the antenna.

The lower bound on Q can be derived in terms of ka [7] through:

$$Q_{lb} = \eta_r \times (\frac{1}{(ka)^3} + \frac{1}{ka}) \tag{1.2}$$

where η_r is the antenna radiation efficiency, $k = \frac{2\pi}{\lambda}$, and a is the radius of a sphere that circumscribes the maximum physical dimension of the antenna.

It is well known that bandwidth and Q are inversely related and for practical purposes, the bandwidth at frequency ω_0 is defined:

$$FBW_V(\omega_0) = \frac{\omega_+ - \omega_-}{\omega_0} \tag{1.3}$$

where ω_+ and ω_- are the frequencies above and below ω_0 at which the VSWR is equal to an arbitrary value denoted by s. The fractional matched VSWR bandwidth and the Q are related [10] through:

$$Q(\omega_0) = \frac{2\sqrt{\beta}}{FBW_V(\omega_0)} \tag{1.4}$$

$$\beta = \frac{s-1}{s\sqrt{s}} \tag{1.5}$$

Combining the equations above, an upper bound on the fractional matched VSWR bandwidth can be written as [10]:

$$FBW_{Vub} = \frac{1}{\eta_r} \frac{(ka)^3}{1+(ka)^2} \frac{s-1}{\sqrt{s}} \tag{1.6}$$

It can be seen that greater radiation efficiency leads to greater value of Q, and consequently a narrower bandwidth. Likewise, smaller radiation efficiency corresponds to wider bandwidth at operation. Obviously, there is a trade-off between the operation bandwidth and the radiation efficiency. Importantly, bandwidth also drops dramatically with the increase of antenna size.

Electrically small antennas can be represented as a simple lumped capacitor or inductor [4,8]. A small electrical monopole can be considered as a cylindrical volume capacitor of an effective height (h), and an effective area (A), where the effective area was defined as the area of an idealized parallel-plate condenser with plates separated by the effective height. The radiation resistance of the electrically small antenna is determined by the antenna effective height and the reactance is a function of both effective height and area [8], as shown in Figure 1.2. The capacitance and inductance introduced by antenna structures can be calculated as in Equations 1.7 and 1.8:

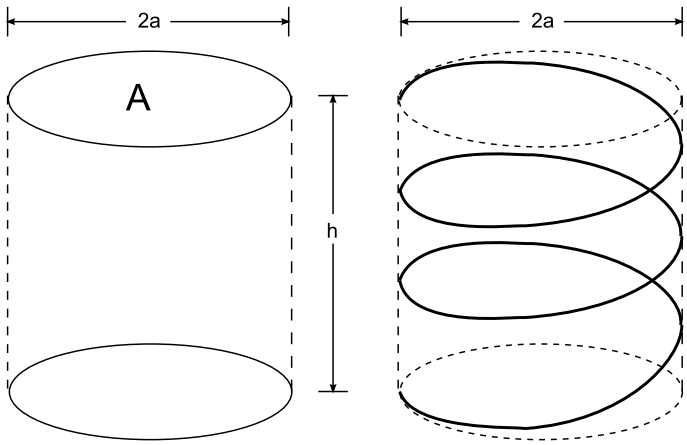

Figure 1.2. Illustrations of capacitor (C) and inductor (L) occupying equal cylindrical volumes.

$$C = \varepsilon \frac{k_a A}{h} = \varepsilon \frac{k_a V}{h^2} \tag{1.7}$$

$$L = \mu n^2 \frac{A}{k_b h} = \mu n^2 \frac{V}{k_b h^2} \tag{1.8}$$

where ε is the electric permittivity in free space, μ is the magnetic permeability in free space, A is the area of base of cylindrical volume, h is the height of cylindrical volume, n is the number of turns of coil, V is the cylindrical volume, k_a is the shape factor of capacitor, and k_b is the shape factor of inductor. In reality, an antenna will be a complex combination of capacitance and inductance.

Therefore, electrically small straight wire antennas, either dipole or monopole, exhibit impedance at the frequency (ω) that can be written as:

$$Z(\omega) = R(\omega) - jX(\omega) \tag{1.9}$$

where $X(\omega)$ is dominated by the equivalent value of $1/\omega C$ established by the antenna structure.

Consequently, to effectively mitigate the dominant $X(\omega)$ term at the antenna feeding point, the structure can be modified to include a capacitive top-hat or inductive loading by additional wire length, or a combination of both [11]. An illustration is shown in Figure 1.3 to indicate the two means [11].

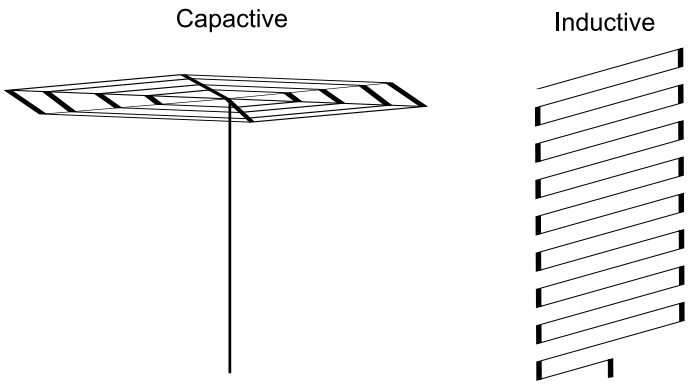

Figure 1.3. Illustrations of top-hat capacitive and inductive loadings on wire antennas.

1.1.3 Multi-Objective Optimization

For most real-world problems, objectives that are competing or even incompatible often co-exist simultaneously. This is certainly the case in antenna design. Optimization in these problems may lead to trade-offs between the objectives, and require the techniques of multi-objective optimization [12]. In contrast to single-objective optimization where the optimal solution is usually clearly defined, multi-objective optimization may have so-called Pareto-optimal solution [12]. These solutions are regarded as optimal in a broader sense that no other solutions in the search space are superior to them when all objectives are considered under certain pre-defined criteria [13]. Alternatively, a

pareto-optimal outcome is one such that no-one could be made better off without making someone else worse off. An optimization problem with M optimization objectives f_1, f_M, ..., f_M can be expressed as [14]:

$$Minimize/Maximize \quad f_m(x),\ m = 1, 2, ..., M, \tag{1.10}$$

where x is the vector of decision variables.

In order to solve multi-objective optimization problem, one classical method is to transform it into a single and parameterized function by analogy to decision making before search [13]. This single objective implicitly includes preference information given by the decision maker. To select a suitable compromise solution, the human decision maker is required to make difficult trade-offs between conflicting objective. In this sense, the multi-objective optimization is converted to an single-objective optimization by generating a linear combination of objectives:

$$Minimize/Maximize \quad f(x) = \omega_1 \times f_1(x) + \omega_2 \times f_1(x) + ... + \omega_M \times f_M(x) \tag{1.11}$$

where ω_i are called weighting factors.

With respect to either single-objective or multi-objective optimization, evolutionary algorithms have been found to be highly effective and have been extensively used in many different fields. Two types of evolutionary algorithms will be introduced and employed in this thesis to pursue the optimal antenna solutions. The genetic algorithm was initially invented based on the natural process of evolution and genetic recombination, and this makes it a robust stochastic search method [15, 16]. Particle swarm optimization is based on the exploration of the problem space according to given rules inspired by the behavior of living organism swarms [17]. GA and PSO are both global search methods, and are effective in finding an approximate global optimum in a high-dimension space.

1.2 Thesis Significance and Contributions

The significance of this thesis comes from three distinct aspects:

I. An antenna design methodology that couples electromagnetic simulation to evolutionary optimization. Conventionally, antennas were designed intuitively and verified

through measurement. With the advent of today's antennas, however, multiple performing features and characteristics are frequently required. In this case, a straightforward approach by trial and error can be inefficient, time-consuming and inaccurate, and thus not feasible for multi-objective antenna design. Evolutionary optimization provides an efficient alternative. In this thesis, the electromagnetic modeling of antenna performance is achieved through numerical techniques. For wire antennas, a method of moments (MoM) [18] is a particularly useful and efficient numerical approach. In the current work, the NEC-2 implementation of the MoM method is employed. Since the optimization process requires the reevaluation of the simulation many times over, the efficiency of the algorithm is important.

An automated tool that controls antenna modeling, electromagnetic simulation and evolutionary optimization, called MATNEC, is developed. This tool uses MATLAB as the primary control, and manipulates the tasks in the flow through the whole process. It first builds the antenna models, drives the process of MoM simulation; then analyzes the simulation results based on appropriate fitness evaluation, applies the evolutionary algorithm correspondingly, and iterates until certain criteria are met. The development of MATNEC plays an essential role, as it is the primary tool for investigating various types of wire antennas in the present investigation. The efficiency of the process, and the accurate results from MATNEC, verify the effectiveness of the methodology proposed in this thesis.

In order to ease the complexity of both electromagnetic simulation and evolutionary optimization, several strategies have been introduced and successfully used in the investigation. Basis function representation of design variations is employed to provide sufficient geometrical variations to the wire antennas without producing an oversized optimization space. In particular, radial basis functions (RBFs) are used (Gaussian RBFs and Multiquadric RBFs). The employment of RBFs makes the optimization more global and robust.

II. The above developments are used to optimize the design of two types of frequently-used wire antennas. Zig-zag antennas with continuous arm-length variations are successfully optimized and shown to exhibit improved performance compared with the traditional uniform variety. A broader operation bandwidth and higher efficiency is obtained with the optimized antennas. Similar results are also obtained for helical type antennas. Both types are important for their common use in portable radios.

III. The impact of non-ideal environments on the optimized design of the above antennas has been investigated. The results of these studies demonstrate the robustness of the design methodology by optimizing the antennas on various platforms, such as portable communication devices. In addition, we have studied the antenna performance in the array environment. The optimized antennas exhibit reduced mutual interactions compared with traditional monopole and hence are particularly useful in this application.

1.3 Thesis Overview

This thesis seeks to explore the optimization of wire antennas for HF and VHF communications. A tree diagram of the thesis is given in Figure 1.4.

In Chapter 1, a brief literature overview is provided in order motivate and illustrate the general approach of the investigation.

Chapter 2 provides the fundamental theory of multi-objective optimization. In this work, evolutionary algorithms are employed, namely the genetic and particle swarm algorithms. These two algorithms are fully described in this chapter, including the theory and applications from the literature. Fitness functions, which are considered as the critical element in the optimization, are also discussed in detail.

Chapter 3 first introduces electromagnetic simulation in a general sense and describes the method of moments in particular. In this work, the NEC-2 electromagnetic simulation is used. NEC-2 is a simulator that is based on MoM and is fully described in this chapter. This introduction to NEC-2 provides information concerning geometry

modeling, environmental modeling, and the modeling constraints. Due to the complexity in both the antenna modeling and optimization, computational load can be very large. Consequently, strategies for reducing this complexity through the use of basis function expansions are described. In particular, the use of radial basis function expansions. MATNEC, an in-house tool which couples NEC simulation and evolutionary optimization, is introduced in this chapter as it plays a big role in the investigations of this thesis.

In Chapter 4, we investigate the impact of distributed inductance along a straight wire antenna, in particular, the impact of lumped inductors. The application of matching networks is also discussed in this chapter.

In Chapter 5, we extend to work on lumped inductance to consideration of continuously distributed inductance in terms of zig-zag windings. Zig-zag antennas are optimized with respect to its linear sections. By continuously varying the length of each section, improved performance with wider operation bandwidth and higher efficiency can be achieved compared with uniform zig-zag. Experimental validation of bandwidth measurement is also described in this chapter.

Chapter 6 provides the extension of distributed inductance by considering antenna composed of helical windings. Variations of both radius and pitch are introduced to the helix to pursue performance improvement. It is shown that either radius or pitch variation can be used to vary the distribution of inductance on the helical antennas, and consequently optimize their performance. The results shown considerable gain in performance for the optimized antennas.

Chapter 7 considers optimization of antennas under non-ideal circumstances. Various types of real ground conditions are investigated, and the impact of near-by objects are also considered. In particular, the application of an antenna that is mounted in a hand-held device is considered. Finally, the effect of optimized antennas in an array environment is considered.

Chapter 8 draws together the conclusion from the work described in this thesis and provides recommendations for future work.

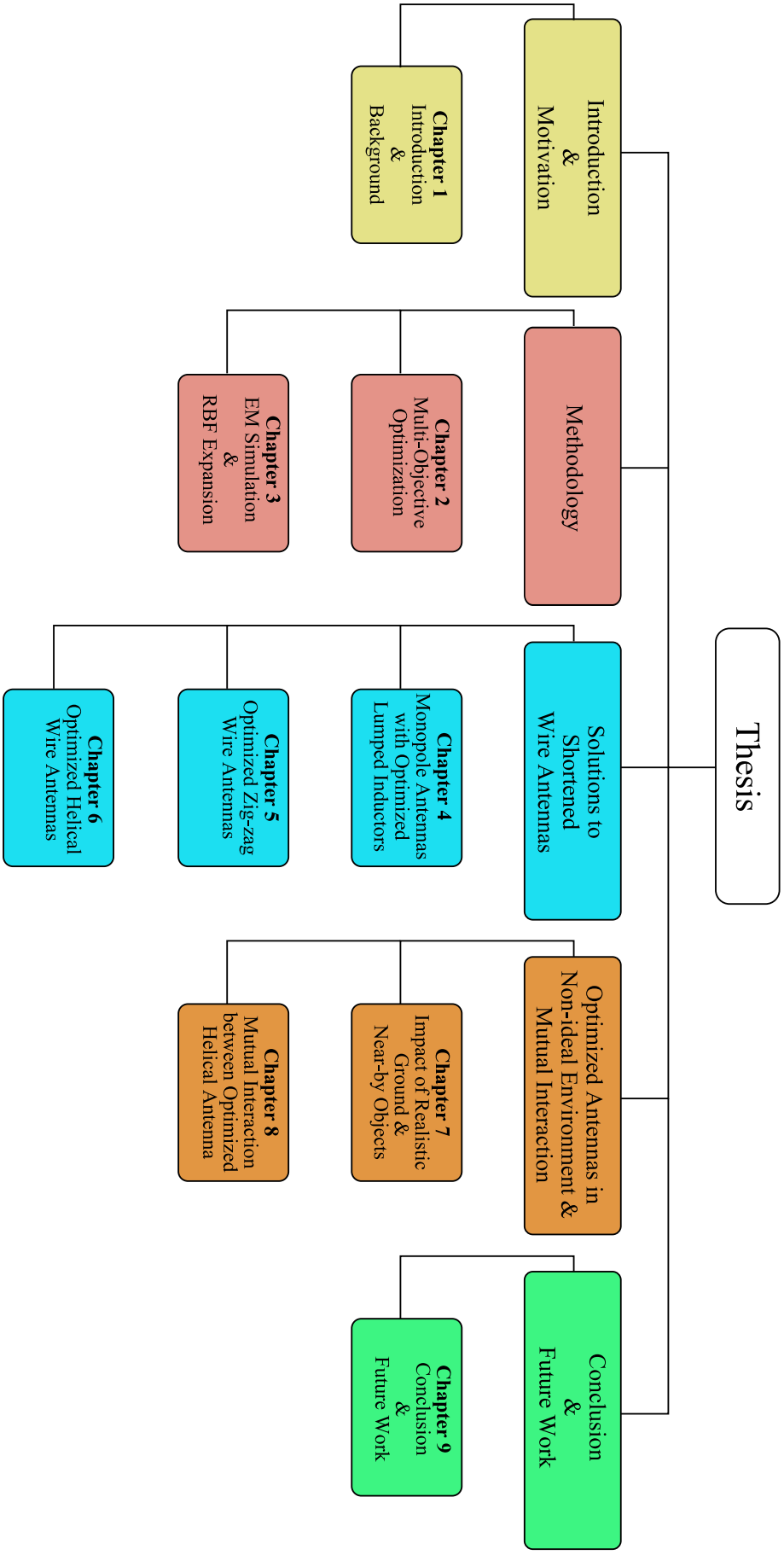

Figure 1.4. The tree diagram of the thesis.

Chapter 2

Evolutionary Algorithms for Multi-Objectives Optimization

THIS chapter reviews the concept of multi-objective optimization, and provides the fundamental theory of two evolutionary algorithms, namely genetic algorithm and particle swarm optimization algorithm. Typical implementations and applications of those two algorithms are introduced. A critical issue in optimization, the influence of the fitness function, is also addressed in this chapter.

2.1 Introduction

For most real-world problems, objectives that are competing or even conflicting often can co-exist simultaneously. An optimal solution to such problems requires trade-offs between these objectives. The term multi-objective optimization describes the process of finding a satisfactory solution to problems with multiple objectives that need to be fulfilled simultaneously [12]. In contrast to single-objective optimization where the optimal solution is usually clearly defined, multi-objective optimization may have a so-called Pareto-optimal solution [12]. These solutions are regarded as optimal in the broad sense that no other solutions in the search space are superior to them when all objectives are considered under certain pre-defined criteria [13]. Alternatively, a Pareto-optimal outcome is one such that an improvement for a particular aspect can only be achieved at the cost of degradation of another condition.

Evolutionary algorithms are well developed and frequently used to solve multi-objective optimization problems. The genetic algorithm is the most popular population-based evolutionary optimization technique. It is based on the survival-of-the-fittest natural principle, where the population pool consists a number of individuals, each representing a possible solution. This technique will be explicitly articulated in the next section of this chapter.

Another evolutionary algorithm frequently used in multi-objective optimization is particle swarm optimization. It is an optimization method that was inspired by a simplified social model. It has roots in two main methodologies: the artificial life and evolutionary computation. The essence of this algorithm is social sharing of information, which provides an evolutionary advantage to a species in the search for food sources. The fundamentals of particle swarm theory will be fully described in this chapter as well.

The beauty of evolutionary algorithms in solving multi-objectives optimization is their ability to locate the global optimum region and avoid the local optimums. Conventional search using gradient methods sweeps the solution linearly from point to point. They stop the search once the gradient changes from positive to negative and regard the corresponding location as the maximum; likewise, a minimum is detected when

the sign of the gradient changes from negative to positive. This, however, makes the starting search point critical in applying this type of methods. Therefore, gradient methods can be easily trapped in local optimum regions, and consequently are prone to failing in finding the global optimum. In contrast, Both GA and PSO probe the whole solution space while iterating through the searching process and therefore can effectively overcome trapping in a local optimum, and thus are more likely to find the global optimum.

2.2 Genetic Algorithm

The well-recognized theory of evolution proposed by Charles Robert Darwin describes biological systems as the product of the ongoing process of natural selection. The GA is a robust, stochastic search method, which is patterned after the natural processes of genetic recombination and evolution [15,16,19,20]. This technique involves mating and mutation, which has been translated to represent the structure of computer programs as introduced in [21]; the term GA was consequently first proposed in [16].

2.2.1 GA Theory

"A genetic algorithm is a form of evolution that occurs on a computer [22]". The name genetic algorithm itself literally tells the story: the first word originates in biological science while the second word is borrowed from computer science. In the evolutionary world, there are two primary processes that allow organisms to evolve, namely natural selection and sexual reproduction [21]. Explicitly, natural selection decides which individuals of a population will survive for reproduction, while the sexual reproduction process provides mixing and recombination among the genes to create offspring. In GA, selection and reproduction are carried out through operations on strings of binary digits. These strings are stored in a computer memory and evolve over time, while their functionality is calculated and processed iteratively according to predefined criteria. This type of iterative algorithms mimics the process through which natural populations of individuals evolve [22].

Before formally introducing the algorithm, a list of some commonly-used genetic algorithm terms are provided in Table 2.1, and related to their natural (biological) counterparts. Fundamentally, the algorithms encode each parameter into binary genes, and a set of genes in a chromosome. Chromosomes from a population and natural selection undergo mating and mutation which produces a new generation and introduces new chromosome into the population, respectively. After been evaluated by fitness function, only the population members with the strongest genes survive and are able to reproduce and create the next generation. After an iteration for many generations, best chromosomes will converge towards a near-optimal solution.

Table 2.1. Terminology in Genetic Algorithm.

Population	Set of trials solution
Parent	Member of current generation
Child	Member of next generation
Gene	Binary encoding of a parameter
Generation	Successively created populations
Chromosome	Coded form of a trial solution vector (string) of genes
Fitness	Value assigned to an individual representing the goodness

2.2.2 GA Procedures

The GA is an exploratory procedure. Therefore, it is often able to locate near optimal solutions to complex problems with nonlinear solution space, which might include local minima and maxima. It involves several operations which are performed iteratively in a sequential and logic order. A flow chart of the GA is provided in Figure 2.1.

I. Representation. This is the very first step in GA where a computer compatible representation or encoding of the parameter space is created in the form of genes and chromosomes [16]. Chromosomes are commonly binary strings in which different parts represent different encoded parameters of a solution, where each parameter in the string is called a gene. Based on this representation, an initial population is created,

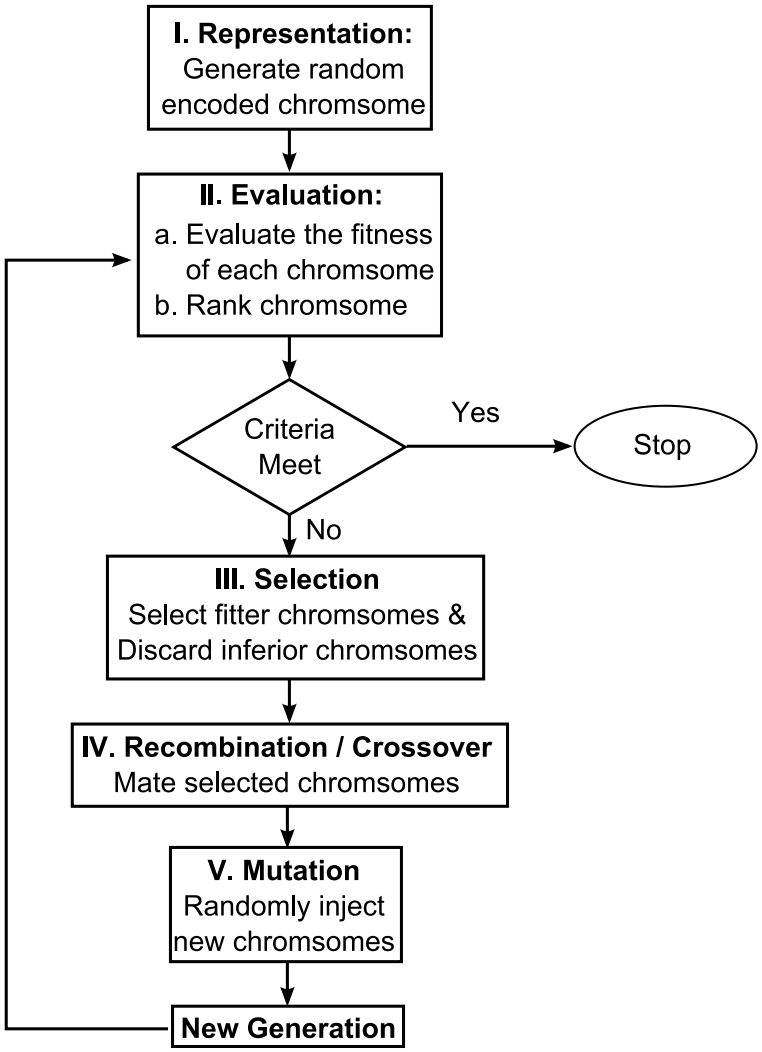

Figure 2.1. Flow chart of genetic algorithm.

composed of randomly selected parameters in the solution region. The numerical resolution of the parameter, and the associated accuracy are determined by the number of bits in the string assigned to a given parameter. The longer the number of bits allocated, the more accurately the parameter can be resolved in the solution, however inevitably at the cost of computation complexity.

II. Evaluation. Evaluation is applied to test the fitness of each chromosome of the population in every generation. This step is accomplished by substituting each chromosome into a pre-defined fitness function (or merit function). The output value from the fitness function is called the fitness of the corresponding chromosome, and is a measure of goodness towards the solution. Chromosomes are then ranked from the

best-performance to the worst-performance based on this fitness function to provide information for the application of the survival-of-the-fittest strategy. In other words, the evaluation process assesses the chromosomes and interprets how well they can solve the problem. Fitness functions are problem specific, and vary with constraints and limits of the problems.

III. Selection. This phase produces a "surviving" generation as parents for the next stage, based upon the evaluation of the current generation, and starts by discarding unacceptable chromosomes. It is an essential step in GA as it connects the fitness function to optimization process. Primarily, selection must be based on the fitness of individual chromosome since it measures the goodness of an individual chromosome. However, the selection by the best-performing individual does not guarantee an optimal solution because the selected best individual may not be close to the global optimal solution region. Therefore, some less-fit individuals can remain to preserve diversity and to prevent the occurrence of premature convergence of the population to a local optimum. There are two most popular methods in the selection stage that can be used to determine the parents from the current chromosomes, namely ratioing and ranking respectively [16].

IV. Recombination/Crossover. In the previous selection stage, a new population has been selected from the members that solve the problem best. In the recombination stage, new chromosomes will be produced from the combination of existing chromosomes. The newly generated offspring do not overlap with the previous population to maintain many features obtained already [16]. The most common method for recombination is crossover, which is described in the following. The crossover operation is carried out in a logic order: two individual chromosomes are either randomly or deterministically selected from the population. The reproduction is governed by a specified crossover probability, where subsections of the two chromosomes are swapped at a randomly or deterministically chosen crossover point. Crossover is the essence of genetic recombination in genetic algorithm, and the choice of the control parameters should be taken with extra care in the programming work. In general, an elitist strategy is applied to keep some of the best individuals in the population, so that crossover operation does not loose the possible best solutions.

V. Mutation. Mutation is a common last step in GA procedures which aims at increasing the exploration component of the algorithm. It adds randomness in the population and widens the search, considering the possibility that the initial randomly generated population may not necessarily contain all the information to solve the problem. Mutation introduces new chromosomes into a generation, which is mostly fulfilled by randomly changing a number of bits in certain chromosomes in the population at a specified mutation probability.

2.2.3 GA Criteria and Considerations

Since the origins of the algorithm, a large experience have been gained, and lessons learnt by researchers to improve its efficiency [15, 20, 23, 24]. It is realized in GA programming that many variables can be manipulated to control the trade-offs, algorithm performance, and parameters flexibility. Extra care and considerations should be taken in selecting those control variables.

- **Number of chromosomes**

  A greater number of chromosomes in the initial random population will provide a better solution space coverage by including more potential good solutions and diversity in the population. However, this will also slow down the speed of convergence due to the heavier computation load. The number of chromosomes define the trade-off between the convergence speed and the accuracy of solutions.

- **Number of bits in the genes**

  More bits assigned to represent a parameter will definitely improve the parameter resolution and thus the accuracy of the solution, but also slow down the convergence. The number of bits for a parameter is problem-specific, and should vary with respect to the requirements. Trade-offs again lie between the convergence speed and the accuracy of solutions.

- **Selection strategy**

  There are two most popular methods that can be used to determine the parents

from the current chromosomes. In ratioing, each individual reproduces in proportion to its fitness. For instance, an individual whose fitness is ten times better than another will produce ten times the number of offspring. An alternative is ranking where the number of offspring each chromosome will generate is determined by how it ranks in the population. For example, the top 50% of the population might generate one offspring each, the bottom 50% of the population generate no offspring.

- **Mutation rate**

  Mutation process introduces new chromosomes or characteristics into the generation which can yield useful information, thus effectively avoiding premature convergence. Mutation is normally applied with a probability, or rate. Increasing the number of mutations will consequently increase the capacity of the algorithm to search outside the current region of parameter space [16], however at the cost of slowing down the convergence speed. It has been found out that mutation with low rate, usually an order of 0.01 - 0.1, is most appropriate.

- **Elitist strategy**

  Due to the probabilistic nature of GA selection, crossover and mutation, it is possible for a next generation to have a best new individual with a lower fitness than preceding generation. This would slow the convergence towards the solution, as a monotonous improvement would not be guaranteed. Elitist strategy is a technique that consists in keeping the best individual from the last generation. Elitist strategy is applied by inserting the best individual from the preceding generation into the new generation, if a decrease in the fitness of the best individual is detected. This will insure that there is a monotonic increase in the best fitness in the population.

2.2.4 GA Advantages and Applications in Electromagnetism

The GA is particularly effective when the goal is to find the approximate global maxima in a high-dimension, multi-parameter function domain in a near-optimal manner [20]. There are several characteristics making the GA more attractive and advanced than other methods.

First, GA does not require knowledge of derivatives unlike calculus-based methods, e.g. gradient descent method. It is clear that, this class of methods can easily be trapped at a local maximum or minimum depending on the starting point of the search [16]. In addition, traditional gradient method can only work on continuous parameters spaces, and are limited to optimizing a few parameters [24]. In comparison, GA operates irrespective of the gradients, and the search extends onto the whole solution space. Consequently, it is far less probable for GA to fall into the trap of local minima or maxima in contrast to gradient search algorithms.

Another advantage of genetic algorithm is that they manipulate representations of potential solutions, rather than the solutions themselves [16]. The codings of the function parameters (chromosomes) are operated in the algorithm instead of the parameters themselves. Therefore, GA does not require a complete understanding or model of the problem, and the only requirement is the availability of a fitness function to evaluate the fitness of potential solutions to the problem.

GA has been used extensively due to its capability of handling common characteristics of electromagnetic problems, where traditional optimization methods are reaching their limits. It has been increasingly and successfully applied to a variety electromagnetic problems, including the optimization of microstrip antennas to pursue the lowest relative side-lobe level [24, 25] ; thinned and non-uniform array antennas to minimize the level of side-lobes [26–29]; light-weight, broadband, multi-layer microwave absorbers backed by a perfect conductor [30]; broadband patch antennas operating with a largely wider frequency range [31]; wire antennas in the forms of Yagi, log-periodic, helix, and spiral antennas under geometrical constraints while covering the goals of antenna efficiency, bandwidth, VSWR, gain, and matching networks [32–39].

2.3 Particle Swarm Optimization

PSO is an optimization method which was inspired by simulation of a simplified social model [40]. It has roots in two main methodologies: 1) the artificial life, e.g. bird flocking, fish schooling, swarm theory; and 2) evolutionary computation that converts evolutionary algorithm (artificial life) to computational program [40]. Explicitly, the PSO algorithm considers that individual members of the group can profit from the discoveries and previous experience of all other members of the group [41]. This indicates that social sharing of information does provide an evolutionary advantage. Importantly, this hypothesis builds the fundamental idea leading to the development of the PSO. It has also been demonstrated that PSO can outperform some other optimization methods in some certain instances. This algorithm can be well understood through an analogy which is articulated in this section. It includes the explanation of the PSO terminology, the description of the optimization procedures, as well as considerations on parameter selection.

2.3.1 PSO Theory

PSO is constructed as an algorithm that models the exploration of a problem space by population of individuals, where the success from individuals can effectively influence the search of the peers [17]. "This algorithm is related to cognition and the representation of schematic knowledge in neural networks, which makes PSO optimize the weights of the network by simulating the sharing of representations among social collaborators [17]". The features originating from the fundamentals of particle swarm intelligence makes PSO effective in optimizing difficult multi-dimensional, nonlinear and even discontinuous problems in various fields of science and engineering [42]. Similarly to other evolutionary optimization algorithms, particle swarm optimization has some key specific terminology. TABLE 2.2 illustrates and explains the descriptions of some key terms:

- **Particle / Agent**

  Each individual in the swarm is called a particle, or sometimes referred to as

Table 2.2. Terminology in Particle Swarm Optimization.

Particle	One single individual in the swarm
Agent	An agent's coordinates which represents a solution to the problem
Swarm	The entire collection of agents
p_{best}	The location of the best fitness returned for a specific agent
g_{best}	The location of the best fitness returned for the entire swarm
v_{max}	The maximum allowed velocity in a given direction

an agent. It is a potential solution assigned with a randomized velocity in the solution space [43], which will change in the iterative search towards the solution. In the PSO, each particle or agent will act individually under the same operating principle [42]. They will accelerate towards their best individual solution and the best entire swarm's solution, while constantly checking the fitness of their current location.

- **Position**

  The position of an agent is represented by the coordinates in solution space of the problem, i.e. in the set of parameter coordinates that describe the solution to the problem [42]. It is an essential and critical representation that reduces the optimization problem to a finite set of parameter values in the solution space.

- **Fitness**

  Much like the fitness in genetic algorithm, it is a function that evaluates the goodness of a position. It takes the coordinates of the position in the solution space and returns a single real value describing the quality of a solution. It is worth pointing out that this fitness function is the only link between the physical problem and the optimization algorithms [17] [42].

- **Best individual position** p_{best}

  p_{best} is the position vector obtained from the fitness function that each particle has achieved individually so far [43]. p_{best} is allocated with an initial randomized value, and is replaced by the current location if it achieves a higher fitness.

- **Best overall position** g_{best}
  In contrast to p_{best}, g_{best} records the vector position with the highest fitness discovered in the entire swarm [43]. Likewise, it will change if any other individuals achieve a better fitness.

- **Maximum velocity** v_{max}
  This parameter defines the maximum velocity allowed for the particle to move in a given direction at each step of the iteration [43].

2.3.2 PSO Procedures

To fully employ particle swarm optimization algorithm in engineering problems, a series of steps need to be implemented as described in the following. A flow chart depicting the PSO algorithm is drawn and provided in Figure 2.2. It can be observed that particle swarm optimization iterates following the algorithm logic.

I. Define the solution space. As in all optimization problems, the very first step to implement the PSO is selecting the relevant parameters that need to be optimized, and defining a reasonable range for the likely position of the optimal solution [42]. A maximum and minimum boundary value for each dimension in the N-dimensional space has to be specified.

II. Define a fitness function. This step generates the link between the physical problem and the optimization algorithm. It is crucial to define an appropriate fitness function which can accurately represent the goodness of a solution is defined.

III. Initialize random swarm locations and velocities. Each particle has to be assigned with a start location in the parameter space and with an initial velocity before commencing the search for the optimal solution. The velocity is a vector with both magnitude and direction. The initial randomized position will be the p_{best} for each particle, and the g_{best} is selected as the best solution from those initial positions [42].

IV. Fly the particles through the solution space. The fundamental of PSO is that each particle moves through the solution space towards the optimal. This will apply to all particles in the swarm moving one by one by a small amount to explore the solution

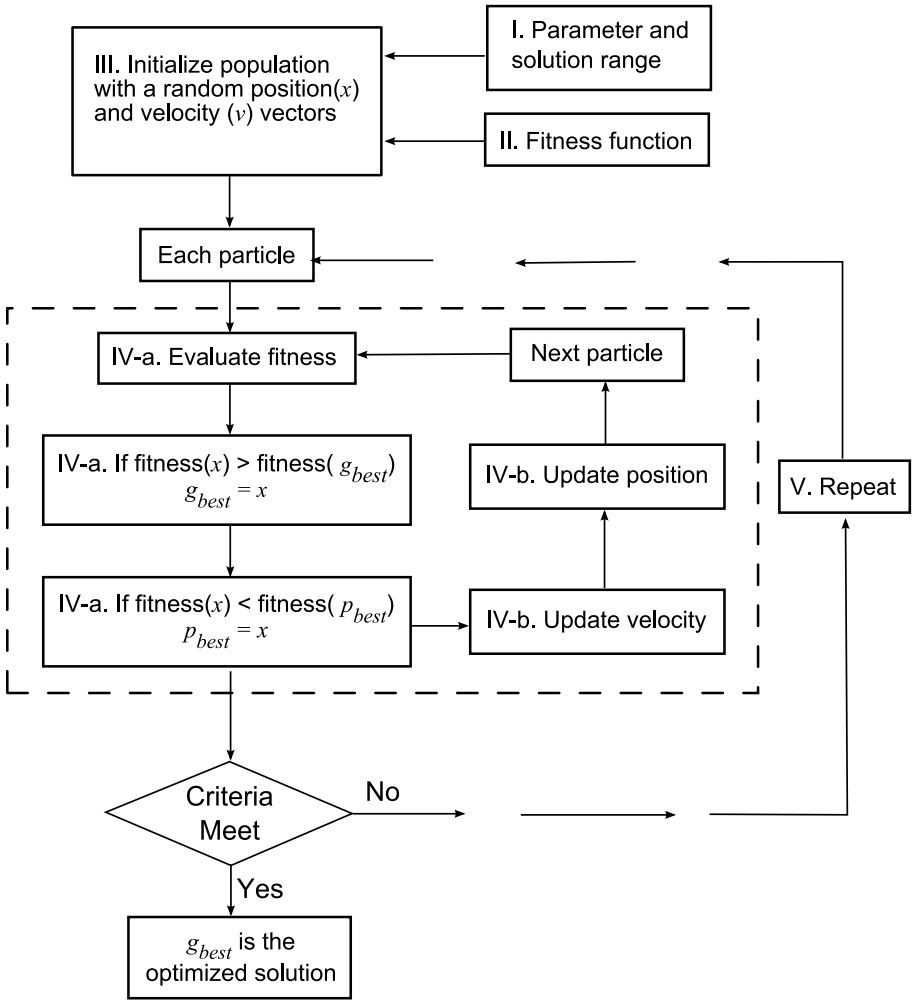

Figure 2.2. Flow chart of particle swarm optimization.

space. Such operation will be carried out in the following steps on each agent individually:

a. Evaluate the considered particle's fitness and compare to p_{best}, g_{best}.

The fitness function defined based on the coordinates of the particle in parameter space allows to evaluate the goodness of the particle. If the fitness value is higher than the respective p_{best} or g_{best}, the appropriate locations are replaced with this particular particle's current location [42].

b. Update the particle's velocity.

The manipulation of the particle velocity is the core operation of the whole optimization [42]. The change of the velocity is determined by the relative locations of p_{best} and

g_{best}, and the acceleration is provided in the formula

$$v_n = \omega \cdot v_n + c_1 \cdot rand(1)(p_{best} - x_n) + c_2 \cdot rand(1)(g_{best} - x_n) \tag{2.1}$$

In this formula, v_n is the velocity of the nth particle and x_n is its coordinate. ω is the inertial weight between 0 and 1, which determines how much the particle remains along its original course. c_1 and c_2 are the scaling factors that determine the relative 'pull' of p_{best} and g_{best}. c_1 determines how much the particle is influenced by its own best location, while c_2 determines how much the particle is influenced by the entire swarm's knowledge of the presently best location. $rand(1)$ is a random number between 0 and 1, this random element $rand(1)$ in the formula introduces slight unpredictable component simulating the natural swarm behavior [42].

c. Move the particle.

The particle will move to the next location once the velocity is determined, and the velocity is applied for a given period of time Δt. The new coordinate will be

$$x_n = x_n + \Delta t \cdot v_n \tag{2.2}$$

V. Repeat. The process of step IV repeats until the pre-defined criteria for a valid solution are satisfied. There are several methods to determine the termination criteria; one of the most-used is a fixed number of iterations [42]. Simply, the optimization will repeat step IV and V for a certain number of times until termination. What can also be used is to set a pre-defined termination fitness, explicitly, the process will cease one the overall performance has met the requirement.

2.3.3 PSO Parameters

In the PSO implementation, there are several parameters critical to control the search of the swarm, e.g. the maximum allowed velocity v_{max}, the inertial weight ω, the scaling factors c_1 and c_2, the time interval Δt, the population size, and the maximum number of iteration. These control parameters values are very important to achieve an efficient algorithm, and are the key factors for achieving a balance between exploration and exploitation. Comparatively, exploration is the search at a large scale in order to find

the optimal region, while exploitation is the probe into the optimal region for more accurate solutions. Exploration is typically preferred at the initial stages of the search but is required to gradually give way to exploitation of promising solutions as the search progresses.

- **Maximum allowed velocity v_{max}**

  v_{max} is the maximum velocity allowed for the particle to move in a given direction after every search. This value controls how fast a particle approaches a location of better fitness. For an efficient search, this velocity should not be either too large or too small. If v_{max} is too large, it will on one hand drive the particle closer to a better location faster, but on the other hand, also possibly make the particle overflow and miss the optimal position. If the v_{max} is too small, it will significantly slow down the movement of particle towards the optimum, and thus slow the convergence of the algorithm. It is found in [44] that v_{max} is best set around 10% - 20% of dynamic range of each dimension.

- **Inertial weight ω**

  The inertial weight ω determines how much the particle remains along its original course, and was developed to better control exploration and exploitation [44]. Greater inertial weight will encourage global exploration as a result of the particle being less moved by the pull of p_{best} and g_{best}, and instead preferring to fly more consistently in their original direction [42]. In the contrast, a smaller inertial weight encourages local exploitation as particles are rapidly pulled towards p_{best} and g_{best} [42]. It is suggested in [45] that varying the inertial weight linearly from 0.9 to 0.4 throughout the course of an optimization run can effectively enhance the algorithm performance, by progressively changing the strategy from exploration to exploitation.

- **Scaling factors c_1 and c_2 / Constriction factor K**

  The scaling factors c_1 and c_2 represent the weighting of the stochastic acceleration terms, which pull the particle towards the p_{best} and g_{best} positions [44]. Changes of these two constants will consequently alter the dynamics in the system. Explicitly, an increase of c_1 encourages exploration of the solution space as each

particle moves towards its own p_{best}, and an increase of c_2 encourages exploitation of the global maximum [42]. The suggestion provided in [44] states the best choice for scaling factors c_1 and c_2 of 2.0 for each, while this figure is 1.494 in [43]. A constriction factor K proposed in [46] reconstructs the velocity equation.

$$v_n = K \cdot [v_n + c_1 \cdot rand(1)(p_{best} - x_n) + c_2 \cdot rand(1)(g_{best} - x_n)] \tag{2.3}$$

where K is the constriction factor determined by c_1 and c_2 according to

$$c = c_1 + c_2; c > 4 \tag{2.4}$$

$$K = \frac{2}{\left|2 - c - \sqrt{c^2 - 4c}\right|} \tag{2.5}$$

The factor K is used in PSO implementation to further constraint the velocity update in the iteration, and insure the convergence of the particle swarm algorithm [44].

- **Time interval Δt and number of iterations**

  The time interval Δt defines the period allocated for each particle to move in Equation 2.2. In the literature, it is suggested to omit this value by setting it to 1.0. Regarding the total time of the procedure, a reasonable number of iterations in PSO is critical. On one hand, if this number is too big, PSO may stagnate until the inertial weight is reduced and exploitation can begin. On the other hand, an insufficient number of iteration leads to an early exploitation of local maxima before the swarm has adequately explored the entire solution space in search of the global maximum [42].

- **Population size**

  The population size is another essential parameter that need to be carefully selected. A large population size can undoubtedly provide a thorough exploration of the solution space, but will also significantly burden the fitness evaluation and increase the computation load. It has been found that a relatively small population can sufficiently explore a solution space and avoid an excessive number of fitness evaluations [42]. Generally, a population size around 20 to 30 is suggested in the literature.

2.3.4 Advantages of PSO and Applications in Electromagnetism

In general, PSO as one of the stochastic global optimizers, is widely recognized as a good candidate to address problems with significant nonlinearity and multi-objectives [47]. One of the major advantages of PSO is the algorithmic simplicity [42]. There is in principle only one operator in PSO, namely the velocity calculation. In comparison, the GA iteration requires several operators, including selection, crossover, and mutation, each associated with several operations and multiple control parameters. A smaller number of operators is beneficial as this leads to a reduction in computation and elimination of necessity to select the best operator for a given optimization [42]. For instance, altering the parameters c_1 and c_2 in PSO is straightforward compared to manipulating multiple operators in GA to increase the efficiency of the optimization procedure.

Another advantage that makes PSO perform efficiently is the ability in controlling the convergence [42]. The control of convergence can be achieved conveniently and adaptively through the manipulation of inertial weight in PSO. The parameter of inertial weight gives the user an opportunity to adjust the convergence rate as well as the stagnation level eventually achieved [42]. In PSO, a great value of inertial weight causes the particles to fly back and forth over the global optimal while still seeking for new locations of greater fitness.

The respective advantages of PSO and GA are based on different conceptions and completely different paradigms. It is however possible to integrate those two techniques to invent a better-performing optimization method, combining the advantages of both algorithms Several attempts have been carried out to hybridize the two optimizations, e.g. in [42, 48].

PSO has been applied to a great number of electromagnetic problems based on its great advantages of simplicity and convenient convergence control, including an multi-band and wideband patch antennas required to cover frequency bands for cellular communication [49, 50]; artificial complex ground planes for low gain wire antennas [51], periodically loaded dielectric slabs to achieve low profile [52]; and reconfigurable phase-differentiated arrays to obtain lowest peak side-lobe level [53].

2.4 Fitness Function

The fitness function, sometimes referred to as merit function, is the key element in the evolutionary optimization. It evaluates the quality of a possible solution based on the objective function value and constraint violations. The fitness function can be composed of many aspects, and each aspect needs an associated weighting factor. Therefore, the selection of suitable fitness function with appropriate weighting factors is a critical aspect of the optimization procedure.

Due to the importance of the fitness function in the process of optimization, it is therefore essential to have a better and deeper understanding of how each meriting or punishing item, as well as its weighting factors, affect the overall performance evaluation. One sample fitness function with two meriting items is used as the investigation body to probe this effect. It is well known that there exists a trade-off between efficiency and bandwidth in the design process of small antennas. Efficiency and bandwidth are taken as the two meriting items, which presumably are both required to be as high as possible. These two items are associated with weighting factors w_1 and w_2 respectively, with both weights considered to be positive.

$$Fitness = w_1 \times \ Efficiency(\%) + w_2 \times \ Bandwidth(MHz) \tag{2.6}$$

If the term w is used to represented the ratio of w_1 / w_2, fitness contour plots can be represented to indicate the trade-offs arising when varying w, as shown in Table 2.3 and Figure 2.3.

Table 2.3. *Fitness Contour at Different Weighting Ratios.*

w_1	1	1	1	1
w_2	50	100	150	200
w	0.02	0.01	0.0067	0.005

It can be observed from the plots where the same fitness is found along the color lines in each subplot, that solutions exist with same fitness but different efficiency and bandwidth. Therefore, the optimization can be improved by modifying one of the weighting factors when the other item has reached a specified objective. In this investigation, this

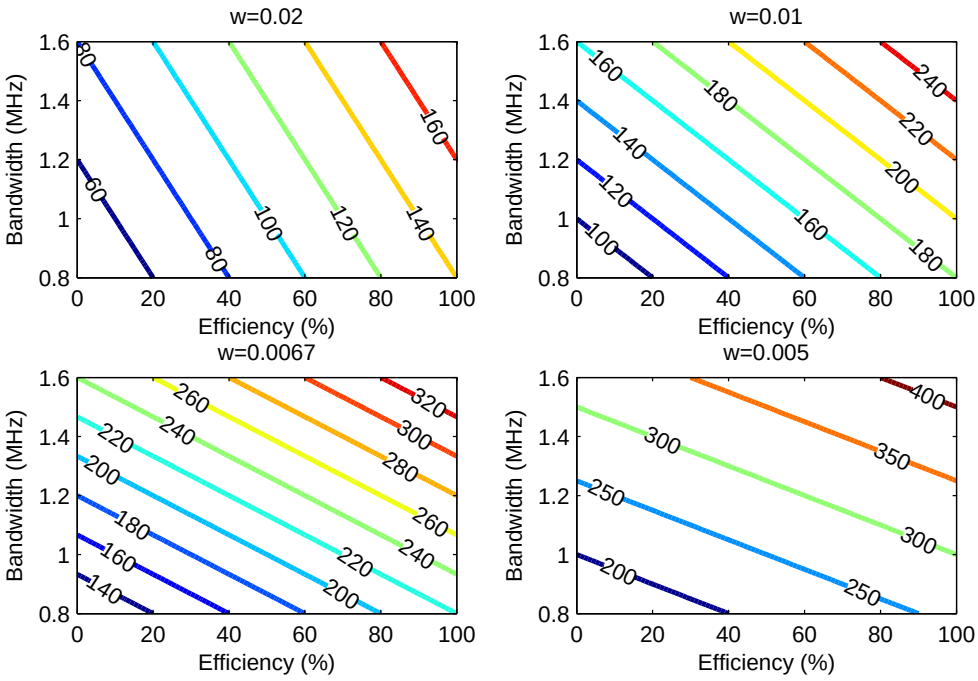

Figure 2.3. Fitness Contour at Different *w*.

strategy is employed in the optimization in the form that when the efficiency reaches a pre-defined value, its associated weighting will be reduced. By this artifice, bandwidth will have more priority in the following optimization iterations, while the efficiency maintains at stable level.

2.5 Conclusions

In this chapter, two types of evolutionary algorithms, genetic algorithm and particle swarm optimization, have been described for the solution of multi-objective optimization problems. After a brief historical description, the procedures of the two algorithms have been explicitly articulated. Parametric considerations have been also addressed in order to provide an effective application of the optimization algorithms. The critical element in the optimization, the fitness function, has also been probed to illustrate its importance and trade-offs between the objectives. Nevertheless, how multi-objective optimization is applied in the antenna design and simulation process remains to be clarified. The next chapter will continue the investigation by introducing the NEC

simulator which uses the method of moments to solve the electromagnetic problem. By integrating multi-objective optimization and NEC, electromagnetic simulation and evolutionary optimization can be effectively applied to antenna optimization. This was resulted in an automated tool (named MATNEC) which is introduced in the next chapter.

Chapter 3

EM Simulation in NEC and Reduction Strategy for Optimization

THIS chapter first provides a brief introduction and the history of the Numerical Electromagnetic Code (NEC). It presents both electric-field integral equation (EFIE) and magnetic-field integral equation (MFIE) that are used in NEC to model the electromagnetic response of a general structure. The method of moments (MoM), which is used in NEC to solve integral equations numerically, is described in this chapter. A version of NEC-2 is applied here to model the structures as well as to provide numerical solutions. The modeling process for physical structure, simulation environments, and constraints are presented in this chapter. In order to reduce the size of the parameter space for optimization, radial basis function (RBF) expansions are applied. The RBF approach can introduce sufficient variation through only a few expansion coefficients. The problem of optimizing antenna structure reduces to optimizing relatively few RBF expansion coefficients. An automated tool that combines electromagnetic simulation and evolutionary optimization, called MATNEC, is developed and employed throughout the whole investigation.

3.1 NEC Introduction

The Numerical Electromagnetic Code, commonly known as NEC, is a user-oriented computer code for the analysis of the electromagnetic behavior of antennas and other metal structures [54]. It is based on the integral equations for electromagnetism, explicitly EFIE and MFIE, where currents are excited on the structure by sources or incident fields. The method provides highly accurate solutions in terms of the current distribution on structures. The method of moments is applied in NEC to solve the integral equations. It is assumed that the structures are made up of wires and the method simply divides these wires into small straight segments and creates a current function along each segment. In NEC, antenna modeling can be composed of two major parts, structure modeling and environment modeling. Basically, the structure modeling part builds the antenna or any structures by wire segments, and the environment modeling defines the simulation settings, including ground conditions, the number of frequency points for sweeping, radiation pattern requirement, etc.

3.1.1 NEC History

The numerical electromagnetic code is well known, and continues to be developed and used widely for the antenna modeling. A brief history of NEC based on [55] is as follows. The interest in the numerical simulation of electromagnetic fields radiated from antennas started in the mid 1960s from the development of suitable integral equations. These equations then became the basis of several computer algorithms that were further developed by U.S. government research laboratories for the next decade. In early 1970s, due to the interest of antenna modeling and electromagnetic scattering, a program called antenna modeling program (AMP) was proposed and developed by several U.S. military laboratories. This had a more user-friendly interface than previous versions. It added the capacity of modeling a structure over a ground plane and the option to use file storage to increase the maximum structure size. However, AMP simulations required a long computation time, and thus had to be modified using approximations when the observation and source points exceeded specified limits. The

so-called AMP-2 used thin-wire approximations at large interaction distances in order to reduce running time. In 1977, the code NEC-1 was developed from AMP and AMP-2 by improving the current basis functions and using extended thin wire kernels. A voltage source option was also added for increased accuracy and efficiency. In 1980, the new version of NEC-2 [54] was released in response to the requirement for modeling antennas on realistic ground and large bodies, such as ships. NEC-2 used Sommerfeld integrals for the interactions of the field with a realistic ground, and interpolation techniques to provide accurate and faster solution for wire antennas. NEC-2, however, could not be used for modeling wires buried or penetrating a ground plane. More versions of NEC, such as NEC-3 and NEC-4, had been developed and released to accommodate the need of antenna modeling for more complicated circumstances. Those versions, however, have been restricted in use due to national security sensitivity. NEC-2 is the version used in this thesis.

3.1.2 Integral Equations & Numerical Solutions in NEC

The NEC-2 program uses both EFIE and MFIE to model the electromagnetic response of general structures. The EFIE is well suited to thin-wire structures of small volume, and MFIE is more efficient for voluminous structures, particularly ones with large smooth surfaces. The EFIE can also be used to model surfaces and is preferred for thin structure with little separation between front and back surfaces. When a structure contains both wires and surfaces, EFIE and MFIE are coupled.

1. **Electric field integral equations**. The form of the EFIE used in NEC-2 follows from an integral representation of the electric field of a volume current distribution $\vec{J}$:

$$\vec{E}(\vec{r}) = \frac{-j\eta}{4\pi k}\int_V \vec{J}(\vec{r'})\cdot\bar{G}(\vec{r}-\vec{r'}))dV' \tag{3.1}$$

where

$$\bar{G}(\vec{r}-\vec{r'}) = (k^2\bar{I}+\nabla\nabla)g(\vec{r}-\vec{r'}) \tag{3.2}$$

$$g(\vec{r}-\vec{r'}) = \frac{e^{-jk\left|\vec{r}-\vec{r'}\right|}}{\left|\vec{r}-\vec{r'}\right|} \tag{3.3}$$

$$k = \omega\sqrt{\mu_0\varepsilon_0} \tag{3.4}$$

$$\eta = \sqrt{\frac{\mu_0}{\varepsilon_0}} \tag{3.5}$$

$\vec{E}(\vec{r})$ is the radiated electric field, $\bar{I}$ is the current on the wire, $\vec{r}$ is the vector to a point on the wire, $\vec{r'}$ is the vector to the observation point, μ_0 is the permeability of free space, and ε_0 is the permittivity of free space.

For thin-wire approximations, several assumptions are applied in order to simplify the Equation 3.1 [54]:

a. Transverse currents can be neglected relative to axial currents on the wire.

b. The circumferential variation in the axial current can be neglected.

c. The current can be represented by a filament on the wire axis.

d. The boundary condition on the dielectric field need only be enforced in the axial direction.

Therefore, the integral equation can be reduced when the surface S is that of a cylindrical thin wire to the form:

$$I(s)\hat{s} = 2\pi a\vec{J}_s(\vec{r'}) \tag{3.6}$$

where s is the distance parameter along the wire axis at $\vec{r'}$, and $\hat{s}$ is the unit vector tangent to the wire axis at $\vec{r'}$.

The integral equation can now be expressed as:

$$-\hat{s}\cdot\vec{E}(\vec{r}) = \frac{-j\eta}{4\pi k}\int_L I(s')\left(k^2\hat{s}\cdot\hat{s}' - \frac{\partial^2}{\partial s\partial s'}\right)g(\vec{r}-\vec{r'})ds' \tag{3.7}$$

Equation 3.7 is well suited for structures with dimensions up to several wavelengths. In order to solve the integral equation numerically, a matrix equation approximation is used.

2. **Magnetic field integral equations**. The magnetic field integral equation is derived from the integral representation for the magnetic field of a surface current distribution $\vec{J}_S$:

$$\vec{H}^S(\vec{r}) = \frac{1}{4\pi}\int_S \vec{J}_S(\vec{r'}) \times \nabla' g(\vec{r} - \vec{r'}) dA' \tag{3.8}$$

where the differentiation is with respect to the integration variable $\vec{r'}$. MFIE are the equations use for modelings in terms of surface patches.

In order to solve the integral equations numerically, the method of moments is used (as introduced in [18]). In the case of wire structures, the method basically divides wires into smaller straight segments, and creates a current function along the segments. This approach parameterizes the current by a series of samples along each wire . NEC-2 uses a three term function as the basis function which is composed of a constant term, a sine term, and a cosine term. This basis function can be written as:

$$I_j(s) = A_j + B_j sin(k(s - s_j)) + C_j cos(k(s - s_j)) \tag{3.9}$$

$$\left|s - s_j\right| < \Delta_j/2 \tag{3.10}$$

where s is the distance along a wire, s_j is the value of s at the center of segment j, Δ_j is the length of segment j, and $k = \omega\sqrt{\mu_0\varepsilon_0}$ of the three constants A_j, B_j and C_j, two are eliminated by local conditions on the current, and the other one is related to the current amplitude that is to be determined by the matrix equation.

3.2 NEC Modeling and Complexity

In NEC-2, the modeling of antennas is constructed in two parts, namely structure modeling and environment modeling. The former part can be considered as the program input that defines the geometry of the structure to be modeled, while the latter one describes the excitation and the properties of the ground underneath the structure.

The basic modeling in NEC-2 uses short straight segments for wires, and flat patches for surfaces [54]. In particular, wire modeling is employed due to the primary focus on

wire antennas in this thesis. The structure to be analyzed is described in terms of wire segments, each of which is simply defined by the coordinates of its two end points and its radius. Modeling a wire structure with segments involves both geometrical and electrical factors. In terms of geometry, the segments should follow the shape of the structure as closely as possible, i.e. the best piece-wise linear approximation. Generally, each segment should be less than one-tenth (10^{-1}) of a wavelength in length; however, a much smaller length is needed to handle some critical regions in order to obtain accurate results. On the other extreme, the segment length should be no less than one-thousandth (10^{-3}) of a wavelength.

In addition, the wire radius (r) is also a major concern and is limited by the nature of approximation in the kernel of the integral equations. In the approximations, only currents in the axial direction on segment are considered, and no variation is allowed for the current around the wire circumference [54]. The acceptability of these approximations depends on both the value of r/λ and the tendency of the excitation to produce circumferential current or current variation. Unless $2\pi r/\lambda$ is much less than 1, the validity of these approximations can be compromised [56].

Another issue that needs to be taken care with is the angle of intersection of wire segments. An acute angle may be too small and place an interpolation point on one wire segment within the volume of another wire segment. A minimum requirement is that one must ensure the angle is large enough to prevent such overlaps.

In NEC-2, structure geometry is defined by using records, which contain the coordinate information of different segments. Some commonly-used structure geometry input records are listed in Table 3.1.

Three groups of data define a simulation. Group 1 contains the frequency, wire geometry, ground conditions, wire material, and loading impedances. Group 2 specifies the excitation details which can be voltage source, current sources, or plane waves. Group 3 defines the specific simulation requests, such as the nature of output (the radiation pattern for example). Groups 1 and 2 are used in all cases, while group 3 suits specific requirements. Environment modeling is controlled by input records. Parameters

Table 3.1. *Structure Geometry Input Cards.*

Cards	Functions
GA	Wire arc specification
GE	End geometry input
GH	Helix/Spiral Specification
GM	Coordinate transformation
GR	Generate cylindrical structure
GS	Scale structure dimensions
GW	Wire specification
SP	Surface patch

that are not set in the input data are given default values by NEC-2, except for excitation which must be specified. During the simulation, all parameters retain their values until changed by subsequent data records. Thus, after parameters have been set and currents or fields computed, selected parameters may be changed and the calculations repeated. Consequently, when the antenna is modified and additional calculations are requested, the order of the records may affect the solution time since the program will repeat only that part of the solution affected by the changed parameters.Some commonly-used structure geometry input cards are listed in Table 3.2.

In this thesis, the input records for the antenna models and environment are constructed and edited in MATLAB. Nevertheless, when the structures are defined in terms of segments, these can require a huge number of variables when optimizing structures. Such large numbers of variables will greatly increase the complexity for the optimization. In order to lower the parametric complexity in the modeling and optimization stages, several strategies will be employed. In particular, radial basis functions expansions will be used.

3.3 Basis Functions

In order to represent structural variation with a limited number of parameters, basis function expansions can be used. In mathematics, a basis function is an element

Table 3.2. *Program Control Cards.*

	Cards	Functions
Group 1	EK	Extended thin-wire kernel
	FR	Frequency specification
	GN	Ground parameter specification
	LD	Structure impedance loading
Group 2	EX	Structure excitation
Group 3	GD	Additional ground parameter specifications
	NE	Near electric field request
	NH	Near magnetic field request
	RP	Radiation pattern request
	EN	End of data

of a particular set of functions that together describe the total variation required [57] [58]. Theoretically, any function in a certain function space can be represented and expressed in a form of a linear combination of such basis functions. Therefore, any continuous function on a compact interval can be interpolated with arbitrary accuracy by a suitable sum of such basis functions. There are several types of basis function which have been widely used in the fields of science and engineering, and three possibilities are described below. They are employed to fit a sinusoidal curve as a demonstration of their use. A sinusoidal function with a unit magnitude and a period of 2π ($f(x) = \sin(x)$), is plotted in Figure 3.1.

- **Polynomial basis function**

  In mathematics, a polynomial is a linear combination of positive powers of the independent variable [57]. A polynomial in x can be written as:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_n x^n \tag{3.11}$$

  For demonstration purposes, a polynomial basis function is applied to represent the sinusoidal curve. To cover the extrema of the sinusoidal, at least six parameters are required in Equation 3.11 (the fitted curve is plotted in Figure 3.2). It is

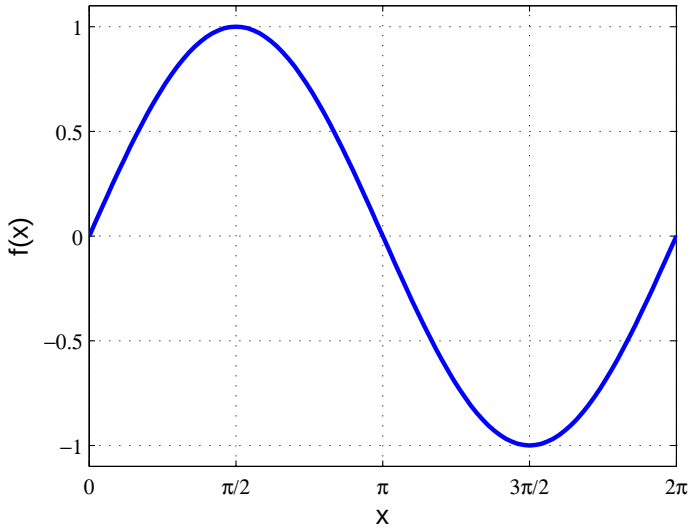

Figure 3.1. Illustrations of sinusoidal curve to be fit.

observed that when 5_{th} order polynomial is applied, the sinusoidal curve can be well represented. Unfortunately, polynomial approximation of the above form can often exhibit strong oscillation above the functions they are meant to represent.

- **Spline basis function**

  In mathematics, a spline is a function defined piecewise by a set of polynomials [58]. For a function taking values within an interval [A, B], interval [A, B] be covered by n ordered, disjoint subintervals $[x_i, x_{i+1}]$, with $A = x_0 \leq x_1 \leq \ldots \leq x_{n-1} \leq x_n = B$, and $i = 0, 1, \ldots n-1$. On interval $[x_i, x_{i+1}]$, the spline is defined by polynomial P_i:

$$\begin{aligned} S_0(x) &= P_0(x), x_0 \leq x < x_1 \\ S_1(x) &= P_1(x), x_1 \leq x < x_2 \\ S_2(x) &= P_2(x), x_2 \leq x < x_3 \\ &\vdots \\ S_i(x) &= P_i(x), x_{n-1} \leq x < x_n \end{aligned} \tag{3.12}$$

  where each polynomial P_i is a cubic equation:

$$P_i(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \tag{3.13}$$

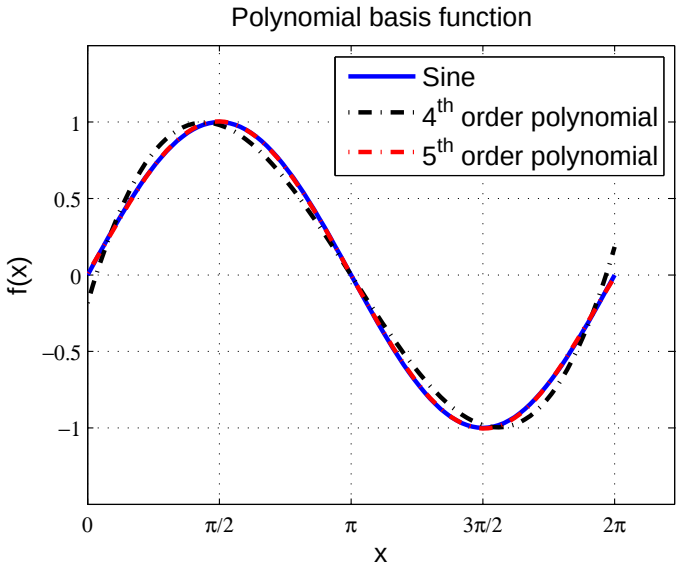

Figure 3.2. Illustrations of polynomial basis function to represent the sinusoidal curve.

Splines have given values at points x_0 to x_n, and are continuous in the second derivative. A spline basis function can be used to represent the curve of Figure 3.1. In this case, a significantly greater number of parameters is required to fill the Equation 3.12 (a total number of 24). The value of the parameters are shown in Table 3.3 and the curve is plotted in Figure 3.3. It is observed that the spline basis function can effectively represent a sinusoidal curve. However, the differentiability requirements are difficult to implement and lead to some difficulties in the optimization, as the optimal structure variation is never pre-defined prior to the optimization process .

Table 3.3. *Parameters of the spline basis function.*

	a_0	a_1	a_2	a_3	Interval
P_1	-0.0314	-0.2961	1.1716	0	$[0,\frac{\pi}{3}]$
P_2	-0.0314	-0.3949	0.4480	0.866	$[\frac{\pi}{3},\frac{2\pi}{3}]$
P_3	0.1571	-0.4936	-0.4824	0.866	$[\frac{2\pi}{3},\pi]$
P_4	0.1571	0	-1	0	$[\pi,\frac{4\pi}{3}]$
P_5	-0.0314	0.4936	-0.4824	-0.866	$[\frac{4\pi}{3},\frac{5\pi}{3}]$
P_6	-0.0314	0.3949	0.4480	-0.866	$[\frac{5\pi}{3},2\pi]$

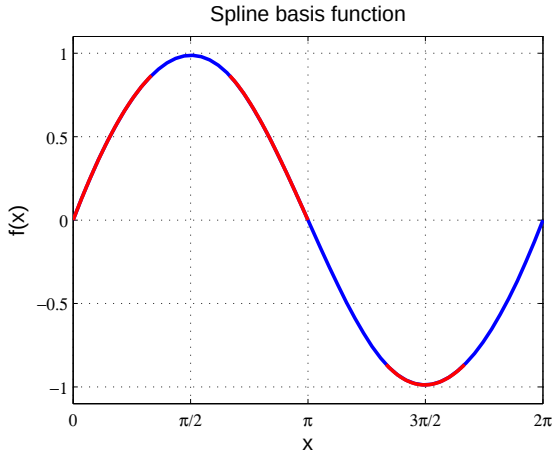

Figure 3.3. Illustrations of Spline basis function to represent the curve.

- **Radial basis function**

  In present work, radial basis functions have been found to be the most effective. A radial basis function is a real-valued function whose value depends only on the distance from the origin, that is $f(x)=(\|x\|)$ [57], or on the distance from a certain point c that is $f(x,c)=f(\|x-c_i\|)$. Linear combinations of RBFs are typically used to build up approximations of the form:

$$y(x) = \sum_{i=1}^{N} \omega_i \cdot f(\|x - c_i\|) \tag{3.14}$$

  where N is the number of RBFs, f is the radial functions, c_i are the control points, and ω_i are scalar weighting parameters.

3.3.1 Gaussian and Multiquadric RFBs

In this thesis, two types of RBFs are employed to represent the variations on the antenna structure, explicitly Gaussian radial basis function (G-RBF) and Multiquadric radial basis function (MQ-RBF). Representations take the form:

$$f_{G-RBF} = \sum_{i=1}^{N} \omega_i \cdot e^{-R(x-c_i)^2} \tag{3.15}$$

for G-RBFs where exponential function is used as the basis, N is the number of RBFs in the combination, ω_i are scalar weighting parameters, c_i are the control points, and R

describes the width of each RBF. Or, in the case of MQ-RBFs:

$$f_{MQ-RBF} = \sum_{i=1}^{N} \omega_i \cdot (1 + \varepsilon^2 (x - c_i)^2)^{\beta} \tag{3.16}$$

where N is the number of RBF in the combination, ω_i are scalar weighting parameters, c_i are the control points, ε describes the width of each RBF, and β is typically taken a value of $-\frac{3}{2}$.

Figure 3.4 shows the RBF representation of the curve in Figure 3.1. It is noticed that a set of six radial basis functions can also roughly represent the sinusoidal curve. More importantly, each subset of the RBFs can be manipulated to be either independent or reliant on its neighbors, which makes RBF easier for implementation.

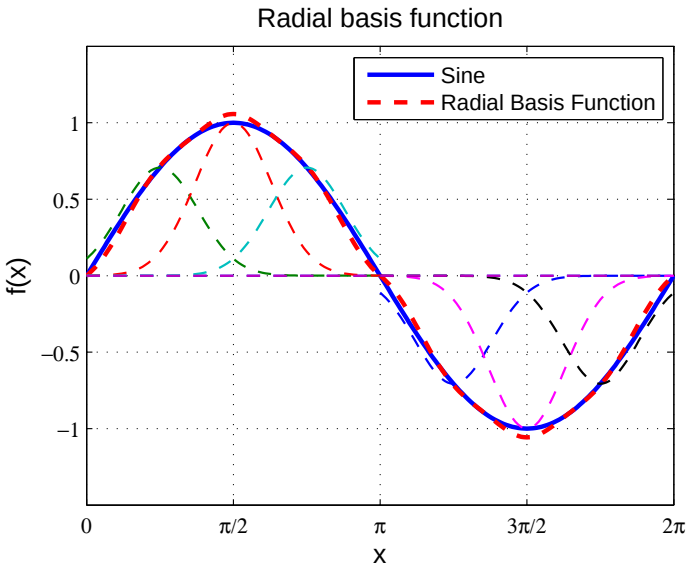

Figure 3.4. Illustrations of radial basis function to represent the curve.

3.3.2 RBF Implementation and Advantages

In this thesis, RBFs are employed and implemented to represent the variations in structure of various types of wire antennas, e.g. zigzag antennas, and helical antennas. Conventionally, variations are represented exhaustively from segment to segment, which

leads to a significantly large number of variables and dramatically increases the computation load and optimization complexity in the design process. Therefore, the purpose of using RBFs is to achieve sufficient shape variations by using a minimum number of optimization variables, and consequently reduce the simulation load and optimization complexity. The implementation of RBFs for different types of wire antennas will be explicitly described in our case studies. Also, comparisons between Gaussian RBF and Multiquadric RBF with respect to convergence will be provided under a specific circumstance in Chapter 4.

3.4 MATNEC

In this thesis, electromagnetic simulation with NEC-2 provides fast estimations of antenna performance for given wire antenna geometries, and the evolutionary optimization guides the available geometric freedom to its optimal value. This combination is built into an in-house automated tool, called MATNEC, that uses MATLAB as the main control. A flow chart of MATNEC is shown in Figure 3.5.

For an initial antenna structure, MATNEC simulates the antenna and evaluates its fitness. If the antenna satisfies a fitness criterion, the design is applied, otherwise it is further optimized using GA or PSO techniques. The fitness is then re-evaluated and further optimized. Such iteration will continue until convergence to a near optimal antenna configuration is achieved. The output of MATNEC is then the optimized antenna geometry and its associated performance.

The development and employment of MATNEC is one of the major milestones and achievements in this thesis. This in-house tool successfully combines simulation and optimization into one automated program with convenient MATLAB control. Two optimization algorithms options are available to target different problems, the settings for each algorithm is highly accessible and controllable. The fitness evaluation process is embedded in MATNEC and can effectively accommodate different fitness functions reflecting different requirements and considerations with respect to different problems.

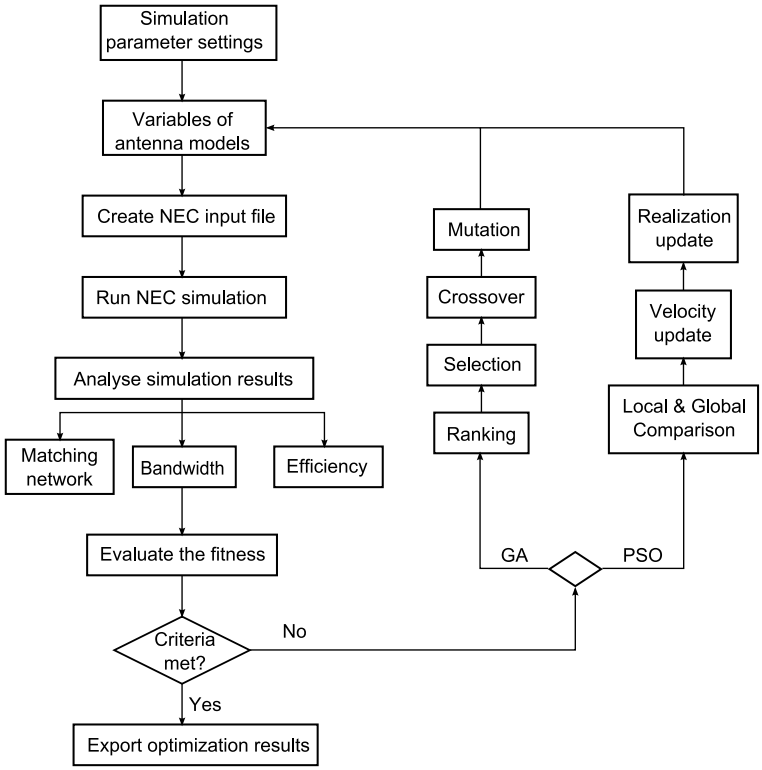

Figure 3.5. Flow chart of MATNEC with GA and PSO.

3.5 Conclusions

This chapter explicitly describes the electromagnetic simulator NEC and its history. The integral equations on which NEC is based are described together with their method of solution, i.e. the method of moments. Basis functions are also introduced in this chapter as these are used to reduce the heavy computational load in optimization. In particular, radial basis functions are employed to provide efficient structural variations to the possible antenna configurations for optimization. MATNEC, an automated in-house tool, is developed to couple the simulation and optimization process. This tool will be intensively used to pursue optimal antenna configurations throughout the rest of this thesis. With MATNEC developed, it is important to test and verify its robustness. The verification will be performed starting from a relatively simple model, i.e. lumped inductors that are to optimally distributed along a straight wire. This will be investigated in the next chapter.

Chapter 4

Case Study 1: Distribution of Lumped Inductances on a Monopole

THIS chapter first revisits the well-known fact that inductive loadings mounted in series on a monopole can shorten the monopole resonant length. To that aim, inductive loadings are aligned along a resonant wire antenna in simulations designed to characterize the radiation performance. It is confirmed that the size of resonant wire antennas can be reduced through this scheme at the cost of performance, both in terms of lower efficiency and narrower bandwidth comparing to a self-resonant quarter-wave monopole. Therefore, the exact location of inductance on the monopole for optimal performance becomes a key issue. In the investigation, relevant issues related to the optimization are considered, including the *L*-section matching network, the impact of Q factor of the inductors, as well as the fitness function selection. Several distribution schemes with variable numbers of inductors are considered, and the optimal solutions in terms of the optimized inductors loading positions and inductance values are obtained through simulation and optimization process.

4.1 Introduction

Size is one of the critical considerations in the design process for HF and VHF antennas. It is a well-known fact that inductive or capacitive loadings in series with a monopole can shorten its resonance length [59]. When the resonance length is successfully shortened, the monopole with a significant size reduction can be categorized as electrically small, as defined in Chapter 1 [4]. Nevertheless, the miniaturization of the monopole is achieved at the cost of performance, in terms of both lower efficiency and narrower operational bandwidth. Thus, the performance considerations needs to take into account the radiation efficiency and operational bandwidth, as done in this chapter. An *L*-section matching network is used to deliver the maximum operational bandwidth around the target frequency, and the total efficiency including radiation efficiency, component loss, and matching loss is used to represent the efficiency performance of the monopole. An appropriate fitness function covering those considerations is then selected for the optimization, which is performed with the evolutionary optimizers introduced in Chapter 2. Investigations are carried out with respect to different numbers of lumped inductors distributed along the monopole, and the optimal solutions and resulting performance effectively verify the roles of inductive loadings in reducing the resonance length of wire antennas.

For demonstration purposes, a quarter-wave monopole resonant at 300 MHz (equivalently with a length of 0.25 m) at its unloaded state is used as the base model to investigate the performance with various loading schemes. For simulation, this monopole is divided into 25 equal-length segments, which are labeled 1 to 25 from the base excitation to the top free end. The configuration and characteristics of the monopole are shown in Figure 4.1 and Table 4.1. The investigation goal is to have an inductively loaded 0.25 m monopole behave resonantly at the reduced frequency of 100 MHz while performing with satisfactory efficiency and bandwidth. Explicitly, if a 0.25 m monopole with inductive loading can have the same resonant frequency as a self-resonant 0.75 m quarter-wave monopole at 100 MHz, the shorter monopole can be claimed to be miniaturized by three times.

Table 4.1. *Characteristics of the resonant monopole with inductive loadings.*

Resonant frequency	100 MHz
Length	0.25 m
Ground condition	Perfect conducting ground
Number of segments	25
Excitation segment number	#1
Excitation voltage	1 V

4.1.1 Matching Networks and Operational Bandwidth

In order to achieve the maximum operation bandwidth for the monopole, an L-section lumped reactive matching network with one inductor and one capacitor is used to conjugately match any source impedance Z_S (usually 50 Ω) to the loading impedance Z_L of the antenna. In general, those two impedances can be complex and composed with both resistance and reactance, i.e.,

$$\begin{aligned} Z_S &= R_S + jX_S \\ Z_L &= R_L + jX_L. \end{aligned} \tag{4.1}$$

The configurations for the L-section matching network depend on whether the resistance of source impedance R_S is greater or smaller than loading resistance R_L [60]. The schematics of the L-section matching network are drawn in Figure 4.2. If $R_S > R_L$, it is referred to as a normal L-section, otherwise it is called reversed L-section. The use of reactive elements minimizes the power losses in the matching network.

For either type, the matching network transforms the load impedance Z_L into the complex conjugate of the source impedance Z_S, that is [60]

$$Z_{in} = Z_S^* \qquad \text{(conjugate match)} \tag{4.2}$$

where Z_{in} is the input impedance looking into the L-section:

$$\begin{aligned} Z_{in} &= \frac{Z_1(Z_2+Z_L)}{Z_1+Z_2+Z_L} \qquad \text{(normal)} \\ Z_{in} &= Z_2 + \frac{Z_1 Z_L}{Z_1+Z_L} \qquad \text{(reversed)} \end{aligned} \tag{4.3}$$

with $Z_1 = jX_1$ and $Z_2 = jX_2$.

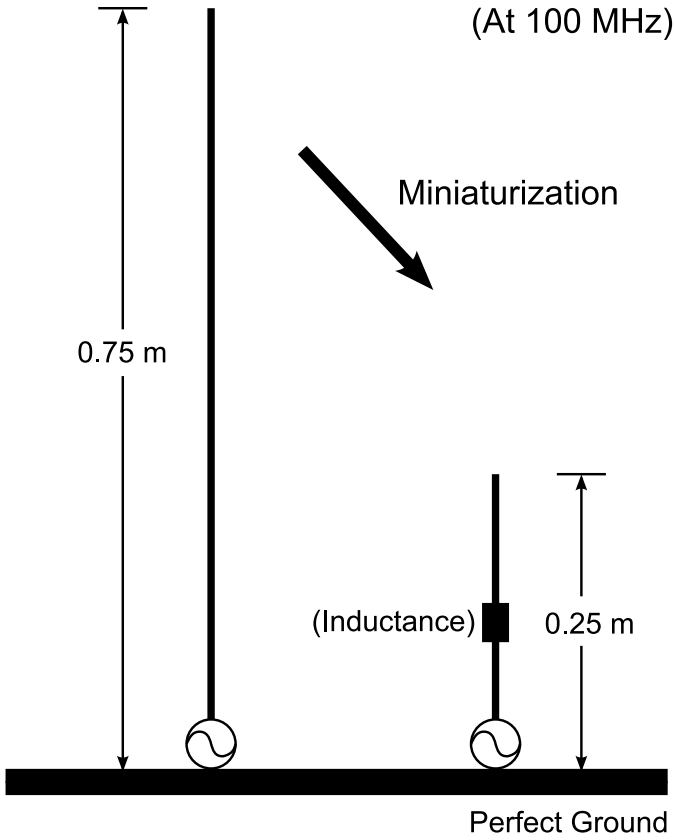

Figure 4.1. Illustration of 0.75 m self-resonant monopole and a miniaturized 0.25 m resonant monopole with inductive loadings at 100 MHz.

Substituting Equation 4.3 into the conjugate match condition Equation 4.2, and equating the real and imaginary parts of the two sides separately, the values for X_1 and X_2 for the matching at the specific frequency can also be obtained for normal and reversed L-section matching network, as

$$\begin{aligned} X_1 &= \frac{X_S \pm R_S Q}{\frac{R_S}{R_L} - 1} \\ X_2 &= -(X_L \pm R_L Q) \qquad \text{(normal)} \\ Q &= \sqrt{\frac{R_S}{R_L} - 1 + \frac{X_S^2}{R_S R_L}} \end{aligned} \tag{4.4}$$

$$\begin{aligned} X_1 &= \frac{X_L \pm R_L Q}{\frac{R_L}{R_S} - 1} \\ X_2 &= -(X_S \pm R_S Q) \qquad \text{(reversed)} \\ Q &= \sqrt{\frac{R_L}{R_S} - 1 + \frac{X_L^2}{R_S R_L}} \end{aligned}$$

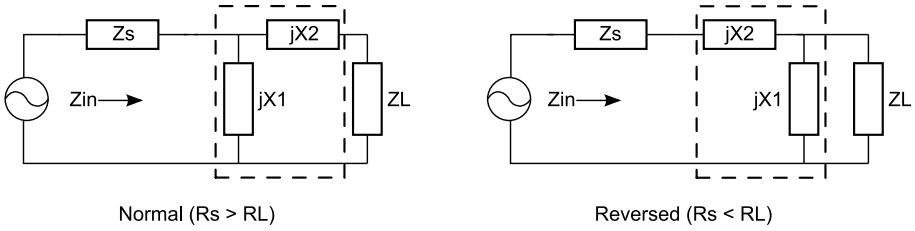

Figure 4.2. L-section reactive conjugate matching network.

In this investigation, the designed antenna is deemed as the load, and the L-section matching network is used to match it to the characteristic impedance of 50 Ω that is R_S = 50 Ω. Thus, the the source impedance is purely resistive so that $X_S = 0$. Equations 4.4 can be simplified as:

$$\begin{aligned} X_1 &= \pm\frac{R_S Q}{\frac{R_S}{R_L}-1} \\ X_2 &= -(X_L \pm R_L Q) \qquad \text{(normal)} \\ Q &= \sqrt{\frac{R_S}{R_L}-1} \end{aligned}$$

$$\begin{aligned} X_1 &= \frac{X_L \pm R_L Q}{\frac{R_L}{R_S}-1} \\ X_2 &= \mp R_S Q \qquad \text{(reversed)} \\ Q &= \sqrt{\frac{R_L}{R_S}-1+\frac{X_L^2}{R_S R_L}} \end{aligned} \tag{4.5}$$

The choice of whether X_1 and X_2 are capacitive or inductive depends on how the circuit is required to perform with respect to the target frequency. The L-section matching network can behave as either a low-pass filter or a high-pass filter [60] [61]. Based on the relationship between R_S and R_L which determines normal or reversed types, the value of X_L defined as capacitive or inductive for the load, and the frequency performance requirement of either low-pass or high-pass, there are total a number of 4 possible L-section configurations. The illustrations of those configurations are shown in Figure 4.3.

With the appropriate L-section matching network, the antenna is conjugate matched to Z_S of 50 Ω at the target frequency. Nevertheless, this antenna will be still imperfectly matched in the rest of the frequency range under the same matching network. This mismatch, while usually undesirable, is very common in antenna transmission. It can

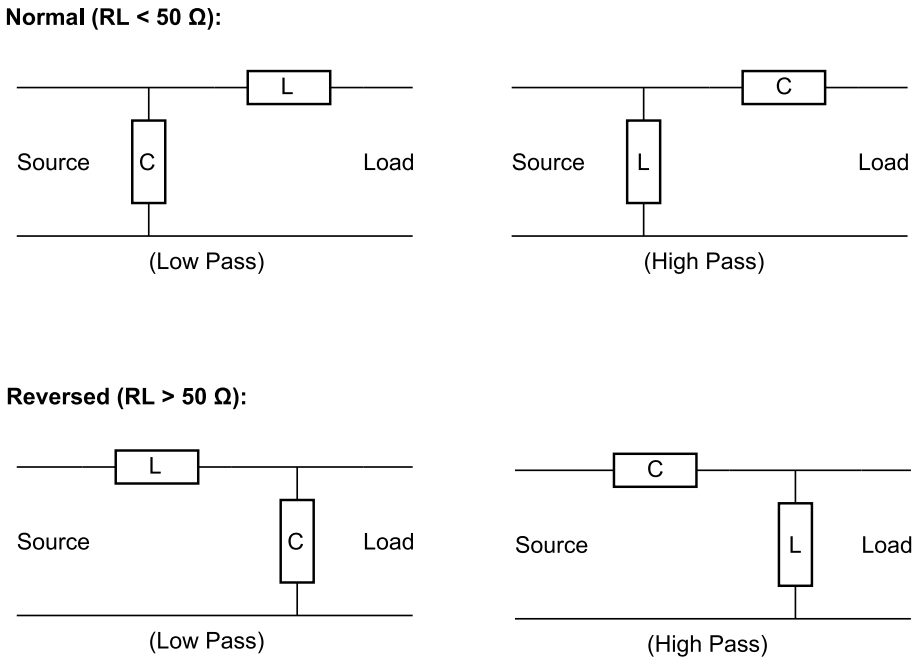

Figure 4.3. L-section reactive conjugate matching network.

be tolerated to a certain extent, and this usually defines the main restriction to the operation bandwidth. The reflection coefficient defined as the ratio between incident and reflected voltage is usually used to characterize this mismatch.

$$\Gamma = \frac{V_r}{V_i} \tag{4.6}$$

It can also be interpreted using Z_S and Z_L:

$$\Gamma = \frac{Z_L - Z_S}{Z_L + Z_S} \tag{4.7}$$

In general, Γ is complex unless both Z_S and Z_L are resistive. It is positive for $Z_S < Z_L$ and negative $Z_S > Z_L$.

When signals are transmitted and reflected, there will exist standing waves of waves and current, and the voltage standing wave ratio (VSWR) can be defined as:

$$VSWR = \frac{V_{max}}{V_{min}} \tag{4.8}$$

where

$$\begin{aligned} V_{max} &= |V_i| + |V_r| \\ V_{min} &= |V_i| - |V_r| \end{aligned} \tag{4.9}$$

Therefore, VSWR can be expressed as:

$$\begin{aligned} VSWR &= \frac{V_{max}}{V_{min}} \\ &= \frac{|V_i|+|V_r|}{|V_i|-|V_r|} \\ &= \frac{1+\frac{|V_r|}{|V_i|}}{1-\frac{|V_r|}{|V_i|}} \end{aligned} \tag{4.10}$$

Explicitly,

$$VSWR = \frac{1+|\Gamma|}{1-|\Gamma|} \tag{4.11}$$

With the VSWR obtained for the frequency range, an operational bandwidth can be calculated under a specific pre-defined VSWR value (for example VSWR < 2.0). It can be seen in Figure 4.4 that the lower and upper frequency boundary with a VSWR of 2.0 are F_1 and F_2 respectively, while the target frequency F_{target} is perfectly matched with a VSWR of 1.0. The operational bandwidth (B) can be calculated as $B = F_2 - F_1$.

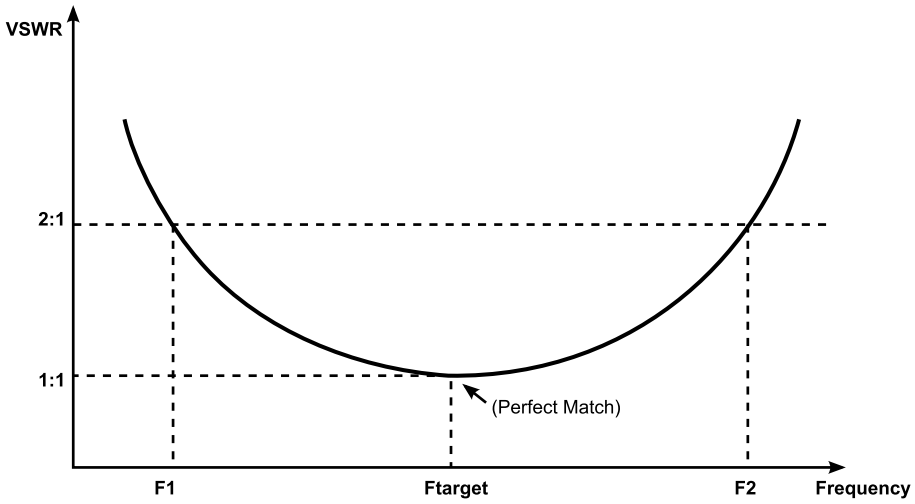

Figure 4.4. Illustrations of operational bandwidth with VSWR < 2.0.

4.1.2 Q Factor of Inductance

In approximation, the lumped inductors that are loaded in series with the monopole can be considered as ideal inductive with an infinite quality factor (Q) for simplicity. Nevertheless, an imperfect Q factor is inevitable because of the inherent resistance within the component [62]. Ideally, a lumped inductor with an infinite Q factor acts

purely reactive, and it opposes current changes based on inductive reaction. In reality, due to the fact that any metal wires possess some amount of resistance, the imperfect conductivity will introduce a resistive element in series with the inductance. Figure 4.5 illustrates the equivalent circuit of a realistic lumped inductor.

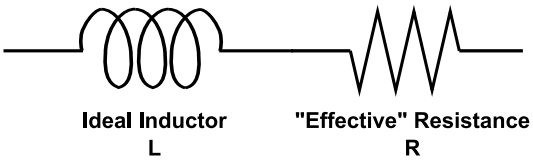

Figure 4.5. Equivalent schematics of a real inductor.

The quality factor of an inductor is the ratio of its inductive reactance (L) to its resistance (R) at a given frequency (f), and is a measure of its efficiency [62]. It can be calculated as:

$$Q = \frac{2\pi f L}{R} \tag{4.12}$$

The Q factor of an inductor is normally determined by its materials, geometry and manufacture procedure, and thus once this quantity is specified, the inherent resistance can be found out:

$$R = \frac{2\pi f L}{Q} \tag{4.13}$$

In this chapter, inductors with various values of Q factor will be used, in order to investigate their impact towards the performance of the designed monopole, and inductors distribution.

4.1.3 Fitness Function

The goal of the optimization presented in this chapter is to find the optimal inductive loading distribution and the corresponding inductance values. The judging criteria to determine the goodness of a solution is derived from the resulting monopole performance. Therefore, the selection of an appropriate fitness function is of critical importance to the success of the optimization. Specifically, the fitness function is composed of two meriting contributions, namely total efficiency and operational bandwidth. Total efficiency includes radiation efficiency at the target frequency and the efficiency of the

matching network, while operational bandwidth is defined as the bandwidth with a VSWR less than 2.0. Each contribution is assigned with an individual weighting factor, and the fitness function can be written as

$$Fitness = w_1 \times Efficiency(\%) + w_2 \times Bandwidth(MHz), \tag{4.14}$$

where the weights w_1 and w_2 are selected such that efficiency and bandwidth are of similar magnitudes. In the test case presented in this chapter, these weights are selected as $w_1 = 1$ and $w_2 = 100$.

4.2 Single Lumped Inductor

When one single lumped inductor is located on the monopole, the effect of the loading inductance can vary significantly at various position along the antenna height, and this can be exploited in order to achieve resonance at the target frequency. Correspondingly, the efficiency and operational bandwidth behave differently for different positions, because the lumped inductance change the current distribution along the antenna height and thus the resonant frequency. Nevertheless, the overall fitness at one or several loading positions may outperform other loading distributions. The first task of the investigation is to find the optimal solution containing the best loading position and inductance value for a single lumped inductor to achieve a monopole resonance at the target frequency of 100 MHz while performing with maximized total efficiency and operational bandwidth.

4.2.1 Parameter Sweep Solution

To produce a reference solution, the optimal position of the inductor along the height is first found for the single inductor case through a parameter sweep procedure. At each loading position from #2 to #25, the inductance is swept logarithmically within the range of 0.1 μH ($1e^{-7}$ H) to 100 μH ($1e^{-4}$ H), where the Q factor is assumed to be ideally infinite. The tuning inductance required to induce a resonance at 100 MHz is obtained at each loading position, and shown in Figure 4.6. It can be seen that the

tuning inductance varies with position from $10^{-6.08}$ H (equivalently 0.82 μH) to $10^{-4.7}$ H (equivalently 19.95 μH).

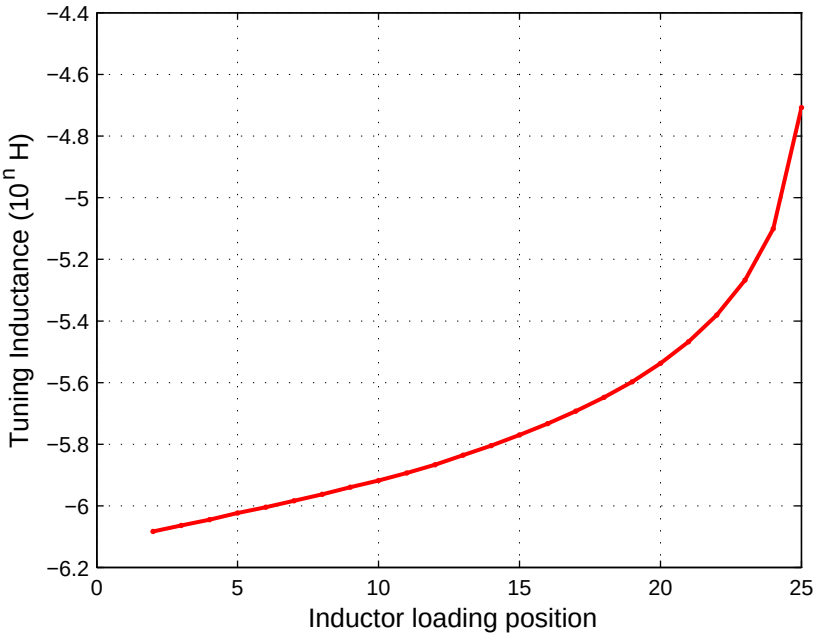

Figure 4.6. Best tuning inductance at each position for a resonance at 100 MHz.

In the next step, lumped inductors with a realistic Q factor are considered with explicit values of 100, 50, and 25 respectively. As shown before, the inductance required to force a resonance varies from one position to another along the monopole, it is shown that it also varies with respect to different Q factors. The resulting performance are compared in Figure 4.7. It is noted that, as can be expected, when the Q factor is increased, 1) the radiation efficiency increases, and 2) the bandwidth decreases. This illustrates the existing tradeoff between the two parameters. Due to the fact that Q =100 is a common value of manufactured inductors, this quantity is used for current and future investigations. Considering specifically the performance with a Q factor of 100 in Figure 4.7, it can easily be found from these simulated data that when an inductor of 1.14 μH is loaded at position #9, it can result in an efficiency of 37.8 % and a bandwidth of 1.012 MHz.

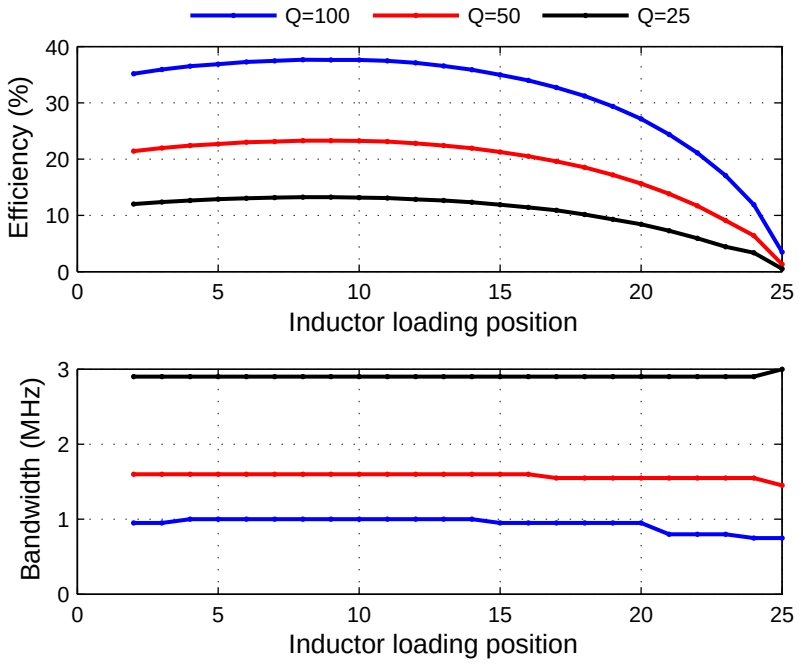

Figure 4.7. Efficiency and bandwidth of a monopole with length 0.25 m and a resonance at 100 MHz with an inductance moved at different height position for different inductor Q factors.

4.2.2 Evolutionary Optimized Solution

As a second step, the evolutionary optimization is carried out on this single inductor problem to validate the methodology. The loading position and inductance values are taken as the two optimizing variables. The optimization settings are listed in Table 4.2. Variable are digitized with 5 bits and 10 bits for position and inductance respectively. A linear scale is used for the position quantization, while the inductance is quantized using a logarithmic scale, so that each position can be represented and the inductance can be finely tuned. A population of 200 genes provide a fairly large group of possible solutions initially, and 50 iterations (generations) provide the genes sufficient freedom to converge to the optimal region. Meanwhile, the elitism selection strategy insures that the best performing genes are kept in the optimization population, and a 10% mutation rate allows some random genes to be included so that trapping in local optimums can be effectively avoided.

The result of the optimization indicates that the achieved solution with the greatest fitness values locates a lumped inductor of 1.14 μH at position #9 from the bottom,

Table 4.2. *GA optimization settings for single inductor model.*

No. of bits for position	5
No. of bits for inductance	10
Position range	[#2, #25]
Inductance range (logarithmic quantization)	[0.1 μH, 100 μH]
No. of genes	200
No. of iterations	50
Selection strategy	Elitism
Mutation rate	10%

which leads to an antenna performance of 37.8% efficiency and a bandwidth of 1.012 MHz. The optimized solution is identical to the one obtained from parameter sweep, validating the effectiveness of evolutionary optimization. By repeating the optimization numerous times, it is observed that the optimal solution is not always found, but that near-optimal solutions always converge to a loading positions at a height between about a quarter (#7) and half (#12) of the total height from the bottom, as shown in Figure 4.8. This phenomenon, i.e. no guarantee of finding the overall optimum, is a typical behavior for evolutionary optimizers, especially in the occurrence of flat region in the merit function.

4.3 Distribution of Multiple Lumped Inductors

In order to further validate the effectiveness and accuracy of the Matlab-controlled NEC simulation and GA optimization, some more sophisticated inductor distribution schemes are considered. First, two inductors are distributed systematically or arbitrarily along the monopole. Then, higher numbers of lumped inductors are allowed to be distributed arbitrarily along the monopole length and their positions and inductance are optimized. For comparison purposes, the fitness function used to assess the performance of each solution remains the same irrespective to the number of inductors. A Q factor of 100 for the inductors and 50 for the capacitors in the matching network are considered.

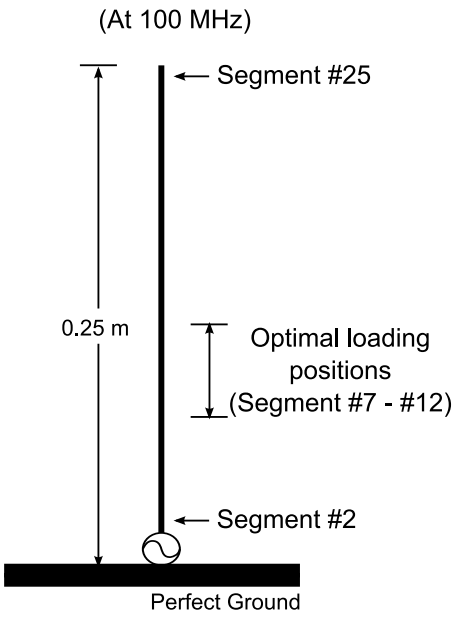

Figure 4.8. Equivalent schematics of a real inductor.

4.3.1 Inductance Distribution Schemes

The first scheme considers two lumped inductors with equal inductance which are positioned symmetrically about the center of the monopole. Explicitly, along the labeled positions from bottom to the top, the two inductors can only be positioned at #2 & #25, #3 & #24, #4 & #23 and so on until #13 & #14. This type of distribution scheme leads to 12 possible location variations. Those possibilities and the associated inductance value are taken as the optimizing variables for MATNEC to process. In a second scheme, the two inductors are given more freedom and can be located arbitrarily anywhere along the monopole, with an inductance value within the pre-defined range. In this case, four variables are involved in the optimization, namely, two loading positions and two inductance values. In the third scheme, the number of loading lumped inductors is increased progressively, up to a number of 10 inductors, which can be located arbitrarily along the monopole. In this last case, the parameter space is dramatically increasing with the number of inductor considered, and the method of parameter sweeping is no long feasible to pursue the optimal solution. Consequently, evolutionary optimization becomes the only way to efficiently solve the problem.

4.3.2 Evolutionary Optimized Solution

To carry out the optimization for the various models consistently, the basic settings are left unchanged from the single inductor investigation. Only the position range and total number of bits for each gene are varied according to the number of lumped inductors. The common settings and the fitness function are shown in Table 4.3.

Table 4.3. *GA optimization basic settings for different models.*

No. of bits for each position	5
No. of bits for each inductance	10
Inductance range	[0.1 μH, 100 μH]
No. of genes	200
No. of iterations	50
Selection strategy	Elitism
Mutation rate	10%
Fitness function	$Fitness = Efficiency(\%)$ $+100 \times Bandwidth(MHz)$

- **Two Symmetrical Inductors**

  When two identical lumped inductors are loaded symmetrically along the center height of the monopole, only two variables are involved in the optimization procedure. The upper loading position (P) is taken as the first of the optimizing variable, and the lower position can then be represented as 27-P. Therefore, the position range for this optimization is [#14, #25], and a number of 10 additional bits are used to define the inductance value in the genes. The solutions are obtained through evolutionary optimization and the achieved performance are plotted in Figure 4.9. It can be seen that the monopole exhibits the best performance when the two identical inductors are loaded at position #17 & #9, #18 & #8, or #19 & #7, in all cases with an efficiency greater than 39%. The overall optimal solution is found for two inductors of 0.849 μH located at segments #18 & #8, with a resulting bandwidth of 1.05 MHz and a total efficiency of 39.4%. It is necessary to point out that a parameter sweep technique can also be applied

in this case, because the number of variable is the same as in the single inductor case. More importantly, it is also noticed that both efficiency and bandwidth are slightly improved in comparison to the single inductor case, which triggers the investigation towards an expansion including larger numbers of inductors. In those cases, the parameter sweep technique in no longer efficient and able to deal with the required numbers of variables.

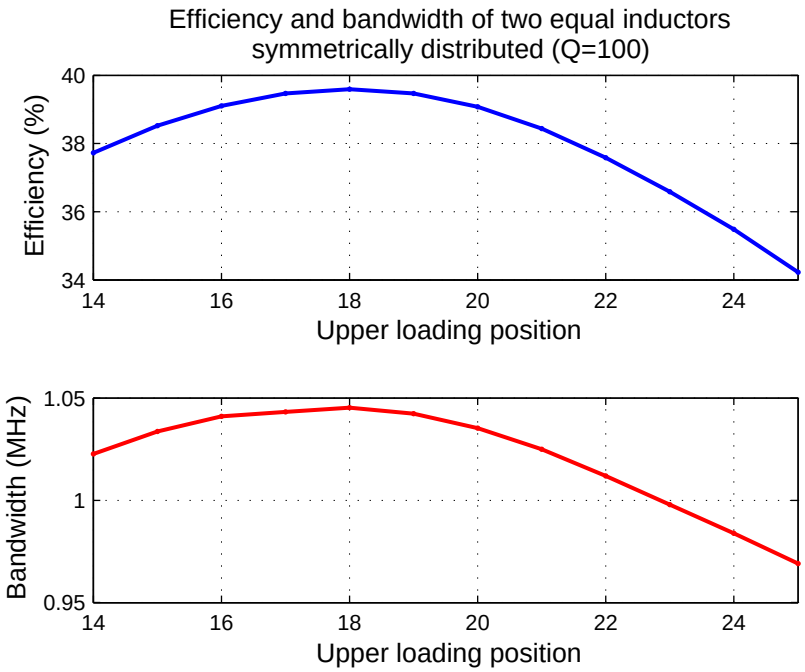

Figure 4.9. Efficiency and bandwidth for all possible symmetrically distributed loading sets with optimized inductance. The loading position describes the position of the top inductance. The bottom inductance is identical and located symmetrically from the center of the monopole.

- **Two Arbitrary Inductors**

  For the case where two inductors can be arbitrarily loaded on the monopole, the number of bits in the optimization increases to 30 where 5 bits are allocated to each position variable and 10 bits for each inductance value. This increase in sophistication significantly burden the load of computation, which translates into an increase of optimization time and slowing down of the convergence. For a Q factor of 100, the optimal simulated solution achieves the maximum efficiency and widest bandwidth with two inductors of 0.709 μH and 0.819 μH loaded at

positions #8 & #15 respectively. The resulting optimized performance indicates an efficiency of 39.6% and a bandwidth of 1.05 MHz. Comparing the performance of two inductors symmetrically or arbitrarily loaded on the monopole, it can be seen that the optimized performance are similar with slight improvement in efficiency, as show in Table 4.4. Nevertheless, this improvement is expected as the symmetrical solution is a sub-set of the arbitrary loading.

Table 4.4. *Performance comparison of two optimized inductors symmetrically or arbitrarily loaded.*

	Symmetric load	Arbitrary load
Positions	#8, #18	#8, #15
Inductance	0.849 μH, 0.849 μH	0.709 μH, 0.819 μH
Efficiency	39.4 %	39.6%
Bandwidth	1.05 MHz	1.05 MHz

- **Three or Four Arbitrary Inductors**

  Based on the achieved improvement from single inductor to two inductors configuration, a further step in complexity is considered by increasing the number of loading elements to three or four arbitrarily located inductors. Consequently, the optimization uses 45 bits and 60 bits respectively, and the computation time grows accordingly. Near optimal solutions are obtained when three inductors of 0.548 μH, 0.539 μH, and 0.579 μH are located at segments #4, #13, and #18 on the three-inductor monopole; and 0.426 μH, 0.319 μH, 0.383 μH, 0.568 μH inductors located at segments #3, #10, #14, #18 respectively on the four-inductor loaded monopole.

  For direct comparison, the optimal solutions and the performance corresponding to the four arbitrary loading schemes are listed in Table 4.5. In all loading scenarios, the antenna size reduction is successfully achieved using lumped inductors along the antenna wire structure. It can be observed that an increasing number of distributed inductors lead to a greater efficiency and a wider bandwidth. Nevertheless, it is also realized that the slope of the increase is becoming smaller when more inductors are added. To an extreme, a number of ten inductors are

arbitrarily applied and their position and inductance is optimized. The results shows an efficiency of 40.97%, a bandwidth of 1.068 MHz, and a total overall fitness of 147.8. This extreme case verifies the diminishing returns in both efficiency and bandwidth arising from the increase in loading inductor numbers, i.e. that a greater number of inductors on the monopole does not guarantee a significant improvement towards the antenna performance, as shown in Figure 4.10.

Table 4.5. *Performance comparison of the best solutions for different arbitrary loading schemes.*

	Single inductor	Two inductors	Three inductors	Four inductors
Positions	#9	#8, #15	#4, #10, #16	#3, #10, #14, #18
Inductance (μH)	1.142	0.709, 0.819	0.548, 539, 0.579	0.426, 0.319, 0.383, 0.568
Efficiency	37.8%	39.6 %	40.4%	40.7%
Bandwidth	1.012 MHz	1.050 MHz	1.058 MHz	1.063 MHz
Fitness	139.0	144.6	146.2	147.0

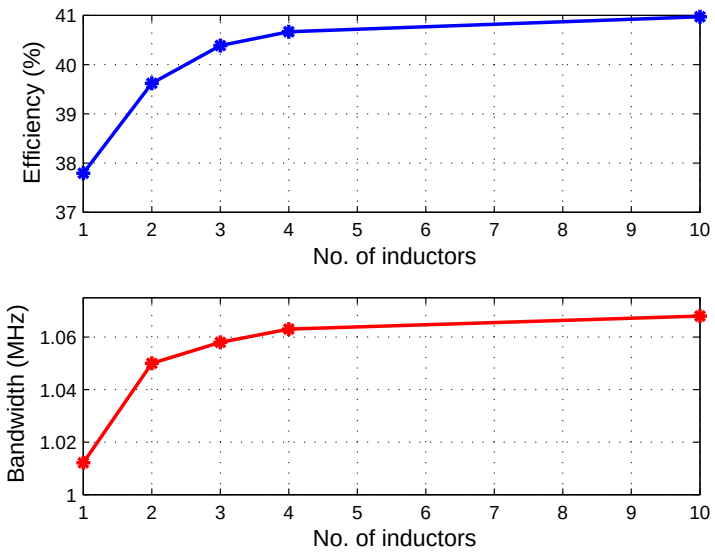

Figure 4.10. Optimized efficiency and bandwidth as a function of the number of loading inductors.

4.4 Conclusions

This chapter has described a strategy to miniaturize the size of HF & VHF wire antennas by using inductive loading along the antenna height. The optimization has been performed using a genetic algorithm implemented in MATNEC. Different loading schemes have been applied to a shortened monopole to verify the effectiveness of the procedure and its ability to find optimal solutions. The achieved optimized solution can effectively reduce the size of a resonant monopole while optimizing efficiency and bandwidth. The results, however, show diminishing return when a large number of inductors are used. It is clear that a different approach is required if further gain are to be achieved. It is possible that the structure of the inductance themselves might provide an avenue for further improvement. The following two chapters will move forward from ideal lumped inductors to realistic distributed inductive loadings.

Chapter 5

Case Study 2: Optimized Zig-Zag Wire Antennas

THIS chapter continues the investigation of inductive loading distribution on wire antennas, by extending the concept using lumped inductors to realistic inductive wire structures. Inductors are often realized with helical structures where the radius and pitch of the helix determine the inductance. As alternative, zig-zag shapes are representative of other structures that can also efficiently add distributed inductance. In contrast to the helix, the zig-zag geometry remains in one plane, which simplifies manufacture and opens the door to planar integration. In this chapter, structural variations of the zig-zag shape are explored to change the distribution of inductance along the monopole height and thus improve the antenna performance. These variations are represented using radial basis functions to lower the number of variables in the optimization process. The implementation that transforms the RBF to the actual zig-zag configuration is also introduced in this chapter. Evolutionary optimizers are applied to obtain the optimal shape variations based on electromagnetic simulations in MATNEC. Prototypes of the optimized zig-zag antennas are fabricated, measured, and compared to a uniform zig-zag monopole. Experiments indicate a good agreement to the optimal simulation results.

5.1 Introduction

HF and VHF antennas have been invented and investigated extensively for decades, and because of the long wavelength, size reduction has always been one of the critical issues. In the previous chapter, the case of lumped inductive loadings have been studied and their roles in shortening the resonant length of wire antennas have been demonstrated. The investigation has also indicated consequent degradation in both efficiency and bandwidth, which have been described and explained theoretically in [4, 59]. This chapter continues the investigation of inductance distribution by replacing the lumped inductances with wire windings. The helical structure is one of most common ways to build inductance, however it significantly increases the geometrical complexity as it is a three-dimensional structure. The zig-zag is another structure that can also add distributed inductance, but contrarily to the helix, it remains in one plane geometrically. This advantage of geometrical simplicity comes at the cost of a slight performance degradation compared to helical structure, but is balanced by the fact that zig-zag structures are amenable to planar integration, and may fit into applications where sufficient space is not given (for example, to implement an antenna into a gap).

Traditional uniform zig-zag antennas were introduced decades ago and had been used frequently since then [63–66]. The objectives of the investigation are to introduce and optimize modifications of the original zig-zag structure, with the aim of an effective improvement of radiation performance. Variations on the zig-zag structure can be made through changes in pitch angle along the length, or through variations in the numbers of zig-zag turns and in the length of the wire segments. In a straightforward approach, the structural variations of zig-zag antennas can be defined for each segment individually. This strategy, however, can lead to a large number of optimization variables, and thus greater complexity in the design process. Alternatively, the variations of qualities such as segment length along the zig-zag can be expanded in terms of radial basis functions. If selected appropriately, RBFs can provide effective description of parameter variations with only a few variables. Two types of radial basis functions,

namely Gaussian radial basis function (G-RBF) and multiquadric radial basis function (MQ-RBF), are employed in the present investigation.

5.1.1 Uniform Zig-Zag Antennas

Zig-zag antennas are of great importance in HF and VHF communications. This kind of antenna is strongly related to the helical antenna [64], but more importantly, the planar nature of the zig-zag wire structure makes it easier and faster to fabricate than helix. Therefore, the zig-zag antenna is a very convenient geometry to verify the small antenna design procedure that is proposed and applied in this thesis.

One of the configurations of zig-zag antennas that has been frequently investigated and extensively used is the uniform zig-zag monopole antenna [65]. The pitch angle (α) which defines the angle between two zig-zag segments, and the segment length (L), are the two essential parameters that fully characterize the structure, as shown in Figure 5.1.

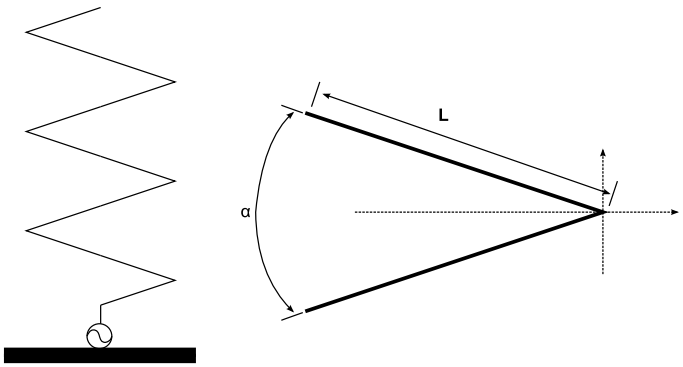

Figure 5.1. Illustration of a uniform zig-zag antenna with definition of the pitch angle (α) and segment length (L).

5.1.2 RBF Implementation

The optimization strategy of the investigation is based on the fact that the segments in the zig-zag antenna do not necessarily have to be all the same along the height. This

brings a great amount of possible variations to the structure, and thus changes the performance correspondingly. Considering the length of each segment in the zig-zag antenna as a variable can, however, significantly increase the complexity of optimization if being defined individually. The alternative is the use of curves that constrain the end points of the zig-zag and an expansion of these curves in terms of a limited number of radial basis functions. Length variations can then be described using a small number of variables. The G-RBF and MQ-RBF expansions used to represent the variations in the zig-zag structure can be written as:

$$f_{G-RBF} = \sum_{i=1}^{N} \omega_i \cdot e^{-R(x-c_i)^2} \tag{5.1}$$

where $N = 5$, $R = 250$, and

$$f_{MQ-RBF} = \sum_{i=1}^{N} \omega_i \cdot (1 + \varepsilon^2 (x - c_i)^2)^{\beta} \tag{5.2}$$

where $\varepsilon = 1$ and $\beta = -\frac{3}{2}$.

In the implementation of shape variations , the weighings ω_i in Equations 5.1 and 5.2 are altered while fixing the center height c_i. With a pre-defined number of segments (M) and a fixed pitch angle ($\alpha = 30^\circ$), the length of each segment (L_i) can be obtained from the RBFs. The accumulated zig-zag height is derived and then scaled to the physical height limitation (H_{max} = 0.25 m). The scaling factor is then applied to the RBF expansion to obtain an antenna that satisfies the length constraint. This process is carried out at each stage of the optimization in order to enforce the length constraint. Figure 5.2 illustrates how the zig-zag is implemented and constrained within the sum of G-RBFs.

Conventionally, a ten-segment (M = 10) zig-zag requires ten variables to describe its shape; whereas in the present investigation, a number of five (N = 5) uniformly spaced RBFs are used to describe the shape variations for both G-RBF and MQ-RBF zig-zag antennas. By limiting the number of variables and their optimizing ranges, the computational burden of the simulation and optimization is bounded. Importantly, increasing the number of zig-zag segments does not require increasing the number of required RBFs. By such artifice, the optimization of the zig-zag shape is successfully

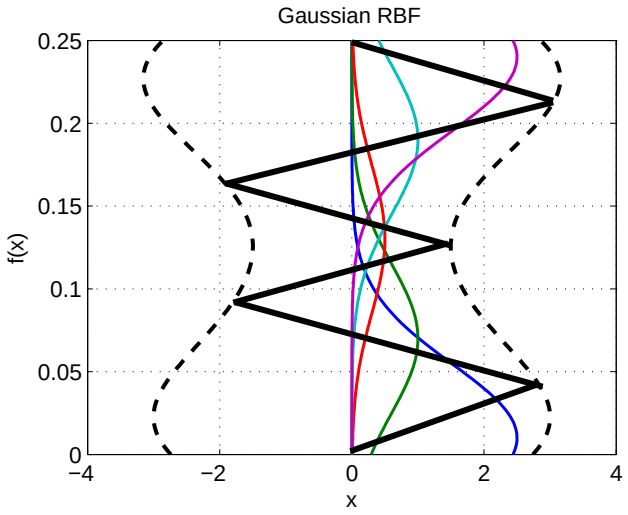

Figure 5.2. Illustration of how the G-RBF implementation constrains the zig-zag antennas. The individual RBFs are represented in color, their sum as black dashed line, and the resulting zig-zag antenna as solid black line.

transformed from optimizing the structural parameters into optimizing the RBFs variables that define and constraint the zig-zag shape.

5.2 Evolutionary Optimized Zig-Zag Structures

The goal of the optimization is an optimized zig-zag antenna with a fixed height of 0.25 m at the resonant frequency of 100 MHz, where maximized bandwidth and efficiency are considered as the performance indicators. A fixed pitch angle of 30° is applied in the zig-zag structure, while the number of segments and their individual length are the variables.

A uniform zig-zag monopole antenna with a height of 0.25 m and a fixed pitch angle of 30° is firstly constructed and simulated as reference. This uniform zig-zag antenna has 10 segments with equal length and a base vertical segment of 0.01 m. It is firstly optimized with respect to the segment length, and the result indicate an optimal value of 97.6 mm, as shown in Figure 5.3. As simulated with NEC-2, this uniform zig-zag antenna has a total efficiency of 81.2% and an operational bandwidth of 1.40 MHz with VSWR < 2.0 when perfectly matched at the target frequency 100 MHz. It is used as the reference for the comparison of optimized zig-zag antennas with varying segment

lengths.

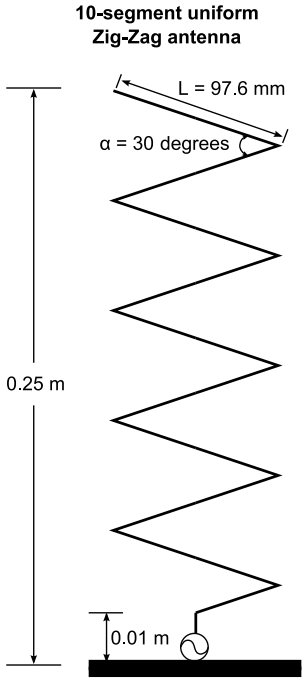

Figure 5.3. Illustration of a 10-segment uniform zig-zag antenna resonant at 100 MHz as the reference model.

5.2.1 G-RBF Zig-Zag Antenna

In the present investigation, the segment-varying zig-zag is described by five ($N = 5$) uniformly spaced RBFs. Therefore, the five weighting factors (ω_i) are the optimizing variables. The value $R = 250$ in Equation 5.1 is delicately selected so that each subset of the five RBFs is separated at an appropriate distance from its neighbors, and the total RBFs reach a balance of internal local impact and possible overall dynamic range. The resulting antenna height is normalized to the physical requirement ($H_{max} = 0.25$ m). The optimized values ω_i are obtained and the resulting curves representing the optimized zig-zag structure are shown in Figure 5.4. After the optimization, an L-section matching network is employed to provide perfect conjugate matching at 100 MHz. It can be observed in Table 5.1 that the optimized zig-zag antenna defined by G-RBF exhibits, in comparison to the uniform reference model, a greater efficiency of

88.2% and wider operational bandwidth of 1.75 MHz. Geometrically, the optimized G-RBF defined zig-zag antenna has comparatively longer segments near the top. More comments will be provided in conjunction with MQ-RBF case.

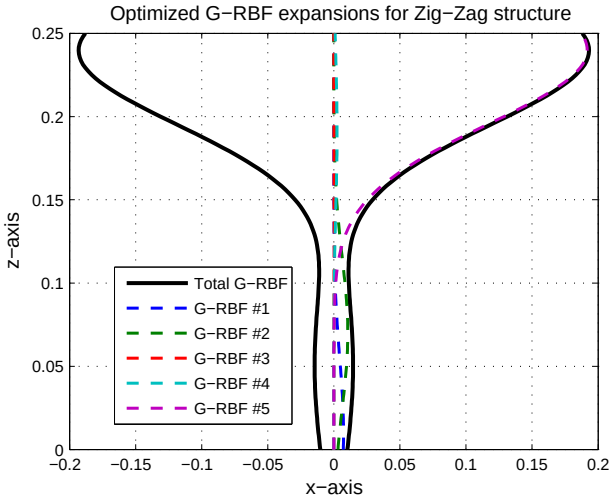

Figure 5.4. Optimized G-RBF curves to represent the zig-zag structure.

5.2.2 MQ-RBF Zig-Zag Antenna

Similarly, when MQ-RBFs are employed to describe the zig-zag shape variations, five weighting factors are optimized with finely tunes value of $\beta = -\frac{3}{2}$ and $\varepsilon = 1$. The optimal zig-zag shape for the antenna is obtained from the optimized MQ-RBF curves shown in Figure 5.5. This optimized antenna also has comparatively longer segment length near the top, as shown in Figure 5.6. It can be seen in Table 5.1 that it outperforms the uniform reference model with an efficiency of 88.3% and an operational bandwidth of 1.75 MHz.

The overall shape comparison of the uniform reference and the optimized models with varying segment lengths are plotted in Figure 5.6. It can be clearly seen that both G-RBF and MQ-RBF optimizations result in a very similar general antenna shape with near identical overall performance. It is also essential to point out that different values for R and β in Equations 5.1 and 5.2 respectively also affect the accuracy and might change the reached optimum of the results.

Table 5.1. *Comparison of uniform and optimized zig-zag antennas with segment length variations.*

10-Segments Uniform	Uniform	G-RBF	MQ-RBF
Segment Length	97.6 mm	N/A	N/A
ω_1	N/A	0.733	0.049
ω_2	N/A	1.056	0.352
ω_3	N/A	0.020	0.029
ω_4	N/A	0.244	0
ω_5	N/A	19.18	19.64
Total Wire Length	0.935 m	0.935 m	0.937 m
L in Matching	245 nH	137 nH	138 nH
C in Matching	112 pF	98.5 pF	97.5 pF
Efficiency with Matching	81.2%	88.2%	88.3%
Bandwidth with Matching	1.40 MHz	1.75 MHz	1.75 MHz
Overall Fitness	221.2	263.2	263.3

5.2.3 Convergence Comparison between G-RBF & MQ-RBF

Furthermore, the convergence of the optimizations for the two types of RBFs are compared with respect to the best individual and overall performances at the same GA optimization settings. For demonstration purpose, G-RBF and MQ-RBF optimizations with $R = 250$ and $\beta = -\frac{3}{2}$ respectively are repeated multiple times. The results show very similar convergence trend, explicitly, G-RBF appears to converge slightly faster than MQ-RBF in terms of total number of simulation runs. In contrast, MQ-RBF achieves a slightly higher total fitness, as shown in Figure 5.7.

5.3 Radiation Pattern of Optimized Zig-Zag Antennas

The optimal non-uniform zig-zag antennas proposed in the present investigation are designed for operation in normal mode, which is typically used for applications where reduced size and omnidirectional radiation are critical operational factors. Due to their small size (in terms of wavelength), the antennas are expected to have radiation

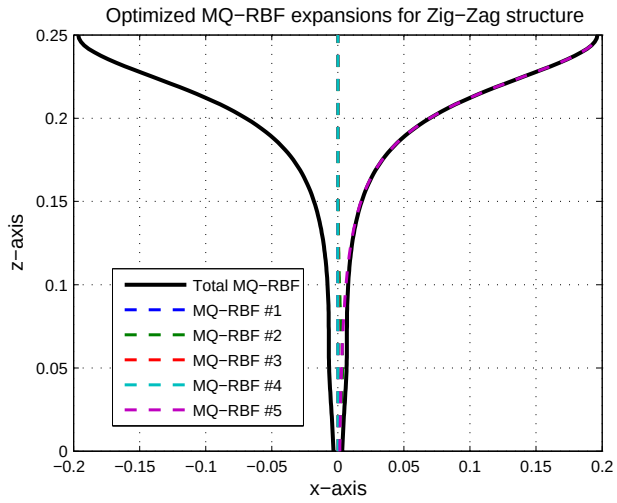

Figure 5.5. Optimized MQ-RBF curves to represent the zig-zag structure.

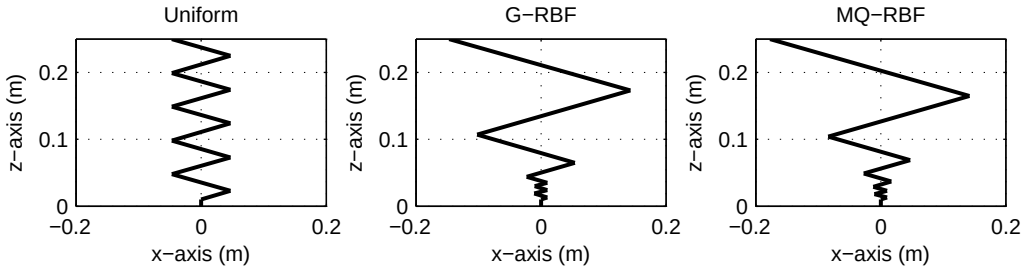

Figure 5.6. Configurations of optimized zig-zag antennas in comparison to the uniform reference.

patterns that are very similar to a short monopole on an infinite ground. For illustration purpose, the simulated radiation patterns of both co-polarization and cross-polarization for the uniform reference and optimized zig-zag antennas at 100 MHz are shown in Figure 5.8.

It can be observed that the three zig-zag antennas have very similar vertical gain (co-polarization), while the optimized models with expanding segment length near the top have greater horizontal gain (cross-polarization) compared to the uniform reference. Explicitly, the amount of cross-polarization appears to scale directly with the transverse extent of the antennas, but remain however generally very small. In particular, the uniform reference has a extreme small cross-polarization under -40 dB, that is similar to a monopole as expected. Comparatively, the other two optimized zig-zag

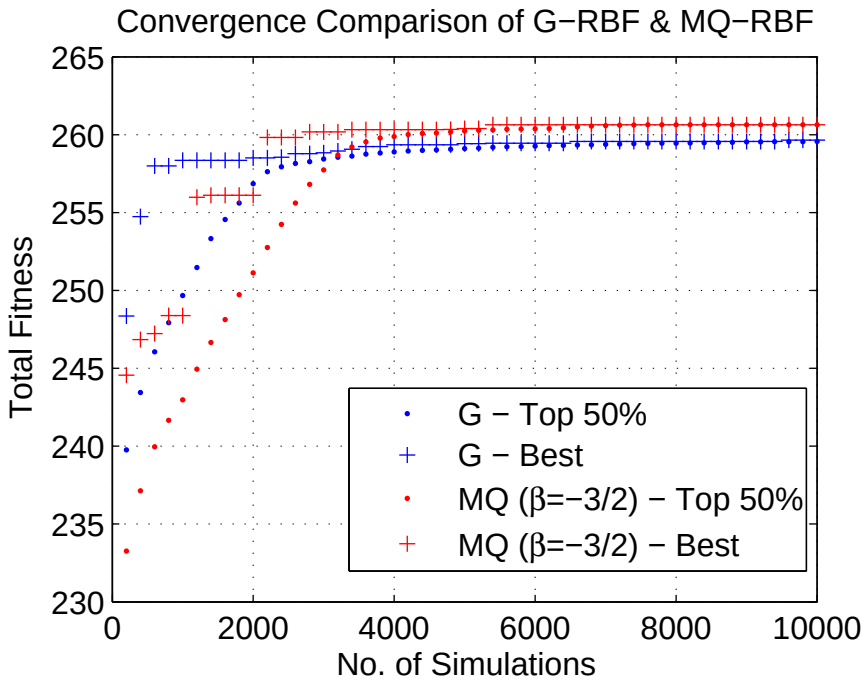

Figure 5.7. Convergence comparison of optimized zig-zag antennas using G-RBF ($R = 250$) and MQ-RBF ($\beta = -\frac{3}{2}$).

antennas with expanding geometrical dimension exhibit higher cross-polarized radiation. As expected, the introduction of a large segment at the top of the antenna introduces a small amount of axial mode radiation. Nevertheless, the cross-polarization level remains clearly more than 20 dB below the co-polarization level. In summary, the optimal zig-zag antennas still operate in normal mode with better performance using their optimized configurations.

5.4 Prototypes and Experimental Validation

For the purpose of experimental validation, the prototypes of the proposed zig-zag antennas have been fabricated according to the optimized parameters in Table 5.1. These fabricated antennas are constructed using aluminum wire with a radius of 1 mm. Comparatively, the uniform reference zig-zag has equal segment length of 97.6 mm, while the other two optimized antennas have varying segment length, ranging from 10 mm (the minimum segment length allowed in NEC simulation) to 330 mm. The three fabricated prototype antennas have a common and fixed pitch angle of 30° along their

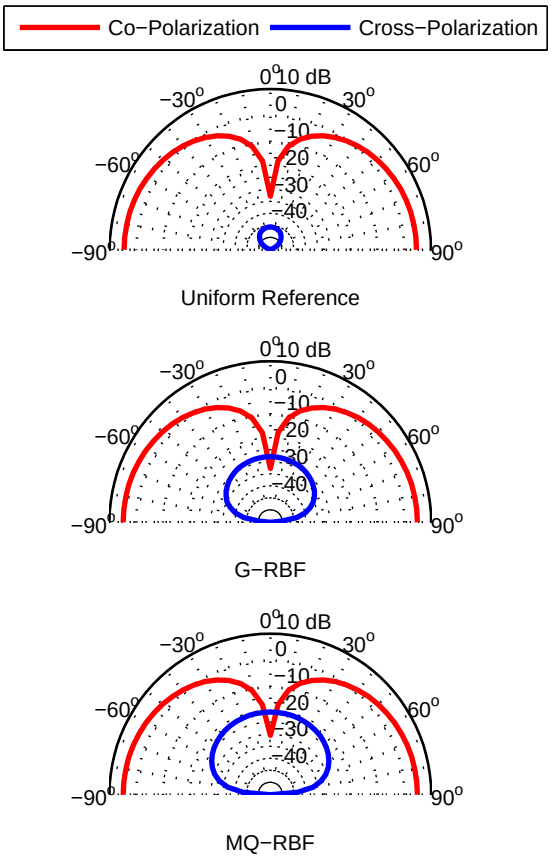

Figure 5.8. Gain patterns for the optimized zig-zag antennas at 100 MHz.

height, and they are shown in Figure 5.9. It is also noticed that there is a slight discrepancy between the originally designed antenna and the realized prototype, taking the form of a small geometrical height offset. Specifically, each fabricated zig-zag antenna is mounted with a 40 mm banana connector at its base whereas the original designs and simulations included only a 10 mm base wire segment. This difference gives the antenna an extra 30 mm wire length. Therefore, new simulations including the wire length offset have been performed and compared to the original simulations, as well as the measurement results. Theoretically, the 30 mm additional wire increases the original size (250 mm) by 12%, which should bring a 12% decrease in natural resonant frequency. Taking this into account, the measurements indicate a good agreement, as shown in Table 5.2.

The measurement has been carried out using an aluminum ground plane of 2.0 m $\times$ 2.0 m with a feeding point at its center, as shown in Figure 5.10. Both efficiency and

Table 5.2. Comparisons of Natural Resonance.

Natural Resonance	Uniform	G-RBF	MQ-RBF
Original Simulation	139 MHz	118 MHz	118 MHz
Prototype Measurement	114 MHz	100 MHz	100 MHz
Difference in Percentage	18%	15%	15%

operational bandwidth have been measured and the results have been obtained with an Agilent 814ET RF network analyzer.

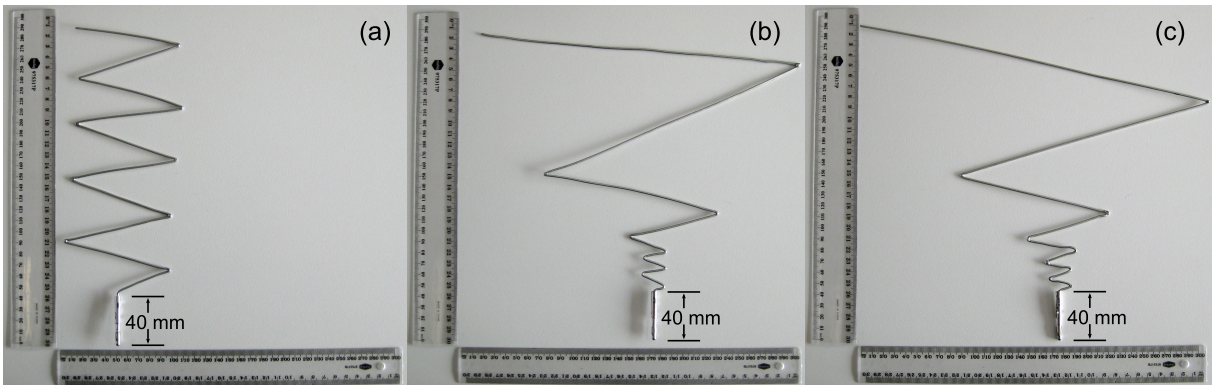

Figure 5.9. Prototypes of the proposed zig-zag antennas, (a) uniform, (b) G-RBF, (c) MQ-RBF, with indication of the 40 mm geometrical height offset due to the connectors

5.4.1 Efficiency Measurement

To measure the efficiency of the proposed antennas, the relative efficiency method is applied. The gain of an antenna is defined as the "ratio of the intensity, in a given direction, to the radiation intensity that would be obtained if the power accepted by the antenna were radiated isotropically" [33]. Generally, the relative gain is measured, which is defined as "the ratio of the power gain in a given direction to the power gain of a reference antenna in its referenced direction" [33]. In our measurement, a linear monopole antenna is used as the transmitting antennas, and each of the zig-zag antennas under test is functioning as the receiving antenna respectively. The receiving antenna and the transmitting antenna are aligned vertically at a fixed distance above the perfect ground plane. Thus, the power strength received at the zig-zag antenna is considered as the power received from the monopole antennas. Gain (G) can be

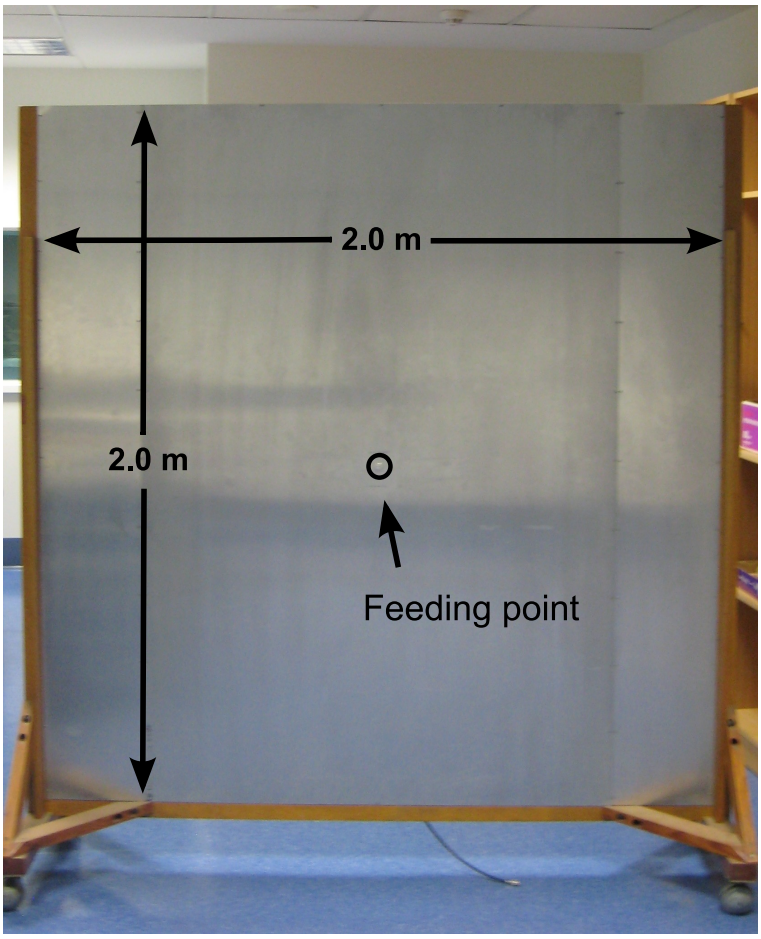

Figure 5.10. The aluminum ground plane used in the measurement.

interpreted the product of efficiency (η) and directivity (D), explicitly:

$$G = \eta \cdot D \tag{5.3}$$

Assuming that the proposed antennas have equal directivity, on the basis of their similar radiation pattern, the signal strength can be regarded as an indication of the relative antenna efficiency. This value is the normalized to one for the reference zig-zag antenna for comparison.

The results are summarized in Table 5.3. The original simulation indicates an efficiency of 81.2%, 88.2% and 88.3% for the uniform, G-RBF, and MQ-RBF zig-zag antennas respectively, which is equivalently a relative efficiency of 1.09 for the optimized zig-zag antennas to the uniform reference. The re-simulation including the base connector offset refines the results to 62.1%, 77.0% and 77.3%, equivalently a relative efficiency of

Table 5.3. Comparisons of original-simulated, re-simulated, and prototype-measured efficiency of three optimal antennas.

Efficiency	Uniform	G-RBF	MQ-RBF
Original Simulation (Norm.)	81.2% (1)	88.2% (1.09)	88.3% (1.09)
Re-Simulation (Norm.)	62.1% (1)	77.0% (1.24)	77.3% (1.24)
Normalized Efficiency from Prototype Measurement	1	1.17	1.17

1.24 for the nonuniform zig-zag antennas compared to the uniform reference. In comparison, the relative efficiency normalized to the uniform structure estimated from the prototype measurements shows a value of 1.17 for both optimized zig-zag structures. In general, a general qualitative agreement can be observed through the relative efficiency comparison, which effectively validates the achievement of the optimal structures.

A more accurate realization of efficiency measurement using a Wheeler-Cap will be presented in later chapters.

5.4.2 Bandwidth Measurement

In order to obtain the operational bandwidth of the proposed zig-zag antennas, a tunable L-section matching network is built and connected to each zig-zag antenna. It is, however, realized that the resulting VSWR may not be perfectly matched to 1.0 at 100 MHz in order to achieve the maximum operational bandwidth with VSWR < 2.0. Therefore, the variable inductor and capacitor used in the matching network are selected and finely tuned so that the maximum bandwidth can be obtained. The sample L-section matching network built for the bandwidth measurement is shown in Figure 5.11.

The operational bandwidths determined through simulations and measurements for each zig-zag antenna are shown in Table 5.4. It can be seen that original simulation

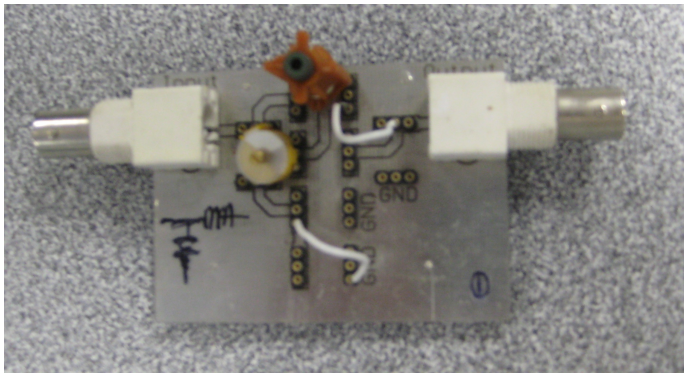

Figure 5.11. Sample of an L-section matching network.

predicts an operation bandwidth of 1.40 MHz and 1.75 MHz respectively for the uniform reference and optimized models, when perfectly matched at 100 MHz. The re-simulation further predicts an increase of the bandwidth to 1.90 MHz and 2.20 MHz because of extra wire length from the base connector. Comparatively, the prototype measurement shows results consistent to the re-simulation, with a bandwidth of 1.90 MHz for the uniform reference and 2.15 MHz for both optimized zig-zag models. The VSWR performance of the original simulation, re-simulation and prototype measurement is also plotted in Figure 5.12. It can be observed in the figure that the original simulation and re-simulation achieve a VSWR very much close to 1.0 at 100 MHz due to their perfect matching network. In contrast, the prototypes with non-perfect matching network can successfully achieve a wider or similar bandwidth with VSWR < 2.0, without necessarily reaching the minimum VSWR of 2.0.

Table 5.4. Comparisons of original-simulated, re-simulated, and prototype-measured bandwidth of three optimal antennas.

Bandwidth	Uniform	G-RBF	MQ-RBF
Original Simulation	1.40 MHz	1.75 MHz	1.75 MHz
Re-Simulation	1.90 MHz	2.20 MHz	2.20 MHz
Prototype Measurement	1.90 MHz	2.15 MHz	2.15 MHz

More importantly, the consistency and agreement between the simulation results and actual prototype measurement successfully validate the optimized zig-zag structures represented by radial basis functions. Consequently, this verifies the effectiveness and

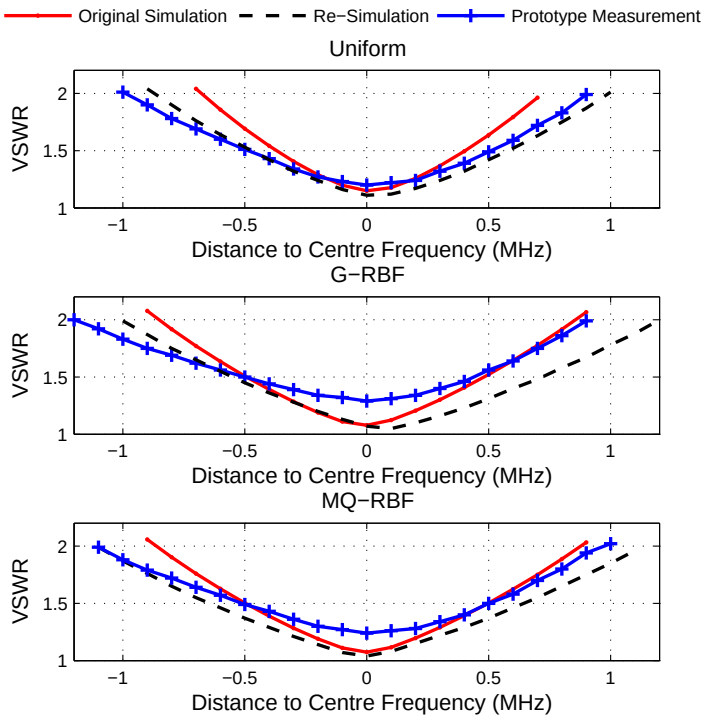

Figure 5.12. Comparisons of original-simulated, prototype-measured, and re-simulated VSWR of three optimal antennas.

accuracy of the antenna design method through electromagnetic simulation and evolutionary optimization.

5.5 Conclusions

This chapter has presented an optimization strategy for zig-zag antennas. The additional wire lengths, provided by a non-uniform zig-zag shape, provide an effective way to introduce distributed inductance along a wire antenna to effectively shorten it. By employing radial basis functions, sufficient structural variations can be made to the zig-zag antennas with fewer optimizing variables, in comparison to a fully parameterized geometrical description. Near-optimal zig-zag configurations are found by applying evolutionary optimization through electromagnetic simulation in MATNEC. The proposed optimal zig-zag antennas were fabricated and measured to obtain their relative efficiency and operational bandwidth. Experimental results indicate a fairly

good agreement between the simulation and measurement, which validates the investigation methodology. It was, however, realized in this chapter that zig-zag remains a planar structure with restricted structural variations and it is important to investigate how structural variation in a more complex form affect the final outcomes. Consequently, the next chapter will extend the investigation further to answer the above mentioned by investigating helical structures.

Chapter 6

Case Study 3: Optimized Helical Wire Antennas

THIS chapter extends the analysis of monopole antennas that are loaded with lumped inductance. In particular, we investigate the situation where inductance is distributed along the antenna by adding lateral structure in the form of helical windings. Section 6.2 describes the modeling of these helices in NEC and then section 6.3 describes optimizations where the lumped inductance of previous chapter is replaced by a realistic inductance consisting of a single layer helical winding. In section 6.4, this work is extended to consider antennas consisting entirely of helical winding. Windings with varying radius and pitch are considered and their construction optimized for efficiency and bandwidth. The antennas are compared with a reference uniform helix. In section 6.5, the advantage of the optimized structures is analyzed, including the implication for other factors such as current distribution and radiation fields. Section 6.6 investigates the impact of matching circuit loss on efficiency and bandwidth. Section 6.7 considers the sensitivity of the final design to the use of reduced optimization space (the use of RBFs). Finally, section 6.8 considers the experimental validation of the results.

6.1 Introduction

At low frequencies, the large size of antennas can be a major problem, especially for mobile and portable operation. The size reduction for wire antennas using lumped loading inductors is well known, but introduces significant degradation in efficiency and bandwidth. One of the alternative approaches to the problem is to increase the wire length through lateral structure while keeping the height fixed. Various techniques have been investigated to introduce longer wire, helical windings being one of the most popular. Helical windings reduce the wave speed on the antenna and hence the size of resonant structure. This approach however, results in antennas with narrow bandwidth in the normal mode.

The helical antenna was first proposed in [67], and has been developed successfully [4, 59, 68, 69] for decades, both theoretically and experimentally. It has gained wide usage as shortened resonant antenna for both HF and VHF communications. Nevertheless, most research on helical antennas has conventionally focused on uniform, conical or spherical helices [69–72]. Several authors have considered optimizing the helix geometry for antennas operating in the axial mode by varying the spacing, pitch angle and other features [73–76] for example.

This chapter, however, considers optimization of the geometry for antennas operating in the normal mode. This is appropriate for the target application since an omnidirectional radiation pattern is required for portable and mobile operation. The application of optimized antennas for HF frequencies have been considered, in [77] and at VHF frequencies in [78]. Such small antennas are characterized by a narrow bandwidth that decreases with size as implied by the Wheeler-Chu limit [4, 8]. Based on previous work [79, 80], this chapter investigates the design of non-uniform helical monopoles at HF and VHF frequencies.

In this chapter, short helical coils are first investigated in order to replace the lumped inductors introduced in Chapter 4 with something more realistic. The effectiveness of helical windings as inductive loadings on straight wire is successfully verified. The schemes to build the helices in MATNEC are described, and also the ability to change

their radius and pitch along the length of a helix. This allows us to optimize the antenna by continuously varying the helix pitch and radius along the length of the antenna. In a straightforward approach to the optimization, the radius and pitch are varied for each helical turn (or even short wire lengths). This, however, can lead to a prohibitive number of variables and greater complexity. To overcome this, the radius and pitch variations are described in terms of a limited number of parameters which are are the coefficients of a radial basis function expansion. With appropriate selection, RBF expansions can effectively provide sufficient variations using only a few variables. Gaussian RBFs are employed in an automated simulation and optimization tool, referred to as MATNEC.

For demonstration purposes, the objective is to optimize the design of 25, 50 and 100 MHz helical antennas with a fixed height of 0.25 m. The aim is the design of antennas operating in normal mode with maximized efficiency and bandwidth, under perfect matching. The ultimate goal of the optimization is to find the best radius and pitch configurations as represented by appropriate RBF expansions. Clearly, there is a compromise between efficiency and bandwidth, and so an important part of the work is the choice of fitness function to be optimized. The fitness function in the investigation takes into account both the efficiency and bandwidth performance. This multi-objective optimization procedure requires suitable trade-offs in determining the optimal solutions.

Using the above techniques, optimal helical configurations are found with just radius varying, then just pitch varying, and then both radius and pitch varying. The adjustment of pitch alone is particularly useful for applications where a large radius is impractical (a portable radio for example). All the optimized results are compared to an optimized uniform reference helix and radiation patterns are simulated in order to check performance in the normal mode. To validate the optimal solutions, prototypes of the proposed antennas at 100 MHz were fabricated and measured. The results successfully verify that helical antennas can be optimized to perform with wider bandwidth while maintaining good efficiency.

6.2 Helix Winding Approximation

In NEC-2, a helix or a spiral can be constructed using a 'GW' card with specified geometrical coordinates. This card can successfully create a helix or a spiral at a constant pitch, and a fixed rate of change of radius. This topology can effectively solve the problem with uniform helices, but will fail when the helix is designed to vary its shape continuously and non-linearly. More freedom is required in the current application and the helix needs to be defined in terms of a set of segments via suitable parametric variables.

Normally, a helix is made from curved conducting wires with a round cross-section. It is approximated as a collection of head to tail segments, i.e. by a stack of polygons, with the number of sides increased to obtain a design that is closer to a perfect helix, as illustrated in Figure 6.1. It should be noted, however, that too many polygon sides will result in an unacceptable burden in NEC-2.

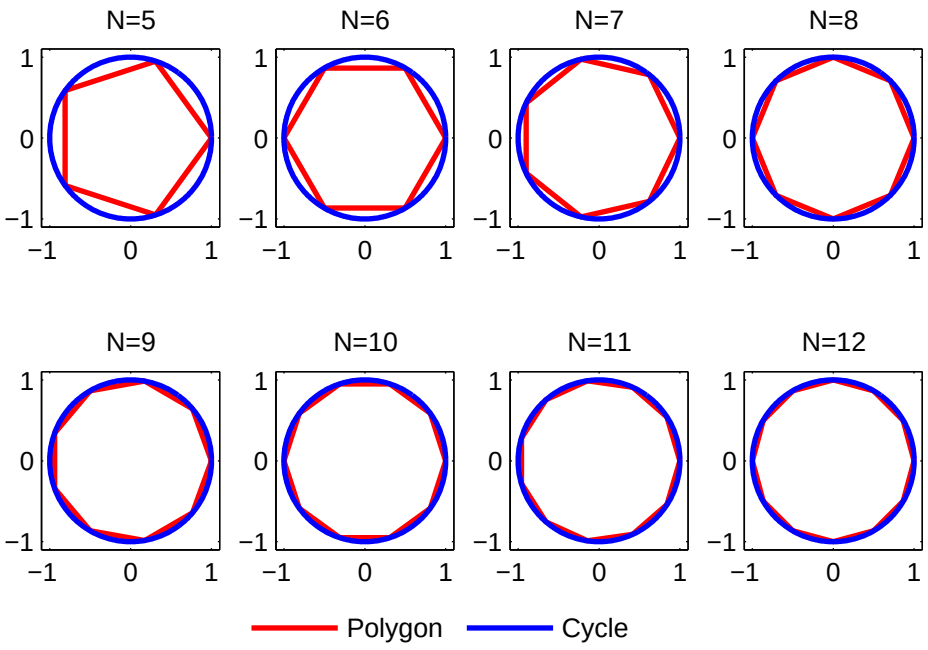

Figure 6.1. Illustration of shape variation between perfect cycle and polygons with different numbers of side (N)

6.2.1 Helix with Fixed Radius

When a helix is constructed with fixed radius along its length, the cross-section for each turn is identical. Taking an octagon (N=8) with unit radius (R=1) as an example, its cross-section is shown Figure 6.2, with the coordinates at each corner indicated in Table 6.1 .

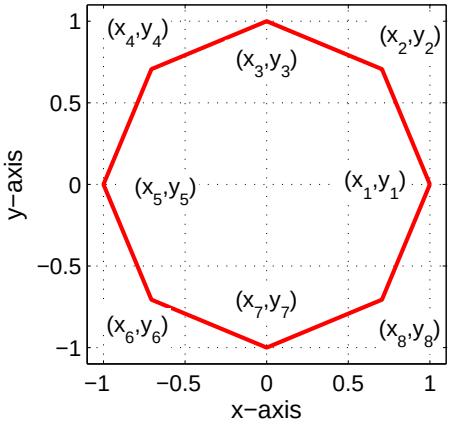

Figure 6.2. Illustration of an octagon with unit radius

Table 6.1. *Coordinates of an octagon with unit radius.*

x_i	x-coordinate	y_i	y-coordinate
x_1	$\cos(0 \times \frac{2\pi}{8})$	y_1	$\sin(0 \times \frac{2\pi}{8})$
x_2	$\cos(1 \times \frac{2\pi}{8})$	y_2	$\sin(1 \times \frac{2\pi}{8})$
x_3	$\cos(2 \times \frac{2\pi}{8})$	y_3	$\sin(2 \times \frac{2\pi}{8})$
x_4	$\cos(3 \times \frac{2\pi}{8})$	y_4	$\sin(3 \times \frac{2\pi}{8})$
x_5	$\cos(4 \times \frac{2\pi}{8})$	y_5	$\sin(4 \times \frac{2\pi}{8})$
x_6	$\cos(5 \times \frac{2\pi}{8})$	y_6	$\sin(5 \times \frac{2\pi}{8})$
x_7	$\cos(6 \times \frac{2\pi}{8})$	y_7	$\sin(6 \times \frac{2\pi}{8})$
x_8	$\cos(7 \times \frac{2\pi}{8})$	y_8	$\sin(7 \times \frac{2\pi}{8})$

Explicitly, the x and y coordinates of a polygon with a number of N sides and a uniform radius of R can be written as:

$$\begin{aligned} x_i &= R \cdot \cos((i-1) \times \tfrac{2\pi}{N}) \\ y_i &= R \cdot \sin((i-1) \times \tfrac{2\pi}{N}) \end{aligned} \tag{6.1}$$

Therefore, the x and y coordinates of each corner remain the same for each turn, while the z coordinates change at a certain rate. For instance, a 5-turn uniform helix with a length of 1 m will correspond to a pitch of 0.2 m, as shown in Figure 6.3.

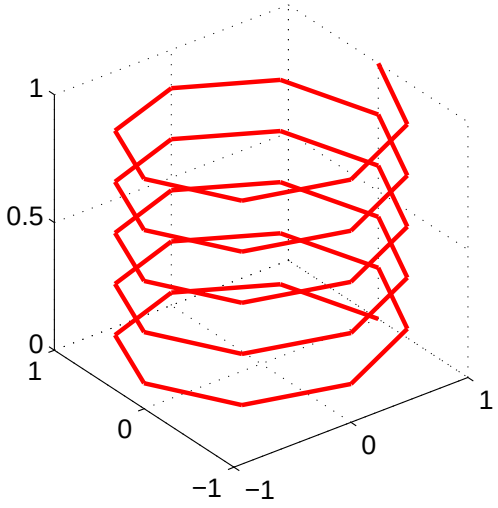

Figure 6.3. Illustration of a 5-turn uniform octagon helix with unit radius.

6.2.2 Helix with Radius Change

In comparison, when a spiral coil with radius change along its length is constructed, the implementation is much more complicated. The radius will differ not only between turns, but also between segments in each turn. Taking a 1-turn octagon spiral as an example, if the starting radius is R_1, and the ending radius is R_2, the radius increment (R_{in}) can be interpreted as:

$$R_{inc} = (R_2 - R_1)/8 \tag{6.2}$$

Thus, the x and y coordinates of the eight corners can be represented as shown in Figure 6.4 and Table 6.2. Therefore, the x and y coordinates of a polygon with a number of N sides and linearly-changing radius between R_1 and R_2 can be written as:

$$\begin{aligned} x_i &= (R_1 + (i-1) \times R_{inc}) \cdot \cos((i-1) \times \tfrac{2\pi}{N}) \\ y_i &= (R_1 + (i-1) \times R_{inc}) \cdot \sin((i-1) \times \tfrac{2\pi}{N}) \end{aligned} \tag{6.3}$$

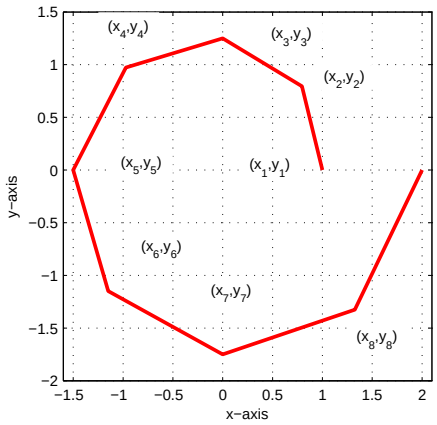

Figure 6.4. Illustration of an octagon with changing radius.

Table 6.2. *Coordinates of an octagon with changing radius.*

x_i	x-coordinate	y_i	y-coordinate
x_1	$(R_1 + 0 \times R_{inc}) \cdot \cos(0 \times \frac{2\pi}{8})$	y_1	$(R_1 + 0 \times R_{inc}) \cdot \sin(0 \times \frac{2\pi}{8})$
x_2	$(R_1 + 1 \times R_{inc}) \cdot \cos(1 \times \frac{2\pi}{8})$	y_2	$(R_1 + 1 \times R_{inc}) \cdot \sin(1 \times \frac{2\pi}{8})$
x_3	$(R_1 + 2 \times R_{inc}) \cdot \cos(2 \times \frac{2\pi}{8})$	y_3	$(R_1 + 2 \times R_{inc}) \cdot \sin(2 \times \frac{2\pi}{8})$
x_4	$(R_1 + 3 \times R_{inc}) \cdot \cos(3 \times \frac{2\pi}{8})$	y_4	$(R_1 + 3 \times R_{inc}) \cdot \sin(3 \times \frac{2\pi}{8})$
x_5	$(R_1 + 4 \times R_{inc}) \cdot \cos(4 \times \frac{2\pi}{8})$	y_5	$(R_1 + 4 \times R_{inc}) \cdot \sin(4 \times \frac{2\pi}{8})$
x_6	$(R_1 + 5 \times R_{inc}) \cdot \cos(5 \times \frac{2\pi}{8})$	y_6	$(R_1 + 5 \times R_{inc}) \cdot \sin(5 \times \frac{2\pi}{8})$
x_7	$(R_1 + 6 \times R_{inc}) \cdot \cos(6 \times \frac{2\pi}{8})$	y_7	$(R_1 + 6 \times R_{inc}) \cdot \sin(6 \times \frac{2\pi}{8})$
x_8	$(R_1 + 7 \times R_{inc}) \cdot \cos(7 \times \frac{2\pi}{8})$	y_8	$(R_1 + 7 \times R_{inc}) \cdot \sin(7 \times \frac{2\pi}{8})$

with $R_{inc} = (R_2 - R_1)/N$.

Therefore, the x and y coordinates change according to the radius change along its length, while the z coordinates changes at certain rate. For instance, a 5-turn spiral with starting and ending of radii R_1 and R_2 of 1 m and 2 m respectively, and a total length of 1 m, is shown in Figure 6.5.

Nevertheless, it is also noticed that when the helix radius is varying significantly along its length, the closeness between the helix and polygon approximation can vary dramatically. Thus, polygons with a fixed number of sides may not effectively represent the entire spiral structure. Take a 5-turn spiral with a radius ranging from 5 mm to 50 mm as an example, due to the limitation in NEC-2 that the minimum segment length

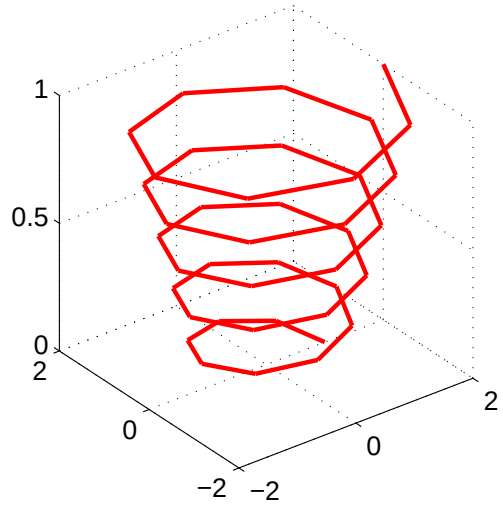

Figure 6.5. Illustration of a 5-turn octagon spiral with linearly-changing radius.

(L_{Seg}) is at least eight times the wire radius which is set to 0.5 mm in this investigation, the minimum length is thus 4 mm. For each turn, the starting and ending radii are indicated in Table 6.3. Thus, if each turn has 6 segments, the resulting approximation is shown in Figure 6.6.

Table 6.3. *Starting and ending radii for each turn on the spiral.*

	Turn 1	Turn 2	Turn 3	Turn 4	Turn 5
Starting radius	5 mm	14 mm	23 mm	32 mm	41 mm
Ending radius	14 mm	23 mm	32 mm	41 mm	50 mm

It can be observed that at the bottom of the spiral, the radius is comparatively small and the hexagon can fairly well represent the helix. However, with the increase of radius the approximation becomes quire crude. One solution is to represent each turn of the spiral with the number of polygon sides varying from turn to turn ($N = \frac{2\pi \times R}{L_{Seg}}$). Explicitly, the number of polygon sides for each turn is based on the starting radius (R), as shown in Table 6.4. The top view and 3D view of this spiral with different types of polygons for each turn is shown in Figure 6.7, and yields a much better approximation.

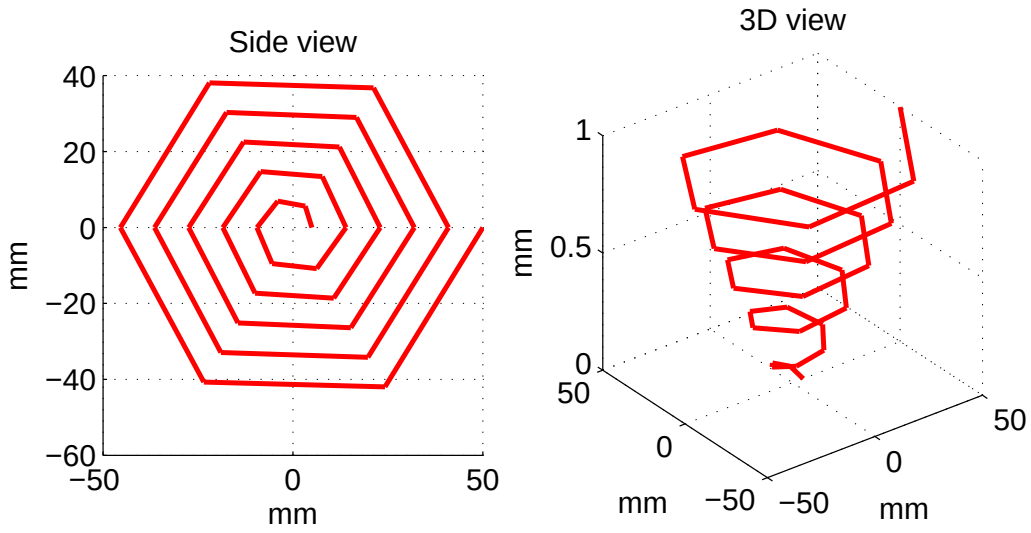

Figure 6.6. Illustration of a 5-turn spiral composed of hexagons.

Table 6.4. *Number of polygon sides for each turn on the spiral.*

	Turn 1	Turn 2	Turn 3	Turn 4	Turn 5
Starting radius	5 mm	14 mm	23 mm	32 mm	41 mm
Number of polygon sides	6	17	28	40	51

6.2.3 Other Considerations and Refinement

There are some other aspects that need to be considered when modeling a helix in NEC-2. These include the size of each side for the polygon approximation and how the helix is connected to the ground plane and other parts of the antenna structure.

- **Number of segments on each straight wire**

  In NEC-2 simulation, straight wires are separated into a number of segments and the method of moments is employed to obtain accurate results. Theoretically, the smaller the discretization segments, the more accurate the result will be, but with an increased cost of simulation time. More importantly, one essential rule strictly applied in NEC-2 is that the minimum segment length should be as eight times the wire radius. Explicitly, once the wire radius is defined (set to 1 mm in this thesis), the minimum segment length (L_{Seg}^{min}) will be set (8 mm). In MATNEC, the

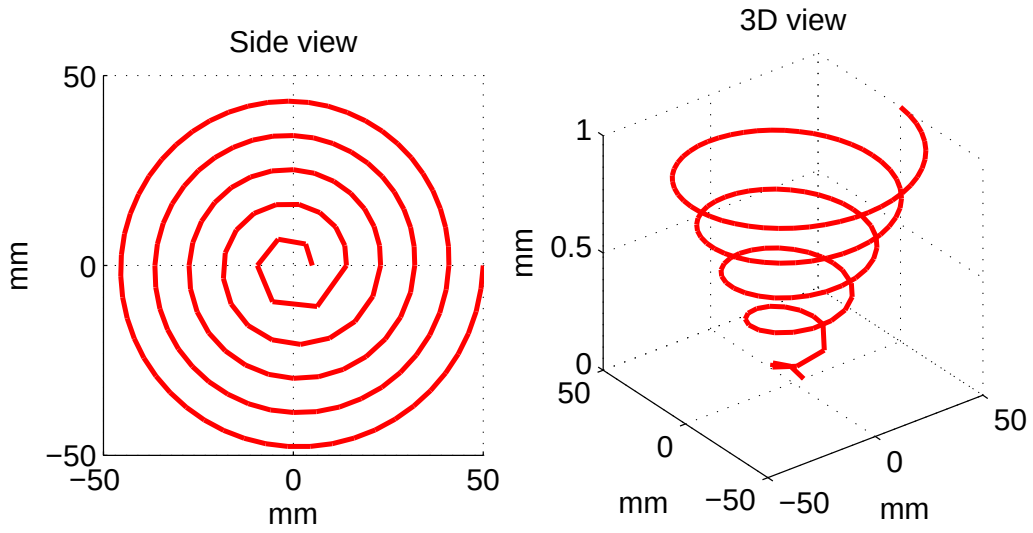

Figure 6.7. Illustration of a 5-turn spiral composed of different types of polygons.

script first calculates the length of each piece of straight wire (L), and applies the minimum segment length and then rounds down to the nearest integer number of segments ($N = Round(\frac{L}{L_{Seg}^{min}})$). This number ($N$) is regarded as the number of segments that a piece of wire is to be segmented into $L_{Seg} = \frac{L}{N}$.

- **Helix connection at bottom and top**

  Another essential issue to be considered is the connection of the two ends of a helix section to the other geometry (usually a straight wire or a ground plane). In this investigation, two horizontal pieces of wires are used to connect the end points at the bottom and top to a wire on the axis of the helix, as illustrated in Figure 6.8. The helix is then connected to the ground through a wire on this axis.

6.3 Straight Wires with a Single Helical Loading Coil

It is well known that inductive loadings can effectively shorten the resonant length of wire antennas. Previous investigations have successfully verified that both lumped inductive components and zig-zag wire structure can be employed

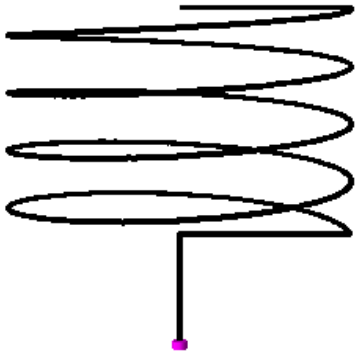

Figure 6.8. Illustration of helix with horizonal connection wires at two ends.

to load wire antennas to reduce the size. In this section, we consider a more realistic inductive loading.

In this investigation, a single uniform helical coil is defined by four parameters, namely its radius (r), its height above the ground (h), the helix length (l), and the pitch (p) (equivalently represented as the number of turns 't'). The structure of the coil is illustrated in Figure 6.9. Those four parameters are taken as the variables in MATNEC, and are optimized through both GA and PSO techniques. The objectives of the investigation is to obtain the optimal helix structure, which is incorporated as part of a monopole antenna with a fixed length of 0.25m. The antenna is optimized for maximum efficiency and operational bandwidth.

Long single-layer coils wound in a helix are often called solenoids, and the inductance of the coil is given in [81]:

$$L = \frac{\mu_0 K N^2 A}{l} \tag{6.4}$$

where $\mu_0 = 4\pi \times 10^{-7}$ H/m is the permeability of free space, K is the Nagaoka coefficient, N is the number of turns, A is the cross-sectional area in square meters, and l is the length of the coil. Alternatively, if the the coil is shorter, and the core is of non-ferromagnetic material, the inductance can be calculated from:

$$L = \frac{\mu_0 K N^2 A}{l + 0.45d} \tag{6.5}$$

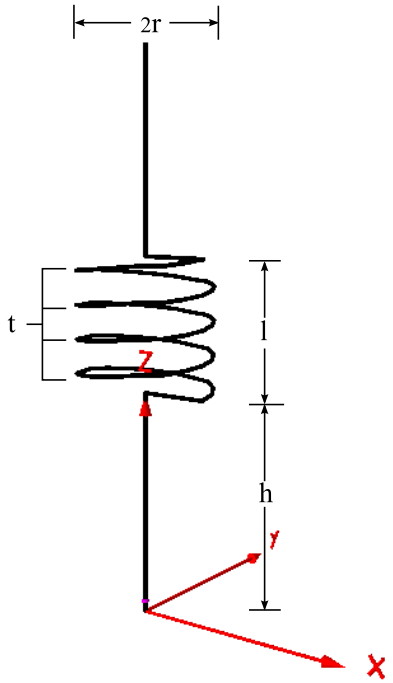

Figure 6.9. Illustration of a uniform helical coil on a monopole.

6.3.1 Optimized Single Coil Using GA

GA is applied in MATNEC to pursue the optimal antenna including the loading coil structure and position on the monopole. The task is very similar to the previous investigation using a lumped inductor on 0.25 m monopole. We now, however, maximize performance at 100 MHz with lumped inductance replaced by uniform helical coil.

The four variables, shown in Figure 6.9, are limited within well-defined ranges and this limits the search space which reduces the optimization burden. Since the monopole is 0.25 m long, the two variable l and h are limited to a total of 0.25 m. The radius of the coil is limited between 0.010 m and 0.030 m, and the number of turns is bounded between 4 and 8. In the GA optimization, a number of 6 bits is used for each variable, the population size is 60, and it runs for 50 iterations. The boundary settings and the GA settings are shown in Table 6.5. The fitness function employed in the optimization can be expressed as:

$$Fitness = 1 \times Efficiency(\%) + 100 \times Bandwidth(MHz) \tag{6.6}$$

Table 6.5. *Boundary and GA settings for optimal single coil on the 0.25 m monopole.*

Variables	Ranges / Settings
Radius (r)	[0.01 m - 0.03 m]
Height (h)	[0.01 m - 0.20 m]
Length (l)	[0.01 m - 0.20 m]
No. of turns (t)	[4 - 8]
No. of bits per variable	6
Population size	60
Iterations	50
Mutation rate	10%

Table 6.6. *Optimized single coil on monopole and performance.*

Radius (r)	0.026 m
Height (h)	0.183 m
Length (l)	0.057 m
No. of turns (t)	4
Radiation efficiency	98.9%
Operational bandwidth	1.487 MHz
Overall fitness	247.6

Near-optimal solutions are obtained for a 4-turn coil with radius of 0.026 m, total length of 0.057 m, and located at 0.183 m high above the base, as shown in Figure 6.10. This optimal antenna configuration exhibits a bandwidth (VSWR $<$ 2.0) of 1.487 MHz, a radiation efficiency of 98.9%, and an overall fitness of 247.6, as listed in Table 6.6. It will be observed this monopole has its inductive coil near the top, similar to the 'top-hat' configuration.

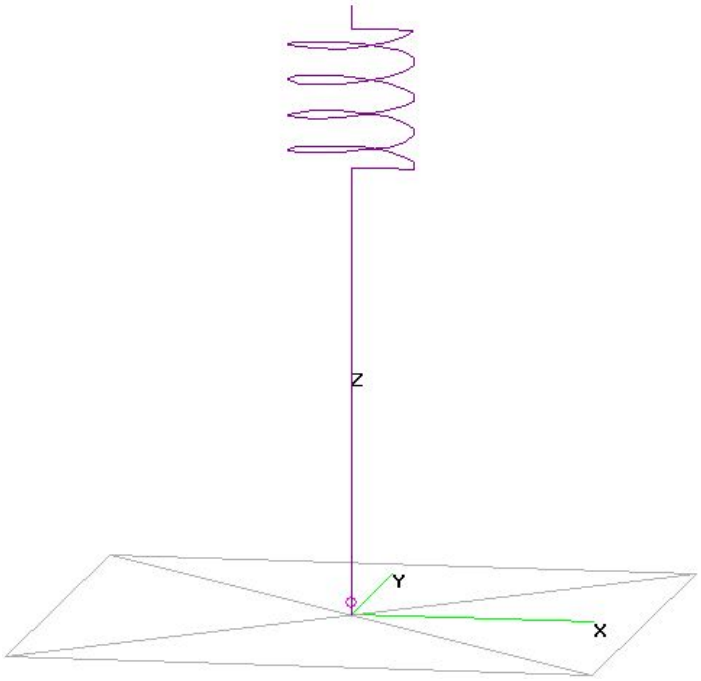

Figure 6.10. Illustration of a uniform helical coil on a monopole.

6.3.2 Optimized Single Coil Using PSO

Similarly, the same monopole with loading is optimized using PSO. Each of the four variables are taken as a swarm with 100 particles. The boundary settings remain the same, and the PSO settings are shown in Table 6.7.

The optimal results occur after 350 iterations, as the particles move around in a fairly narrow region. It can be observed in Figure 6.11 that most of the particles have their optimized radius between 0.025 m and 0.0265 m, optimized coil height between 0.165 m and 0.180 m, optimized coil length between 0.075 m and 0.09 m, and optimized number of turns for either 4 or 5. Explicitly, the PSO does provide a group of outperforming particles which will result in greatest overall performance in terms of both efficiency and bandwidth.

In comparison, GA and PSO both converge to solutions that share the similar structural characteristics, as shown in Table 6.8.

Table 6.7. *Boundary and PSO settings for optimal single coil on the 0.25 m monopole.*

Variables	Ranges / Settings
Radius (r)	[0.01 m - 0.03 m]
Height (h)	[0.01 m - 0.20 m]
Length (l)	[0.01 m - 0.20 m]
No. of turns (t)	[4 - 8]
No. particles in a swarm	100
Iterations	400
Time interval	0.01
Inertial weight	0.7
Scaling factors c_1, c_2	0.2, 2

Table 6.8. *Comparison of optimized solution of single coil on 0.25 m monopole between GA and PSO .*

	GA	PSO
Radius (r)	0.026 m	[0.025 m - 0.0265 m]
Height (h)	0.183 m	[0.165 m - 0.180 m]
Length (l)	0.057 m	[0.075 m - 0.090 m]
No. of turns (t)	4	[4 - 5]

6.4 Helices with Non-linear Structural Variations

Inductance can also be distributed along the antenna to shorten its length, usually in the form of a helical winding. In this section, non-linear structural variations are introduced to the helices and these are optimized using GA techniques. The optimization is considered at three different frequencies (25, 50, and 100 MHz). All antennas are consistently constrained to be 0.25 m high, which corresponds to $\frac{1}{48}\lambda$, $\frac{1}{24}\lambda$, and $\frac{1}{12}\lambda$ at those frequencies. L-type matching networks are included in the design with Q factors of 100 and 1000 for the inductance and capacitance, respectively. The total efficiency and operational bandwidth are considered as the performance indicators. They are

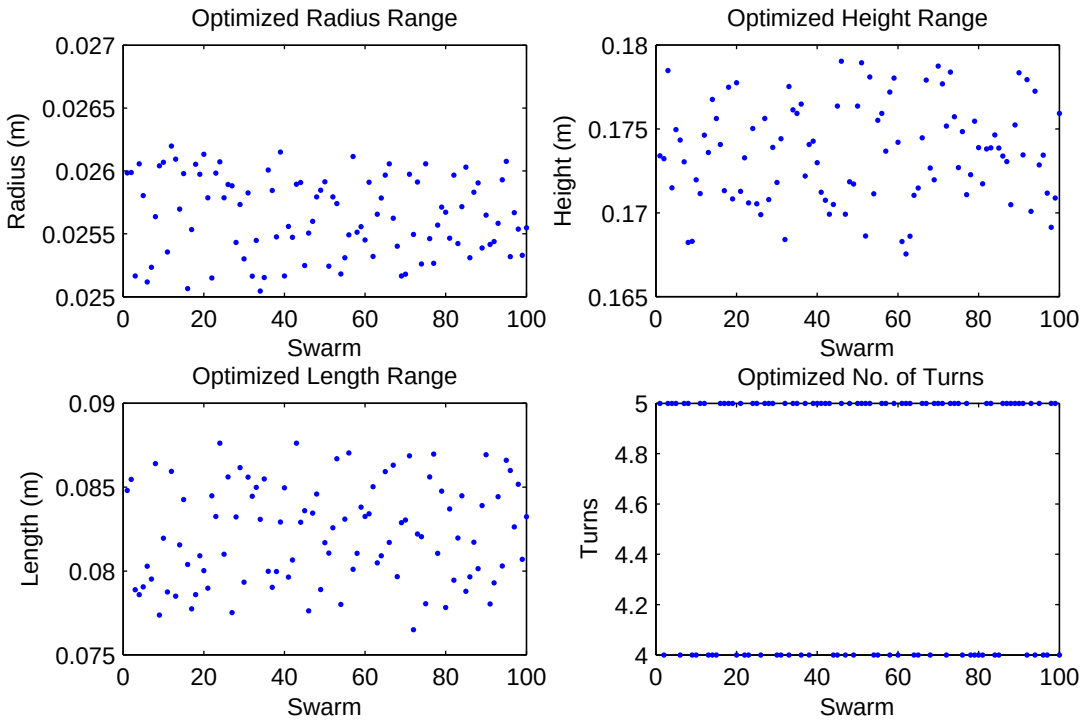

Figure 6.11. PSO convergence of single uniform coil on 0.25 m monopole.

optimized with respect to the pre-defined fitness function:

$$Fitness = Efficiency(\%) + w \times Bandwidth(MHz) \tag{6.7}$$

where the weight ratio (w) balances the compromise between the performance requirements. With suitably chosen w, this formula gives a good compromise between efficiency and bandwidth, and thus is used for all optimizations in this investigation.

Uniform cylindrical-shaped helices are used as the reference for comparison purpose, and antenna structure is optimized in terms of radius and/or pitch.

6.4.1 Uniform Helix without Structural Variations

MATNEC parametric simulations are performed for a variety of uniform helices made out of 1 mm-diameter copper wire with radius (r) between 0.01 and 0.20 m, and pitch (p) between 0.10 and 0.70 m. The radius sweeping range constrains the horizonal dimension at a similar level to the vertical dimension, and the pitch range limits the helices to have at least three winding turns. The most efficient uniform resonant antennas

Table 6.9. *Optimized uniform helices for reference at 25, 50, and 100 MHz .*

Uniform reference	25 MHz	50 MHz	100 MHz
Radius (r)	0.10 m	0.07 m	0.04 m
Pitch (p)	0.07 m	0.07 m	0.07 m
Wire length	1.93 m	1.34 m	0.81 m
Efficiency	48.1%	77.1%	77.9%
Bandwidth	0.025 MHz	0.17 MHz	1.50 MHz
Overall fitness	73.1 ($w = 1000$)	94.1 ($w = 100$)	227.9 ($w = 100$)

for frequencies 25, 50, and 100 MHz, their structural features, performance, and configurations are illustrated in Table 6.9, side view in Figure 6.12 and 3D view in Figure 6.13 respectively.

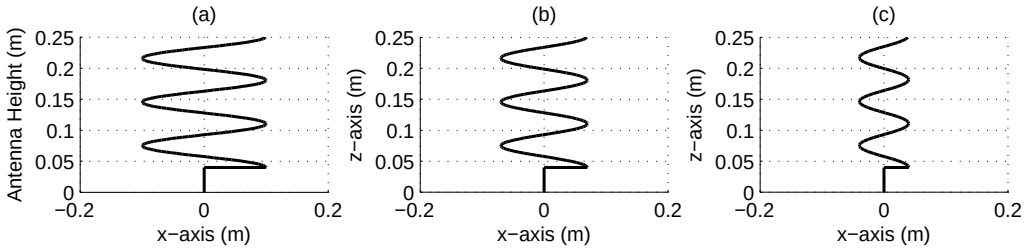

Figure 6.12. Side view of optimized configurations of uniform helices at (a) 25 MHz, (b) 50 MHz, and (c) 100 MHz.

It can be seen that the three optimized uniform helical antennas have constant pitch of 0.07 m, and radius of 0.10 m, 0.07 m, 0.04 m respectively, which contribute to a total wire length of 1.93 m, 1.34 m and 0.81 m. At 25 MHz, it exhibits an efficiency of 48.1%, and an operational bandwidth of 0.025 MHz. The performance improves at 50 MHz and 100 MHz, with the efficiency increasing to 77.1% and 77.9%, while a bandwidth of 0.17 MHz and 1.50 MHz can be obtained with appropriate matching networks. The performance of the three sets of uniform helical antenna will be used as the reference, in order to verify the success of optimized helical antennas having structural variations along their axis.

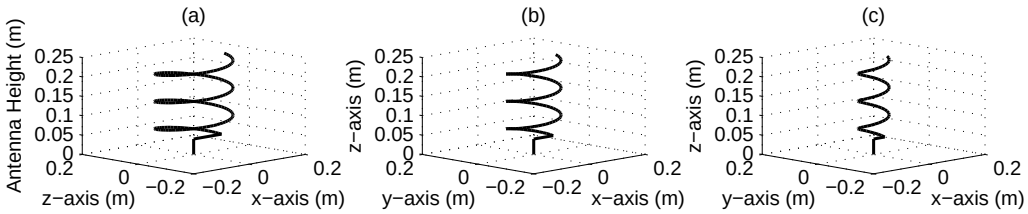

Figure 6.13. 3D view of optimized configurations of uniform helices at (a) 25 MHz, (b) 50 MHz, and (c) 100 MHz.

6.4.2 Non-Linear Structural Variations in Radius

When the radius feature is allowed to vary along the antenna height, a segment-to-segment description would need a prohibitive number of variables in the optimization. The solution is to use RBFs to describe the variation of radius. Specifically, a set of five Gaussian RBFs with a values of $A = 50\ 1/m^2$ are used to represent the radius variation, and the five assigning weights (w_i) are allowed to vary from 0.01 to 0.10 m in 6.8. The value A is finely tuned to yield a compromise between independence and overlapping between two adjacent RBFs. Once the number of turns is defined, the radius at the beginning and end of each turn is obtained from the RBF expansion, the radius of the segments is then set to change linearly along each turn, and the height increment (pitch) of the segment is kept fixed along the whole structure ($p = 0.07m$). A genetic algorithm is applied in the MATNEC optimization and a solution space with sufficiently large population size ($Pop = 400$) is used. The number of bits for each gene is 10, the mutation rate R_m is 10%, and the iteration is run for 50 steps. Ranking, elitist selection and crossover strategies are adopted, as listed in Table 6.10.

$$f_{G-RBF} = \sum_{i=1}^{N} \omega_i \cdot e^{-A(x-c_i)^2} \tag{6.8}$$

An illustration of radius and contribution from the RBFs is shown in Figure 6.14. It can be observed in the left-hand graph that the total RBF solution (solid black curve) representing helix radius increases with respect to antenna height, Figure 6.14 shows the side view of the actual helix. The discontinuous connections is caused by the discretization used in NEC-2 due to minimum segment length requirement.

Table 6.10. *Genetic algorithm settings for optimization.*

Population size	400
No. of bits per gene	10
Mutation rate	10%
No. of iterations	50
Selection & crossover strategies	Ranking & elitist

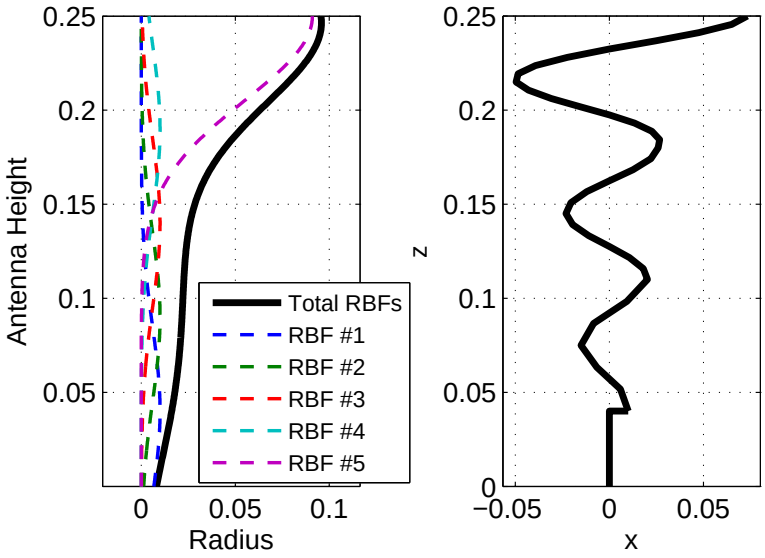

Figure 6.14. Illustration of radius implementation from Gaussian RBFs, Left-hand side graph: RBFs for the radius variation, Right-hand side image: Antenna side view in the discretization used in NEC-2.

The optimized helical antennas for frequencies 25, 50 and 100 MHz are shown in Figure 6.15 and 6.16. The optimized RBFs weighting variables (ω_i) and the performance are listed in Table 6.11.

It can be seen that all the three helices have increasing radius along their length, and expand towards the top. It is also noticed that they achieve wider bandwidth than the uniform references at all three different frequencies. Explicitly, the bandwidth increases from 0.025 MHz to 0.04 MHz at 25 MHz, 0.17 MHz to 0.25 MHz at 50 MHz, and 1.50 MHz to 1.95 MHz at 100 MHz respectively. It can be stated that the antenna performance can be significantly improved when the radius characteristics are allowed

Table 6.11. *Optimized helices with radius variation at 25, 50, and 100 MHz.*

Radius variation	25 MHz	50 MHz	100 MHz
Radius	$\omega_1 = 0.01$ m $\omega_2 = 0.01$ m $\omega_3 = 0.10$ m $\omega_4 = 0.10$ m $\omega_5 = 0.10$ m	$\omega_1 = 0.01$ m $\omega_2 = 0.01$ m $\omega_3 = 0.01$ m $\omega_4 = 0.07$ m $\omega_5 = 0.10$ m	$\omega_1 = 0.01$ m $\omega_2 = 0.01$ m $\omega_3 = 0.01$ m $\omega_4 = 0.01$ m $\omega_5 = 0.09$ m
Pitch (p)	0.07 m	0.07 m	0.07 m
Wire length	2.32 m	1.44 m	0.83 m
Efficiency	49.2%	81.6%	85.0%
Bandwidth	0.04 MHz	0.25 MHz	2.05 MHz
Overall fitness	89.2 ($w = 1000$)	106.6 ($w = 100$)	290.0 ($w = 100$)

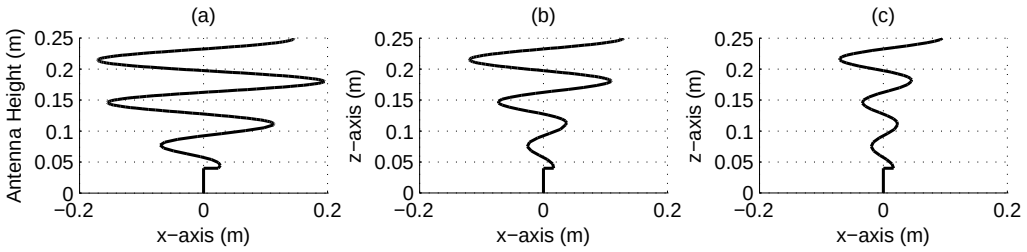

Figure 6.15. Side view of optimized configurations of helices with radius variation at (a) 25 MHz, (b) 50 MHz, and (c) 100 MHz.

to vary along the helices. This poses a convincible guideline in designing helical antennas, i.e. allow the helical body to expand towards the top when the space permits, in the pursuit of antenna performance improvement.

6.4.3 Non-Linear Structural Variations in Pitch

For pitch variations, a different RBFs implementation strategy has to be applied in order to fulfill the fixed height requirement. Explicitly, the height increment (pitch) for each segment is first obtained from the RBFs by locking the number of segments

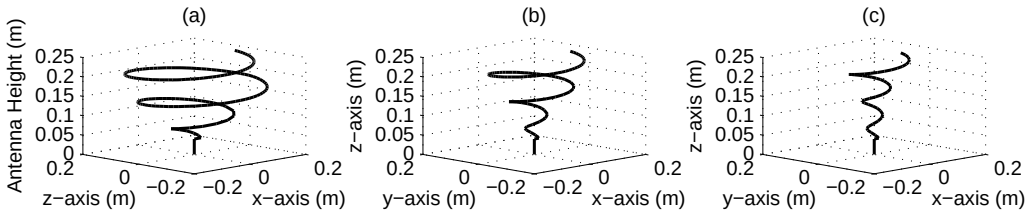

Figure 6.16. 3D view of optimized configurations of helices with radius variation at (a) 25 MHz, (b) 50 MHz, and (c) 100 MHz.

(M). The accumulated height is derived and scaled to the physical height limit ($H_{max} = 0.25m$). The scaling factor is then applied to the RBFs expansion to obtained an antenna that satisfies the height constraint. This process is performed at every stage of the modeling and optimization. An illustration of pitch and contribution from RBFs is shown in Figure 6.17. It can be seen that larger RBF values at lower height contribute to large pitch on the antenna near the bottom.

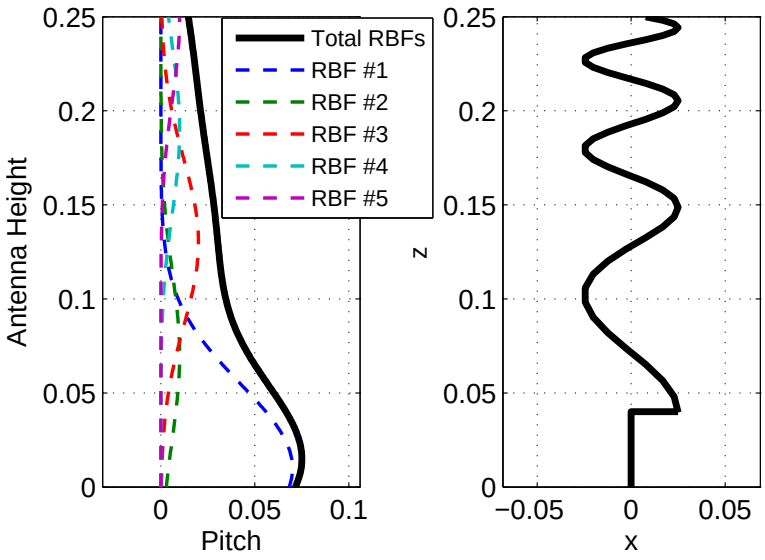

Figure 6.17. Illustration of pitch implementation from Gaussian RBFs, Left-hand side graph: RBFs for the pitch variation, Right-hand side image: Antenna side view in the discretization used in NEC-2.

The optimizing variables ω_i are allowed to vary from 0.01 to 0.07 m, which limits helix to have a minimum of three turns. The radius of the helix is set constantly and consistently based on the uniform models at different frequencies, i.e. r = 0.10 m at 25 MHz, r = 0.07 m at 50 MHz, and r = 0.4 m at 100 MHz respectively. The optimized helical antennas with constant radius and pitch variation are obtained, as shown in Figure 6.18 and 6.19. The optimized RBFs weighting variables (ω_i) and the performance are listed in Table 6.12.

Table 6.12. *Optimized helices with pitch variation at 25, 50, and 100 MHz.*

Pitch variation	25 MHz	50 MHz	100 MHz
Radius (r)	0.10 m	0.07 m	0.04 m
Pitch	ω_1 = 0.07 m ω_2 = 0.07 m ω_3 = 0.01 m ω_4 = 0.01 m ω_5 = 0.01 m	ω_1 = 0.07 m ω_2 = 0.07 m ω_3 = 0.01 m ω_4 = 0.01 m ω_5 = 0.01 m	ω_1 = 0.07 m ω_2 = 0.07 m ω_3 = 0.01 m ω_4 = 0.01 m ω_5 = 0.01 m
Wire length	1.98 m	1.42 m	0.83 m
Efficiency	60.9%	84.0%	74.3%
Bandwidth	0.041 MHz	0.27 MHz	1.95 MHz
Overall fitness	101.9 ($w = 1000$)	111.0 ($w = 100$)	269.3 ($w = 100$)

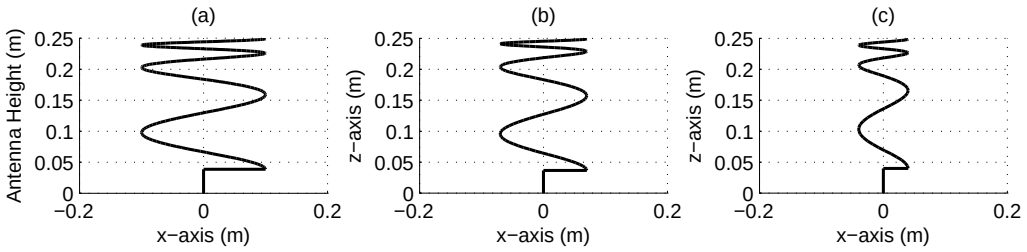

Figure 6.18. Side view of optimized configurations of helices with pitch variation at (a) 25 MHz, (b) 50 MHz, and (c) 100 MHz.

It is seen that the three optimized pitch-varying helices exhibit increasing density of winding close to the top end. They exhibit improved performance in comparison to the

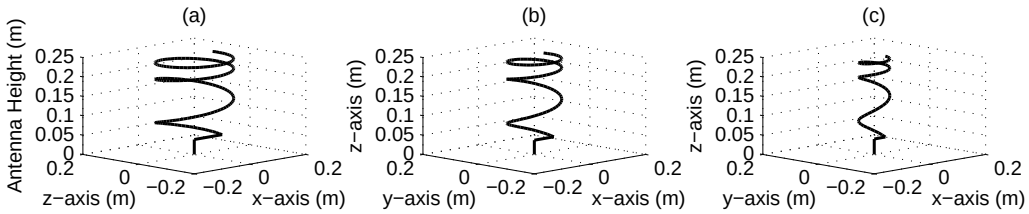

Figure 6.19. 3D view of optimized configurations of helices with pitch variation at (a) 25 MHz, (b) 50 MHz, and (c) 100 MHz.

uniform references. This finding also directs a guideline for antenna designing process that denser winding towards the top of helix can effectively improve the efficiency and bandwidth performance. This indication is particular useful for real-world applications when space is constrained for the antenna to grow. Pitch-varying technique will be used in the next chapter for the design of a portable communication device.

6.4.4 Non-Linear Structural Variations in Radius & Pitch

In order to implement helical antennas with both radius and pitch variations, two RBF expansions are required. One expansion is used to describe the radius variation and determine the total number of segments (M), whereas the other describes the pitch of each segment when normalized to achieve the specified antenna total height. All RBF parameters are optimized simultaneously to yield an optimal helix where both radius and pitch vary continuously along the antenna length. The optimizing ranges for radius and pitch satisfy $\omega_r \in [0.01m, 0.20m]$ and $\omega_p \in [0.01m, 0.07m]$ respectively.

The optimizations are carried out at 25, 50, and 100 MHz, and the optimized helical antennas with both radius and pitch variations are obtained, as shown in Figure 6.20 and 6.21. The optimized RBFs weighting variables (ω_r and ω_p) and the performance are listed in Table 6.13.

Obviously, when both radius and pitch variations are applied to the helix, the optimized helix combines the features of increasing radius and denser pitch along the

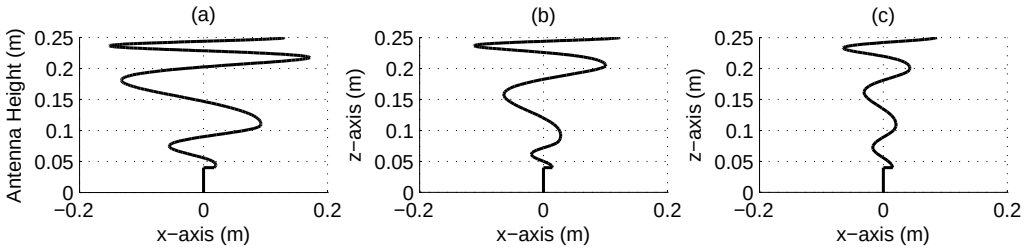

Figure 6.20. Side view of optimized configurations of helices with radius and pitch variations at (a) 25 MHz, (b) 50 MHz, and (c) 100 MHz.

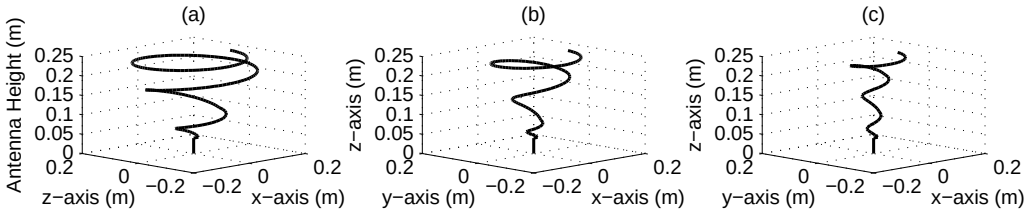

Figure 6.21. 3D view of optimized configurations of helices with radius and pitch variations at (a) 25 MHz, (b) 50 MHz, and (c) 100 MHz.

height towards the top end. The consequent performance can also be improved compared to the uniform references, and maximized total fitness can be achieved.

6.5 Overall Analysis

For demonstration purposes, the optimized helical antennas with uniform and optimized non-linear structures are compared at 100 MHz.

6.5.1 Performance Comparison at 100 MHz

At 100 MHz, a fixed overall height $h = 0.25$ m corresponds to one twelfth of the wavelength ($h = \lambda/12$). Optimized helices are shown in Figure 6.22 with their simulated performance listed in Table 6.14. All optimized antennas out-perform the uniform helices by utilizing radius and/or pitch variation along their length. The geometrical

Table 6.13. *Optimized helices with radius and pitch variations at 25, 50, and 100 MHz.*

Radius & pitch variations	25 MHz	50 MHz	100 MHz
Radius (ω_r)	$\omega_1 = 0.01$ m $\omega_2 = 0.01$ m $\omega_3 = 0.10$ m $\omega_4 = 0.10$ m $\omega_5 = 0.10$ m	$\omega_1 = 0.01$ m $\omega_2 = 0.01$ m $\omega_3 = 0.01$ m $\omega_4 = 0.08$ m $\omega_5 = 0.10$ m	$\omega_1 = 0.01$ m $\omega_2 = 0.01$ m $\omega_3 = 0.01$ m $\omega_4 = 0.02$ m $\omega_5 = 0.08$ m
Pitch (ω_p)	$\omega_1 = 0.07$ m $\omega_2 = 0.07$ m $\omega_3 = 0.01$ m $\omega_4 = 0.01$ m $\omega_5 = 0.01$ m	$\omega_1 = 0.07$ m $\omega_2 = 0.07$ m $\omega_3 = 0.01$ m $\omega_4 = 0.01$ m $\omega_5 = 0.01$ m	$\omega_1 = 0.07$ m $\omega_2 = 0.07$ m $\omega_3 = 0.01$ m $\omega_4 = 0.01$ m $\omega_5 = 0.01$ m
Wire length	2.19 m	1.40 m	0.80 m
Efficiency	62.3%	84.7%	77.8%
Bandwidth	0.042 MHz	0.28 MHz	2.30 MHz
Overall fitness	104.3 ($w = 1000$)	112.7 ($w = 100$)	307.8 ($w = 100$)

characteristics of the optimized antennas indicate the trend that 1) the optimized radius varying helix has an increasing radius with height, 2) the optimized pitch varying helix exhibits an increasing density of the winding with height, and 3) the radius and pitch varying helices combine the features of the above two to maximize the performance.

6.5.2 Antenna Q Factor at 100 MHz

As described in Chapter 1, the properties of electrically small antennas can be interpreted as the quality factor (Q), and the fractional bandwidth (FBW). Taking the helical antennas at 100 MHz as examples, they are designed with fixed height of 0.25 m, which corresponds to a wavelength $\lambda = 3$ m, thus $k = \frac{2\pi}{\lambda} = \frac{2\pi}{3}$. The term a, which is defined as the radius of a sphere that circumscribes the maximum physical dimension

Table 6.14. Optimized uniform and parameter-varying helices at 100 MHz.

50 MHz	Uniform reference	Radius variation	Pitch variation	Radius & pitch variations
Wire Length	0.81 m	0.85 m	0.83 m	0.80 m
Efficiency	71.9%	85.0%	74.3%	77.8%
Bandwidth	1.50 MHz	2.05 MHz	1.95 MHz	2.30 MHz
Overall Fitness (w=100)	221.9	290.0	269.3	307.8

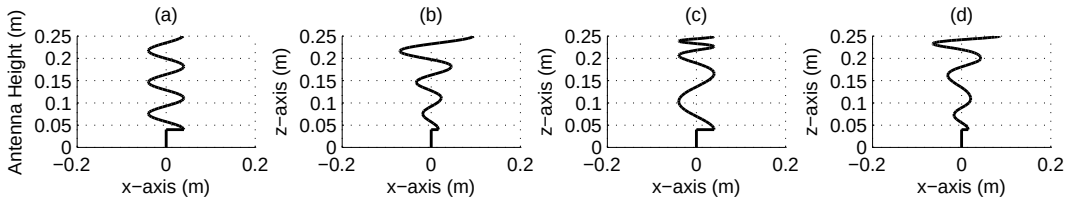

Figure 6.22. Optimized helical antennas at 100 MHz:(a) uniform reference, (b) radius variation, (c) pitch variation, (d) radius and pitch variations.

of the antenna is the height of the antenna, is equal to $a = 0.25$ in this case. This leads to $ka = 0.524$, which satisfies the definition and requirement of being electrically small.

Assuming the radiation efficiency (η_r) be ideally 100%, the lower bound for the Q factor (Q_{lb}) can be defined:

$$\begin{aligned} Q_{lb} &= \eta_r\left(\frac{1}{(ka)^3} + \frac{1}{ka}\right) \\ &= 100\% \times \left(\frac{1}{(0.52)^3} + \frac{1}{k0.52}\right) \\ &= 8.876 \end{aligned} \tag{6.9}$$

For bandwidth, the definition of bandwidth is set to fractional matched VSWR bandwidth (FBW_V):

$$FBW_V(\omega_0) = \frac{\omega_+ - \omega_-}{\omega_0} \tag{6.10}$$

where ω_+ and ω_- are the frequencies above and below ω_0, and at those frequencies, the VSWR is equal to an arbitrary value denoted by s.

The fractional matched VSWR bandwidth and the Q are related:

$$\begin{aligned} Q(\omega_0) &= \frac{2\sqrt{\beta}}{FBW_V(\omega_0)}) \\ \sqrt{\beta} &= \frac{s-1}{2\sqrt{s}} \end{aligned} \tag{6.11}$$

where s is the VSWR value defining the bandwidth.

Thus the upper bound of FBW_{Vub} can be calculated:

$$\begin{aligned} FBW_{Vub} &= \frac{1}{\eta}\frac{(ka)^3}{1+(ka)^2}\frac{s-1}{\sqrt{s}} \\ &= \frac{1}{100\%}\frac{(0.52)^3}{1+(0.52)^2}\frac{2-1}{\sqrt{2}} \\ &= 0.0797 \end{aligned} \tag{6.12}$$

For the optimized helical antennas at 100 MHz, the bandwidth for the uniform, just radius varying, just pitch varying, both radius and pitch varying helices have matched bandwidth (VSWR < 2.0) of 1.50 MHz, 2.05 MHz, 1.95 MHz, and 2.30 MHz respectively. The corresponding $FBW_V(\omega_0)$, β, s, and $Q(\omega_0)$ are calculated and listed in Table 6.15 .

Table 6.15. Q factor of the optimized uniform and parameter-varying helices at 100 MHz.

100 MHz	Uniform reference	Radius variation	Pitch variation	Radius & pitch variations
Bandwidth	1.50 MHz	2.05 MHz	1.95 MHz	2.30 MHz
$FBW_V(\omega_0)$	0.015	0.0205	0.0195	0.023
$\sqrt{\beta}$	$\frac{1}{2\sqrt{2}}$			
$Q(\omega_0)$	47.1	34.5	36.3	30.7

It can be seen that either pitch or radius variation along the helix can effectively decrease the Q factor of this types of electrically antennas, in comparison to the conventionally uniform winding strategy. The fractional bandwidth (FBW_V) is also significantly increased from 0.015 of the uniform reference to around 0.20 with structural variations. Particularly, when both radius and pitch variations are combined and applied on the helix, the bandwidth and Q factors are significantly improved. This successfully verifies the robustness of the design and methodology.

6.5.3 Current Distribution at 100 MHz

With the optimized helical antennas obtained at 100 MHz, the current distribution along those antennas are investigated. For comparison purpose, a 0.75 m long quarter-wave monopole is simulated at 100 MHz, and used as the reference model. Their configurations and current distributions are shown in Figure 6.23.

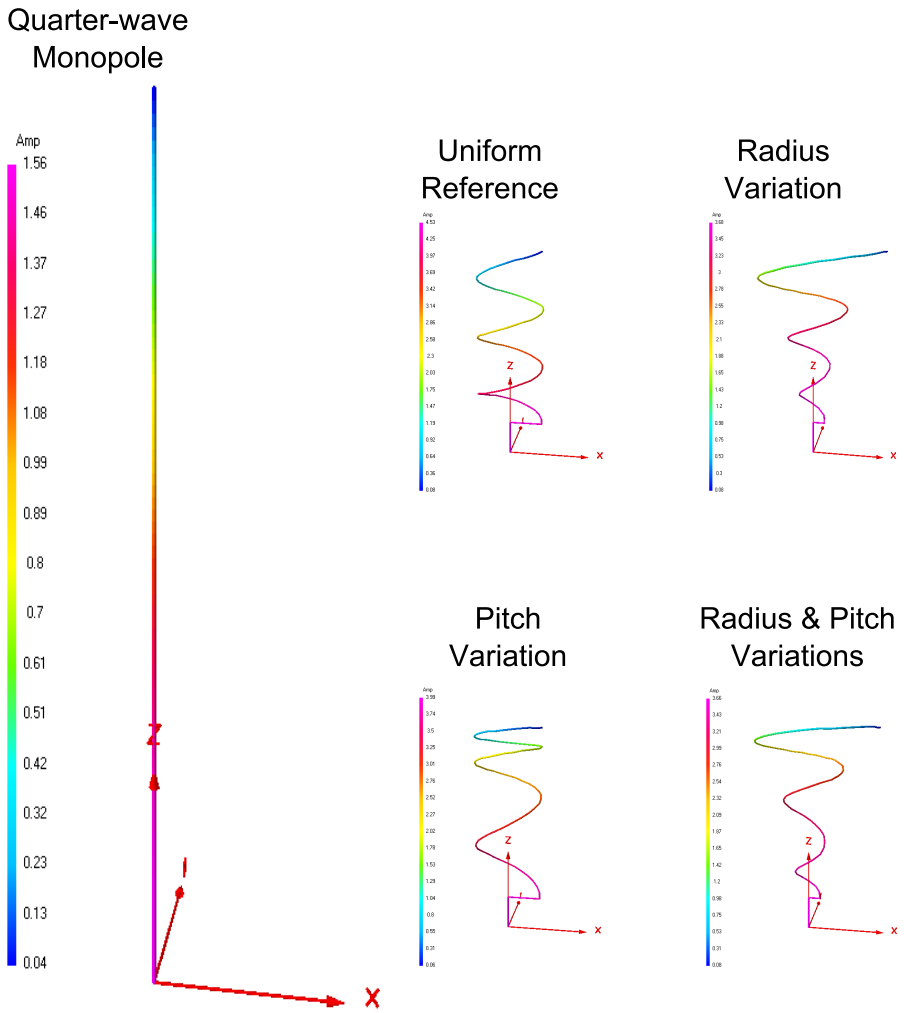

Figure 6.23. Current distribution of optimized helical antennas at 100 MHz.

Furthermore, the current magnitude is plotted with respect to antenna height in order to distinguish the difference in distribution, shown in Figure 6.24. It can be observed that the optimized helical antennas have lower current magnitude near the bottom where the excitation is located, and comparatively higher magnitude to the top end in comparison to the uniform reference case. This indicates that the current is more

evenly distributed along the antenna height, and thus provides the potential for the matching networks to yield a wider operational bandwidth. At 25 and 50 MHz, the current distributions look very similar to the ones at 100 MHz.

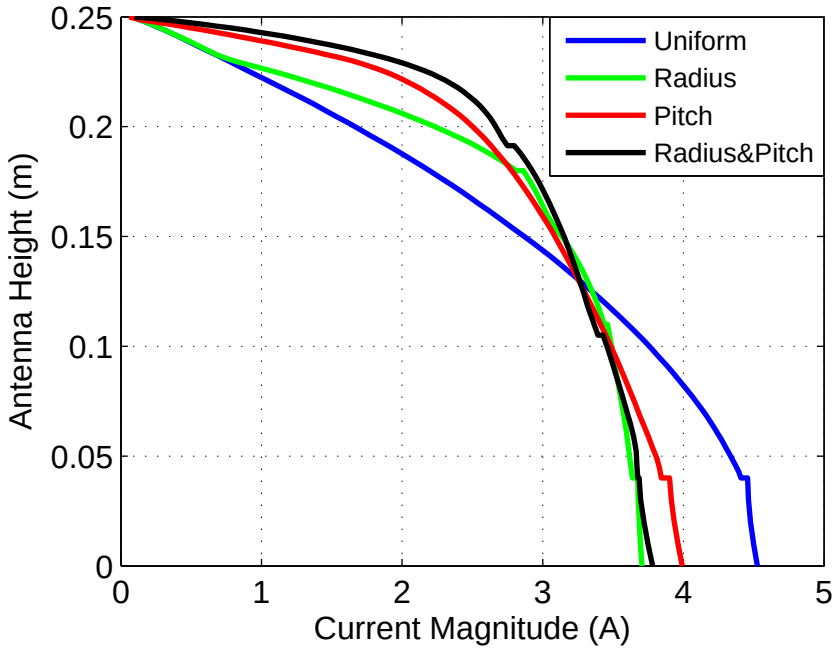

Figure 6.24. Current distribution of optimized helical antennas with respect to antenna height at 100 MHz.

6.5.4 Radiation Patterns at 100 MHz

The optimized helical antennas are designed for operation in normal mode which is typically used for applications where reduced size and omni-directional radiation are the critical operational factors. Owing to their small size (with respect to wavelength), the antennas will have a radiation pattern that is very similar to a short monopole on an infinite ground plane. For illustration purpose, the radiation patterns of both co-polarization and cross-polarization for the optimized antennas at 100 MHz are shown in Figure 6.25.

It can be observed that the cross-polarization scales directly with the transverse extent of the antenna, and is generally very small. In particular, the uniform and pitch variation versions have the same cross-polarization (under -30 dB). Comparatively,

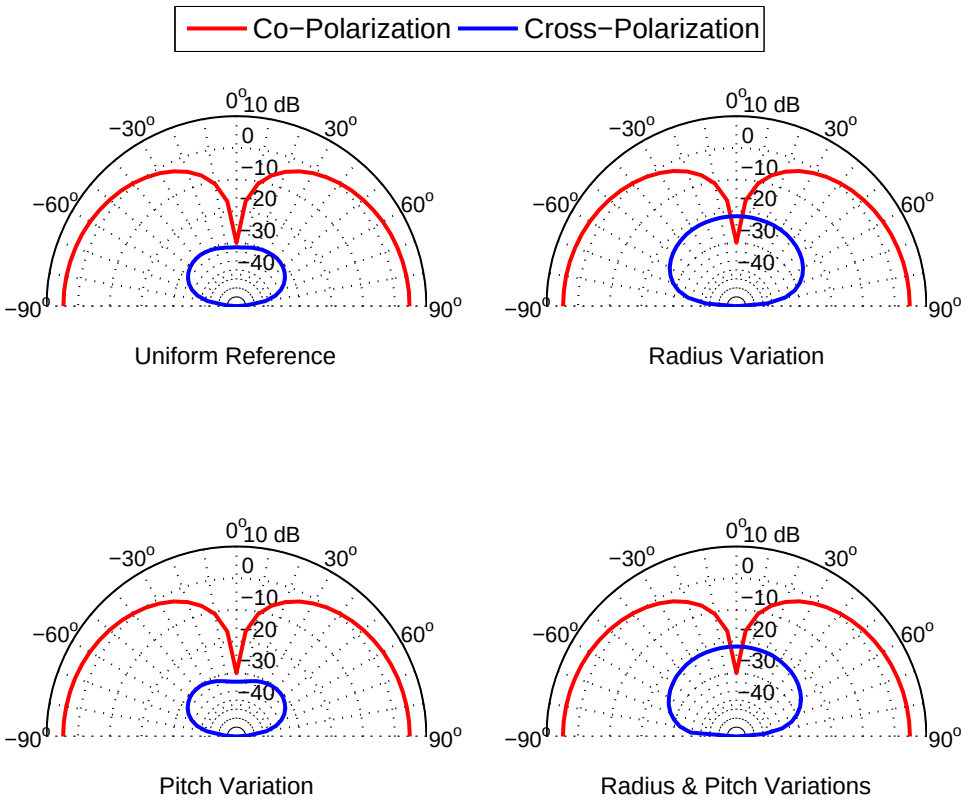

Figure 6.25. Gain pattern for the optimized helical antennas at 100 MHz.

the other two antennas with radius variations exhibit higher cross-polarized radiation. However, the cross-polarization level remains clearly more than 20 dB below the co-polarization level. As expected, the introduction of a large winding radius at the top of the antenna introduces a small amount of axial mode radiation. At 25 and 50 MHz, the gain patterns look very similar to the ones at 100 MHz.

6.6 The Effect of the Matching Network

A matching network is usually added to provide an exact match at the center frequency of the operation band. This network, however, will have an effect on both efficiency and bandwidth of the total antenna system and this section investigates its impact.

6.6.1 Efficiency

One of the major optimization objectives is efficiency. This is the total efficiency including the antenna and the matching network, and can be written as:

$$\eta_{total} = \eta_{rad} \times \eta_{mn} \tag{6.13}$$

where η_{rad} is the radiation efficiency of the antenna including the material loss, and η_{mn} is the efficiency of the matching network. The radiation loss can be calculated from NEC, but the matching network loss needs to be calculated separately.

- **Matching Network Loss**

  The L-type matching network used in this work can inevitably cause some power loss to the total antenna system. The lumped components used in the matching network are attributed Q factors, explicitly Q_L and Q_C are the quality factors for inductor and capacitor respectively. Finite Q factors of lumped components imply resistance and therefore lead to power loss in the matching network.

  Taking the normal L-type matching network as an example, the parasitic resistance of the inductor and the capacitor are R_L and R_C, while the reactance are X_L and X_C respectively (shown in Figure 6.26). The quality factor of the matching network is:

$$Q = \sqrt{\frac{R_{Source}}{R_{Load}} - 1} \tag{6.14}$$

  The quality factors for the series-leg (inductor) and the shunt-leg (capacitor) are:

$$\begin{aligned} Q_L &= \frac{|X_L|}{R_L} \\ Q_C &= \frac{R_C}{|X_C|} \end{aligned} \tag{6.15}$$

  The efficiency of the matching network can written as:

$$\eta_{mn} = \frac{|P_{Load}|}{|P_{Source}|} \times 100\% \tag{6.16}$$

  where $|P_{Load}|$ and $|P_{Source}|$ can be calculated as:

$$\begin{aligned} P_{Load} &= \frac{1}{2} I_L^2 R_{Load} \\ P_{Source} &= \frac{V_{Source}^2}{2R_{Source}} \end{aligned} \tag{6.17}$$

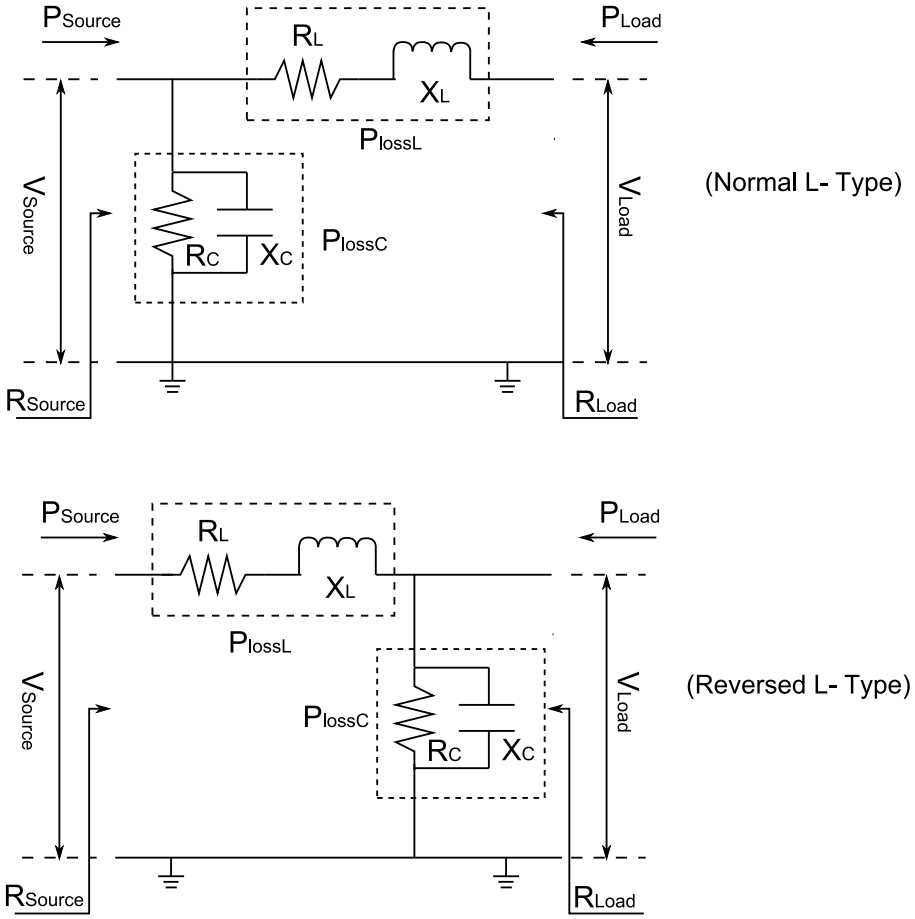

Figure 6.26. Illustration of normal and reversed L-type matching network.

The power losses in the inductor and capacitor are:

$$\begin{aligned} P_{lossL} &= \tfrac{1}{2} I_L^2 R_L \\ P_{lossC} &= \tfrac{V_{Source}^2}{2R_C} \end{aligned} \tag{6.18}$$

Substitute Equations 6.15 and 6.17 into 6.18, it can be obtained:

$$\begin{aligned} P_{lossL} &= \tfrac{1}{2} I_L^2 R_L \\ &= \tfrac{1}{2} \tfrac{X_L}{Q_L} I_L^2 \\ &= \tfrac{Q}{Q_L} |P_{Load}| \end{aligned}$$

$$\begin{aligned} P_{lossC} &= \tfrac{V_{Source}^2}{2R_C} \\ &= \tfrac{|X_C|}{2Q_C} \left(\tfrac{V_{Source}}{|X_C|}\right)^2 \\ &= \tfrac{V_{Source}^2}{2Q_C|X_C|} \\ &= \tfrac{Q}{Q_C} |P_{Source}|. \end{aligned} \tag{6.19}$$

The total power (P_{Source}) is dissipated on the load (P_{Load}), the inductor (P_{lossL}), and the capacitor (P_{lossC}):

$$P_{Source} = P_{Load} + P_{lossL} + P_{lossC}. \tag{6.20}$$

Substituting Equations 6.19 into 6.20 yields

$$P_{Source} = P_{Load} + \frac{Q}{Q_L}|P_{Load}| + \frac{Q}{Q_C}|P_{Source}|. \tag{6.21}$$

Therefore, the efficiency of the normal L-type matching network can be derived:

$$\begin{aligned} \eta_{mn} &= \frac{|P_{Load}|}{|P_{Source}|} = \frac{1-\frac{Q}{Q_C}}{1+\frac{Q}{Q_L}} \\ Q &= \sqrt{\frac{R_{Source}}{R_{Load}} - 1}. \end{aligned} \tag{6.22}$$

Similarly, the efficiency of the reversed L-type matching network can be expressed as

$$\begin{aligned} \eta_{mn} &= \frac{|P_{Load}|}{|P_{Source}|} = \frac{1-\frac{Q}{Q_L}}{1+\frac{Q}{Q_C}} \\ Q &= \sqrt{\frac{R_{Load}}{R_{Source}} - 1}. \end{aligned} \tag{6.23}$$

In this thesis, the Q factors for the inductor and the capacitor are set consistently to 100 and 1000 respectively (based on typical values for realizable matching networks). Assuming a source impedance to be 50 Ω, an efficiency plot can be drawn with respect to various loading impedance (R_{Load}) from 0 Ω to 100 Ω, as shown in Figure 6.27. It can be observed that the efficiency of the L-type matching network remains high with the input impedance lower than 100 Ω. Explicitly, the L-type matching network does not have significant impact upon the total efficiency.

6.6.2 Bandwidth Sensitivity

In this thesis, L-type matching networks are used. The component values of the inductor and the capacitor are first calculated based on the input impedance at the target frequency (e.g. 100 MHz) in order to achieve a VSWR of 1.0 for a 50 Ω load. The obtained L-type matching network is then applied at the rest of the frequency points, and the bandwidth is derived as that frequency range for which there is a VSWR less then

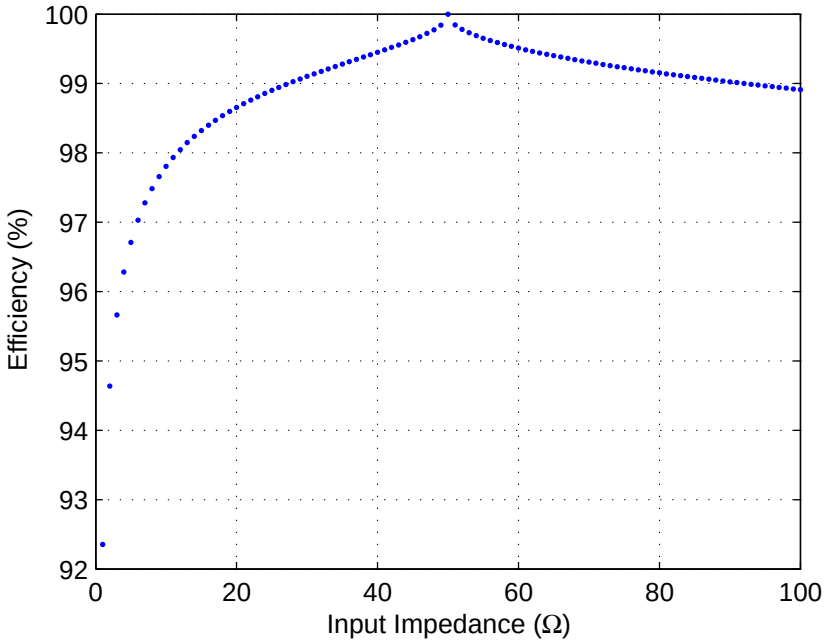

Figure 6.27. The efficiency of normal L-type matching network at various input impedance.

2.0. Nevertheless, it should be noted that a VSWR of 1.0 is not necessary at the target frequency, because the ultimate goal is a bandwidth that is as wide as possible. Therefore, it is investigated whether some variations to the LC components can be used to effectively broaden the bandwidth by sacrificing the VSWR at the target frequency. Explicitly, the VSWR at the center frequency increases but remains under 2.0, and a wider operational bandwidth can be obtained.

To carry out the investigation, both inductor and capacitor are given a variation of one-tenth to twice their original values. The resulting network is applied, and the impact is observed with respect to the components variations. The sample selected in this investigation is a matching network with an inductor of 37.3 nH at a Q factor of 100, and a capacitor of 54.5 pF at a Q factor of 1000. It is applied to an antenna with an input impedance of $Z = 12.7\Omega - j1.65\Omega$, a bandwidth of 1.023 MHz with perfect matching at target frequency.

- **Impact of Inductor (L)**

  In Figure 6.28, the resulting bandwidth is shown for the cases where the inductance in the matching network is varied. The peak is obtained with an inductor

0.86 times the original value, while the resulting bandwidth is 1.024 MHz. The VSWR is also plotted for both models. It can be stated that the variation in the inductor does not have an obvious impact on the bandwidth.

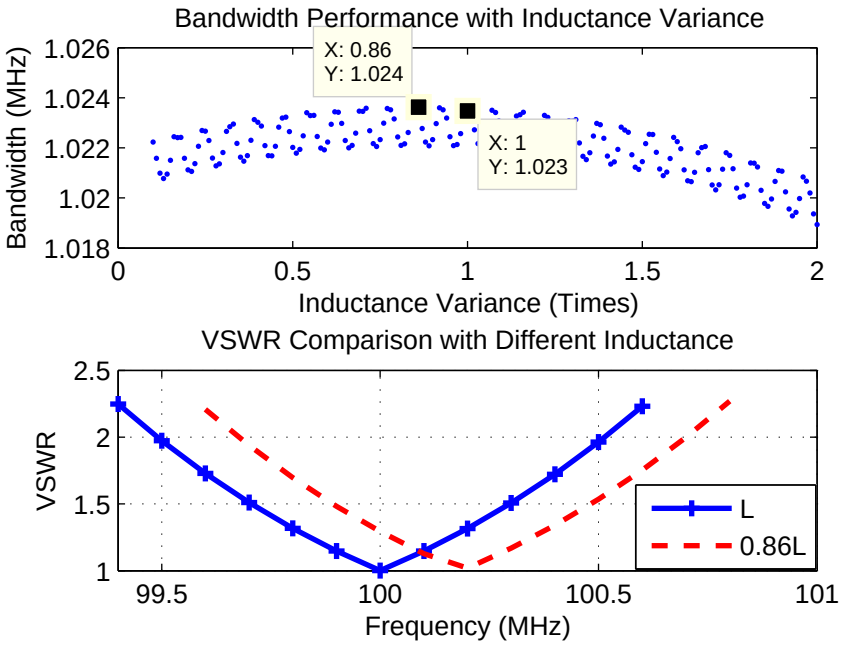

Figure 6.28. Bandwidth performance with inductor variations.

- **Impact of Capacitor (C)**

  Likewise, the impact of the capacitor is investigated alone with respect to variations. The bandwidth and VSWR performance are plotted in Figure 6.29. Basically, a reduction of 15% in the capacitance can lead to a light improvement, but relatively insignificant.

- **Impact of Inductor and Capacitor (L&C)**

  Further, a more sophisticated test is carried out by varying both inductor and capacitor simultaneously. It can be ascertained from Figure 6.30 that the capacitor comparatively plays a more important role in determining the bandwidth, but the overall effect is small. As a result, the perfect matching network scheme is applied throughout the entire investigation in this thesis. The study does, however indicate, that the results are stable to small variations in matching network components.

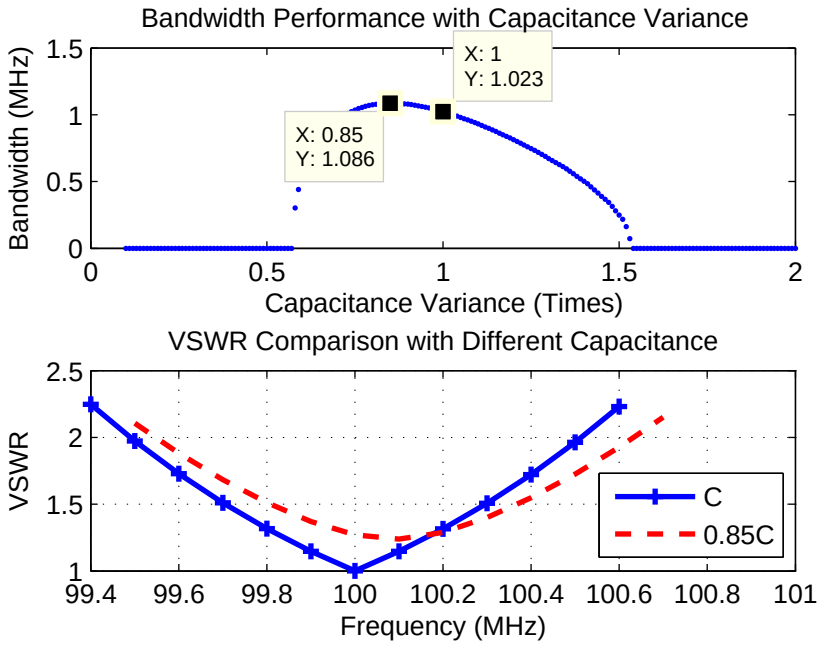

Figure 6.29. Bandwidth performance with capacitor variations.

6.7 RBF Sensitivity

Another critical issue that affects the complexity and accuracy of the simulation and optimization is the selection of an appropriate number of radial basis functions in the investigation. A greater number of RBFs can bring more flexibility to structural variations and thus more accurate results, but also at the cost of slower optimization convergence and increased simulation load. Therefore, a sensitivity test is carried out with respect to RBFs when the number of subset of RBF varies and all other settings remain the same. The example used is a helical antenna with a constant pitch of 30 mm, and the radius varies along the length. The settings of the simulation and optimization are tabulated in Table 6.16.

The optimal solutions are converged and obtained with different numbers of RBFs respectively, the comparisons are shown in Table 6.17. Based on the overall fitness, it can be seen that when a number of 5, 6, and 7 RBFs are applied, the overall fitness is greater than the case of 4 RBFs. Nevertheless, it is also noticed that the overall fitness saturates when this number increases from 5 to 7, which indicates a diminishing return.

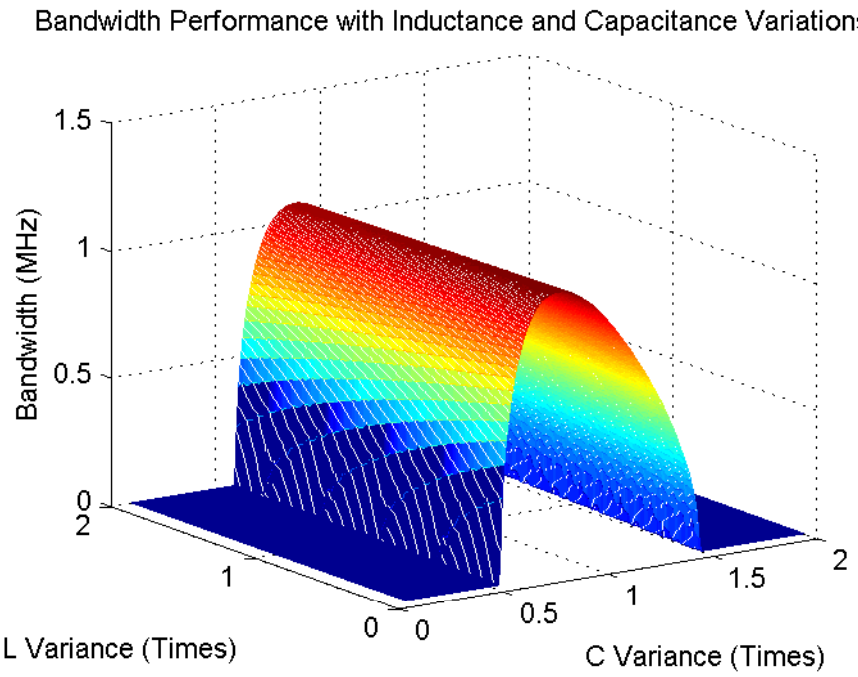

Figure 6.30. Bandwidth performance with both inductor and capacitor variations.

Table 6.16. *Simulation and optimization settings for RBF sensitivity test.*

Geometry		Optimization	
Antenna height	0.25 m	No. of genes	200
Radius variable range (ω_i)	0.01 m - 0.03 m	No. of iterations	50
Pitch	0.03 m	No. of bits for variable ωi	6
No. of RBFs	4, 5, 6, 7	Mutation rate	10%
Fitness function	$Efficiency(\%) + 100 \times Bandwidth(MHz)$		

Explicitly, a number of 5 RBFs can effectively provide sufficient structural variations to the helical antenna optimization.

6.8 Experimental Validation

For the purpose of experimental validation, prototypes of the proposed helical antennas at 100 MHz were fabricated according to their optimized parameters, as shown in Figure 6.31. Helices were wound using an aluminum wire of radius 1 mm, and mounted with a base connector with a length of 0.04 m. Experiments were carried

Table 6.17. *Optimized solutions and performance with various numbers of RBFs.*

	4 RBFs	5 RBFs	6 RBFs	7 RBFs
Radius	$\omega_1 = 0.01$ m $\omega_2 = 0.01$ m $\omega_3 = 0.03$ m $\omega_4 = 0.03$ m	$\omega_1 = 0.01$ m $\omega_2 = 0.01$ m $\omega_3 = 0.01$ m $\omega_4 = 0.03$ m $\omega_5 = 0.03$ m	$\omega_1 = 0.01$ m $\omega_2 = 0.01$ m $\omega_3 = 0.01$ m $\omega_4 = 0.01$ m $\omega_5 = 0.03$ m $\omega_6 = 0.03$ m	$\omega_1 = 0.01$ m $\omega_2 = 0.01$ m $\omega_3 = 0.01$ m $\omega_4 = 0.01$ m $\omega_5 = 0.01$ m $\omega_6 = 0.03$ m $\omega_7 = 0.03$ m
Efficiency	93.1%	93.5%	93.4%	93.2%
Bandwidth	1.45 MHz	1.70 MHz	1.70 MHz	1.70 MHz
Overall fitness	238.1	263.5	263.4	263.2

out with respect to both matched bandwidth and radiation efficiency, using an Agilent 814ET RF network analyzer.

6.8.1 Efficiency

For efficiency measurement, a Wheeler cap [5] was used to fully enclose the antenna and suppress the entire energy. The Wheeler cap was a 0.75 m cubic aluminium box covering the antenna on top of the ground with edge fully sealed, shown in Figure 6.32.

For a transmitting antenna, the radiation efficiency can be defined as:

$$\eta = \frac{P_r}{P_o} = \frac{P_r}{P_r + P_l} \tag{6.24}$$

where P_r is the total power radiated, P_i is the total power input, and P_l is the total power lost. P_l includes ohmic losses in the antenna wires and the losses in any tuning or matching networks.

Equivalently, this equation can be written as:

$$\eta = \frac{R_r}{R_r + R_l} \tag{6.25}$$

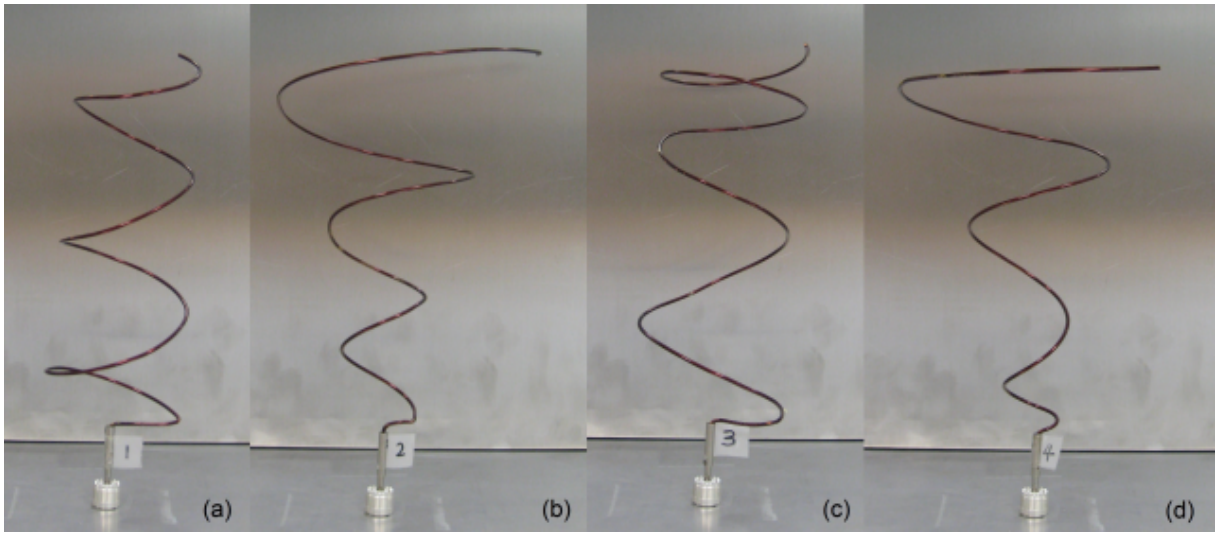

Figure 6.31. Fabricated helical antennas: (a) uniform reference, (b) radius variation, (c) pitch variation, (d) radius and pitch variations.

where R_r is the radiation resistance and R_l is the loss resistance.

When the Wheeler cap is not used, the real part of the input impedance measured is composed of the radiation resistance R_r and the loss resistance R_l, explicitly $R_p = R_r + R_l$. It is suggested in [5] that when an antenna is enclosed with a conducting sphere of a radius equals to a radian length $\frac{\lambda}{2\pi}$, the radiation resistance R_r will be eliminated from the the input impedance without changing the loss resistance significantly. Therefore, when the antenna is enclosed by the Wheeler cap, the real part of the input impedance measured is purely the loss resistance R_l. The input impedance measured with or without the cap can be listed in Table 6.18.

Table 6.18. *Input impedance measured with and without the cap.*

	Input impedance (Real part)
Without cap	$R_r + R_l$
With cap	R_l

With Equation 6.25, the radiation efficiency of an antenna can be determined using R_r and R_l. The matching networks were not included in the measurement, and the efficiency was taken at the self-resonant frequency. The comparison between simulations

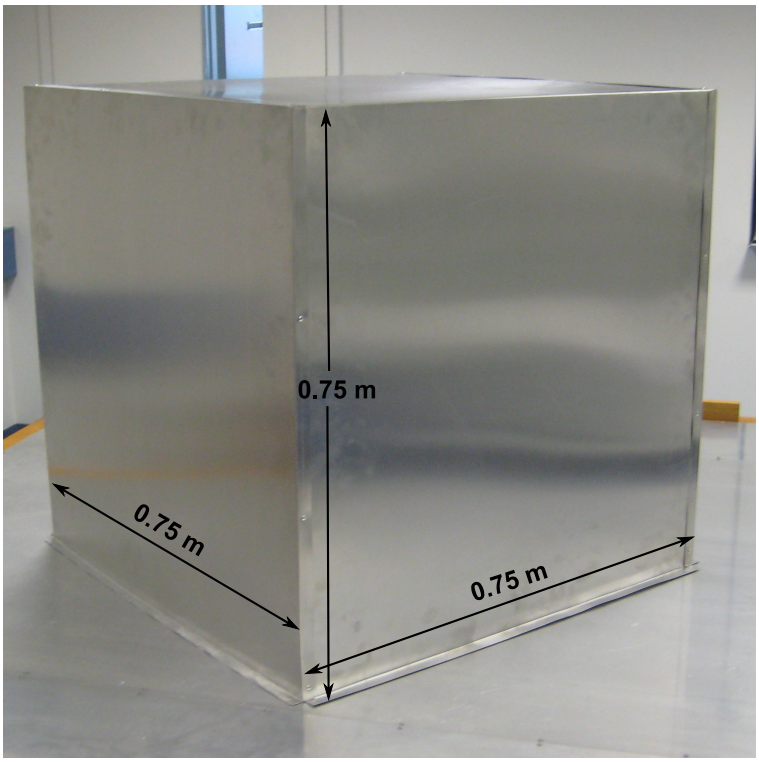

Figure 6.32. Illustration of a Wheeler cap for efficiency measurement.

and measurements are listed in Table 6.19. Much of the difference between measurement and simulation can be attributed to measurement inaccuracy and the wire model representation.

Table 6.19. Efficiency comparison of simulations and measurements at 100 MHz.

Efficiency	Uniform Reference	Radius Variation	Pitch Variation	Radius and Pitch Variations
Simulation	83.6%	82.9%	79.2%	83.2%
Measurement	80.4%	79.5%	75.7%	85.2%

Importantly, the optimized parameter-varying helical antennas exhibit reasonable and consistent performance as predicted in simulation.

6.8.2 Bandwidth

For bandwidth measurement, experiments were performed on top of an aluminium sheet acting as the ground plane, shown in Figure 5.10. Appropriate L-type matching networks were connected and finely tuned between the antennas and the the network analyzer. In order to obtain the operational bandwidth of the proposed helical antennas, an L-section matching network is built for each helical antenna. The VSWR, however, is not necessarily matched to 1.0 at 100 MHz when the maximum operational bandwidth with VSWR < 2.0. In the following work, the variable inductor and capacitor used in the matching network is finely tuned such that the maximum bandwidth can be obtained. The sample L-section matching network built for the bandwidth measurement is shown in Figure 5.11.

The measured bandwidth for the optimized helical antennas are compared to the simulated results, as shown in Table 6.20. It can be seen that the original simulation indicates matched operational bandwidths of 1.50 MHz, 2.05 MHz, 1.95 MHz, and 2.30 MHz respectively for the four helical antennas. Comparatively, the prototype measurement results are consistent with results from the simulation. The VSWR performance of the simulation and prototype measurement is also plotted in Figure 6.33. The consistency between the measured results and those predicted through simulations validates the approach. The discrepancies are attributed to the underestimated losses in the matching networks as well as the higher ohmic losses of the aluminium wire.

Table 6.20. Bandwidth comparison of simulations and measurements at 100 MHz.

Bandwidth	Uniform Reference	Radius Variation	Pitch Variation	Radius and Pitch Variations
Simulation	1.50 MHz	2.05 MHz	1.95 MHz	2.30 MHz
Measurement	1.80 MHz	2.55 MHz	2.50 MHz	2.50 MHz

6.9 Conclusions

This chapter has considered the optimization of helical antennas operating at HF and VHF frequencies. Structural variations in terms radius and pitch were proved to be

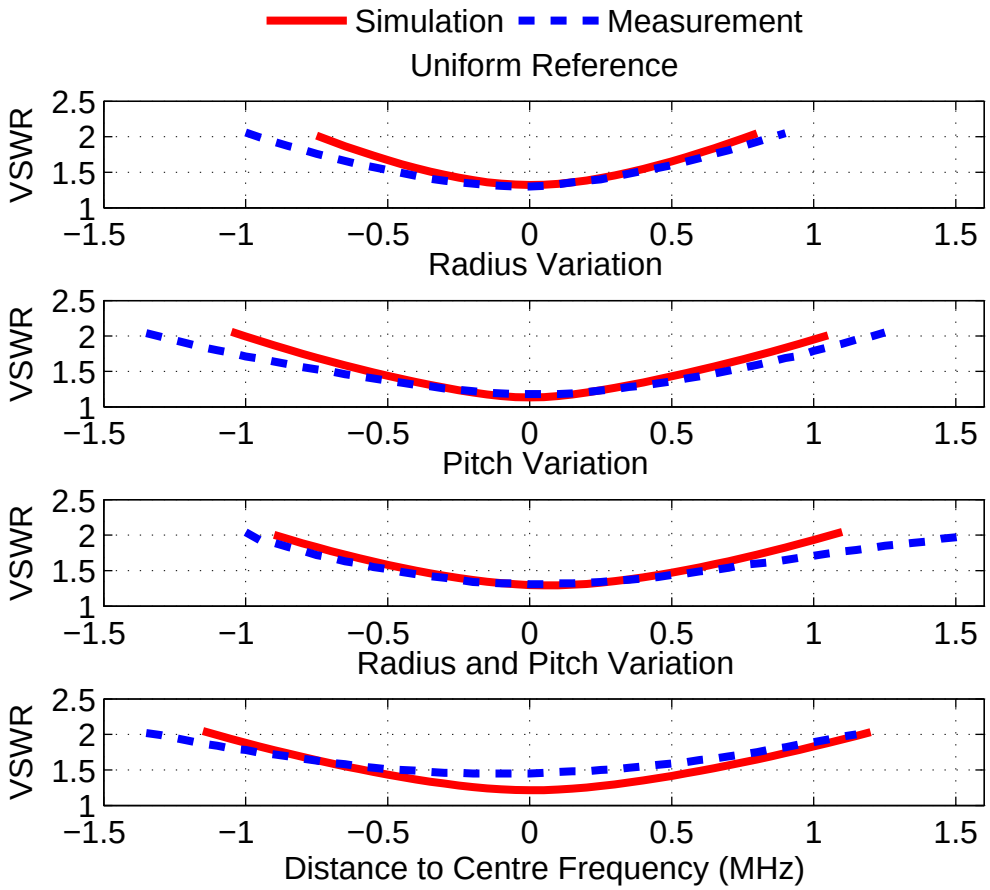

Figure 6.33. Bandwidth of optimized helical antennas from simulation and measurement at 100 MHz.

an effective and practical way to introduce distributions of inductance along a helical antenna. By employing radial basis expansions, sufficient structural variations can be investigated on the antennas with few optimization variables. Near-optimal helical configurations were found at three different frequencies by applying an in-house developed tool coupling an electromagnetic simulator to an evolutionary optimizer. A set of the proposed optimal helical antennas were fabricated and measured, with experimental results indicating a good agreement between simulation and measurement. It was successfully verified that helical antennas can be optimized to perform with wider bandwidth while maintaining good efficiency. With this satisfactory achievement, the question now arises as to how the optimal helical antennas perform in a non-ideal environment. Explicitly, when the infinite conducting ground plane is replaced by more

complex material and/or structure (e.g. a hand-held device used as the ground plane). Therefore, the next chapter will continue to investigate those relevant issues.

Chapter 7

Impact of Non-ideal Environment & Mutual Interaction

THIS chapter extends the investigation of optimized helical antennas to include the impact of non-ideal environments, by probing a series of realistic operation scenarios. Different types of ground conditions are first explored, including realistic soil ground and radial metallic wire ground. Radial wire screens are often used as an alternative to realize a conducting ground, and have shown great effectiveness in mimicking a solid planar metallic ground. The impact on the antenna performance of metallic objects in its vicinity is also investigated. A case study is performed in order to design an optimal VHF helical antenna on a hand-held device operating at 300 MHz. The achieved results successfully verify the antenna design methodology and the role of non-ideal ground plane in realistic applications. The mutual interaction between two identical optimal helical antennas is also investigated. Scattering parameters of the whole antenna system including the matching networks and transmission between antennas are explicitly described and cascaded. The results indicate that the optimal helical antennas do not increase the mutual interaction when used in an array condition, in comparison to the reference case of a quarter-wave monopole array.

7.1 Introduction

In the previous chapters, helical antennas have been investigated and optimal designs have been obtained by varying their structural features along the length. This approach has demonstrated enhanced performance for the helical antenna operating in the normal mode as shortened resonant omnidirectional antennas for HF and VHF communications. Nevertheless, most helical antennas are mounted and used on non-ideal ground planes, e.g. on portable devices. In contrast, the optimized helical antennas obtained in the previous chapter were designed, simulated, and optimized under the assumptions of ideal environment, i.e. an infinite perfectly conducting ground plane, and no near-by objects. Such environments, however, do not exist in the real world. Consequently, the impact of non-ideal environments upon the robustness of the proposed optimizations need to be considered, and this builds the subject of this chapter.

First, we consider the impact of having the perfect infinite conducting plane replaced by a realistic soil ground or a planar system of radial wires with finite length. Realistic soil grounds can be categorized as ”good”, ”average”, and ”poor” quality based on their dielectric constants and conductivities. The conductivity quality of the realistic soil ground is dominantly determined and influenced by the water content [82]. The bandwidth and radiation efficiency are taken as the performance indicators for comparison. As expected, on an imperfect ground the radiation efficiency decreases significantly due to the attenuation of surface waves, and bandwidth increases as a tradeoff. A radial wire screen is a commonly used alternative ground plane configuration for HF and VHF antennas, and it can effectively increase the efficiency in comparison to realistic soil ground. Other advantages of a radial wire screen also include its lower weight, and the simplicity for assembly compared to a solid metal ground.

Second, the impact of near-by metallic objects on the antenna performance is investigated in terms of objects geometry, dimensions, distance, orientation. The mesh size of the objects in the computation is also considered. In NEC-2, metal sheets are modelled as wire meshes, and thus metal objects can be constructed using this approximation. A cubic metal box with various dimensions is located and aligned horizontally to the optimized antenna, in order to find out the influence on the radiation performance.

When a re-optimization is carried out including a metallic box, the similarity of the results for the optimal helical shape compared to the optimal antenna in free-space, demonstrates the robustness of the antenna design methodology.

Furthermore, we consider the case where the ground plane is replaced by the metal casing of a portable device, to consider the operation of the antenna from a more practical viewpoint. A pitch-varying helical antenna with a fixed height of 0.08 m is to be optimized on a hand-held metal case for operation at 300 MHz. The goal of the design is to maximize the efficiency and bandwidth under perfect matching by continuously varying the pitch only, as the application of the radius expansion strategy would be impractical. An optimized antenna is obtained that exhibits a satisfactory omnidirectional radiation pattern.

Finally, the mutual interaction between two identical optimal helical antennas is explored. The scattering parameters between two quarter-wave monopole antennas are used as the reference for comparison purpose. The whole antenna system is described using S parameters in terms of S-matrices of the matching networks, and the transmission path between antennas. The results show that the operation of our optimized helical antenna do not increase the interference compared to the reference monopoles, when applied in an array mode. This further proves the effectiveness and reliability of the optimization procedure.

7.2 Impact of Ground Conditions

In the present work, NEC-2 is used to model and simulate the environment as well as the antenna. A 3-turn uniform helical antenna resonant at 100 MHz and with total height of 0.25 m is taken as the reference model for comparison purpose. The optimized uniform pitch (P) and radius (R) are obtained by parameter sweeping, which yields P = 0.07 m and R = 0.04 m. This reference helix is simulated above an infinite conducting ground plane in NEC-2, and simulation results yield a radiation efficiency of 71.9% and bandwidth of 1.50 MHz with a center frequency of 100 MHz.

7.2.1 Realistic Soil Ground

When the infinite conducting ground is replaced by a realistic soil ground, the surface wave is attenuated, which results in a dramatic drop in the radiation efficiency. Importantly, The term 'radiation efficiency' is defined as the integration of the energy of all fields radiated from the antenna in relation to the power input to the antenna. With respect to different electromagnetic properties of the soil, the ground can be categorized as good, average and poor based on its relative permittivity ϵ_r and dc-conductivity σ, as shown in Table 7.1.

Table 7.1. Realistic soil property of various types.

	Good	Average	Poor
Relative Permittivity ϵ_r	4	13	5
Conductivity σ	0.010 S/m	0.005 S/m	0.001 S/m

The four types of optimized helical antenna, namely uniform reference, radius variation, pitch variation, radius and pitch variations, are simulated with a realistic soil ground. The resulting radiation efficiency and bandwidth performance are listed in Table 7.2.

It is important to point out that any energy loss whether it is in the elements or in the ground will contribute to the resistive element of the impedance measured at the drive point. Therefore, it is observed that when realistic soil grounds are used, the efficiency decreases significantly as predicted, and the operation bandwidth increases dramatically as a trade-off. Such findings can be verified using the methods introduced in [83]. It is therefore realized that a realistic soil ground is not applicable for practical operation, and thus a ground with metallic conducting properties is essentially required.

7.2.2 Radial Wire Ground

A radial wire screen is an alternative practical way to replace the perfect conducting ground in simulation and operation. The uniform reference antenna mounted on a sample radial wire screen composed of 16 wires is illustrated in Figure 7.1. Radial

Table 7.2. *Performance comparison of optimized helical antennas between perfect ground and realistic soil ground at 100 MHz.*

	Uniform Reference		Radius Variation	
	Efficiency	Bandwidth	Efficiency	Bandwidth
Perfect Ground	71.9%	1.5 MHz	85.0%	2.1 MHz
Good Soil	0.1%	45.0 MHz	0.1%	45.0 MHz
Average Soil	2.0%	7.1 MHz	0.1%	45.0 MHz
Poor Soil	1.2%	4.0 MHz	1.8%	4.1 MHz
	Pitch Variation		Radius and Pitch Variations	
	Efficiency	Bandwidth	Efficiency	Bandwidth
Perfect Ground	74.3%	1.95 MHz	77.8%	2.30 MHz
Good Soil	0.094%	45.0 MHz	0.144%	45.0 MHz
Average Soil	2.49%	7.30 MHz	2.91%	7.50 MHz
Poor Soil	1.50%	4.10 MHz	1.78%	4.30 MHz

wires can effectively sustain the propagation of surface waves and thus increase the radiation efficiency. A screen of radial wires can be defined with several parameters including the number of wires (N), the radius of the screen (R_{screen}), and the wire thickness (R_{wire}). In this section, these three parameters are explicitly investigated in terms of their impact towards performance. Specifically, in the investigation, the number of wires takes the values of 4, 8, 16, 64, 256 for screen radii of 0.25 m and 1 m, with wire thicknesses of 1 mm and 10 mm. The uniform reference antenna is simulated with the different radial screens, to test the radiation efficiency difference as shown in Table 7.3. When the number of radial wires or the wire radius increases beyond a certain limit, the radial wires start to overlap, and NEC-2 cannot provide estimation, and thus the results are denoted as 'N/A' in the table.

Particularly, the impact of the radius of the plane is further probed by comparing the results to the ones using a 0.50 m radial screen with 8 and 16 wires. The comparison

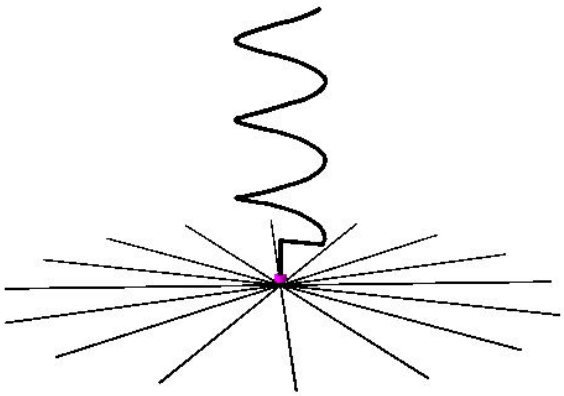

Figure 7.1. Illustration of a radial wire screen composed of 16 wires with a uniform reference helical antenna.

Table 7.3. *Radiation efficiency comparison of a radial wire screen of different features with a uniform reference antenna at 100 MHz.*

Radiation efficiency	R_{screen} = 1 m R_{wire} = 1 mm	R_{screen} = 0.25 m R_{wire} = 1 mm	R_{screen} = 0.25 m R_{wire} = 10 mm
$N = 4$	11.3%	11.1%	29.7%
$N = 8$	25.4%	24.4%	45.9%
$N = 16$	46.2%	40.8%	53.3%
$N = 64$	75.1%	54.8%	N/A
$N = 256$	80.6%	N/A	N/A

indicates a good and consistent trend by enlarging the size of the radial screen, as shown in Table 7.4 and Figure 7.2.

It can be seen that, a) the greater the numbers of wires, the higher radiation efficiency can be achieved; b) when the same number of wires are used, a larger screen yields higher radiation efficiency; c) thicker wires lead to higher efficiency for screens of the same size. This test successfully verifies the effectiveness of a radial wire screen for use as the conducting ground plane for helical antennas.

Table 7.4. Efficiency comparison of a uniform helix on a radial wire screen with different radius.

R_{screen}	0.25 m	0.50 m	1.00 m
$N = 8$	24.4%	25.0%	25.4%
$N = 16$	40.8%	43.1%	46.2%

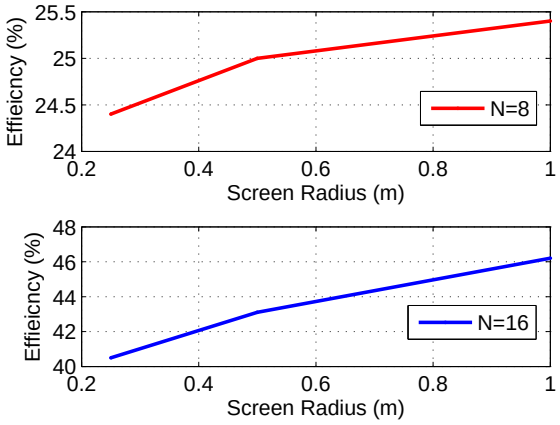

Figure 7.2. Efficiency of a uniform helix on a radial wire screen with different radius.

7.3 Impact of Near-by Metallic Objects

To probe the impact of near-by objects, a metallic box is placed in the vicinity the optimal antenna and simulated in NEC-2. This metallic box is represented using meshed wires, and therefore, the sensitivity of the arrangement of meshed wires to model metallic plate is firstly investigated. The metallic box is then varied in terms of its dimension and distance to the antenna. The antenna placement above the box is also optimized for an optimal loading position. With the optimized mesh size and loading position, the optimized helical antennas obtained from previous investigations are re-optimized on top of a metallic box. The original optimal antennas are compared to the re-optimized helical antennas, and the similarity of the obtained helical geometries verifies the robustness and reliability of the design methodology.

7.3.1 Modeling of Near-by Metallic Objects

- **Mesh Size**

  It is generally accepted that when the mesh size (d_{mesh}) is equal or less then one-tenth of a wavelength ($\frac{\lambda}{10}$), i.e. $d_{mesh} \leq \frac{\lambda}{10}$, the mesh is dense enough to model a metallic plate when the method of moments (MoM) is applied in NEC-2. In the current investigation, at the operating frequency of 100 MHz, $\frac{\lambda}{10}$ is correspondingly 300 mm. Nevertheless, in the present application, a mesh size of 300 mm appears not sufficiently fine to model the metal box accurately. Therefore, a mesh size of 25 mm, 50 mm, 100 mm, and 150 mm is tested to observe the sensitivity of the result to the wire-mesh discretization fineness. The uniform reference helical antenna is used for this purpose and mounted at the center point on top of a metal box. The metal box has dimensions of Width $\times$ Length $\times$ Height = 0.3 m $\times$ 0.3 m $\times$ 0.1 m. The radiation efficiency and bandwidth performance are compared to the case of an infinite conducting ground. The results are listed in Table 7.5, and the configurations are shown in Figure 7.3 .

Table 7.5. *Sensitivity test of the mesh size with the reference uniform helix at 100 MHz.*

d_{mesh}	No Box	150 mm	100 mm	50 mm	25 mm
Efficiency	71.9%	59.4%	60.0%	60.9%	60.5%
Bandwidth	1.5 MHz	1.0 MHz	1.0 MHz	1.0 MHz	1.0 MHz

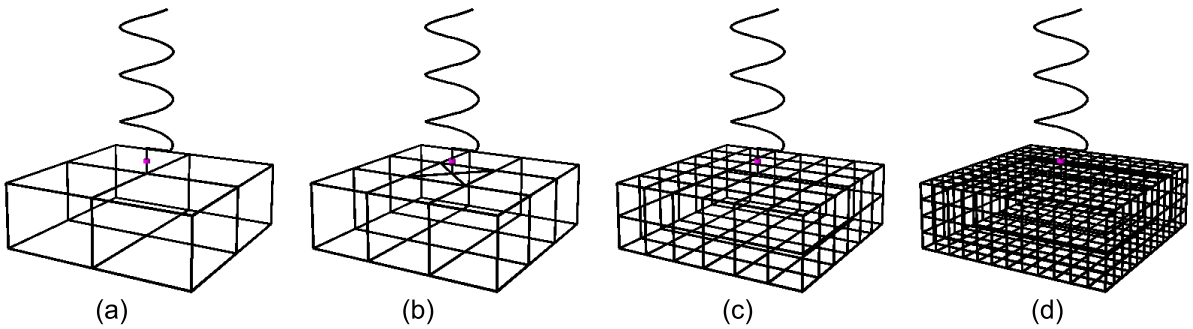

Figure 7.3. Illustrations of a metal box of different mesh size with a uniform reference helical antenna: (a) 150 mm, (b) 100 mm, (c) 50 mm, (d) 25 mm.

From the results in the table, it is verified that a mesh size equal or less than 100 mm can lead to stable results. However, considering the possibility of a further

size reduction of the metal box, 50 mm (or $\frac{\lambda}{60}$) is used as the standard mesh size at 100 MHz in the following investigation.

- **Dimension**

  The impact of near-by objects for different overall dimensions is then investigated. The cubic metal box is aligned 10 mm next to the uniform helical antenna, and 10 mm above the perfect conducting ground. This increased distance is necessary, as NEC-2 does not allow wires to touch the ground. The dimension of the box (d_{box}) is varied to take the values d_{box} = 50 mm, 100 mm, 200 mm, and 400 mm. The radiation efficiency and bandwidth performance are compared to the case with no near-by objects. The results are listed in Table 7.6, and the configurations are shown in Figure 7.4. An alternative practical scenario would consider a lossy dielectric object, e.g. a head or hand, which however, NEC-2 can not provide a solution for. The metallic box can in this sense be considered as a worst case scenario.

Table 7.6. *The impact of a metal box due to different dimensions towards the performance of the reference uniform helix at 100 MHz.*

d_{box}	No Box	50 mm	100 mm	200 mm	400 mm
Efficiency	71.9%	71.5%	70.8%	72.4%	82.8%
Bandwidth	1.50 MHz	1.45 MHz	1.30 MHz	0.95 MHz	1.00 MHz

It is realized that when the dimension of the metal box increases, it starts to act as a reflector, and thus improves the radiation efficiency at the cost of degradation in bandwidth. It is therefore critical to consider the distance between the antenna and the object, which will be discussed in the next section. More importantly, a metallic body with a considerable dimension will affect the radiation pattern of the antenna as shown in Figure 7.5, where the gain patterns for the uniform reference with various dimensions of nearby metallic box are displayed. 7.5.

- **Distance**

  In this section, the dimension d_{box} of the metal box is fixed at 200 mm, which is considered as a reasonable size representing an HF transceiver. It is still located

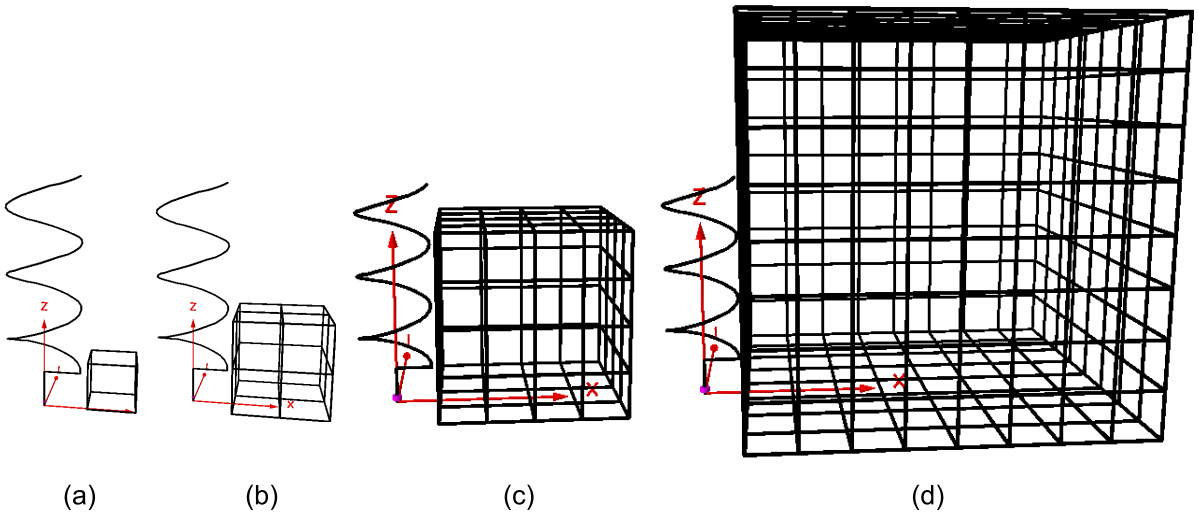

Figure 7.4. Illustrations of a metal box of different dimensions d_{box} next to a uniform reference helical antenna: (a) 50 mm, (b) 100 mm, (c) 200 mm, (d) 400 mm.

10 mm above the perfect ground, and the horizontal distance (d_{hor}) to the uniform reference helical antenna is set to 10 mm, 50 mm, 100 mm, and 200 mm. The radiation efficiency and bandwidth performance are compared the case with no near-by objects. The comparisons are listed in Table 7.7 with the configurations shown in Figure 7.6.

Table 7.7. *The impact of a 200 mmm cubic metal box towards the performance of the reference uniform helix at 100 MHz due to different horizontal distance.*

d_{hor}	10 mm	50 mm	100 mm	200 mm	No Box
Efficiency	72.4%	69.1%	69.2%	70.3%	71.9%
Bandwidth	0.95 MHz	1.10 MHz	1.30 MHz	1.35 MHz	1.50 MHz

It is noticed that when the metal box is aligned further away from the helical antenna, it has less impact towards the performance. Vise versa, closer placement will decrease the bandwidth and increase the radiation efficiency. Furthermore, the gain pattern of those four models are shown in Figure 7.7, indicating a negligible impact for larger distances as expected.

- **Loading Position Above the Box**

  When a helical antenna is mounted on top of a metal box, the loading position

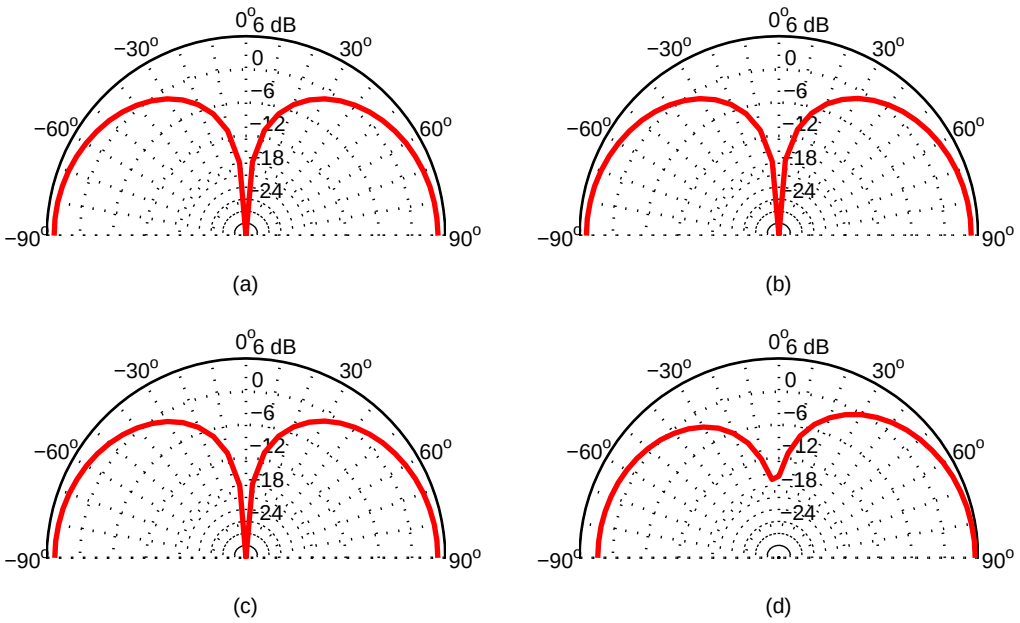

Figure 7.5. Gain patterns of a uniform reference helical antenna with a metal box of different dimensions d_{box} nearby: (a) 50 mm, (b) 100 mm, (c) 200 mm, (d) 400 mm.

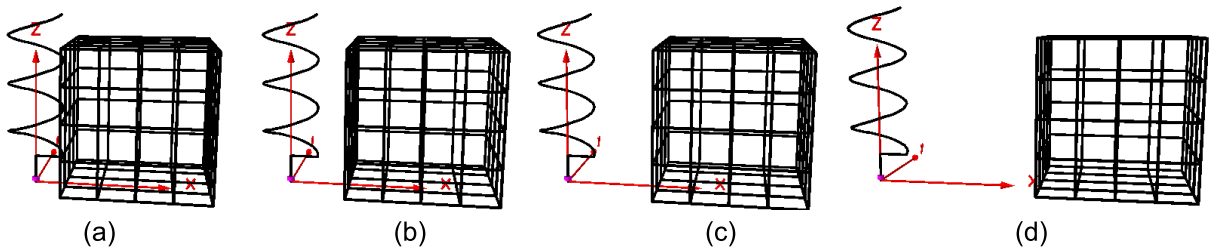

Figure 7.6. Illustrations of a metal box next to a uniform reference helical antenna at different horizontal distance d_{hor}: (a) 10 mm, (b) 50 mm, (c) 100 mm, (d) 200 mm.

may also have a strong impact towards its performance [84]. Therefore, the optimal loading position for an antenna above a rectangular metal box is investigated. To simplify this problem, the square top metal sheet is represented as a $4L \times 4L$ meshed grid and several nodes are selected as connection point for the antenna, as shown in Figure 7.8. The center is labeled as #0, and the other five nodes are labeled based on their distance to the center, as listed in Table 7.8. Considering the symmetries, the six positions can fully represent all the positions on the sheet.

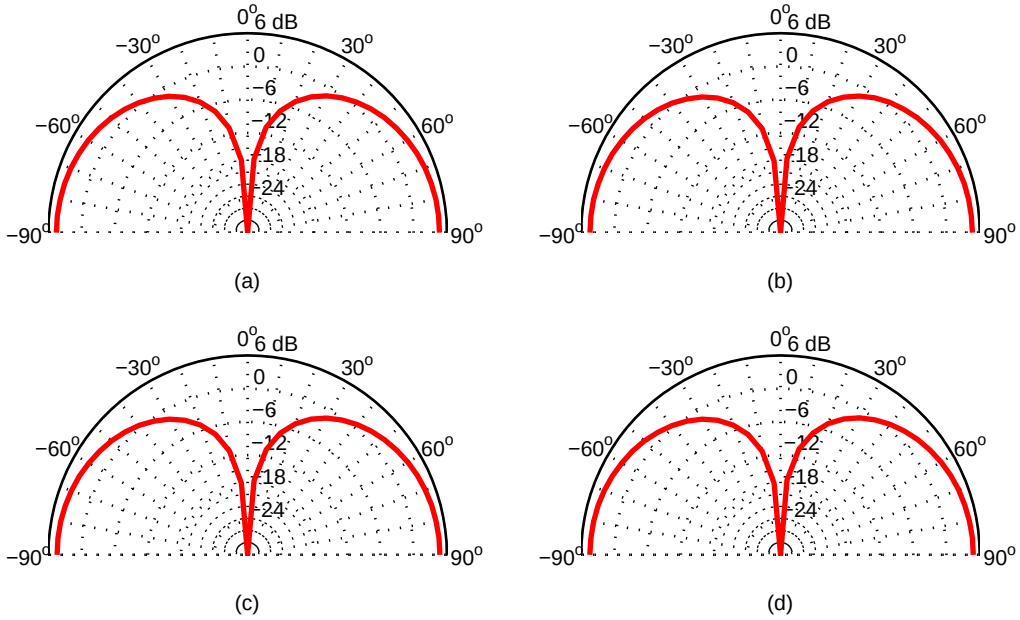

Figure 7.7. Gain patterns of a uniform reference helical antenna with a metal box nearby at different horizontal distance d_{hor}: (a) 10 mm, (b) 50 mm, (c) 100 mm, (d) 200 mm.

Table 7.8. *The positions selected on the square sheet and their distance to center.*

Loading position	#0	#1	#2	#3	#4	#5
Distance to center	0	L	$\sqrt{2}L$	$2L$	$\sqrt{5}L$	$2\sqrt{2}L$

In the simulation, the dimension of the metal box is set to $Width \times Length \times Height = 0.3m \times 0.3m \times 0.1m$, and thus one grid is equal to L = 75 mm. The radiation efficiency and bandwidth performance are compared at all positions and listed in Table 7.9.

It is observed that the further away the antenna is mounted from the center, the better performance it can achieve. Particularly, it exhibits maximized radiation efficiency and bandwidth with appropriate matchings when the antenna is mounted at the corner of the metal sheet. Therefore, the corner position is taken as the preferred loading position for antenna in the later investigation.

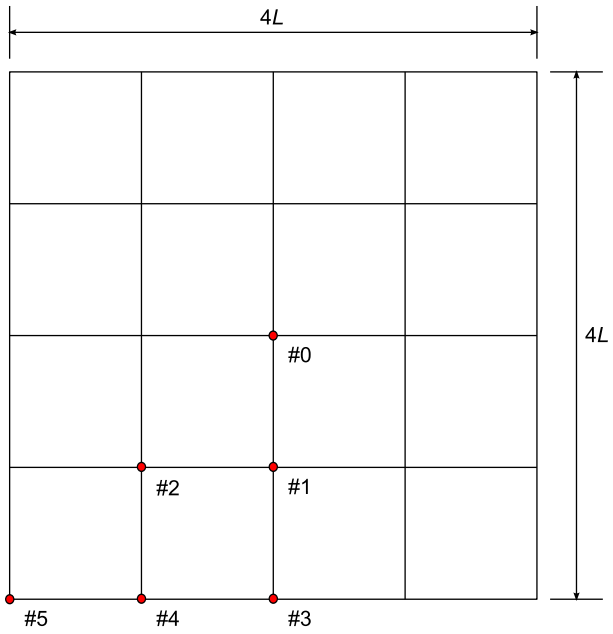

Figure 7.8. Illustrations of the possible loading position on a square metal sheet.

Table 7.9. *The impact loading position towards the performance of the reference uniform helix at 100 MHz.*

	#0	#1	#2	#3	#4	#5
Efficiency	60.9%	63.3%	65.6%	69.4%	71.3%	76.0%
Bandwidth	1.0 MHz	1.2 MHz	1.2 MHz	1.5 MHz	1.6 MHz	2.0 MHz

7.3.2 Re-optimization on Top of Metallic Box

In order to probe the robustness of the optimization procedure, the proposed optimal helical antennas are re-simulated and re-optimized with a metal box involved. In this process, the infinite ground is replaced by a metal box (Width $\times$ Length $\times$ Height = 0.3 m $\times$ 0.3 m $\times$ 0.1 m) with a mesh size of 50 mm, while the top surface is used as the ground. For comparison purpose, each optimized helical antenna is firstly mounted on top of the metal box and simulated to obtain the performance. This performance is then compared to the original performance on an infinite perfect ground plane, i.e. where no box is included, and then to the performance from the helical antenna re-optimized with the metal box included.

- **Reference Uniform Helix**

  The reference uniform helix is designed through a parameter sweeping process carried out with respect to radius and pitch. The only difference with previous results is the replacement of infinite conducting ground by the metal box. The procedure results in identical optimal antenna configuration with or without the metal box, as listed in Table 7.10. The configurations are illustrated in Figure 7.9.

Table 7.10. *The comparison the optimized reference uniform helix with and without a metal box at 100 MHz.*

	Original optimization without a metal box	Simulation of original helix on a metal box	Re-optimization with a metal box
Radius	40 mm		40 mm
Pitch	70 mm		70 mm
Efficiency	71.9%	60.9%	60.9%
Bandwidth	1.50 MHz	1.02 MHz	1.02 MHz

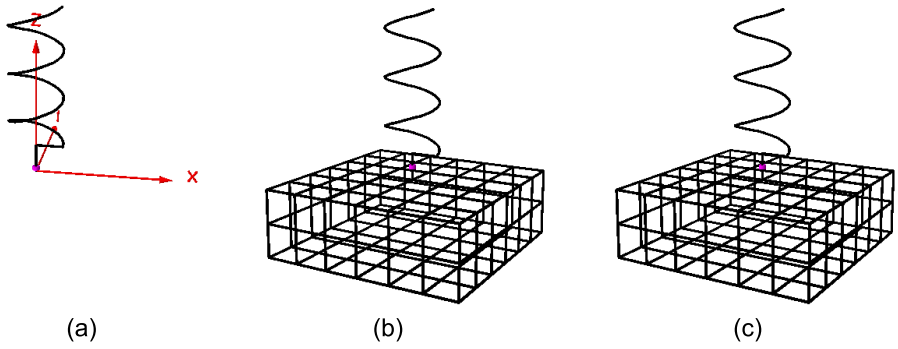

Figure 7.9. Configurations comparison of the original and re-optimized uniform helices.

- **Helix with Radius Variation**

  Similarly, the helical antenna with radius variation is re-optimized with the metal box included. The optimization settings remain identical to previous investigations in Chapter 6. The results are listed in Table 7.11 with the configurations shown in Figure 7.10. Comments and discussion will be provided later together with other types of helical variations.

Table 7.11. *The comparison the optimized radius-varying helix with and without the metal box at 100 MHz.*

	Original optimization without a metal box	Simulation of original helix on a metal box	Re-optimization with a metal box
Radius (ω_r)	$\omega_1 = 0.010$ m $\omega_2 = 0.010$ m $\omega_3 = 0.010$ m $\omega_4 = 0.010$ m $\omega_5 = 0.091$ m		$\omega_1 = 0.010$ m $\omega_2 = 0.010$ m $\omega_3 = 0.010$ m $\omega_4 = 0.020$ m $\omega_5 = 0.080$ m
Pitch	70 mm		
Efficiency	85.0%	73.7%	76.1%
Bandwidth	2.05 MHz	1.50 MHz	1.45 MHz

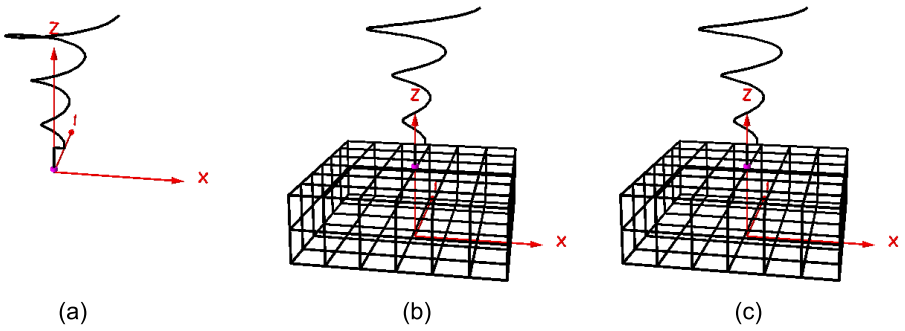

Figure 7.10. Configurations comparison of the original and re-optimized radius-varying helices.

- **Helix with Pitch Variation**

  The helical antenna with pitch variation is also re-optimized with the metal box included, while optimization settings are same as previous investigation. The results are listed in Table 7.12, and the configurations are shown in Figure 7.11. The similarity in the results will be discussed later.

- **Helix with Radius & Pitch Variations**

  Finally, the helical antenna with both radius and pitch variations is re-optimized with the metal box included. The results are listed in Table 7.13 with the configurations shown in Figure 7.12.

Table 7.12. *The comparison the optimized pitch-varying helix with and without the metal box at 100 MHz.*

	Original optimization without metal box	Simulation of original helix on metal box	Re-optimization with metal box
Radius	40 mm		
Pitch (ω_p)	$\omega_1 = 0.070$ m $\omega_2 = 0.070$ m $\omega_3 = 0.010$ m $\omega_4 = 0.010$ m $\omega_5 = 0.010$ m		$\omega_1 = 0.070$ m $\omega_2 = 0.070$ m $\omega_3 = 0.014$ m $\omega_4 = 0.010$ m $\omega_5 = 0.010$ m
Efficiency	74.3%	62.6%	62.6%
Bandwidth	1.95 MHz	1.30 MHz	1.28 MHz

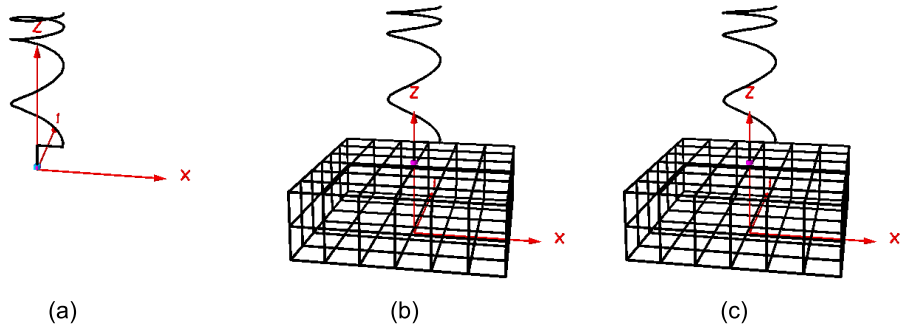

Figure 7.11. Configurations comparison of the original and re-optimized pitch-varying helices.

It can be seen that when the metal box is used as the ground plane, the antenna performance is degraded to some extends. More importantly, it can be also realized from those four types of helical antennas that when the metal box is included in the optimization procedure, the resulting optimal antenna configurations are very similar to those obtained in the scenario where an infinite ground plane is used. This similarity effectively proves the robustness and reliability of the optimal design previously achieved, and also, verifies the success of the methodology and optimization applied in this thesis.

Table 7.13. *The comparison the optimized radius-pitch-varying helix with and without the metal box at 100 MHz.*

	Original optimization without metal box	Simulation of original helix on metal box	Re-optimization with metal box
Radius (ω_r)	$\omega_1 = 0.010$ m $\omega_2 = 0.010$ m $\omega_3 = 0.010$ m $\omega_4 = 0.022$ m $\omega_5 = 0.080$ m		$\omega_1 = 0.010$ m $\omega_2 = 0.010$ m $\omega_3 = 0.015$ m $\omega_4 = 0.024$ m $\omega_5 = 0.078$ m
Pitch (ω_p)	$\omega_1 = 0.040$ m $\omega_2 = 0.040$ m $\omega_3 = 0.020$ m $\omega_4 = 0.010$ m $\omega_5 = 0.010$ m		$\omega_1 = 0.070$ m $\omega_2 = 0.070$ m $\omega_3 = 0.010$ m $\omega_4 = 0.010$ m $\omega_5 = 0.010$ m
Efficiency	77.8%	67.4%	71.2%
Bandwidth	2.30 MHz	1.50 MHz	1.45 MHz

7.4 Case Study: Optimized VHF Antennas on a Hand-held Device

From a more practical point of view, VHF antennas are often used on portable communication devices (a walkie-talkie for example) where the supporting metal box is much smaller than the box considered in previous section. In this section, a VHF antenna operating at 300 MHz with a fixed height of 8 cm is to be mounted on a hand-held device which has a dimension of Width $\times$ Length $\times$ Height = 5 cm $\times$ 6 cm $\times$ 10 cm. This helical antenna can be designed with optimal continuous pitch variation, while the radius is fixed at 1 cm, considering that an expanding radius helix is impractical for portable usage. The configuration of the antenna and hand-held device case is illustrated in Figure 7.13. The length of the antenna corresponds to $\frac{\lambda}{12.5}$ at 300 MHz. The goal of the optimization is to obtain a helical antenna with maximized efficiency and bandwidth.

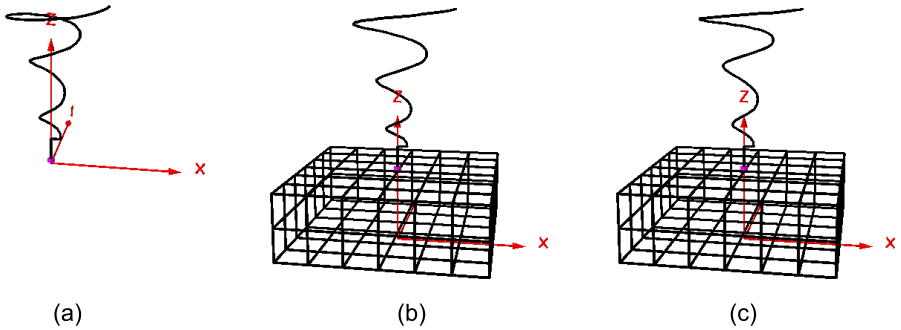

Figure 7.12. Configurations comparison of the original and re-optimized radius-pitch-varying helices.

In an initial step, a uniform helix is optimized with respect to its number of turns (determined by the pitch) on an infinite ground plane. This optimized uniform helix has a pitch of 2.33 cm (correspondingly 3 turns) with a radius of 1 cm. The same helix antenna when mounted at the center on the top of the metal case is used as the reference, as shown in Figure 7.14(a). From the NEC-2 simulation, it is expected to exhibit an efficiency of 94.1% and a bandwidth of 0.75 MHz, as listed in Table 7.14.

In the second step, the optimization process is now carried out by allowing the pitch to continuously vary for the infinite ground. The optimized helix is then also positioned on top of the metal box to find its corresponding performance, as shown in Figure 7.14(b). This optimized helix with pitch variation outperforms the uniform reference helix when mounted on the metal case, with a maintained efficiency of 92.0% and improved bandwidth of 1.14 MHz, as listed in Table 7.14. Comparatively, a uniform three-turn helix will exhibit an efficiency of 75.6% and a bandwidth of 0.06 MHz when used in free space.

Table 7.14. *Comparison of uniform and optimized helical antennas on a metal case.*

	Original optimization without metal case	Simulation of original helix on metal case	Re-optimization with metal case
Efficiency	94.1%	92.0%	92.0%
Bandwidth	0.75 MHz	1.14 MHz	1.14 MHz

In the third step, a re-optimization is performed for the antenna on the metal case as a whole unit. The result shows a very similar optimal antenna configuration to the one

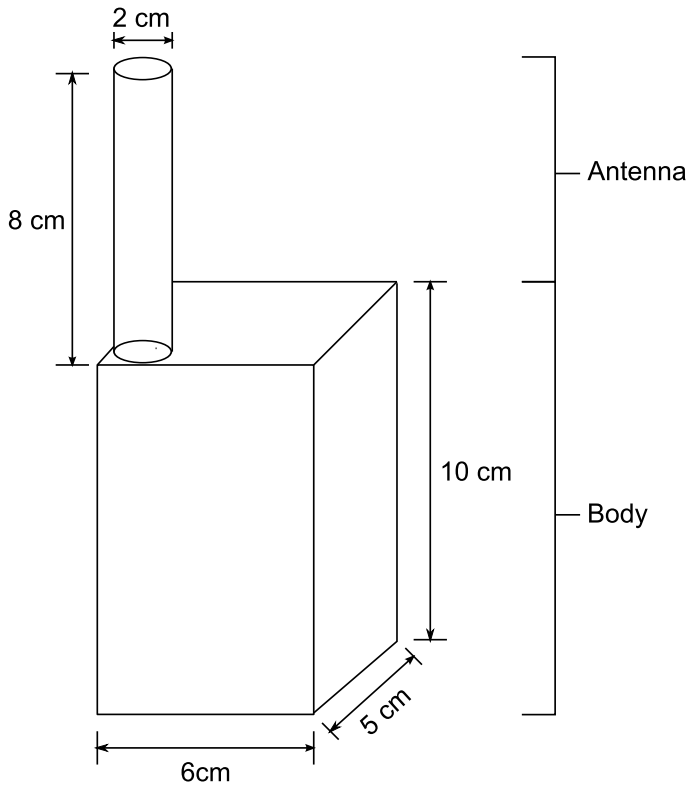

Figure 7.13. Illustration of the antenna on a hand-held device.

for infinite ground in Figure 7.14(c). Such similarity effectively proves the the robustness of the optimization, and provides a guideline in designing pitch-varying helix. Specifically, by altering the pitch distribution along the antenna height and introducing denser winding at the top of the monopole and looser winding near its bottom, the bandwidth can be increased.

Finally, the optimized helix is positioned on the side and at the corner of the metal case. It is found that altering the loading position can change the antenna performance significantly, especially in bandwidth (shown in Table 7.15). By mounting the antenna on the side or at the corner of the metal case, the resulting bandwidth can be greatly improved to 4.09 and 10.7 MHz respectively. The gain pattern of the optimized example with the metal case is shown in Figure 7.15. The light tilt of the pattern axis is caused by the placement of the antenna at the corner of the case. Such findings have been mentioned in [84].

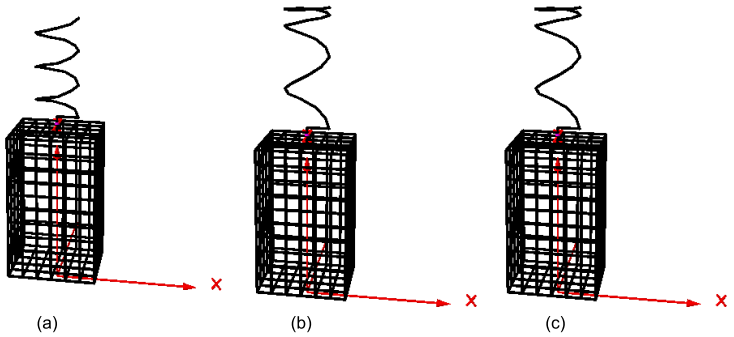

Figure 7.14. Illustration of uniform and optimized helical antennas on a metal case.

Table 7.15. *Comparison of optimized helix at different loading position on the metal case.*

	Center	Side	Corner
Efficiency	92.%	94.9%	97.0%
Bandwidth	1.14 MHz	4.09 MHz	10.7 MHz

7.5 Mutual Interaction Between Optimized Helical Antennas

When two antennas are used in a row, they can be related to the case of mutual coupling in the near field or the transmission in the far field, depending on their separation distance. This section will investigate both cases of 1) near-field mutual coupling, and 2) far-field transmission in according to various separation as a function of wavelength. When two identical optimized helical antennas are operated with perfect matching, one is used as the transmitting element and the other for receiving, as illustrated in Figure 7.16. The whole system can be described by three 2-port subsystems, namely the matching network #1 which connects the input port (50 Ω) to the transmitting antenna, the transmission between the two antennas, and the matching network that connects the antenna to the output (50 Ω). The characteristics of the three subsystems can be interpreted using scattering parameters S_{mn1}, S_{tx}, and S_{mn2} respectively. By cascading the S matrices of the three subsystems, the S parameters of the whole system can be computed.

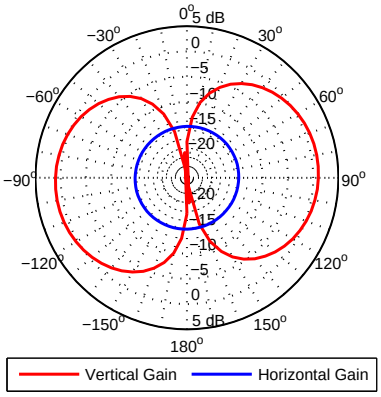

Figure 7.15. Gain pattern of the optimized helix with pitch variation on a metal case.

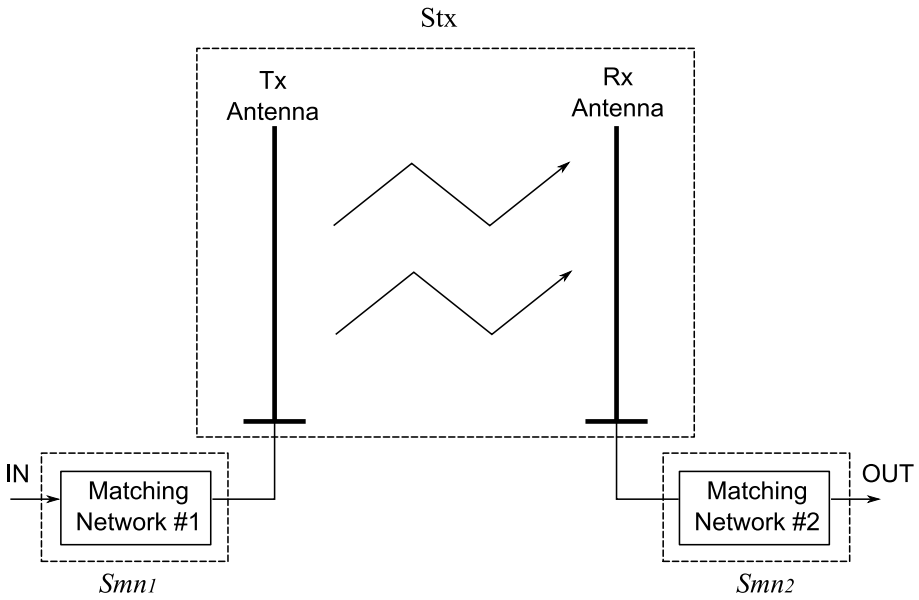

Figure 7.16. Illustration of dual antenna system with three subsystems.

7.5.1 Scattering Parameters of Transmission between Antennas

In NEC-2, the two identical antennas are simulated with one of the antenna excited with a 1V voltage source, as illustrated in Figure 7.17. Therefore, the Z parameters of the system can be calculated as $Z_{11} = \frac{V_1}{I_1}$ and $Z_{21} = \frac{V_1}{I_2}$. Due to symmetry arising from the use of two identical antennas , the other parameters in the Z-matrix (z matrix) can be obtained as $Z_{11} = Z_{22}$ and $Z_{21} = Z_{12}$.

Considering that the two identical antennas are both matched at their terminals, the impedance at the two terminals Z_{01} and Z_{02} can be written as $Z_{01} = Z_{02} = Z_0$, and the

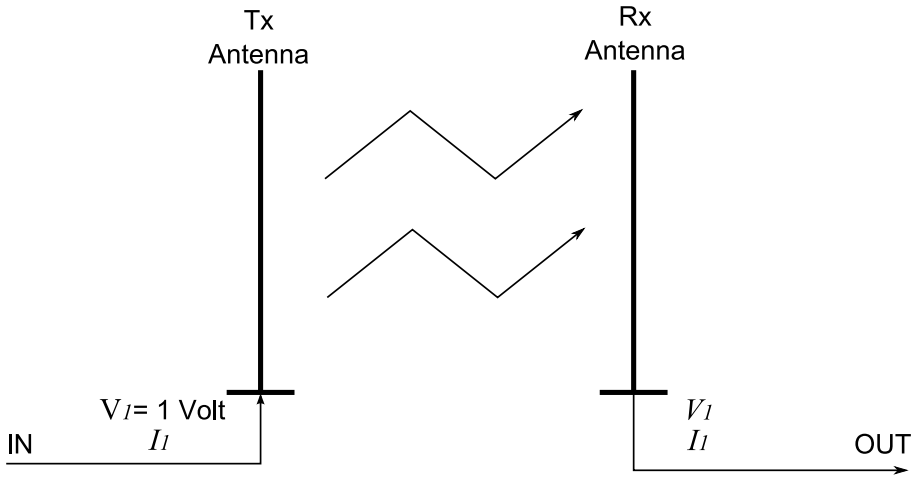

Figure 7.17. Illustration of dual antenna system with excitation.

Z-matrix for the two-port system can be converted to an S-matrix according to [61]:

$$\begin{cases} S_{11} = \frac{(Z_{11}-Z_0)(Z_{22}+Z_0)-Z_{12}Z_{21}}{\Delta} \\ S_{12} = \frac{2Z_0Z_{12}}{\Delta} \\ S_{21} = \frac{2Z_0Z_{21}}{\Delta} \\ S_{22} = \frac{(Z_{11}+Z_0)(Z_{22}-Z_0)-Z_{12}Z_{21}}{\Delta} \end{cases} \tag{7.1}$$

where $\Delta = (Z_{11} + Z_0)(Z_{22} + Z_0) - Z_{12}Z_{21}$.

Therefore, the S-matrix (S_{tx}) for the transmission in the free space between the two antennas $\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$ can be computed from the Z-parameters.

7.5.2 Scattering Parameters of Matching Networks

Apart from the free space transmission between the two antennas, the matching networks on each antenna also contribute to the system scattering matrix. This section investigates the S parameters of typical L-type matching networks used in previous sections. As illustrated in Figure 7.18, a 2-port L-type matching network has one parallel component and one series component with impedance of Z_1 and Z_2 respectively.

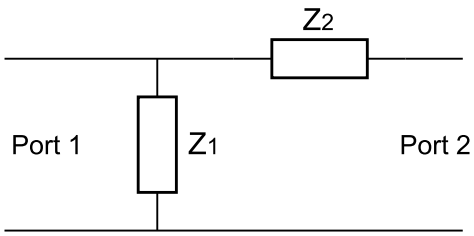

Figure 7.18. Illustration of a two-port L-type matching network.

- **Z-Matrix of Matching Networks**

  To derive the Z-matrix of a two-port matching network, we start from the network representation as shown in Figure 7.19. The voltage at port 1 and 2 can be

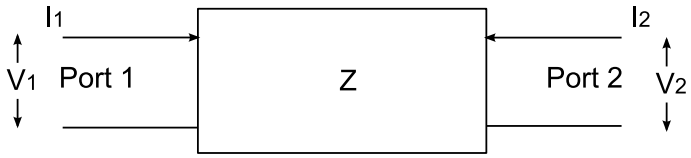

Figure 7.19. Illustration of a two-port matching network for Z-matrix.

represented as

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{7.2}$$

or explicitly,

$$\begin{cases} V_1 &= Z_{11}I_1 + Z_{12}I_2 \\ V_2 &= Z_{21}I_1 + Z_{22}I_2 \end{cases} . \tag{7.3}$$

The impedance between port 1 and 2 can be found as follows:

$$\begin{cases} Z_{11}=\frac{V_1}{I_1}|I_2 = 0 \\ Z_{12}=\frac{V_1}{I_2}|I_1 = 0 \\ Z_{21}=\frac{V_2}{I_1}|I_2 = 0 \\ Z_{22}=\frac{V_2}{I_2}|I_1 = 0 \end{cases} \tag{7.4}$$

where $I_1 = 0$ and $I_2 = 0$ refer to open circuit condition at port 1 and 2 respectively. These two cases are considered successively in the following.

Let us consider first the case where port 2 is open ($I_2 = 0$). Equation 7.3 can then be simplified to

$$\begin{cases} V_1 &= Z_{11} I_1 \\ V_2 &= Z_{21} I_1 \end{cases} . \tag{7.5}$$

Since Z_1 is loaded, and its voltage is fully delivered to port 2

$$\begin{aligned} V_2 &= Z_{21} I_1 \\ &= V_1 = Z_{11} I_1 \end{aligned} \tag{7.6}$$

which leads to

$$Z_{21} = Z_{11} . \tag{7.7}$$

Due to the fact that $Z_{11} = Z_1$ when looking into port 1 with port 2 open, it can be found that

$$Z_{21} = Z_{11} = Z_1 . \tag{7.8}$$

Now let us consider the second case, when port 1 is open ($I_1 = 0$). Equation 7.3 can be then written as

$$\begin{cases} V_1 &= Z_{12} I_2 \\ V_2 &= Z_{22} I_2 \end{cases} . \tag{7.9}$$

The L-type network of Figure 7.18 can then be regarded as a voltage divider with

$$V_1 = V_2 \frac{Z_1}{Z_1 + Z_2} . \tag{7.10}$$

Substituting V_1 and V_2 from Equation 7.9 into 7.10 yields

$$Z_{12} I_2 = Z_{22} I_2 \frac{Z_1}{Z_1 + Z_2} . \tag{7.11}$$

Eliminating the common element I_2 from both sides yields

$$Z_{12} = Z_{22} \frac{Z_1}{Z_1 + Z_2} . \tag{7.12}$$

Since $Z_{22} = Z_1 + Z_2$ when looking into port 2 with port 1 open, it is derived that

$$Z_{12} = Z_1 . \tag{7.13}$$

Therefore, the Z-matrix of the matching network is obtained

$$\begin{cases} Z_{11} &= Z_{12} = Z_{21} = Z_1 \\ Z_{22} &= Z_1 + Z_2 \end{cases} . \tag{7.14}$$

- **S-Matrix of Matching Networks**

  The scattering parameters of a two-port system can also be interpreted using S-matrix, as illustrated in Figure 7.20, where V_1^+ and V_2^+ are the ingoing voltages at port 1 and 2, while V_1^- and V_2^- are outgoing voltage.

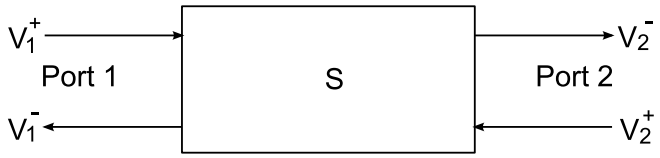

Figure 7.20. Illustration of a two-port L-type matching network for S-matrix.

The S-matrix can be found from V^+ and V^- since $V^- = SV^+$, and thus can be expanded as

$$\begin{bmatrix} V_1^- \\ V_2^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \end{bmatrix} . \tag{7.15}$$

Explicitly, this matrix operation can be written as

$$\begin{cases} V_1^- &= S_{11}V_1^+ + S_{12}V_2^+ \\ V_2^- &= S_{21}V_1^+ + S_{22}V_2^+ \end{cases} . \tag{7.16}$$

When a two-port network is used, each side can be matched as illustrated in Figure 7.21. First, when a matched load is used at port 2 ($V_2^+ = 0$), as shown in Figure 7.21(a).

The impedance looking into port 1 is

$$\begin{aligned} Z_{in1} &= Z_1 // (Z_2 + Z_0) \\ &= \frac{Z_1Z_2 + Z_0Z_1}{Z_0 + Z_1 + Z_2} \end{aligned} . \tag{7.17}$$

Thus, S_{11} is the reflection coefficient Γ_{in1} at port 1

$$S_{11} = \Gamma_{in1} = \frac{Z_{in1} - Z_0}{Z_{in1} + Z_0} . \tag{7.18}$$

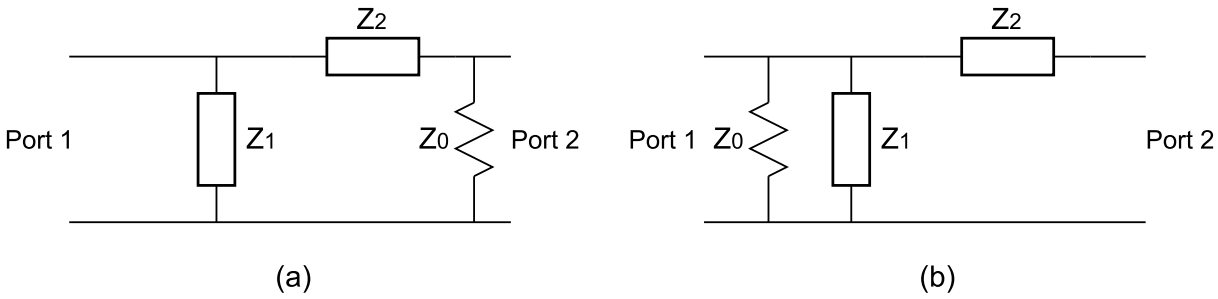

Figure 7.21. Illustration of a two-port L-type matching network with matched load at (a) port 2, and (b) port 1.

Substituting Z_{in1} from Equation 7.17 into 7.18, the reflection coefficient becomes

$$S_{11} = (\frac{Z_1Z_2 + Z_0Z_1}{Z_0 + Z_1 + Z_2} - Z_0)/(\frac{Z_1Z_2 + Z_0Z_1}{Z_0 + Z_1 + Z_2} + Z_0) \tag{7.19}$$

which can be expanded and simplified into

$$S_{11} = \frac{Z_1Z_2 - Z_0Z_2 - Z_0^2}{2Z_0Z_1 + Z_0Z_2 + Z_1Z_2 + Z_0^2} \quad . \tag{7.20}$$

Since port 2 is matched i.e. $V_2^+ = 0$, thus $V_2 = V_2^-$. Based on the voltage divider, the voltage V_2^- can be written as

$$V_2^- = V_2 = V_1\frac{Z_0}{Z_0 + Z_2} \quad . \tag{7.21}$$

At port 1, the voltage can be written as

$$\begin{aligned} V_1 &= V_1^+ + V_1^- \\ &= V_1^+ + S_{11}V_1^+ \quad . \end{aligned} \tag{7.22}$$

V_1^+ can be expressed as

$$V_1^+ = \frac{V_1}{1 + S_{11}} \quad . \tag{7.23}$$

Therefore, S_{21} can be obtained

$$S_{21} = \frac{V_2^-}{V_1^+} = (V_1\frac{Z_0}{Z_0 + Z_2})/(\frac{V_1}{1 + S_{11}}) \tag{7.24}$$

which can be simplified by eliminating V_1 and yields

$$S_{21} = \frac{Z_0}{Z_0 + Z_2}(1 + S_{11}) \quad . \tag{7.25}$$

The S_{11} from Equation 7.20 can be then substituted in the equation above resulting in

$$S_{21} = \frac{Z_0}{Z_0 + Z_2}(1 + \frac{Z_1Z_2 - Z_0Z_2 - Z_0^2}{Z_1Z_2 + 2Z_0Z_1 + Z_0Z_2 + Z_0^2}) \tag{7.26}$$

which can be rearranged as

$$S_{21} = \frac{Z_0}{Z_0 + Z_2}(\frac{2Z_1Z_2 + Z_0Z_1}{Z_1Z_2 + 2Z_0Z_1 + Z_0Z_2 + Z_0^2}) \quad . \tag{7.27}$$

Secondly, port 2 is matched ($V_1^+ = 0$) as shown in Figure 7.21(b). The impedance looking into port 2 is

$$\begin{aligned} Z_{in2} &= Z_2 + (Z_0 // Z_1) \\ &= \frac{Z_0Z_1 + Z_0Z_2 + Z_1Z_2}{Z_0 + Z_1} \quad . \end{aligned} \tag{7.28}$$

Thus, the reflection at port 2 can be written as

$$S_{22} = \Gamma_{in2} = \frac{Z_{in2} - Z_0}{Z_{in1} + Z_0} \quad . \tag{7.29}$$

Substituting Z_{in2} from Equation 7.28 yields

$$S_{22} = \frac{Z_0Z_2 + Z_1Z_2 - Z_0^2}{2Z_0Z_1 + Z_0Z_2 + Z_1Z_2 + Z_0^2} \quad . \tag{7.30}$$

Since port 1 is matched that $V_1^+ = 0$, thus $V_1 = V_1^-$, the voltage at port 1 is

$$V_1^- = V_1 = V_2\frac{Z_0 // Z_1}{(Z_0 // Z_1) + Z_2} \tag{7.31}$$

or explicitly,

$$V_1^- = V_2\frac{Z_{in2} - Z_2}{Z_{in2}} \quad . \tag{7.32}$$

At port 2, the voltage can be written as

$$\begin{aligned} V_2 &= V_2^+ + V_2^- \\ &= V_2^+ + S_{22}V_2^+ \quad . \end{aligned} \tag{7.33}$$

V_2^+ can be expressed as

$$V_2^+ = \frac{V_2}{1 + S_{22}} \quad . \tag{7.34}$$

Therefore, S_{12} can be obtained

$$S_{12} = \frac{V_1^-}{V_2^+} = (V_2 \frac{Z_{in2} - Z_2}{Z_{in2}}) / (\frac{V_2}{1 + S_{22}}) \quad . \tag{7.35}$$

Eliminating V_2 yields

$$S_{12} = \frac{Z_{in2} - Z_2}{Z_{in2}} (1 + S_{22}) \quad . \tag{7.36}$$

Substituting S_{22} from Equation 7.30 in the equation above, S_{12} can be obtained

$$S_{12} = \frac{Z_0 Z_1}{Z_0 Z_1 + Z_0 Z_2 + Z_1 Z_2} (1 + \frac{Z_0 Z_2 + Z_1 Z_2 - Z_0^2}{2Z_0 Z_1 + Z_0 Z_2 + Z_1 Z_2 + Z_0^2}) \quad . \tag{7.37}$$

Therefore, the four items in the S-matrix (S_{mn}) of the matching networks $\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$ are obtained as

$$\begin{cases} S_{11} &= \frac{Z_1 Z_2 - Z_0 Z_2 - Z_0^2}{2Z_0 Z_1 + Z_0 Z_2 + Z_1 Z_2 + Z_0^2} \\ S_{21} &= \frac{Z_0}{Z_0 + Z_2} (1 + S_{11}) \\ S_{22} &= \frac{Z_0 Z_2 + Z_1 Z_2 - Z_0^2}{2Z_0 Z_1 + Z_0 Z_2 + Z_1 Z_2 + Z_0^2} \\ S_{12} &= \frac{Z_0 Z_1}{Z_0 Z_1 + Z_0 Z_2 + Z_1 Z_2} (1 + S_{22}) \quad . \end{cases} \tag{7.38}$$

7.5.3 S-Matrix Cascade

In previous sections, the S-matrices of both the transmission (S_{tx}) and matching networks (S_{mn}) have been obtained. To obtain the scattering parameters of the whole systems, it is then necessary to cascade the S-matrix of the sub-systems, including two matching networks and the free-space transmission, as shown in Figure 7.16. When two 2-port networks A and B, with scattering parameters S^A and S^B respectively, are connected in series (port 2 of network A to port 1 of network B) to form a new 2-port network with scattering matrix (S_{int}), this matrix can be represented as [61]

$$S_{int} = \begin{bmatrix} S_{11}^A + \frac{S_{12}^A S_{21}^A S_{11}^B}{1 - S_{11}^B S_{22}^A} & \frac{S_{12}^A S_{12}^B}{1 - S_{11}^B S_{22}^A} \\ \frac{S_{21}^B S_{21}^A}{1 - S_{11}^B S_{22}^A} & S_{22}^B + \frac{S_{21}^B S_{12}^B S_{22}^A}{1 - S_{11}^B S_{22}^A} \end{bmatrix} \quad . \tag{7.39}$$

By applying recursively this equation into the present investigation the overall scattering parameters of the whole system (S_{total}) can be obtained. S-matrices of the matching

network #1 (S_{mn1}) and the transmission (S_{tx}) can be first obtained, and then cascaded to the matching network #2 (S_{mn2}).

$$S_{total} = S_{mn1} \times S_{tx} \times S_{mn2}. \tag{7.40}$$

To provide a reference for the investigation, two identical quarter-wave monopoles are positioned half a wavelength apart operating at 100 MHz, and their mutual coupling is determined. The interaction between the two monopoles can be calculated based on the current distribution of a receiving element with appropriate matching networks for a perfect matching to 50 Ω. The cascaded scattering parameters of the monopoles and matching networks are used as the references for comparison purpose. The same process is then carried out for the four optimized helical antennas, i.e. two identical helical antennas are positioned half a wavelength apart with one of them excited at 1 volt, while the other one is used as the receiving element. The S parameters of the optimal helical antennas and the reference monopole antennas are listed in Table 7.16 including both transmission and matching networks.

Table 7.16. *Scattering parameters of transmission between two identical antennas and the appropriate matching networks at 100 MHz.*

S Parameters (dB)	Transmission	Matching Network
Monopole	$S_{tx} = \begin{pmatrix} -0.19 & -23.5 \\ -23.5 & -0.19 \end{pmatrix}$	$S_{mn} = \begin{pmatrix} -11.3 & -0.33 \\ -0.33 & -11.3 \end{pmatrix}$
Uniform	$S_{tx} = \begin{pmatrix} -0.08 & -12.7 \\ -12.7 & -0.08 \end{pmatrix}$	$S_{mn} = \begin{pmatrix} -0.73 & -8.13 \\ -8.13 & -0.73 \end{pmatrix}$
Radius	$S_{tx} = \begin{pmatrix} -0.02 & -19.8 \\ -19.8 & -0.02 \end{pmatrix}$	$S_{mn} = \begin{pmatrix} -1.89 & -8.13 \\ -8.13 & -1.89 \end{pmatrix}$
Pitch	$S_{tx} = \begin{pmatrix} -0.07 & -12.3 \\ -12.3 & -0.07 \end{pmatrix}$	$S_{mn} = \begin{pmatrix} -0.73 & -8.10 \\ -8.10 & -0.73 \end{pmatrix}$
Radius-Pitch	$S_{tx} = \begin{pmatrix} -0.09 & -13.1 \\ -13.1 & -0.09 \end{pmatrix}$	$S_{mn} = \begin{pmatrix} -0.95 & -7.09 \\ -7.09 & -0.95 \end{pmatrix}$

It is noticed that the quarter-wave monopoles have lower S_{11} in the matrix of their matching networks. This is because the quarter-wave monopole has about 35 Ω input

impedance, whereas the others are quite low due to their low radiation resistance. This means the other four optimal antennas require the matching networks to provide resonance at the target frequency, and thus have significantly higher S_{11} in the matching S-matrices.

By cascading the S matrices, the total scattering parameters for the dual antennas operating half a wavelength apart are achieved, as listed in Table 7.17. It can be observed that the reflection coefficients (S_{11}) decrease which indicate better matching for the optimal helical antennas. Also, the S_{12} performance is lower than the reference monopole, demonstrating weaker mutual interaction between the optimal antennas.

Table 7.17. *Scattering parameters of the integrated system.*

S Parameters (dB)	Integrated System
Monopole	$S_{total} = \begin{pmatrix} -0.17 & -24.2 \\ -24.2 & -0.17 \end{pmatrix}$
Uniform	$S_{total} = \begin{pmatrix} -0.01 & -35.4 \\ -35.4 & -0.01 \end{pmatrix}$
Radius	$S_{total} = \begin{pmatrix} -0.04 & -27.5 \\ -27.5 & -0.04 \end{pmatrix}$
Pitch	$S_{total} = \begin{pmatrix} -0.01 & -36.2 \\ -36.2 & -0.01 \end{pmatrix}$
Radius-Pitch	$S_{total} = \begin{pmatrix} -0.02 & -34.3 \\ -34.3 & -0.02 \end{pmatrix}$

7.5.4 Impact of Distance & Winding Direction on Mutual Coupling

To further investigate the mutual interaction between two antennas, the impact of their distance is considered. Scattering parameters are computed as a function of the distance between the two antennas considered. In the second part of the section, the effect of winding direction of the helical antennas on their coupling is investigated.

- **Impact of Distance**

  In this section, two identical antennas operating at 100 MHz are located on a ground plane with a distance between their feed varying from one-tenth of a wavelength ($\frac{\lambda}{10}$) to twice the wavelength (2λ). The scattering parameters are calculated from simulation based on the current distribution. The S parameters are plotted with respect to distance in Figure 7.22.

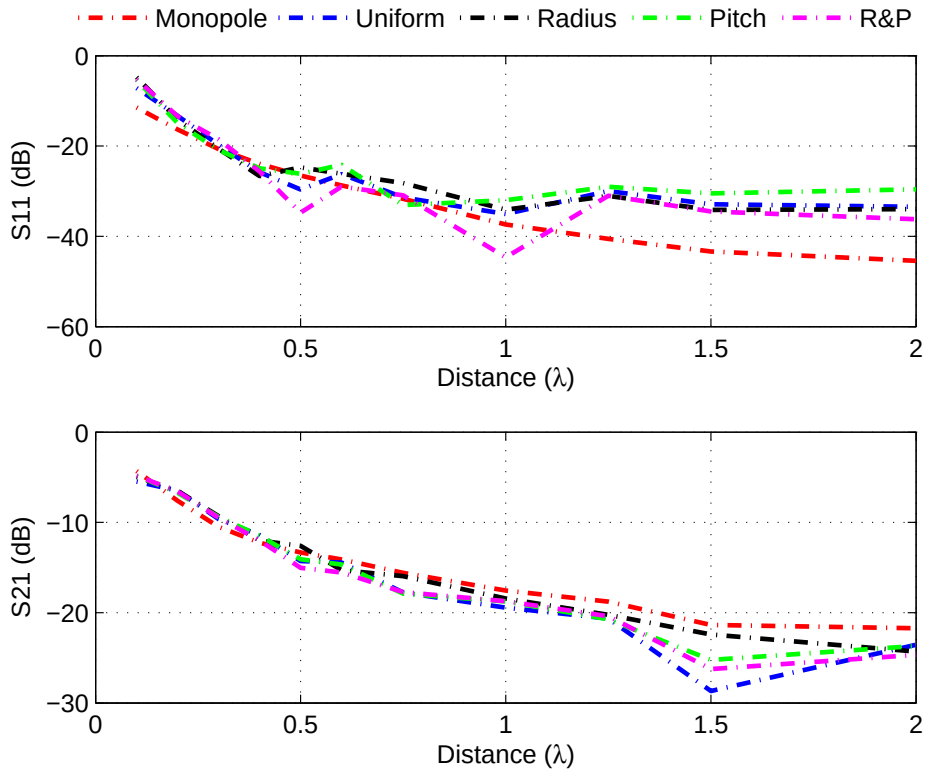

Figure 7.22. Scattering parameters of dual antennas operating at 100 MHz as a function of the distance.

It can be seen that all the antennas have similar trend of S_{11} with respect to distance. More reflection can be detected when two antennas are located close to each other. For perfect matching, there will be no reflection. However, because of the mutual coupling to the nearby antenna, the matching is no more perfect. More specifically, in the near-field, the coupling is desired to be minimized. It is realized the optimal antennas have very much similar values to the uniform reference helix, which indicates no significant impact of coupling introduced by

their optimized winding configurations. In the far field, the transmission is desired to be maximized. It can be observed that S_{21} decreases as expected when the distance increases. The optimal antennas do not decrease faster or more than the uniform reference helix, which demonstrates their normal operation.

- **Impact of Winding Direction**

  In this section, the impact of winding direction on the mutual coupling between nearby helix antennas is investigated. Specifically, the winding of the receiving antenna is mirrored and becomes the opposite of the transmitting antenna. An illustration can be provided in Figure 7.23.

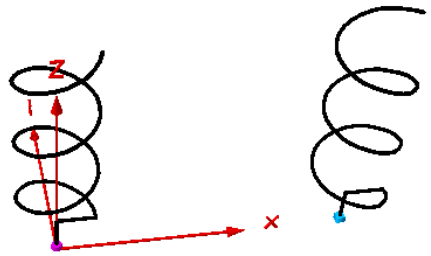

Figure 7.23. Illustration of two antennas with opposite winding directions.

The scattering parameters are investigated in the near field and the distance between the transmitting and receiving antenna is varied from 0.1λ to 0.5λ. The S_{11} and S_{21} performance are shown in Figure 7.24. It is observed that the winding direction does not appear to have a significant impact towards the performance of the optimal helical antennas.

7.6 Conclusions

This chapter has investigated the impact of non-ideal practical environments upon the operation of the optimal helical antennas obtained using the approach presented in previous chapters. Firstly, the design was tested on a system of radial wires instead of the PEC ground plane. Four wires were found sufficient to achieve similar performance to the ground plane, hence verifying the stability of the design. The influence of metallic objects on the optimal helical antennas has also been investigated by changing

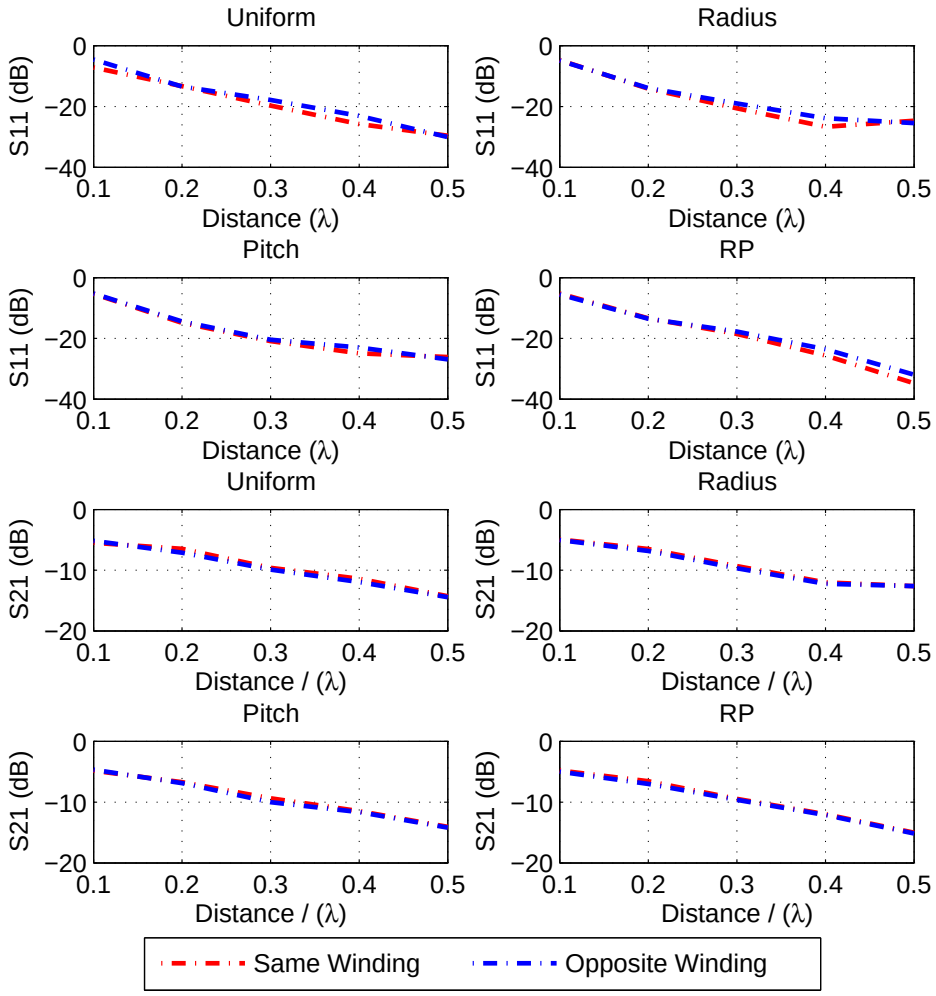

Figure 7.24. S_{11} and S_{21} performance of two antennas with the same and opposite winding directions.

the dimension of the interfering object and its distance to the helix. Re-optimization has been carried out, including the non-ideal environment, and was shown to produce similar helical configurations to the infinite plane, thus verifying the robustness of the antenna design methodology. A realistic case study has been applied to find optimal antenna configuration and shape on a hand-held device for VHF communication and with the same conclusion. Finally, the mutual interactions between optimal helical antennas have been studied. The scattering parameters indicate that the optimized helices do not introduce significant additional mutual coupling compared to standard

monopoles in the near-field, which shows the antenna to be suitable for array configurations. The next chapter will wrap up and draw together the conclusion of this thesis.

Chapter 8

Conclusion & Future Work

THIS chapter draws together the conclusions from the work described in this thesis and provides recommendations for future work. It summaries all the factors that are important in the design of wire antennas for HF and VHF communications. The methodology that couples electromagnetic simulation and evolutionary optimization has been employed, with an in-house tool being successfully developed and applied in this thesis. The results have proved such method to be effective and reliable in providing robust configurations of different types of wire antennas. A realistic application has been applied using the proposed design method and shown great success in providing an optimal antenna for maximized performance. In addition, the mutual interaction between optimized antennas has shown weak interference when used in an array mode.

8.1 Conclusion

The concept of electrically small antennas has been investigated, and developed extensively since its inception decades ago. When electrically small antennas are used for HF and VHF communications, size reduction is usually the most important goal in the design process. Particularly, inductive loading has been adopted as the primary approach in miniaturizing antennas in a way that retains good performance. This thesis focuses on optimizing inductive loading distributions on various types of wire structures, in the pursuit of size reduction with optimal efficiency and bandwidth.

Chapter 1 has provided a brief overview on the fundamentals of HF and VHF communication including both advantages and drawbacks. Electrically small antennas have been defined conceptually and their limitations discussed. Multi-objective optimization has been also introduced in this chapter as it later applied to antenna optimization.

Chapter 2 has described the multi-objective optimization and evolutionary algorithms in more details. Two types of evolutionary algorithms, genetic algorithm and particle swarm optimizations, have been explicitly described including both fundamentals and applications from the literature. Parametric antenna descriptions have been also considered as they lead to a more effective optimization process. The fitness function has been discussed to illustrate its importance in balancing the trade-offs between multiple objectives.

Chapter 3 has described the electromagnetic simulator (NEC-2) that is used in this work. It has also described the method of moments which is the solution method of the electric field integral equations and magnetic field integral equations, that lie at the heart of NEC-2 simulator. In order to ease the complexity from antenna modeling and computational load from multi-objective optimization, the antennas are parameterized using radial basis function expansions (RBFs) and these have been described in this chapter. More importantly, an in-house tool that couples NEC simulator and evolutionary optimizer has been introduced. This simulator plays an essential role in the investigations of this thesis.

Chapter 4 has described a strategy to miniaturize the size of HF and VHF wire antennas by using lumped inductive loadings along the antenna height. The optimization has been performed using a genetic algorithm implemented in MATNEC. Different loading schemes have been applied to a shortened monopole to verify the effectiveness of the procedure and its ability to find optimal solutions. The application of matching networks has been discussed in order to provide appropriate matching in the pursuit of maximized operational bandwidth. The achieved optimized solution can effectively reduce the size of a resonant monopole while optimizing efficiency and bandwidth. The results, however, show diminishing return when a large number of inductors are used. The next chapter considers the use of the antenna structure itself to provide the inductive loading.

Chapter 5 has continued the investigation of distributed inductance using the zig-zag structure. The additional wire lengths, provided by the zig-zag shape, provide an effective way to introduce distributed inductance along a wire antenna and to effectively shorten it. Radial basis functions were used to introduce sufficient structural variations to the zig-zag with relatively few optimization variable. In this approach, structural variation is described in terms of overall parameters such as width and winding density. Near-optimal zig-zag configurations were found in MATNEC with maximized efficiency and bandwidth. The proposed optimal zig-zag antennas were fabricated and measured to obtain their relative efficiency and operational bandwidth. Experimental results have indicated a fairly good agreement between the simulation and measurement, which has validated the investigation methodology.

Chapter 6 has further extended the investigation of distributed inductance by employing antenna composed of helical windings. A description of structural variation in terms radius and pitch were found to be an effective and practical way to introduce distributions of inductance along the antenna. By employing radial basis expansions, sufficient structural variations can be implemented on the antennas with few optimization variables. Near-optimal helical configurations were found at three different frequencies by applying MATNEC. A set of the proposed optimal helical antennas were fabricated and measured, with experimental results indicating a good agreement between simulation and measurement. The work has shown that helical antennas can be

optimized to perform with wider bandwidth while maintaining good efficiency. Peripheral investigations were also carried out to consider such matters as the impact of matching circuit loss on efficiency and bandwidth, the sensitivity of the final design, and the use of a reduced optimization space. These investigations have further demonstrated the robustness and reliability of the optimized design and the overall the design methodology.

Chapter 7 has considered the optimized helical antenna from a more practical point of view, i.e. the impact of non-ideal environment. Different types of ground plane were tested in terms of radiation efficiency and bandwidth using a uniform reference helix. The influence of near-by objects upon the optimal helical antennas was also investigated. Re-optimization was carried out with the non-ideal environment included, and only slight changes in predicted antenna structure successfully verified the stability of the design. Finally, the mutual interactions between optimal helical antennas have been studied. The scattering parameters indicate that the optimized helices do not pose significant interference when operating in an array mode, which also attests the robustness and reliability of the optimal helical antennas previously achieved.

8.2 Future Work

This thesis has provided some feasible optimal solutions for several types of wire antennas for HF and VHF communications. Nevertheless, due to the time constraint imposed by the duration of this research and the limitation of resources, some topics and ideas that have the potential to further improve the quality of the solutions have not been carried out. They are listed as follows:

- **Hybrid Evolutionary Algorithm**

  In this thesis, two types of evolutionary algorithms, i.e. genetic algorithm and particle swarm algorithm, have been introduced and employed in the optimizer in MATNEC. Nevertheless, each of the algorithms is operated alone without connections to other types of optimization algorithms. It was also realized in the

investigation that the convergence slows down when most of the candidate solutions gather to the global optimal region in the later stage of the optimization. This phenomenon indicates that the candidate solutions are wandering around near optimal region, and thus lowers optimizing efficiency.

Therefore, it is of great interest to employ some calculus-based methods (gradient method for example) in the later stage of the optimization. Explicitly, the program monitors the convergence closely and detects the closeness of the candidate solutions. Once the optimization has converged to a pre-defined level, for instance 95% similarity of the whole solutions, evolutionary algorithm is replaced by gradient method. Further optimization is thus performed in the global near-optimal region so that it overcomes the risk of been trapped in the local optimum. This hybrid optimization algorithm effectively takes the advantages of both evolutionary and calculus-based algorithm, so that it is able to converge faster by avoiding local optimum, as well as to achieve more accurate absolute global optimal solution.

Furthermore, it is also possible to integrate GA and PSO by taking their unique features and implement into a hybrid evolutionary algorithm, e.g. hybrid-GA-PSO algorithm. In the investigation, it was also noticed that PSO can be manipulated to exploit the whole solution region more efficiently than GA, e.g. by increasing the velocity to move the particles faster. Consequently, it is also potentially valuable to apply the features of both GA and PSO at different stages of the optimization.

- **Other Antenna Structures / Types**

  In chapters 5 and 6, two types of wire structures, i.e. zig-zag and helix, were investigated in detail. They have shown great success in providing distributed inductance along their bodies, by varying the geometrical parameters. The optimal configurations have also demonstrated improved performance with respect to efficiency and bandwidth, thus the robustness and reliability of the design methodology. Such design methodology can also be applied to other types of wire structures in the pursuit of improved performance. The potential wire structures may

include, 1) Yagi-Uda antennas with optimized element lengths and separations, 2) spiral antennas with optimized varying arm length, arm width and winding radians, 3) fractal antennas with optimized space-filling curve, and so on.

More importantly, this method is applicable, but not restricted to wire structures. With the assistance of appropriate simulation tools, this method can be applied to optimize other types of antenna, such as micro-strip antennas (PIFA for example), reflector antennas (dish antenna for example), or aperture antennas (slot antenna for example).

- **Higher Frequency Band**

  This thesis primarily has investigated wire antennas operating at HF and VHF bands (3 - 300 MHz) with corresponding wavelength ranges from 1 to 100 meters. Specifically, the monopole antennas with lumped inductors and the zig-zag antennas were optimized at 100 MHz, while the helical antennas were optimized at 25, 50 and 100 MHz, the portable antenna for hand-held device was targeted for 300 MHz. Since the design methodology has been proved to be robust and reliable for a number of times, it is believed that it can be applied to higher frequency bands in future investigations.

- **Complex Base Structure**

  In chapter 7, a cubic metal box is used to act as a finite conducting ground for the operation of helical antennas. Other non-ideal ground planes are also of importance and merit investigation. Most often, HF and VHF antennas are mounted on top of a mobile vehicle with the metal roof acting as the ground. The remaining part of the vehicle can, however, also influences the antenna performance, and the investigation of its impact is of extreme importance.

- **MIMO Array Mode**

  Chapter 7 of this thesis has carried out an investigation of the mutual interaction between two identical optimized helical antennas when operated in an array mode, and the results indicate negligible interference. MIMO (multiple-input-multiple-output) is a technology that requires multiple antennas at both receive and transmit sides, and their mutual interference is of extreme importance to this

technology. MIMO is increasingly employed to improve spectral efficiency as well as operational reliability. Traditionally, it was regarded that radio signals reflect off objects, create multiple paths and consequently cause interference and fading in a rich multi-path environment. MIMO, however, turns such disadvantage into benefits [85]. It sends different data each path, and hence increases the amount of information the system carries.

Crucial to MIMO is the use of array antennas for both receive and transmit. At lower frequencies, antenna size reduction of the type considered in this thesis is crucial. For MIMO application, the mutual interactions between array elements is extremely important and so future work on mutual interactions for more complex array configurations will be an important future study.

Bibliography

[1] N. Maslin, *HF communications: a systems approach*. New York: Plenum Press, 1987.

[2] L. W. Barclay, *Propagation of radiowaves*, 2nd ed. Institution of Electrical Engineers, 2003.

[3] A. A. Kolosov, *Over-the-horizon radar*. Boston: Artech House, 1987.

[4] H. A. Wheeler, "Fundamental limitations of small antennas," *IRE Proc.*, vol. 35, pp. 1479–1484, 1947.

[5] H. A. Wheeler, "The radiansphere around a small antenna," *IRE Proc.*, vol. 47, pp. 1325–1331, 1959.

[6] H. A. Wheeler, "Small antennas," *IEEE Trans. Antennas Propag.*, vol. 23, pp. 462–469, 1975.

[7] H. A. Wheeler, "A re-examination of the fundamental limits on the radiation Q of electrically small antennas," *IEEE Trans. Antennas Propag.*, vol. 44, pp. 672–675, 1996.

[8] L. Chu, "Physical limitations of omnidirectional antennas," *Journal of Applied Physics*, vol. 19, pp. 1163 – 1175, 1948.

[9] S. R. Best, "A discussion on the quality factor of impedance matched electrically small wire antennas," *IEEE Trans. Antennas Propag.*, vol. 53, pp. 502–508, 2005.

[10] A. D. Yaghjian and S. R. Best, "Impedance, bandwidth, and Q of antennas," *IEEE Trans. Antennas Propag.*, vol. 53, pp. 1298–1324, 2005.

[11] S. R. Best and D. L. Hanna, "A performance comparison of fundamental small-antenna designs," *IEEE Antennas Propag. Mag.*, vol. 52, pp. 47–70, 2010.

[12] K. Deb, *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons, 2001.

[13] E. Zitzler, *Evolutionary algorithms for multiobjective optimization: Methods and applications*. Zurich: Ph.D. Dissertation, Swiss Federal Inst. Technol. (ETH), 1999.

[14] S. H. Yeung and K. F. Man, "Multiobjective optimization," *IEEE Microwave Mag.*, vol. 12, pp. 120–133, 2011.

[15] D. E. Goldberg, *Genetic algorithms in search, optimization and machine learning*. New York: Addison-Wesley, 2001.

[16] J. F. Frenzel, "Genetic algorithms," *IEEE Potentials*, vol. 12, pp. 21–24, 1993.

[17] J. Kenedy, "The particle swarm: social adaptation of knowledge," in *IEEE International Conf. on Evolutionary Computation*, 1997, pp. 303–308.

[18] R. F. Harrington, *Field computation by moment methods*. New York: Macmillan Company, 1968.

[19] G. F. Uler and O. A. Mohammed, "Utilizing genetic algorithms for the optimal design of electromagnetic devices," *IEEE Trans. Magnetics*, vol. 30, pp. 4296–4295, 1994.

[20] J. M. Johnson and V. Rahmat-Samii, "Genetic algorithms in engineering electromagnetics," *IEEE Antennas Propag. Mag.*, vol. 39, pp. 7–21, 1995.

[21] J. Holland, "Genetic algorithms," *Scientific American*, vol. 267, pp. 66–72, 1992.

[22] S. Forrest, "Genetic algorithms: principles of natural selection applied to computation," *Science*, vol. 260, pp. 872–878, 1993.

[23] D. S. Weile and E. Michielssen, "Genetic algorithm optimization applied to electromagnetics: a review," *IEEE Trans. Antennas Propag.*, vol. 37, pp. 343–353, 1997.

[24] R. L. Haupt, "An introduction to genetic algorithms for electromagnetics," *IEEE Antennas Propag. Mag.*, vol. 37, pp. 7–15, 1995.

[25] R. L. Haupt, "Comparison between genetic and gradient-based optimization algorithms for solving electromagnetics problems," *IEEE Trans. Magnetics*, vol. 31, pp. 1932–1935, 1995.

[26] R. L. Haupt, "Thinned arrays using genetic algorithms," *IEEE Trans. Antennas Propag.*, vol. 42, pp. 993–999, 1994.

[27] R. L. Haupt, "Optimization of subarray amplitude tapers," in *Proc. IEEE AP-S Int. Symp. Antennas Propag.*, vol. 4, 1995, pp. 1830 – 1833.

[28] D. Marcano and F. Duran, "Synthesis of antenna arrays using genetic algorithms," *IEEE Antennas Propag. Mag.*, vol. 42, pp. 12–20, 2000.

[29] F. J. Ares-Pena, "Genetic algorithms in the design and optimization of antenna array patterns," *IEEE Trans. Antennas Propag.*, vol. 47, pp. 506–510, 1999.

[30] E. Michielssen, "Design of lightweight, broad-band microwave absorbers using genetic algorithms," *IEEE Trans. Microw. Theory and Tech.*, vol. 41, pp. 1024–1031, 1993.

[31] J. M. Johnson and Y. Rahmat-Samii, "Genetic algorithms and method of moments (GA/MOM) for the design of integrated antennas," *IEEE Trans. Antennas Propag.*, vol. 47, pp. 1606–1614, 1999.

[32] E. E. Altshuler and D. S. Linden, "Wire-antenna designs using genetic algorithms," *IEEE Antennas Propag. Mag.*, vol. 39, pp. 33–43, 1997.

[33] C. A. Balanis, *Antenna theory: analysis and design*, 3rd ed. New Jersey: John Wiley & Sons, 2005.

[34] W. L. Stutzman and G. A. Thiele, *Antenna theory and design*, 2nd ed. New York: John Wiley & Sons, 1998.

[35] H. Choo, "Design of electrically small wire antennas using a pareto genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 53, pp. 1038–1046, 2005.

[36] A. Boag, "Design of electrically loaded wire antennas using genetic algorithms," *IEEE Trans. Antennas Propag.*, vol. 44, pp. 687–695, 1996.

[37] E. E. Altshuler, "Electrically small self-resonant wire antennas optimized using a genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 50, pp. 297–300, 2002.

[38] Z. Altman, "New designs of ultra wide-band communication antennas using a genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 45, pp. 1494–1501, 1997.

[39] K. Yegin and A. Q. Martin, "On the design of broad-band loaded wire antennas using the simplified real frequency technique and a genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 51, pp. 220–228, 2003.

[40] J. Kennedy and R. Eberhart, *Swarm Intelligence*. San Francisco: Calif: Morgan Kaufmann, 2001.

[41] E. O. Wilson, *Sociobiology: The new synthesis*. Cambridge, MA: Belknap Press, 1975.

[42] J. Robinson and Y. Rahmat-Samii, "Particle swarm optimization in electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 52, pp. 397–407, 2004.

[43] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. IEEE Int. Symp. Micro Machine and Human Science*, 1995, pp. 39–73.

[44] R. Eberhart and Y. Shi, "Particle swarm optimization: developments, applications and resources," in *Proc. IEEE Int. Conf. Evolutionary Computation*, 2001, pp. 81–86.

[45] R. Eberhart and Y. Shi, "Comparing inertia weights and constriction factors in particle swarm optimization," in *Proc. IEEE Int. Conf. Evolutionary Computation*, 2000, pp. 84–88.

[46] M. Clerc and J. Kennedy, "The particle swarm - explosion, stability, and convergence in a multidimensional complex space," *IEEE Trans. Evolutionary Computation*, vol. 6, pp. 58–73, 2002.

[47] N. Jin and Y. Rahmat-Samii, "Particle swarm optimization for antenna designs in engineering electromagnetics," *Journal of Artificial Evolution and Applications*, pp. 1–10, 2008.

[48] S. J. Robinson and J. Kennedy, "Particle swarm, genetic algorithm, and their hybrids: optimization of a profiled corrugated horn antenna," in *Proc. IEEE AP-S Int. Symp. Antennas Propag.*, 2002, pp. 314–317.

[49] N. Jin and Y. Rahmat-Samii, "Parallel particle swarm optimization and fnite difference time-domain (PSO/FDTD) algorithm for multiband and wide-band patch antenna designs," *IEEE Trans. Antennas Propag.*, vol. 55, pp. 3459–3468, 2005.

[50] X. Y. F. Yang, X.-X. Zhang and Y. Rahmat-Samii, "Wide-band e-shaped patch antennas for wireless communications," *IEEE Trans. Antennas Propag.*, vol. 49, pp. 1094–1100, 2001.

[51] F. Yang and Y. Rahmat-Samii, "Wire antennas on artificial complex ground planes: a new generation of low gain antennas," in *Proc. IEEE AP-S Int. Symp. Antennas Propag.*, 2004, pp. 309–312.

[52] F. Yang and Y. Rahmat-Samii, "Reflection phase characterizations of the EBG ground plane for low pro?le wire antenna applications," *IEEE Trans. Antennas Propag.*, vol. 51, pp. 2691–2703, 2003.

[53] D. Gies and Y. Rahmat-Samii, "Particle swarm optimiza-tion for reconfigurable phase-differentiated array design," *Microwave and Optical Technology Letters*, vol. 38, pp. 172–175, 2003.

[54] G. Burke, *Numerical Electromagnetic Code (NEC) - Method of Moments*. Lawrence Livermore National Laboratory, 1981.

[55] E. M. G. Burke and A. Poggio, "The numerical electromagnetics code (NEC) - a brief history," in *Proc. IEEE AP-S Int. Symp. Antennas Propag.*, 2004, pp. 2871 – 2874.

[56] L. B. Cebik, "A beginner's guide to modeling with NEC," *QST -NEWINGTON-*, vol. 85, pp. 31–38, 2001.

[57] M. D. Buhmann, *Radial basis function: theory and implementation*. Cambridge University Press, New York, 2003.

[58] J. O. Ramsay, *Functional data analysis*. Springer, New York, 1997.

[59] S. R. Best, "A discussion on the properties of electrically small self-resonant wire antennas," *IEEE Antennas Propag. Mag.*, vol. 46, pp. 9–22, 2004.

[60] C. Bowick, *RF circuit design*. H. W. Sams, Indianapolis, 1982.

[61] C. Coleman, *An introduction to radio frequency engineering*. Cambridge University Press, New York, 2004.

[62] R. J. Smith, *Electronics: circuits and devices*. Wiley, New York, 1987.

[63] W. Cumming, "A nonresonant endfire array for VHF and UHF," *IRE Trans, Antennas and Propagation*, vol. 3, pp. 52–58, 1955.

[64] D. Sengupta, "The radiation characteristics of a zig-zag antenna," *IRE Trans, Antennas and Propagation*, vol. 6, pp. 191–194, 1958.

[65] L. Sing and K. Mei, "Analysis of zigzag antennas," *IEEE Trans, Antennas and Propagation*, vol. 18, pp. 760–764, 1970.

[66] P. Bhatnagar and S. Sachan, "Analysis of infinite zig-zag antenna," *IERE Proceedings*, vol. 14, pp. 44–46, 1976.

[67] J. D. Kraus, "Helical beam antenna," *Electronics*, vol. 20, pp. 109–111, 1947.

[68] H. Nakano, "Axial mode helical antennas," *IEEE Trans, Antennas and Propagation*, vol. 34, pp. 1143–1148, 1986.

[69] S. Rao and J. Kralovec, "Design of a printed quadrifilar-helical antenna on a dielectric cylinder by means of a genetic algorithm," *IEEE Antennas Propag. Mag.*, vol. 53, pp. 262–268, 2011.

[70] J. Wong and H. King, "Broadband quasi-taper helical antennas," *IEEE Trans. Antennas Propag.*, vol. 27, pp. 72–78, 1979.

[71] E. K. N. H. T. Hui, K. Y. Chan, "The input impedance and the antenna gain of the spherical helical antenna," *IEEE Trans. Antennas Propag.*, vol. 49, pp. 1235–1237, 2001.

[72] H. T. Hui, K. Y. Chan, and E. K. N.Yung, "Optimization of helical antennas [antenna designer's notebook]," *IEEE Antennas Propag. Mag.*, vol. 48, pp. 107–115, 2006.

[73] P. C. Day, "Some characteristics of tapered helical beam antennas," *MS Thesis, Ohio State University, USA*, 1950.

[74] C. H. Chen, E. K. N. Yung, and B. J. Hu, "Axial mode helix antenna with exponential spacing," *Microwave and Optical Technology Letters*, vol. 49(7), pp. 1525–1530, 2007.

[75] F. Yang, P. Zhang, and C. J. Guo, "Axial mode elliptical helical antenna with variable pitch angle," *Electronics Letters*, vol. 44, 2008.

[76] T. Peng, S. Koulouridis, and J. L. Volakis, "Miniaturization of conical helical antenna via optimized coiling," *ACES Journal*, vol. 26, pp. 452–458, 2011.

[77] A. Gupta, D. Grant, and T. Fickenscher, "An efficient electrically small antenna at HF band ," in *Proc. Asia-Pacific Microw. Conf., Melbourne, Australia*, Dec 2011, pp. 856–859.

[78] X. Ding, B. Zhong, G. Zheng, and X. Li, "Design and realization of a GA-optimized VHF/UHF antenna with "on-body" matching network," *IEEE Antennas Wireless Propag. Lett.*, vol. 9, pp. 1536–1225, 2010.

[79] S. Zhao, C. Fumeaux, and C. Coleman, "Optimal helical antenna with continuously varying radius using evolutionary optimizers," in *Proc. IEEE AP-S Int. Symp. Antennas Propag., Spokane WA, USA*, Jul 2011, pp. 757–760.

[80] S. Zhao, C. Fumeaux, and C. Coleman, "Evolutionary optimization of zig-zag antennas using gaussian and multiquadric radial basis functions," in *Proc. Asia-Pacific Microw. Conf., Melbourne, Australia*, Dec 2011, pp. 1594–1597.

[81] H. Nagaoka, "The inductance coefficients of solenoids," *Journal of the College of Science*, vol. 27, pp. 18–33, 1909.

[82] J. McNeill, *Electromagnetic terrain conductivity measurement at low induction numbers*. Geonics Limited Mississauga, Ontario, Canada, 1980.

[83] C. J. Coleman, "The effect of an imperfectly conducting plane upon an incident electromagnetic field," *IEEE Antennas Propag. Mag.*, vol. 45, pp. 1445–1450, 1997.

[84] J. Carr, *Practical antenna handbook*. McGraw-Hill, 2001.

[85] D. Gesbert and M. Shafi, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE Communications*, vol. 21, pp. 281–302, 2003.