Title

Nature-Inspired Optimization Techniques Applied to Antennas for Wireless Communications and Radar

Author

Kovitz, Joshua Michael

Publication Date

2012

University of California

Los Angeles

Nature-Inspired Optimization Techniques Applied to Antennas for Wireless Communications and Radar

A thesis submitted in partial satisfaction

of the requirements for the degree

Master of Science in Electrical Engineering

by

Joshua Michael Kovitz

2012

Abstract of the Thesis

Nature-Inspired Optimization Techniques Applied to Antennas for Wireless Communications and Radar

by

Joshua Michael Kovitz
Master of Science in Electrical Engineering
University of California, Los Angeles, 2012
Professor Yahya Rahmat-Samii, Chair

In this work, two nature-inspired optimization techniques, namely Particle Swarm Optimization (PSO) and Covariance Matrix Adaptation Evolution Strategy (CMAES), are presented. Some comparisons are made between the two algorithms in different applications and different optimization problems. First, a comparison of each algorithm in resource limited problems is demonstrated using mathematical functions. Next, a comparison is shown between the two algorithms for a real-world antenna design problem for radar systems. In particular, a weather radar antenna array element is optimized using two different approaches, and some suggestions are made for antenna designers hoping to implement dual polarized antenna arrays for new use in weather radar systems. In the last half of this work, the Particle Swarm algorithm is applied to two other antenna systems. The first application of PSO investigates the use of a smooth contour septum design in circular waveguide for possible use in high power microwave systems. Similar antenna performance compared to a stepped septum in circular waveguide is demonstrated using the Sigmoid function contour. PSO is also applied on two newly proposed reconfigurable E-shaped patch antenna designs. Two reconfigurability mechanisms are introduced, namely a polarization (RHCP/LHCP) reconfigurable design and a frequency reconfigurable design. Both designs are optimized using a simple MEMS circuit model for fast optimization, and possible bias network implementations are discussed. Wideband patch designs are realized with these optimizations, and prototypes are fabricated and measured to validate these designs.

The thesis of Joshua Michael Kovitz is approved.

Tatsuo Itoh

Lieven Vandenberghe

Yahya Rahmat-Samii, Committee Chair

University of California, Los Angeles

2012

To my sweet wife Anel, who encouraged me in the difficult times
and rejoiced with me in my accomplishments

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

I would like to acknowledge and thank Shubhendu Bhardwaj, Ilkyu Kim, and Harish Rajagopalan from the UCLA Antenna Research, Analysis, and Measurement Laboratory for their help in preparing this thesis as well as their collaboration on the various projects conducted during my Masters at UCLA. A special thanks also to Timothy Brockett for his help with many of the measurements presented in this thesis.

I am sincerely grateful to Prof. Tatsuo Itoh and Prof. Lieven Vandenberghe for their time and energy spent as part of my thesis committee. I am also forever indebted to Prof. Yahya Rahmat-Samii's support, guidance, wisdom, and time given towards furthering my education and teaching me how to become a better researcher.

CHAPTER 1

Introduction

The continuous technological advancement of wireless communication and radar systems has increased the complexity of design constraints as well as performance evaluation for antenna designs. For practicing engineers there are often two options to deliver a final antenna design when the classical textbook designs are not sufficient. The first is to develop an antenna using a canonical shape thereby allowing exact mathematical analysis. This can be quite painstaking and rigorous, and it is often impossible to solve for non-canonical geometries, therefore limiting its applicability. The second option is to use *trial-and-error* techniques, which can be cumbersome and time-consuming. This approach can often lead to a characterization of the relationship between the design objectives and its design parameters, which is useful to engineers and can even aid in developing a theoretical analysis of the antenna structure. Yet, without any *a priori* knowledge on the antenna structure and its physical mechanisms, one may need to conduct an extraordinarily large number of simulations, which can be time consuming. In many cases, there can exist objectives which may have no easily seen correllation to the design objectives as will be seen in Chapter 4.

Therefore, in order to address the challenges and requirements presented by these sophisticated electromagnetic systems, there is a need to develop novel antenna design solutions and strategies. One of the recently emerging approaches that has been proposed to tackle this problem in a systematic fashion is the amalgamation of nature-inspired optimization techniques with simulation models which output the performance evaluation for each particular test design. By combining electromagnetic simulation tools with global optimization techniques, one can reach an antenna design solution in an autonomous manner. These methods also do not require any *a priori* knowledge about the design. Certainly, nature-inspired optimization has been one of the forefronts within electromagnetics research due to its widely proven abilities to solve complex problems that do not lend themselves to simple solutions.

This chapter is devoted to providing a brief motivational overview on the role of antennas in communication systems and radar. This serves to familiarize the reader with the significance of antennas in the overall system, thereby emphasizing the importance in conducting optimization. The next section then discusses the role of optimization within electromagnetics and provides some optimization examples. We conclude this chapter by defining many of the performance parameters that are important in optimizations for those that may be less familiar with electromagnetics and antenna engineering.

1.1 Motivation

Electromagnetics has given rise to the development of engineering systems over a wide range of applications. Communication and radar systems are among the leading technologies that have made a significant impact on society. With worldwide communication now readily available, one's local community in the surrounding area has now expanded to a global community, thus enabling more interaction across cultures. With this new enhancement a multi-cultural society begins to thrive as people learn more about other cultures that were previously inaccessible. Commerce is expanded to a global level that is more available to people of all socioeconomic status. The Internet, an innovative means of communication, has brought about revolutionary concepts such that it is now commonplace to have conversations with others halfway across the globe in real time for a low cost. The influence of these relatively newer means of communication have been far-reaching.

Wireless communication systems have been gaining more popularity due to their added versatility and flexibility. Recent advances in signal processing, microelectronics fabrication, and microwave system design have permitted the development of wireless systems operating at low power levels with minimal interference towards other devices under the same infrastructure. Among these are the cellular phone, WiMAX, WiFi, and 4G LTE systems. Indeed, newly proposed systems such as LTE-advanced aim to support systems with nearly 1 Gb/s downlink and

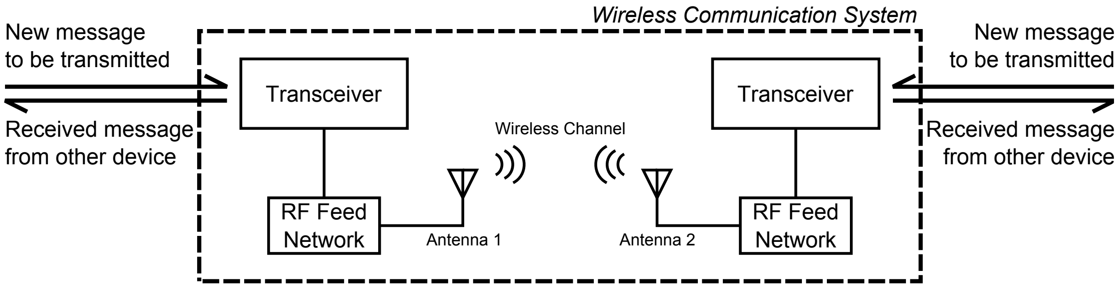


Figure 1.1: Depiction of the basic blocks of a wireless communication system

500 Mb/s uplink [1]. These systems could not have been developed without the use of signal processing, communication theory, and ultimately electromagnetics.

Wireless systems are also implemented for other non-multimedia applications that have aided in saving lives. Some systems provide a means of communication in distaster areas requiring immediate attention and help responders pinpoint the location with the most need. The development of systems that can operate under harsh conditions or even in remote areas is a challenging task [2], and several organizations such as UNICEF have created projects to develop new systems that can increase communications in isolated areas [3]. Others have approached the problem by developing low altitude balloon-based WiFi systems that are capable of covering a 5.5 km radius [4]. These are only some of the many interesting examples of applications using wireless communication systems. There still remains a much greater variety of applications that would take a whole volume of books to cover every possible wireless communication system that has been developed.

As engineers, we are interested in the technical aspects of the communication system. A high level block diagram of a duplex wireless communication system is shown in Figure 1.1. It is often the role of computer engineers, communication engineers, and embedded system designers to create a user interface to the transceiver. With this interface, one can then either transmit or receive messages via the wireless communication system. The transceivers within this system block perform several functions which often include message encoding/decoding, channel encoding/decoding, error coding, and modulation/demodulation for a digital system [5]. While the previously mentioned components are key to the success of a communication system, the propagating electromagnetic wave still serves to carry the message from the transmitter to the

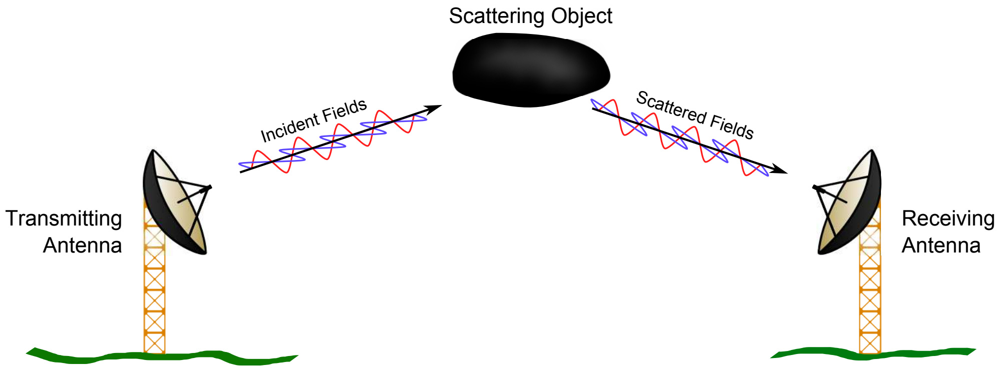


Figure 1.2: Basic bistatic radar system detecting a scattering object

receiver. Therefore it is extremely important to design properly a system which manipulates the electromagnetic fields in a way such that the SNR is maximized, thus emphasizing the significance of a proper antenna design.

A radar system's primary goal is to identify or characterize objects or materials in a given area of interest. It accomplishes this in various ways, but the most widely recognized form is through the measurement of the backscattered fields, or radar cross section (RCS). In the most basic sense, a radar system transmits electromagnetic waves which impinges upon an object if present. This object scatters the incident waves, and these scattered fields can be received by a receiving antenna (bistatic radar) or by the original transmitting antenna (monostatic). This basic system is depicted in Figure 1.2, where the transmitting antenna radiates an electromagnetic wave and the receiving antenna measures the backscattered fields. If no object is present (or if the RCS is small), then the measured fields will be minimal. For military applications, the aim is to identify possible incoming threats or targets. For scientific and space applications, radars can provide new scientific data, giving scientists valuable information about world climate, fresh water supply, sea level, sea-ice mapping, precipitation, and storms.

Several space missions sponsored by the National Aeronautics and Space Administration (NASA) provide interesting examples of radar applications. One example is CloudSat, a NASA mission deployed to investigate cloud behavior and its effect on climate. It is equipped with a 94 GHz nadir-looking radar that observes the backscattered radiation from the clouds. With these measurements scientists believe that they will be able to advance our knowledge on cloud abundance, structure, and radiative properties. Ultimately this will enable researchers to better understand the exchange of solar and thermal energy between the atmosphere, hydrosphere, land surface, biosphere, and space [6]. Another mission that was joint between NASA in the United States and the Deutsche Forschungsanstalt fr Luft und Raumfahrt (DLR) in Germany was the Gravity Recovery and Climate Experiment (GRACE) mission. This mission aimed to finely map the Earth's gravitational fields using both a GPS and microwave ranging system. This data can then be used to study changes in the Earth's gravity field due to surface and deep currents in the ocean, variations of mass within the Earth, exchanges between glaciers and the ocean, and ground water storage on land masses [7, 8]. Some other future and current NASA missions using radar technologies also include Aquarius [9], Soil Moisture Active Passive (SMAP) [10], and the Gravity Recovery and Interior Laboratory (GRAIL) mission [11].

From these brief discussions, one can recognize the significance of antennas to wireless communication and radar system performance. It is common for these systems to operate with strict design specifications on parameters such as high bandwidth and low cross-polarization radiation. This can make it difficult to meet with standard antenna designs, and therefore new designs are a must. It is common for the RF feed networks and the antennas to operate at relatively high frequencies in the 100 MHz to 100 GHz range. Operation at these frequencies necessitates the use of advanced electromagnetic theory in order to properly characterize the microwave feed network. The radiative properties of the antenna also require the use of these advanced theories for analysis. This theoretical framework does not always lead to a solution that is easily identifiable, and therefore optimization can be applied to provide a final design solution which provides optimal performance for the system.

Modern electromagnetic theory is embodied by Maxwell's equations, which are the characteristic equations that define electromagnetic waves and radiation and are shown in the following equations. These equations have been discussed in [12] as well as many other textbooks.

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{1.1a}$$

$$\nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} \tag{1.1b}$$

$$\nabla \cdot \vec{D} = \rho \tag{1.1c}$$

$$\nabla \cdot \vec{B} = 0 \tag{1.1d}$$

In Equation 1.1, $\vec{E}$ represents the electric field, $\vec{B}$ is the magnetic flux density, $\vec{H}$ is the magnetic field, $\vec{J}$ is the electric volume current density, $\vec{D}$ is the electric flux density, and ρ is the electric charge density. Assuming linear, isotropic, and homogenous media, we also have the constituent equations given as

$$\vec{B} = \mu \vec{H} \tag{1.2a}$$

$$\vec{D} = \varepsilon \vec{E} \tag{1.2b}$$

where μ is the permeability of the medium and ε is the permittivity of the medium. With equations 1.1 and 1.2 we have vector partial differential equations in which the variables of interest, $\vec{E}$ and $\vec{H}$, are also coupled. From these equations we can observe the analytical challenges inherent in the analysis of radiating antenna structures. With such complex equations, scientists are often limited to solve only canonical geometries using exact analysis such as dipole or loop antennas [13, 14, 15]. At times approximations can be made if the antenna structure only slightly deviates from these geometries. For the most general structures, researchers forego exact analysis and utilize numerical algorithms in order to characterize the antenna performance. While numerical analysis can significantly aid in evaluating the performance of a structure, it cannot provide exact analysis which relates the geometrical parameters to the antenna performance. Therefore researchers often resort to *trial-and-error* techniques in order to find a good design (if one exists).

These *trial-and-error* techniques can provide a great deal of information to the investigator, but it can often be quite time-consuming in order to properly search the space with uniformly

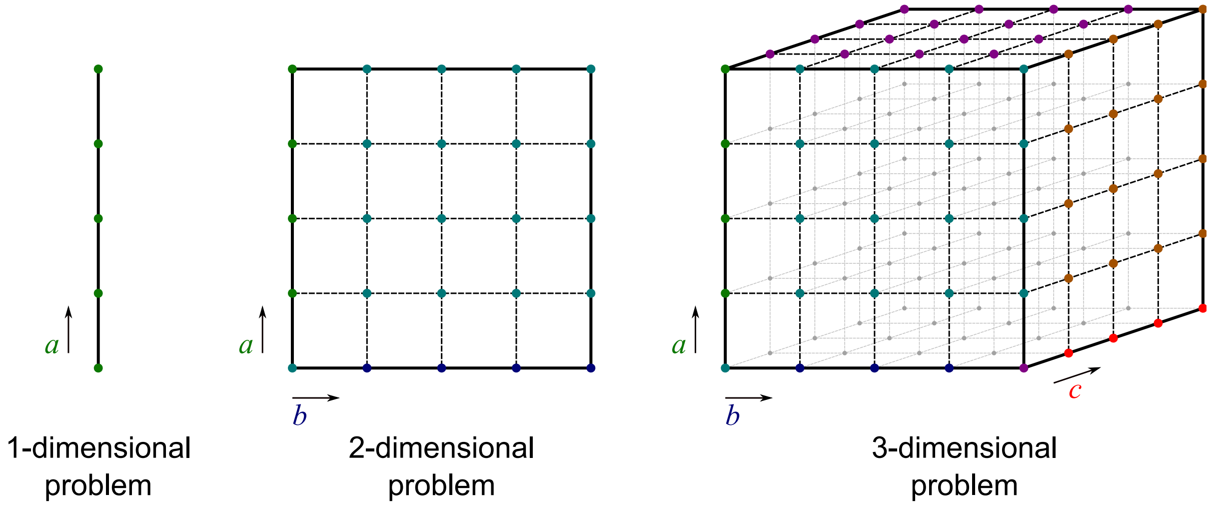


Figure 1.3: Representation of design parameters with uniformly spaced testing points for 1 dimension, 2 dimensions, and 3 dimensions

spaced testing points. This technique only works well for cases involving a small number of parameters, but a different approach is needed for highly dimensional problems involving tens or hundreds of variables. Figure 1.3 shows how an increase in the dimensionality increases the number of points to be tested. With highly dimensional problems, the number of test points increases exponentially, therefore making the overall time required very lengthy. For example, suppose an antenna design had 5 variables that needed to be given values, and *trial-and-error* was used to search the possible solution space. Assuming that 10 points per dimension was used, then a total of 10^5 solutions would be tested. A typical simulation time for numerical electromagnetic simulation tools is on the order of a few minutes, and therefore we assume a 3 minute simulation time for demonstration purposes. The total time needed for the *trial-and-error* testing would be 3×10^5 minutes, which translates to roughly 208 days! Note that this is only for a 5-dimensional problem, which is not necessarily a highly-dimensional problem. Imagine the length of time for a 25-dimensional or even 100-dimensional problem. The length of time needed to search the solution space uniformly could be well beyond one's lifetime, which can hinder further research.

In this context, optimization can offer a significant advantage for these types of design problems where no *a priori* knowledge is given. The primary difference between an optimization versus *trial-and-error* techniques is that the optimization guides the new points to be tested based on previous performance evaluation. This knowledge can significantly reduce the amount of time to get to a good solution, and there exist many different types of optimization techniques discussed in Section 1.2 which use this information in different ways. Many times a design has a specific requirement that indirectly relates to the more common antenna performance parameters. These parameters may be non-intuitive and might not provide a clear relationship between the design parameters and the non-intuitive objective. In many cases designs have multiple criterion to meet, and this can make it difficult for designers to decide between certain solutions. This case is often refered to as a multi-objective optimization problem and are quite challenging to find a final design. For these reasons, one might choose optimization over *trial-and-error* and exact analysis.

1.2 Optimization in Electromagnetics

It should be emphasized that optimization is not a numerical solver for electromagnetics problems. It cannot be applied to solve Maxwell's equations for an arbitrary problem. Rather, it can be used in conjuction with other solvers that employ numerical techniques to solve Maxwell's equations such as the Method of Moments (MoM) [16, 17, 18], the Finite Element Method (FEM) [19, 20, 21], or the Finite-Difference-Time Domain (FDTD) method [22, 23, 24]. Each of these algorithms have their shortcomings as well as advantages over the others, but a discussion on the topics in numerical electromagnetics is out of the scope of this thesis. Instead, we approach design problems with the black box assumption; these algorithms have already been well established for computing antenna parameters of interest in the optimization and no further details are necessary for the construction of the optimization run. There may be some occasional discussions on the choice of numerical algorithm due to speed considerations, but this will usually be the farthest extent of our discussion on these algorithms.

Some researchers have presented numerical techniques involving nature-inspired optimization having the appearance of a numerical solver for Maxwell's equations, but this impression is not the most accurate one. These new techniques involving nature-inspired optimizers provide simplified models of antennas using infinitesimal dipole models [25]. They can provide a fast analytical alternative in comparison to the techniques previously mentioned, and many parameters which may have been difficult to simulate are now achievable at low computational cost. This technique has been applied to design problems for aperiodic arrays [26] as well as MIMO systems [27]. The approach is to place infinitesimal dipoles in the same volume containing the antenna. The near fields along an observation plane are assumed to have been computed by a MoM simulator for the original antenna of interest. An optimizer is then used to find the coefficients of the dipoles and their locations in space by comparing the near field data given by MoM and the dipole near fields. While this might suggest that it is another approach to solving far-fields/near-fields of an antenna, the approach's use of optimization is more reminiscent of solving inverse problems $\mathcal{L}[f] = g$, where the source vector f is found by optimization techniques. If the dipoles' location is also being optimized then both the operator $\mathcal{L}$ and the source vector f are being optimized to finally equate to the simulated observation vector g, and this operation conducted by the optimization algorithm has the same function as a matrix inversion, which is still a difficult problem to solve. This hereby reinforces our original statement that these optimization techniques do not solve Maxwell's equations but rather can be used as tools in the process.

By necessity we incorporate these techniques (MoM, FEM, FDTD) into the optimization algorithm in order to link Maxwell's equations to the optimizer. Figure 1.4 demonstrates the typical work flow for an electromagnetic optimizer. The steps below give a foundation for developing an optimization run for a given design [28]. Steps 1-3 are encompassed by the *Establish parameters* box, and steps 4-5 are embodied by the three box cycle which continues until the design meets the objectives. Note that the bird, people population, and DNA strands are often logos representing the Particle Swarm Optimization (PSO), Evolutionary Strategies (ES), and Genetic Algorithms (GA) techniques, which are popular nature-inspired techniques for electromagnetics applications.

1. **Defining the optimization problem.** The first step in the optimization is to define the problem, which can involve defining the antenna topology as well as the parameters of interest. These parameters can vary from application to application, and Section 1.3 will discuss more on the various aspects of antennas that are of interest. At this step, many assumptions can be made which can greatly enhance the convergence of the optimization,

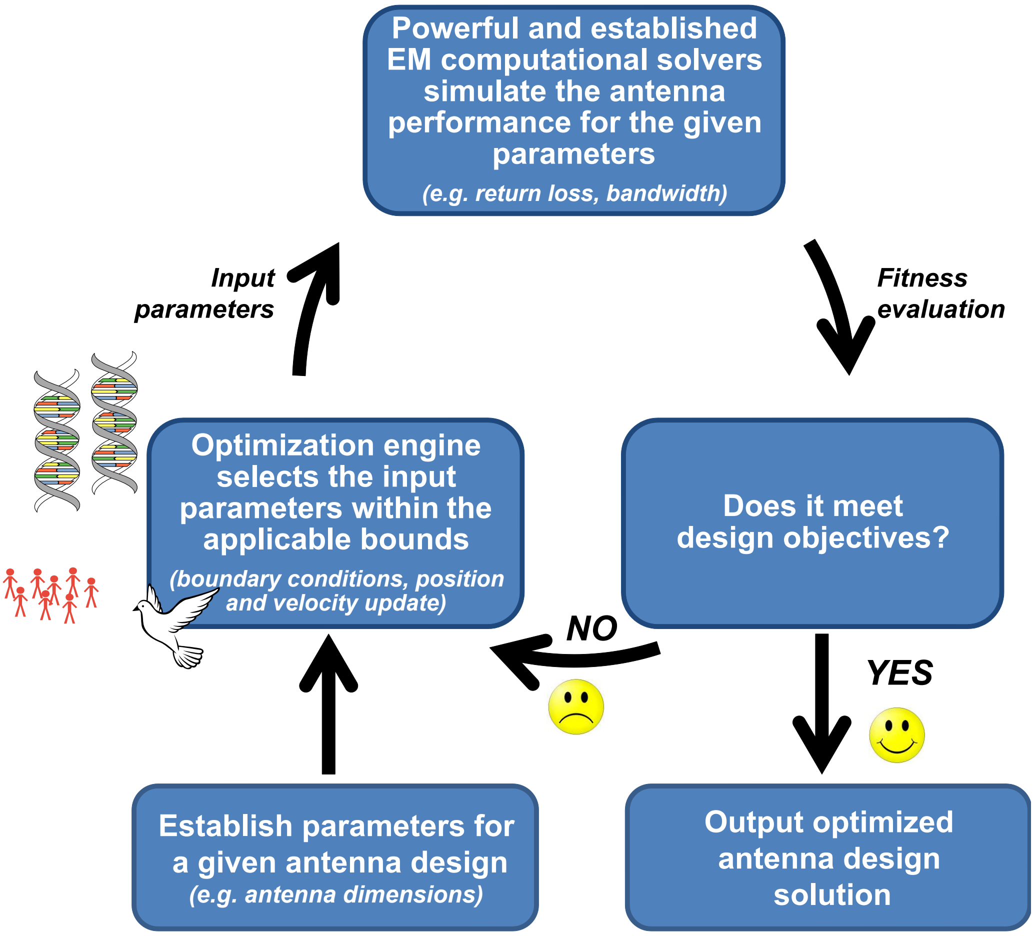


Figure 1.4: Typical flowchart for an optimization algorithm for electromagnetic design problems

and it is the antenna designer's task to minimize the number of variables to optimize. For instance, one could use a completely generalized topology defined by pixels, but this represents a highly multidimensional binary optimization problem which is extremely difficult to optimize. These types of runs are only encouraged when high performance demands are required and no topology demonstrates adequacy. Sometimes these optimizations are used because of their generality, and it can demonstrate limits for possible applications.

2. **Defining the algorithm parameters.** Many optimization techniques have intrinsic parameters that must be specified in order to proceed. Some of these can be solution space boundaries, mutation operators, particle velocities, and many others. Researchers in the evolutionary computation field have been working to develop codes where the intrinsic parameters are predefined (or are adaptive) and do not require any user input. Other techniques still require these parameters, and thus this step still remains an important one in an optimization.

3. **Define a *fitness function.*** In order to optimize a design, one must characterize the parameter(s) of interest that are important in the design specifications. In the optimization community, this is known as either a cost function or fitness function, and it represents a performance measurement. For multi-objective optimizations, this can be represented by a vector function, and for single-objective optimizations this is represented by a scalar function. In general, multi-objective optimizations can take longer than their single-objective counterparts, and therefore one can use a linear transformation, $f = \vec{w}_g^T \vec{g}$, to make a multi-objective optimization a single-objective one. In this equation, f represents the fitness function, $\vec{w}_g^T$ is the weighting vector which properly weights the optimization goals, and $\vec{g}$ is a vector representing the goals (parameters) of interest in the optimization.

4. **Initialize test points.** There are different techniques on how to initialize the test solutions, and most nature-inspired optimization techniques randomly generate starting points based on a given solution space. Some use uniformly distributed initialization points whereas others might use gaussian distributed initialization points. Derivative-based methods typically require some *a priori* knowledge which places the test points near the global optima, otherwise the algorithm might converge upon local optima.

5. **Systematically search the solution space.** This step and its substeps distinguish each algorithm from the others. This section defines how each algorithm uses the previous points to choose for the next iteration. Nature-inspired optimizers typically generate the next set of test points by a randomly generated shift in space with the distribution controlled by previous observations. Gradient based methods typically are directed towards zero-derivative points from quadratic approximations given by a Hessian matrix.

    (a) *Evaluate the fitness function.* The fitness function for each testing point is computed by numerical techniques (MoM, FEM, FDTD) and then output as a single number (single-objective) or as a vector (multi-objective). This is often the most time-consuming component depending on the antenna geometry of interest. If there are additional engineering constraints, then these are sometimes incorporated into the fitness function by adding penalty functions [29].

    (b) *Test convergence criterion.* There are many stop criterion for optimization. A simple one that is often used is a maximum number of iterations. One can either stop the optimization if the performance parameters satisfy the criterion desired or if the optimization shows no significant progress over the past few iterations. For the latter, a certain tolerance δ is often provided, and the optimizer stops when the fitness does not change by more than δ over a certain number of iterations.

    (c) *Select the next values to be tested.* Choosing the next solutions to be tested is always dictated by the optimization algorithm. For this particular thesis, we will go over two nature-inspired optimization algorithms, and Chapter 2 will discuss the details on how these two algorithms choose the next points.

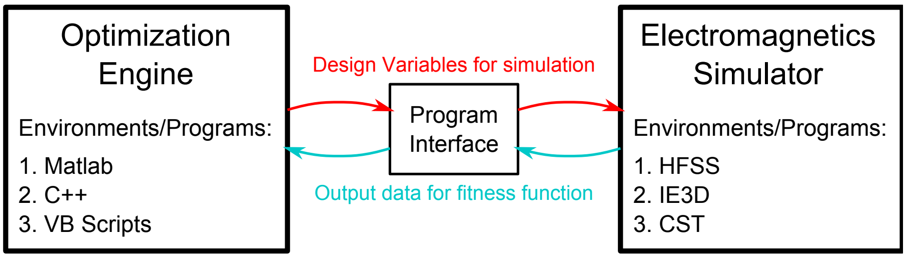


Figure 1.5: Interface configuration for the simulator and the optimization engine

The flowchart in Figure 1.4 and the generalized steps shown above provide only a high level understanding to the optimization process as a whole. With this big picture in mind, we can proceed to discuss the details of each step in the proceeding chapters. We also see from this generalized framework that it is rather straightforward to apply optimization to antenna designs. Once we have a numerical model that accurately outputs the desired antenna parameters, we can establish a program interface which connects the numerical solver to the optimizer as shown in Figure 1.5. The red connections depicted represent the optimizer providing the design variable values to the simulation tool, and the cyan colored arrows represent the output data being given to the optimizer for the computation of the fitness. Once this connection is made then one can proceed forward with the optimization.

There are infinitely many applications where optimization could improve designs for electromagnetic systems, and it also has a rich history leading up to the use of nature-inspired optimization techniques. Optimization has played a major role in electromagnetics, and in the early days the standard gradient based techniques such as the generalized Newton-Rhapson or the Steepest Descent techniques were primarily used [30]. Other techniques which did not require derivatives include the nonlinear simplex method and the pattern search technique [30]. Several examples of common optimization problems in electromagnetics during the 1960-1970s include microwave circuit design [31] and antenna array design [32]. Of course these are only a few representative examples of applications using classical optimization techniques, and by no means is this list exhaustive. In Chapter 2 there will be a further discussion on the categorization of the different optimization techniques. One of the first published uses of nature-inspired algorithms in electromagnetics began with the use of Simulated Annealing (SA) in 1983 for image reconstruction in coded-aperture imaging and its following applications to optical image reconstruction [33, 34, 35]. Later these techniques were applied to microwave imaging in 1991 [36]. Nature-inspired optimization techniques started being applied to antenna design problems when SA was used in the design of correlation antenna arrays [37] in 1988. Around the same time, Evolutionary Strategies (ES) was being applied in 1989-90 to the design of magnets used in Nuclear Magnetic Resonance applications [38, 39, 40]. Research in nature-inspired optimization for antenna applications lay dormant for a few years until Genetic Algorithms (GA) was introduced to the electromagnetics community in its application to broadband, lightweight microwave absorber [41] in 1993. After this awakening, a few significant papers were written on the application of GA to general electromagnetics problems [42, 43, 44].

This began the era where nature-inspired optimization techniques were gradually accepted into the electromagnetics community as a viable option to provide final antenna design solutions, and many other techniques have been introduced other than SA and GA since then. At this present time, it would be impractical to list all applications that have been researched in the literature. Therefore, the following bullets provide several major areas to which researchers have applied nature-inspired optimization algorithms.

- Antenna array nulling [45, 46]
- Thinned arrays [47, 48]
- Ultra-Wideband (UWB) antennas [49, 50]
- Reflector Shaping [51, 52]
- Yagi-Uda antennas [53, 54]
- Radar absorbing material (RAM), microwave absorbers [41, 48, 55]
- Pixelated patch antennas [56, 48]
- Frequency selective surfaces (FSS) [57, 58]
- Antennas for wireless communications [59, 60]

This list provides a brief glimpse into many of the applications which optimization has aided researchers in electromagnetics, and there exist many other projects which would take many book volumes to list. Yet, there still remain many new exciting discoveries yet to be made through the application of optimization towards electromagnetic problems.

1.3 Performance Evaluation in Antenna Design

It is important to remember that the most time consuming component in the optimization algorithms (and even in the *trial-and-error* technique) is the *fitness function* evaluation. The fitness function defines the link between the physical system and the optimizer, and it describes the performance of a given design with a single number (single objective) or a vector (multi-objective). This function represents the performance for a given set of parameter values, and one must develop a fitness function that allows the optimizer to distinguish better designs from unsuitable ones [28]. Therefore, it is imperative to understand the antenna performance characteristics. Otherwise an erroneous understanding can lead to designs that do not meet the specifications desired. It is also the case that a good understanding of the antenna properties can lead one to develop performance evaluation functions which can be less time-consuming. This can be done by a clever formulation or by an assumption which holds for a particular design.

In wireless communication systems, the antenna link and electromagnetic wave propagation represents the system channel, as shown in Figure 1.1. Therefore, the primary goal is to maximize this link despite the poor quality of typical wireless channels. Typically the parameters of interest to the system designers include the operational frequency, bandwidth, and the expected received signal-to-noise (SNR) ratio at the receiver. From the system perspective, the operational carrier frequency is typically determined by the spectrum allocated to that system and often is a predetermined value. The system bandwidth directly affects the available data rate, and the maximum bandwidth in the available frequency band is typically desired. The received SNR directly determines the bit error rate (BER), which in most cases decreases as the SNR increases [5, 61]. By decreasing the BER, one can achieve less communication errors and better overall system performance, which is the general goal of the systems engineer.

For many antenna designs, there is a bi-directional relationship between the parameters and the antenna design. For instance, the operational frequency typically determines the smallest antenna dimensions possible from the classical antenna design perspective. For a given frequency f, the antenna design dimensions are often near $\lambda/2$, where $\lambda = c/f$ is the wavelength in the antenna medium and c is the speed of light. In the reverse sense, one can find the operational frequency of an antenna design for a given resonant dimension $\ell \approx \lambda/2$. To this end, a knowledge of one parameter will typically dictate the other's value.

The SNR is often the most important parameter that guides the antenna designer, and electromagnetic theory must be utilized to characterize the communication link that results from a given design. For line-of-sight (LOS) applications where the electromagnetic radiation from the transmitter has a clear unobstructed path to the receiver, this can be done using the Friis transmission equation as shown below [15].

$$\frac{P_r}{P_{in}} = \left(\frac{\lambda}{4\pi r}\right)^2 G_{abs,t}\left(\phi_t, \theta_t\right) G_{abs,r}\left(\phi_r, \theta_r\right) \left|\hat{\rho}_t \cdot \hat{\rho}_r\right|^2 \tag{1.3}$$

In this equation we can find the ratio of the power received by the load, P_r, to the power input to the antenna, P_{in}. This ratio is defined by the Friis equation above and can be computed with the knowledge of the free space wavelength λ, the distance r between the antennas, the receiving antenna's absolute gain $G_{abs,r}$, the transmitting antenna's absolute gain $G_{abs,t}$ [15], and the polarization loss factor (PLF) $|\hat{\rho}_t \cdot \hat{\rho}_r|^2$. This is depicted by the antennas in Figure 1.6.

It should be noted that this equation is only valid for free space, but it can be used as an rough approximation to the antenna connectivity in the case where ground (often represented by a lossy dielectric [62]) and other objects are included in the scenario. Objects such as the ground can often be modeled as simple scattering structures, and a geometrical optics and geometrical theory of diffraction approach can be used to compute the interference and total signal level at a particular point. Despite the improvement in accuracy that this technique provides, it is extremely environment dependent, and even slight changes can alter the results. One could always optimize for the most typical scenarios encountered, but from the optimization viewpoint this technique is far too profuse for consideration. Rather, one could approach this problem by optimizing the system parameters that can improve the signal level for the average scenario. One begins from this perspective by asking what parameters in equation 1.3 could improve the

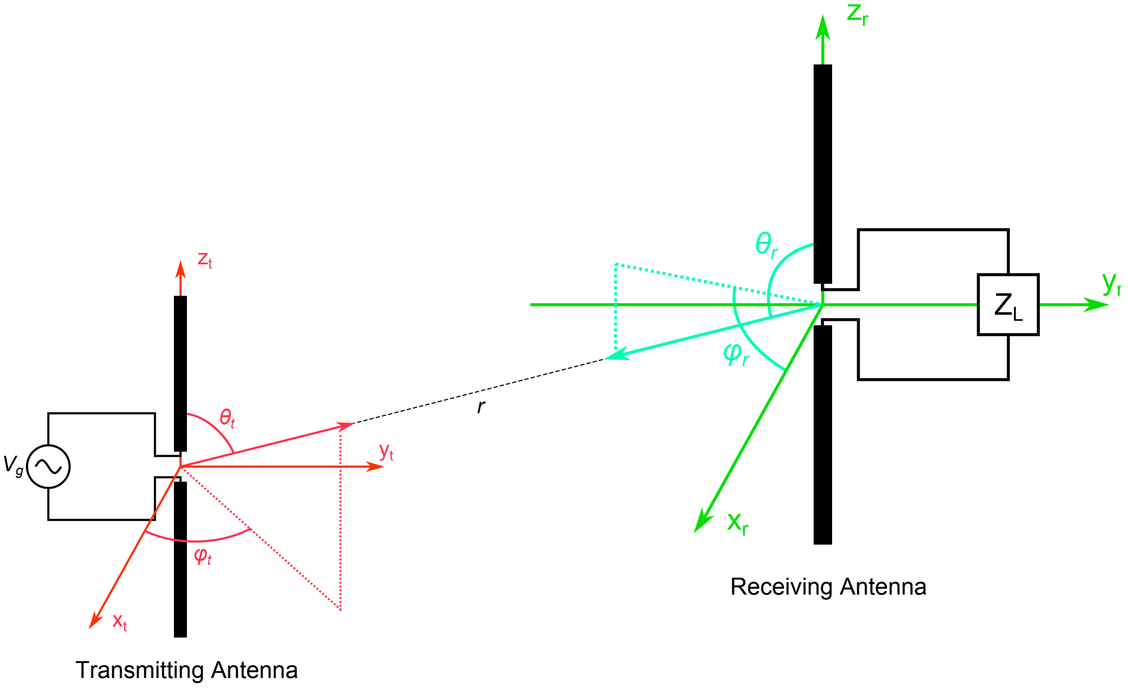


Figure 1.6: Transmitting and receiving antennas in free space

signal accordingly for a given arrangement of objects as well as antenna position and orientation. The parameters λ and P_{in} are often dictated by the system regulations and assigned frequency band, and the PLF $(|\hat{\rho}_r \cdot \hat{\rho}_t|^2)$ and the distance r are defined by the given antenna positions and orientation. The only parameters left for optimization are the absolute gains, $G_{abs,r}(\phi_r, \theta_r)$ and $G_{abs,t}(\phi_t, \theta_t)$, and it has been standard practice in the antenna engineering community to optimize the absolute gain for the best guaranteed performance on the average.

At this point we will only focus on one antenna at a time for optimization purposes. The absolute gain for an antenna can be broken down into several major components as shown in the equation below [15].

$$G_{abs}(\phi,\theta) = e_{cd}\left(1 - |\Gamma|^2\right) D(\phi,\theta) \tag{1.4}$$

Here e_{cd} represents the conductor-dielectric efficiency which represents the power lost to losses in the conductor or dielectric materials. The term Γ represents the impedance matching of the antenna given by

$$\Gamma = \frac{Z_{in} - Z_0}{Z_{in} + Z_0} \tag{1.5}$$

where Z_{in} is the antenna input impedance and Z_0 is the characteristic impedance of the transmission line providing electromagnetic power to the antenna [15]. This quantifies the magnitude of the reflected waves which appear in order to satisfy the impedance boundary condition at the load. These reflected waves represent losses in the power transmitted by the antenna with the term $(1 - |\Gamma|^2)$ [63]. The last term in equation 1.4 is $D(\phi, \theta)$, which is the antenna directivity in the (ϕ, θ) direction. The antenna directivity describes the amount of power is radiated in a particular direction in comparison to an isotropic antenna, which radiates equally in all directions. This provides a number which describes the nature of the radiation for that particular antenna for all angles in the spherical space $(\phi \in [0, 2\pi], \theta \in [0, \pi])$. The quantities D_0 and $G_{0,abs}$ commonly refer to the maximum directivity and absolute gain, respectively, and this notation will be used throughout this thesis. The maximums provide useful information about the radiation pattern in one number. More specifically, it specifies whether the antenna radiates in a few particular directions or more uniformly over the entire spherical space. We will often refer to e_{cd}, Γ, and Z_{in} as the antenna port parameters (or specifications), and $G_{abs}(\phi, \theta)$ and $D(\phi, \theta)$ are referred to as radiation parameters. This may be confusing since the absolute gain incorporates both port and radiation parameters, but this is a common notation since G_{abs} is simply the directivity scaled by its overall efficiency.

Each of these components have a unique dependence on frequency, and it is important to understand this relationship to frequency in order to obtain a useful criterion for antenna bandwidth. Antenna bandwidth can be stated in terms of the various components listed above, and in general bandwidth can be specified by all sorts of antenna parameters. The quote below from C. A. Balanis describes this rather well.

> *The bandwidth of an antenna is defined as the range of frequencies within which the performance of the antenna, with respect to some characteristic, conforms to a specified standard. The bandwidth can be considered to be the range of frequencies, on either side of a center frequency (usually the resonance frequency for a dipole), where the antenna characteristics (such as input impedance, pattern, beamwidth, polarization, side lobe level, gain, beam direction, radiation efficiency) are within an acceptable value of those at the center frequency.*
>
> C. A. Balanis, *Antenna Theory: Analysis and Design [15]*

For many antenna designs, the impedance bandwidth typically defines the antenna's upper and lower limits of operation and is often the most restrictive bandwidth among the other components. Impedance bandwidth refers to the bandwidth in which the reflection coefficient is $\Gamma < 0.316$. This reflection coefficient is often provided using a decibel scale, which is computed by the following.

$$\Gamma_{dB} = 20 \log_{10} |\Gamma| \tag{1.6}$$

In the dB scale, the impedance bandwidth is then guided by the $\Gamma_{dB} < -10$ dB rule of thumb. Using the formula $(1 - |\Gamma|^2)$, a reflection coefficient $\Gamma_{dB} = -10$ dB implies that roughly 90% of the power is accepted by the antenna . It should be noted that hereafter the subscript dB will be dropped and will be implied by the units provided for the rest of the thesis.

As stated in the quote from Balanis, the bandwidth can be restricted by other parameters as well including polarization, radiation efficiency, directivity, etc. However, for certain computational electromagnetics programs, such as the High Frequency Structure Simulator (HFSS) tool [64], it can be computationally expensive to calculate any far-field parameters such as axial ratio or directivity versus frequency. Therefore, one tries to avoid these as much as possible in the fitness function evaluation. Yet, there are times where these parameters are required to meet a specification, and one must incorporate them into the fitness function. Some examples include antennas with wide impedance bandwidths whose radiation patterns can change versus frequency. Since the directivity directly affects the final power received as seen in equation 1.3 and 1.4, one must account for the directivity in that particular direction versus frequency, therefore one would include this in their fitness function.

Of the three components seen in equation 1.4, the impedance matching is almost always incorporated into the fitness functions. For antennas in wireless communications, the radiation pattern is not as much a concern as the impedance bandwidth and antenna radiation efficiency. In radar applications, the radiation pattern can be just as important as the impedance matching. In a nutshell, it is extremely important that all information concerning the scattered fields is extracted properly, and this requires good radiation and port characteristics. The conductor-dielectric efficiency is often not of high concern, and this term is typically above 90% unless lossy materials are present. As we can see from the discussion above, our primary concern will be the impedance matching and the radiation patterns of the design antenna, and these are the parameters on which we will focus in the optimization problems. The next chapters will provide several illustrative and illuminating examples of antenna design optimizations which utilize either port parameters, radiation parameters, or a mixture of both.

1.4 Outline of Work

In this thesis, two nature-inspired algorithms will be introduced and compared, and several optimization problems will be solved using these techniques. Figure 1.7 provides an illustration of the work flow in this document. First some brief background is given in this particular chapter (Chapter 1) with the goal to introduce the subject of optimization within the electromagnetics context and give a brief history on the many possible applications of antenna designs for communication systems and radar. Some of the parameters of interest to the antenna designer are also explained from the connectivity perspective using the well known Friis equation. These parameters will then be used in conjunction with numerical solvers in order to characterize the fitness a given design.

The next chapter will begin by defining the optimization framework to be used for the design problems. It provides definitions for the common terms to be used in connection with the global optimizers and describing the problem in general. The Particle Swarm Optimization and Covariance Matrix Adaptation algorithms will then be described, and some pseudocode is provided for a better picture on the implementation of the two algorithms. The last sections will discuss the additional features required for constrained optimization problems and also discuss the idea of convergence applied towards the nature-inspired algorithms.

Chapters 3 and 4 will compare the two algorithms for several different types of optimization problems. First, a comparison will be made using mathematical fitness functions which allow a quick simulation because the fitness function evaluation is computationally inexpensive. However, the optimization will be done in a resource limited setting, where the number of iterations/function evaluations is limited. An antenna array problem will also be provided in order to compare these two problems using an electromagnetics problem whose fitness can be evaluated

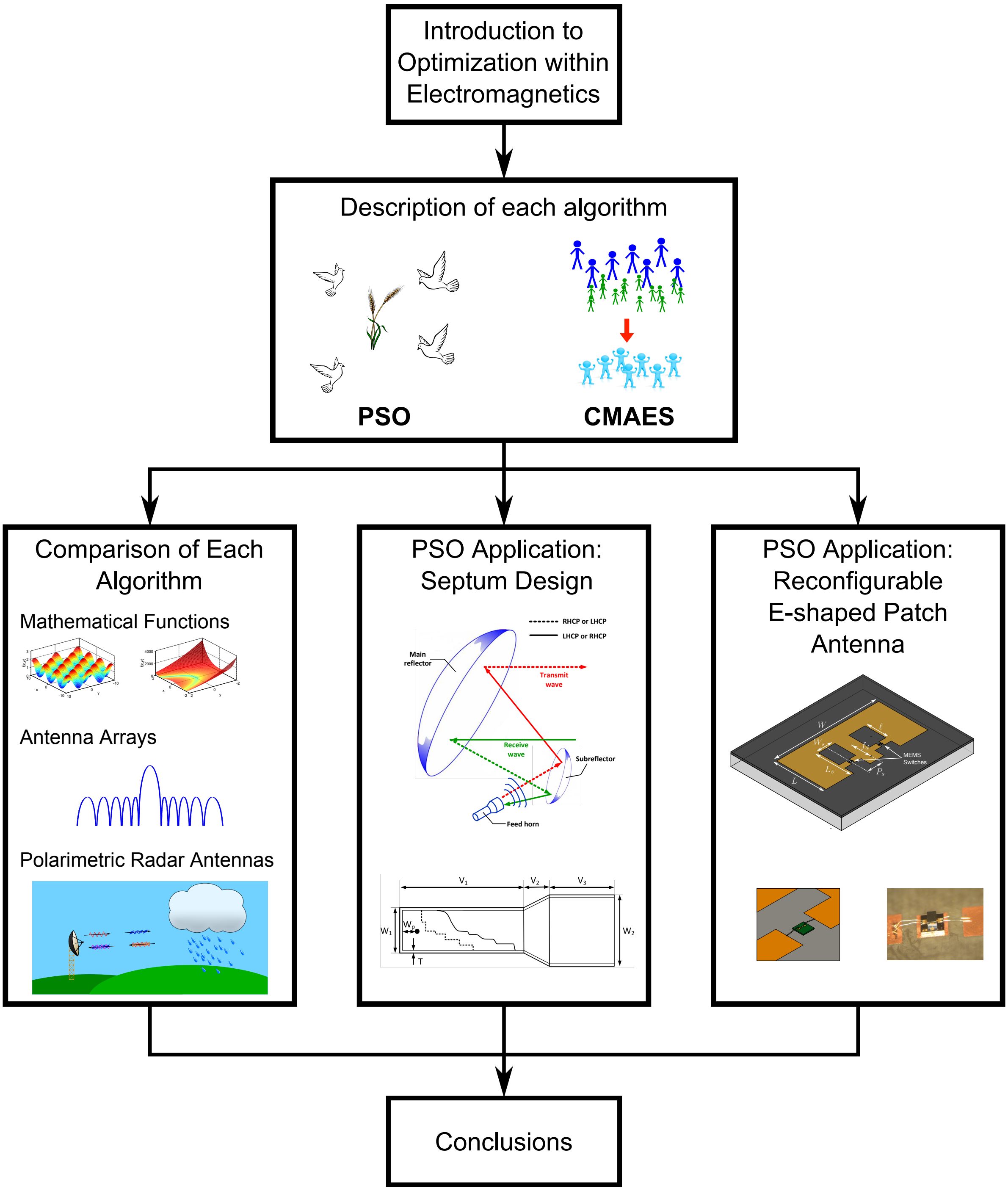


Figure 1.7: Outline of work in this thesis. The introduction to optimization within electromagnetics is given in Chapter 1. The comparison of the algorithms is given in Chapters 3 and 4. The septum design and the reconfigurable E-shaped patch designs are given in Chapters 5 and 6, respectively. The final conclusions are given Chapter 7.

quickly. The goal of this array problem is to minimize the sidelobe level in the sidelobe region by using a non-uniform array. In Chapter 4, the typical parameters in a dual polarize weather radar system are introduced and described, and the connection between the radar performance evaluation and the antenna parameters is given. Two different optimization formulas are used and each algorithm is used to optimize each type. In this chapter there are four optimization runs total, and comparison of the algorithms as well as the optimization formulation is provided.

Two septum designs are optimized using PSO in Chapter 5. The first is a standard stepped septum design commonly seen in the literature. A smoothed septum design using the Sigmoid function is also introduced for possible use in high power applications. These are both challenging highly dimensional optimization problems, and PSO is applied to provide good antenna performance over a specified frequency band in terms of the parameters discussed in Section 1.3. In Chapter 6, the E-shaped patch antenna concept is introduced and two possible reconfigurable E-shaped patch antennas are shown. A simple MEMS circuit model is given to simulate the performance of the reconfigurable antenna using MEMS switches. PSO is applied to both antenna designs to optimize the antenna performance. Bias network solutions are also provided to demonstrate the full implementation of the MEMS reconfigurable E-shaped patch antenna.

CHAPTER 2

Nature-Inspired Optimization Techniques

The general nonlinear programming problem can be stated by

$$\begin{aligned} &\text{Optimize } f(\vec{x}), \quad \vec{x} \in \mathbb{R}^N \\ &\text{Subject to } \vec{g}(\vec{x}) \leq 0 \end{aligned} \tag{2.1}$$

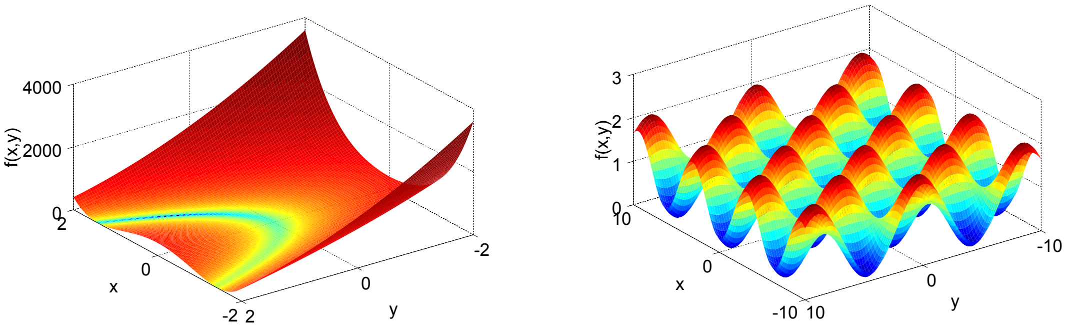
where the function $f(\cdot)$ is the fitness function, $\vec{x}$ is the design variable vector, and $\vec{g}(\vec{x})$ are possible constraint equations [65]. This mathematical problem is often seen in many scientific circles, and therefore its widespread applicability has stirred interest in many different communities. Over the past few decades there have been many approaches proposed for solving this problem, but to this date there has been no such algorithm that is able to solve every problem. The classical techniques are quite useful in solving this problem, and they converge rapidly onto the optima. One of the primary issues is that they often require the fitness function gradient, which may not be available. These techniques are also highly dependent on their initialization [65]. Typically these techniques will converge onto the optimum in the neighborhood of their initial test point. This may or may not be the global optimum, and therefore one must have *a priori* knowledge of the fitness function in order to have global convergence with these techniques. These techniques are often termed *local* optimization techniques due to their likelihood to find local optima. On the other hand, nature-inspired optimization techniques are placed in the category of *global* optimizers. They often mimic particular operations observed in nature, hence the name. They also are classified as stochastic optimization techniques due to their use of random numbers within the algorithm, which aids them in conducting a global search. These nature-inspired optimization techniques have gained interest due to their demonstrated robustness for the global optimization problem in many different research areas. It has only been recently that they have become popular in the microwave and antenna engineering community.

The explosive growth of computing technology and the development of numerical methods in electromagnetics has enabled the use of these nature-inspired optimization techniques to provide final antenna design solutions. Naturally, one might ask what their advantages are in comparison to *trial-and-error* techniques and the classical optimization techniques such as Newton's method or the simplex technique. The typical advantages that these techniques provide are given in the following:

- Derivatives are not required
- Naturally suited for parallel processing
- Converge on a global extrema
- Both continuous or discrete parameters can be used
- *A priori* knowledge of the fitness function topology is not necessary
- Highly multi-dimensional and multi-objective problems can be solved

In electromagnetic problems the fitness functions are typically multimodal, non-differentiable, highly dimensional, non-convex, nonlinear, discontinuous, and ill-conditioned, which make the problem difficult for any optimization technique. Unimodal functions typically have one optimum point where $\nabla f = 0$, whereas multimodal problems can have a multitude of local optima. These multimodal functions can be quite difficult to optimize globally, and many techniques can have issues of premature convergence where the optimizer settles on a local optimum as opposed to the global optimum. Functions such as the 2-dimensional Rosenbrock function are unimodal [66], while other functions such as the Griewank function are multimodal [67]. A 2D version of these functions are plotted in Figure 2.1. As illustrated, there are many more local minima and maxima for the multimodal case whereas there exists only one minimum in the unimodal case. More detail on these functions will be provided in Chapter 3.

While the Rosenbrock function happens to be unimodal, it is still a difficult problem to



(a) Unimodal Rosenbrock function **(b)** Multimodal Griewank function

Figure 2.1: Illustration in the differences between Unimodal and Multimodal functions for optimization

solve due to the narrow ridge-like topology. This condition is often known as an ill-conditioned optimization problem, where the Hessian matrix has a high condition number [68]. For many problems, an ill-conditioned Hessian matrix implies that the gradient does not supply enough information to predict the location of the optimum. This can make it difficult for gradient based algorithms which depend solely on the gradient of the function. Even further, most fitness functions are non-differentiable and make it even more difficult on gradient based optimization techniques to find the global optimum. Therefore, these nature-inspired techniques have become popular among many different research fronts, especially electromagnetics. In order to overcome these issues, the nature-inspired algorithms often have two phases: global optimization versus local optimization. The whole search space must be properly sampled if the algorithm is to find the global optima. Without a good global optimization phase, the optimizer will likely converge to a local optimum point as opposed to a global optimum. Once the full search space has been properly sampled, the algorithm starts decreasing the amount of change given in the next test point compared to the last iteration. By decreasing these step sizes, the optimizer is effectively performing a fine search among the local area. This is the point where most of the test points are in the neighborhood of the global optimum. The transition between global versus local optimization is typically smooth; there does not exist a threshold in which the optimizer suddenly switches over to local optimization. Each algorithm has its own way of shifting from a global to a local one, and each have their own advantages. The main difficulty for the algorithm is in finding the best parameters which work for all fitness function topologies, and many years of research has been spent in developing the most robust optimization algorithm by fine-tuning their intrinsic parameters. Some researchers have also discussed the possibility of hybridizing the global optimizers with the classical (local) optimization techniques in order to exploit the rapid convergence of the classical techniques when the optimization run is in the local optimization stage.

Since a few algorithms will be discussed throughout this thesis, a proper terminology must be established. Therefore we provide the following list of terms below which we will use to describe certain aspects of the optimization problem shown at the beginning of this chapter.

Intrinsic Parameters

Parameters that are used by the optimization algorithm and characterize the algorithm's performance and convergence. This includes parameters that change throughout the optimization run as well as those that remain constant.

Design Variables

These are the N variables that characterize the antenna design geometry that is being optimized (e.g. the length and width of a simple patch antenna). Each set of values for the design values represents a possible design, and we will be representing this with the design vector $\vec{x}$.

Design Boundaries

For bounded optimization techniques, one must provide the lower and upper bounds of each design variable. These bounds are denoted by the $\vec{x}_{min}$ and $\vec{x}_{max}$ variables.

Design Constraints

For many optimization problems, constraints are required in order to abstain from simulating physically unrealizable solutions. Constraints may also be given as part of a specification and therefore must be incorporated into the optimization algorithm. These constraint equations are denoted by $\vec{g}(\vec{x}) \leq 0$ as seen in the problem definition and are discussed more in detail in Section 2.3.

Solution Coordinates

This is a N-dimensional coordinate system whose components are the design variables. For example, the length, width, and height of a patch antenna are represented by the 3-tuple (L, W, h) which is the location in the solution coordinate system.

Solution Space

The solution space is a N-dimensional hypercube in the solution coordinate system which is defined by the limits of each design variable. This is only applicable to bounded optimization techniques. For unbounded algorithms, the solution space is infinitely large. For most bounded optimization techniques, no solution outside of the solution space will be tested. This space will be denoted mathematically as the set $\mathcal{S} = \{\vec{x} | \vec{x}_{min} \leq \vec{x} \leq \vec{x}_{max}\}$.

Feasible Space

While the design boundaries define the solution space $\mathcal{S}$, the design constraints define the feasible space $\mathcal{F} \subseteq \mathcal{S}$. This is defined as the space where the design constraints are satisfied, i.e. $\mathcal{F} = \{\vec{x} \,|\, \vec{g}(\vec{x}) \leq 0, \vec{x}_{min} \leq \vec{x} \leq \vec{x}_{max}\}$. More detail will be given in Section 2.3.

Fitness Function

As mentioned previously in Section 1.2, the fitness function defines the link between the antenna system and the optimizer. This function maps the quality and performance into a single number which allows the optimizer to decide whether a given design is better than others.

Penalty Function

This term is introduced into the fitness function in order to account for constraints or for boundaries. The constraint penalty function will be denoted at $p_c(\vec{x})$ and the boundary penalty function will be denoted as $p_b(\vec{x})$. These are also discussed more in detail in Section 2.3.

This given terminology will be used throughout the rest of the text in order to remove any ambiguities. It should be noted that not all optimizations have constraints $\vec{g}(\vec{x})$, and they are referred to as unconstrained optimization problems. However, if constraints are included then the problem is designated as a constrained optimization problem.

The idea of constraints is slightly different than that of bounded optimization techniques. Some techniques such as Particle Swarm Optimization require upper ($\vec{x}_{max}$) and lower ($\vec{x}_{min}$) boundaries on the design variables. These are denoted as *bounded* optimization techniques. The other case are those that do not require upper and lower boundaries, which are the *unbounded* optimization techniques. While bounded techniques are limited to the specified solution space, it is often the case in electromagnetic optimization problems that boundaries are provided by the nature of the problem. Therefore, the global optimum within those boundaries is the only one of interest because other possible optimum points outside the solution space are unusable designs.

Multi-objective optimization is also a topic of interest in the optimization community, and this presents an even more difficult optimization problem. In the case of multi-objectives, a

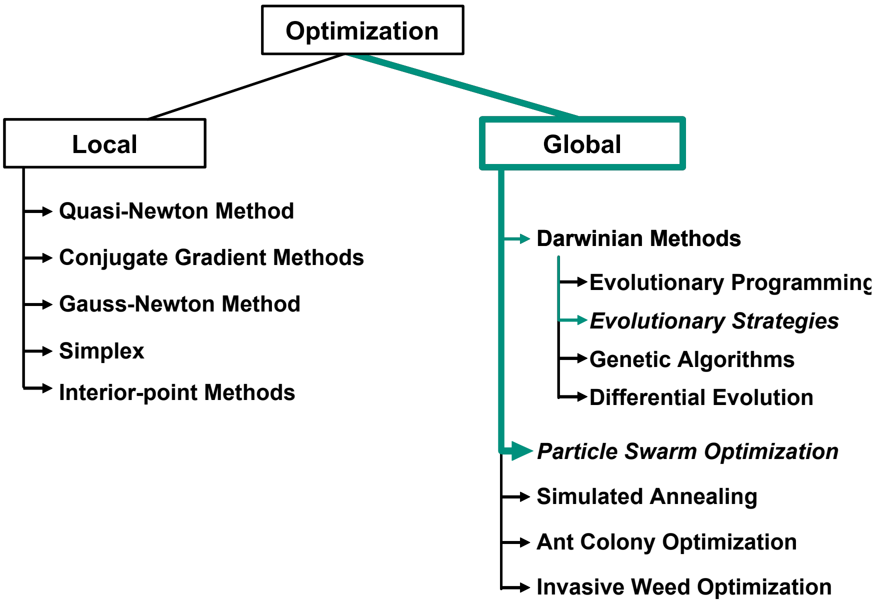


Figure 2.2: A short list of different optimization algorithms and their classification

vector is returned rather than a single number. This characterization makes it more complex to characterize better designs from others. For multi-objective problems a final set of superior designs, or *Pareto front*, is provided as the final output of the optimizer. The *Pareto front* concept allows designers to make a choice when faced with multi-objective designs. The only issue with multi-objective optimization is that it can take a significant amount of time to solve in comparison to its single-objective counterpart. Therefore, our approach to the optimization problems in this thesis are to wrap the multi-objectives into one fitness function by

$$f(\vec{x}) = \sum_{p=1}^{P} \alpha_p f_p(\vec{x}) \tag{2.2}$$

where $f(\cdot)$ is the final fitness function, $f_p(\cdot)$ is the pth objective, and α_p is the proportionality constants. The only difficulty in this approach is to find the appropriate weights $\alpha_p, \forall p \in 1, \ldots, P$. There is no hard and fast rule to provide these coefficients, and our approach has been to equate the ratio of each weight to the ratio of the expected average values of the objectives. This will be demonstrated in later chapters.

Some of the first nature-inspired optimization algorithms were first investigated in the 1960's, which included Genetic Algorithms, Evolutionary Strategies, and Evolutionary Programming.

Since then, many more algorithms have been proposed, and these algorithms are shown in Figure 2.2. Some of these have their merits, and the techniques Particle Swarm Optimization (PSO) and Evolutionary Strategies (ES) will be covered in this thesis. More specifically, the Covariance Matrix Adaptation Evolutionary Strategies (CMAES) will be examined in detail. Afterwards, a detailed discussion on constrained optimization problems will be given. The convergence of nature-inspired optimization techniques will be covered, and then this chapter will close with implementation of these algorithms.

2.1 Particle Swarm Optimization (PSO)

While many nature-inspired optimization techniques have been proposed to the scientific community, most of these algorithms rely on the use of complicated operators (or mechanisms) which mimic naturally occuring processes. However, the Particle Swarm Optimization (PSO) technique uses very simple operators. In particular, PSO exploits the power of social interactions as its primary operator, and the use of this mechanism lends to an inherent algorithmic simplicity. This simplicity implies that only a minimal number of intrinsic parameters need to be defined by the user. The recommended values for the fastest and guaranteed global convergence are also typically more obvious. For PSO, no *a priori* knowledge of the fitness function landscape is necessary to have global convergence, and therefore one does not have to choose arbitrary values for its intrinsic parameters. This and its widely proven use in electromagnetics problems make PSO one of the leading candidates for global optimization techniques in electromagnetics applications.

Many scientists have made observations and experimental models aiming to predict the social behavior of large groups of animals seen in nature. The most prominent groups include bird flocks and bee swarms. Russell and Eberhart, among many other scientists, tried to develop working models that would graphically demonstrate similar properties to the dynamics of a bird flock formation [69]. Their revolutionary idea came when implementing a model for a bird flock which incorporated two forces acting on each individual in the flock. Both of these forces are derived from the individual's memory, which includes a memory of the best point visited by that particular individual as well as a memory of the best point visited by the whole bird flock. The first force drove each individual back to its own best observed point, and this has often been termed the *cognizant drive*. The second drives each individual to the best seen point of the flock, which has been termed the *social drive*. These driving forces are depicted in Figure 2.3. In the figure, each marked point shows the visited points for each bird. Note that the motion of the birds is simulated by finite jumps in space. In other words, the birds are shifted by randomized increments such that its flight (and tested points) are not continuous lines but rather points in space. The figure also provides the location of each birds own best personally visited point ($pBest$) and the best visited point of the flock ($gBest$) with blue circles and green stars, respectively. One may ask what characteristic defines one point better than another and how do these simulated birds know the difference. In their simulation they used the function

$$f\left(x, y\right) = \sqrt{\left(x - x_{food}\right)^2} + \sqrt{\left(y - y_{food}\right)^2} \tag{2.3}$$

in order for each bird (or agent) to evaluate its current position (x, y) in space. The (x_{food}, y_{food}) point represents the location of food that the birds are flying towards. They did this to simulate a bird flock being driven towards food without prior knowledge of its location. For bird 4, one can see its total velocity $\vec{v}_{total}$ decomposed into the two driving forces. The vector $\vec{v}_p$ points towards bird 4's previously best visited point, and $\vec{v}_g$ points towards the best visited point of the flock. By combining these two vectors, one can see that it redirects bird 4 to explore new appealing territory. Russell and Eberhart started their experiments with the intention to model a bird flock's movement, but in the end they had discovered a remarkable technique that seemed to optimize general nonlinear functions such as the one in equation 2.3. Interestingly, the authors chose the name Particle Swarm Optimization because the group reflected more swarm-like characteristics with each individual moving in a quasi-random fashion. Therefore, the group of individuals is typically referred to as the swarm. Also, the term particle (or agent) is often used to describe the individuals in the swarm due to their point-like nature which retains velocity and acceleration.

From observations made to describe this social behavior, it has been conjectured that the

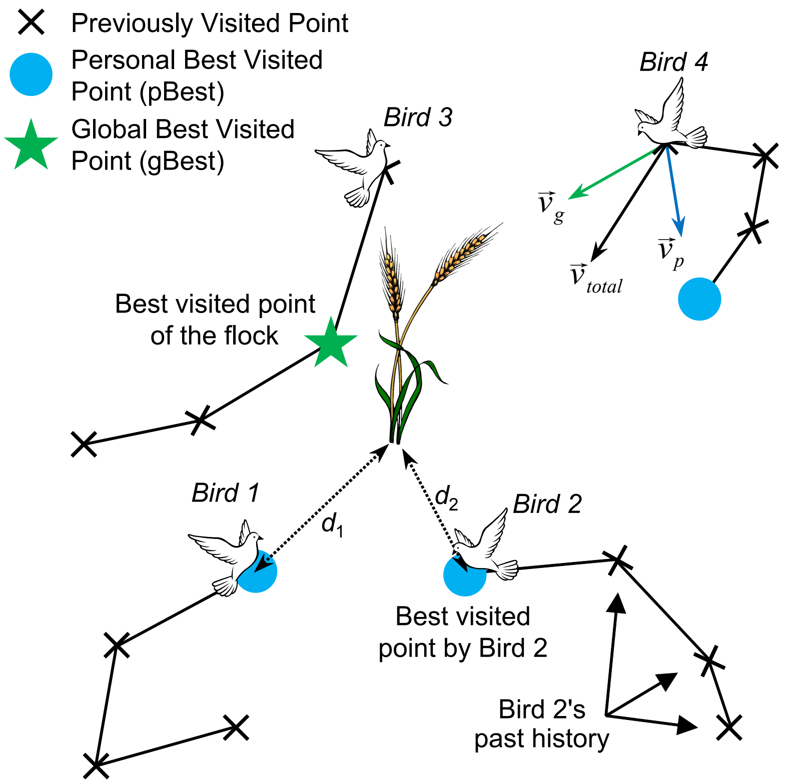


Figure 2.3: A graphical depiction of Kennedy and Eberhart's original simulation model which inspired Particle Swarm Optimization

collective intelligence of the swarm directs them towards the most prominent feeding location. With the gradual acceptance of these conclusions, one might be able to apply these social interactions to more general nonlinear, multimodal, and even non-differentiable functions. Research since the original discovery in 1995 towards progressing the algorithm has skyrocketed due to its proven use in many types of problems. Later, this algorithm was formally introduced to the electromagnetics and antenna community in 2004 [70]. That particular paper explained the PSO algorithm using a bee swarm metaphor, which is another well-used analogy seen in the literature. Since this time, many students within the Antenna Research, Analysis, and Measurement laboratory at UCLA have researched different applications of PSO [51, 71] as well as comparisons to other algorithms, such as GA. There have been much improvements to the PSO algorithm in recent years, but this particular optimization technique is still in its infancy and there exists many more areas of exciting research in its improvement. There now exist several popular variations of PSO which can handle different types of problems encountered. The next subsection will cover the algorithms for the standard Real-valued PSO (RPSO) technique for continuous parameters. Recommended intrinsic parameter values will also be provided for the reader's benefit.

2.1.1 Real-valued Particle Swarm Optimization

The defining equations and thought process for this version of PSO are the most natural, and this is an improved version of the original algorithm presented in [69]. We will use the same terminology of the Section 2.1 and follow suit from its explanations. Each particle in the swarm has an associated velocity, location in the solution coordinate system, personal best visited location ($\vec{p}$), and global best visited location ($\vec{g}$). As stated previously, each particle in the swarm has two memories: a *cognizant memory* and a *social memory* which are affiliated with the $\vec{p}$ and the $\vec{g}$ vectors, respectively. Naturally, each particle would like to revisit the area near its previously best seen point. However, the particle is also aware of the best seen point of the swarm and is torn between the two locations. This is reflected in the equations

$$\vec{v}_i^{k+1} = w^k \vec{v}_i^k + c_1 \vec{r}_{1,i}^k \circ \left(\vec{p}_i^k - \vec{x}_i^k\right) + c_2 \vec{r}_{2,i}^k \circ \left(\vec{g}_i^k - \vec{x}_i^k\right) \tag{2.4}$$

$$\vec{x}_i^{k+1} = \vec{x}_i^k + \vec{v}_i^{k+1} \Delta t \tag{2.5}$$

where equation 2.4 defines the velocity of particle i at iteration $k+1$ and equation 2.5 describes the position of particle i at iteration $k+1$. It also should be noted that the binary operation $\vec{a} \circ \vec{b}$ in equation 2.4 represents the element-wise vector multiplication for $\vec{a}, \vec{b} \in \mathbb{R}^N$. We list out the definitions of each component below and provide a simple description of each intrinsic parameter.

$\vec{v}_i^{k+1}$ The N-dimensional velocity vector of particle i at iteration $k+1$. This represents the speed at which particle i is traveling in the solution coordinate system. As the particles progress in the optimization, the velocity will decrease on average in order to facilitate local exploration.

$\vec{x}_i^{k+1}$ The N-dimensional position vector of particle i at iteration $k+1$. The values of this vector literally represent the values of the design variables being optimized, and the final gBest vector $\vec{g}$ represents the final design values of the optimization.

w^k The inertial weight. By forcing the particle to continue along its previous trajectory, this weighting factor forces the particle to overshoot its target. It perpetuates each particle's original velocity similarly to inertia observed in physics. With a large inertia weight, the particle is forced to explore the surrounding area which promotes more global exploration of the solution space, while a smaller inertial weight allows the particle to make finer adjustments, promoting more local exploration of the solution space [72, 73, 74].

c_1 The nostalgia weight. This is the weight of proportionality which drives the particle to return to its pBest location $\vec{p}$. Increasing this relative to c_2 results in a swarm of isolated individuals which have little social interaction, and the end result is quick stagnation [69].

c_2 The social weight. This weight controls each particle's drive to explore regions where other particle's have had success. Increasing this relative to c_1 motivates the particles to explore the areas of their neighbors, which in turn can result in premature local convergence.

$\vec{r}^{\,k}_{1,i}, \vec{r}^{\,k}_{2,i}$ N-dimensional random vectors. These parameters manifest the stochastic nature of PSO, and each component of the N-dimensional vector has a uniform distribution from $[0, 1]$.

Δt The finite time step. In order to describe the new position of a particle with constant velocity, one must have knowledge of the time travelled. This is included in equation 2.5 as a formality such that the equation would appear similarly to those in elementary mechanics describing particle trajectories. In the literature, it is standard to set $\Delta t = 1$ [70]. Changing its value simply scales the velocity, and the other velocity parameters should be scaled accordingly in order to have a similar performance if one so desires to change this parameter.

These intrinsic parameters are critical to determine the convergence performance for PSO applied to general multimodal optimization problems. If not set correctly, then it is possible that premature convergence upon a local optimum will ensure, and this is to be avoided at all costs. Table 2.1 provides the recommended values for every parameter in PSO. There are other parameters not mentioned in the definition list above due to their exclusion in equations 2.4 and 2.5. One parameter is *swarm size*, and this parameter is typically recommended to be at least equal to the number of dimensions N. There has not been much research devoted to characterizing the best choice of swarm size, but a few sources have reported good performance with this guideline [75]. Another parameter that needs to be set is the maximum number of iterations i_{max}. In order to have the inertial weight linearly decrease from 0.9 to 0.4, one must provide i_{max}, as seen in the formula in Table 2.1. We recommend $i_{max} = 500$ iterations as a starting point for typical optimization problems in electromagnetics. For extremely multimodal problems

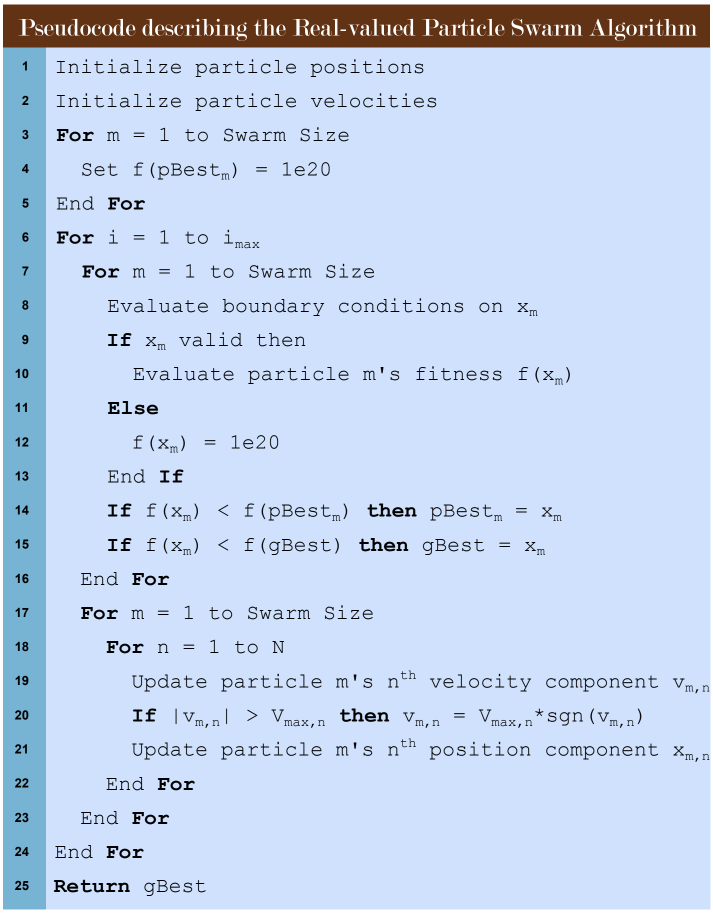
Table 2.1: Recommended Values for the Intrinsic Parameters of PSO when used in Electromagnetics problems

PSO Parameter	*Recommended Values*
c_1	2.0
c_2	2.0
Swarm Size	$(N, 2N)$
Δt	1.0
Max Iterations (i_{max})	500
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

or highly-dimensional problems, it may be recommended to use more iterations. Lastly, v_{max} is used to clip the particle velocity if it gets too high. This ensures that particles do not fly out of the solution space by an extremely large distance. It is interesting to note that the original algorithms before 1998 used v_{max} in order to ensure global convergence, and much smaller values were used in order to tune the performance of PSO. However, it was later found that by linearly decreasing w^k one could generalize v_{max}, which was difficult to tune, to all problems by simply setting v_{max} to half the solution space [74].

There have been many different variations of PSO proposed, but many of them either do not improve performance or they detract from it. Some examples of possible variations tested include a momentum-less PSO where the inertial weight is set to zero, but this example reportedly had poor results [69]. Another style of PSO used a local best instead of a global best, where each particle would remember the best seen position by their immediate neighbors. This provided good results and seemed more resistant to local optima in comparison to the original versions. However this version also took much more time to converge [76]. There has been much more extensive research into improving PSO by adding other operators onto it [77], but the original algorithm still prevails as the most simple and most applicable to all optimization problems. The pseudocode for the original algorithm is given in Figure 2.4 for a better understanding of the full algorithm.

In Figure 2.4, we begin by initializing the particles in the solution space by assigning them a random location with a uniform distribution from $[x_{min}, x_{max}]$. Their velocities are also randomly



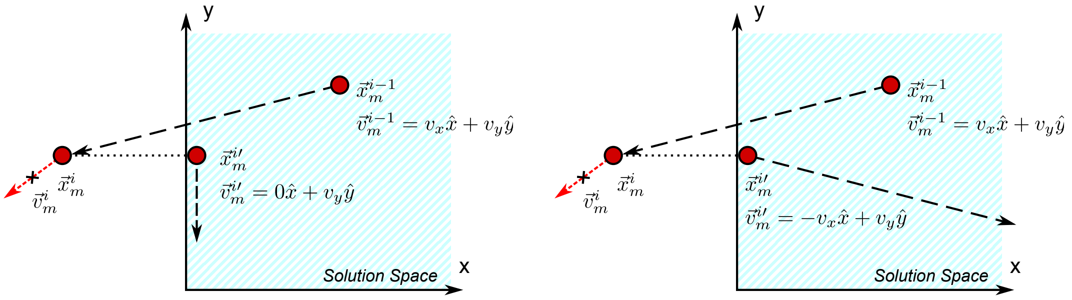
Pseudocode describing the Real-valued Particle Swarm Algorithm

```
Initialize particle positions
Initialize particle velocities
For m = 1 to Swarm Size
  Set f(pBest$_m$) = 1e20
End For
For i = 1 to i$_{max}$
  For m = 1 to Swarm Size
    Evaluate boundary conditions on x$_m$
    If x$_m$ valid then
      Evaluate particle m's fitness f(x$_m$)
    Else
      f(x$_m$) = 1e20
    End If
    If f(x$_m$) < f(pBest$_m$) then pBest$_m$ = x$_m$
    If f(x$_m$) < f(gBest) then gBest = x$_m$
  End For
  For m = 1 to Swarm Size
    For n = 1 to N
      Update particle m's n$^{th}$ velocity component v$_{m,n}$
      If |v$_{m,n}$| > V$_{max,n}$ then v$_{m,n}$ = V$_{max,n}$*sgn(v$_{m,n}$)
      Update particle m's n$^{th}$ position component x$_{m,n}$
    End For
  End For
End For
Return gBest
```

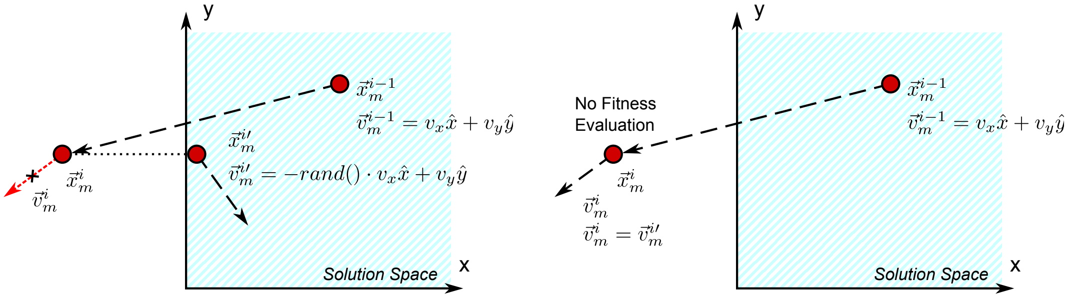
Figure 2.4: Pseudocode implementation of the Real-valued Particle Swarm Optimization technique which minimizes the fitness function

assigned with a uniform distribution from $[-\vec{v}_{max}, +\vec{v}_{max}]$. Since this is a minimization problem, we set the initial fitness for each particle extremely high. Ideally, we would set it at ∞, but this number is not storable in finite sized memory and therefore is set to a high number, 10^{20}.

The next step is to evaluate whether each particle is within the given limits $[\vec{x}_{min}, \vec{x}_{max}]$. If not, there exist several different boundary conditions to keep the particle within these appropriate boundaries. These conditions are often necessary to avoid physically unrealizable systems or to avoid physically insignificant systems. Some examples might include patch antenna with a negative width or a patch antenna where the probe feed does not connect to the patch antenna. These systems either have no meaning or might even force errors in the simulation tools. Therefore they must be avoided by imposing these boundary conditions. The most typical boundary conditions are demonstrated in Figure 2.5, and this includes the absorbing, reflecting, damping,



(a) Absorbing **(b)** Reflecting



(c) Damping **(d)** Invisible

Figure 2.5: Boundary conditions applied to a two-dimensional problem

and invisible boundary conditions. Out of these conditions, the invisible boundary condition has been recommended due to its demonstrated ability to converge upon the global optima for the general optimization problems [70].

The RPSO boundary conditions have often been separated into two categories: *restricted* and *unrestricted* boundary conditions. The restricted boundary conditions contain all particles within the solution space by manipulating their position and velocity, while the unrestricted conditions allow the particle to fly out of the solution space but assign a bad fitness to those particles. The absorbing, reflecting, and damping cases are all considered restricted boundary conditions because they keep the particle within the solution space, as seen in Figure 2.5. The invisible boundary condition allows the particle to fly outside the solution space but sets the fitness value for that point at a particularly high number, e.g. 10^{20}. For certain cases where the global optimum is located near the solution space edge, these boundary conditions can also decelerate the convergence upon the global optima as shown in [78]. If the global optimum is located near the edge of the solution space, then using the invisible boundary condition may slow down the convergence due to a large number of particles flying outside the solution space. There have been other boundary conditions investigated for the standard PSO algorithm including hybrid boundaries such as the invisible reflecting and the invisible damping boundary condition [78], but these are not as popular as the four shown above in Figure 2.5.

The pseudocode shown in Figure 2.4 is configured such that it can handle the implementation of any of the boundary conditions. If one of the restricted boundary conditions is chosen, then the boundary evaluation alters the position and velocity accordingly. If the invisible boundary condition is used, then the particle is considered invalid which forces the algorithm to assign a high fitness to the particle. With this organization one can apply any of the aforementioned boundary conditions.

The next step in the pseudocode updates the pBest and gBest if the newly tested point has a better fitness than the current locations. The last block of pseudocode goes through the whole swarm to update the nth velocity component of particle m using equation 2.4. If a particular component goes above the nth component of the velocity threshold $\vec{v}_{max,n}$, then the velocity magnitude is set to $\vec{v}_{max,n}$ and the direction is that of the original velocity. We use this velocity to find the new location of particle m, where equation 2.5 is used (assuming $\Delta t = 1$).

2.2 Covariance Matrix Adaptation Evolutionary Strategies (CMAES)

The Covariance Matrix Adaptation Evolutionary Strategies (CMAES) technique lies under the umbrella of Evolutionary Strategies (ES), which has a rich history in its development. ES was one of the frontrunners in the early developments of Evolutionary Computation (EC), i.e. nature-inspired optimization. Its development began in the 1960's when working to design bodies with minimal drag per volume [79]. When designing a 2D joint plate in turbulent air flow, the researchers demonstrated that this stochastic procedure outperformed the classical optimization techniques [80]. With those exciting results, researchers went on to further develop the ES technique.

As with PSO, the Evolutionary Strategies technique also takes a heuristic approach to optimization. The Evolutionary Strategies works by evolving a population of individuals, where each

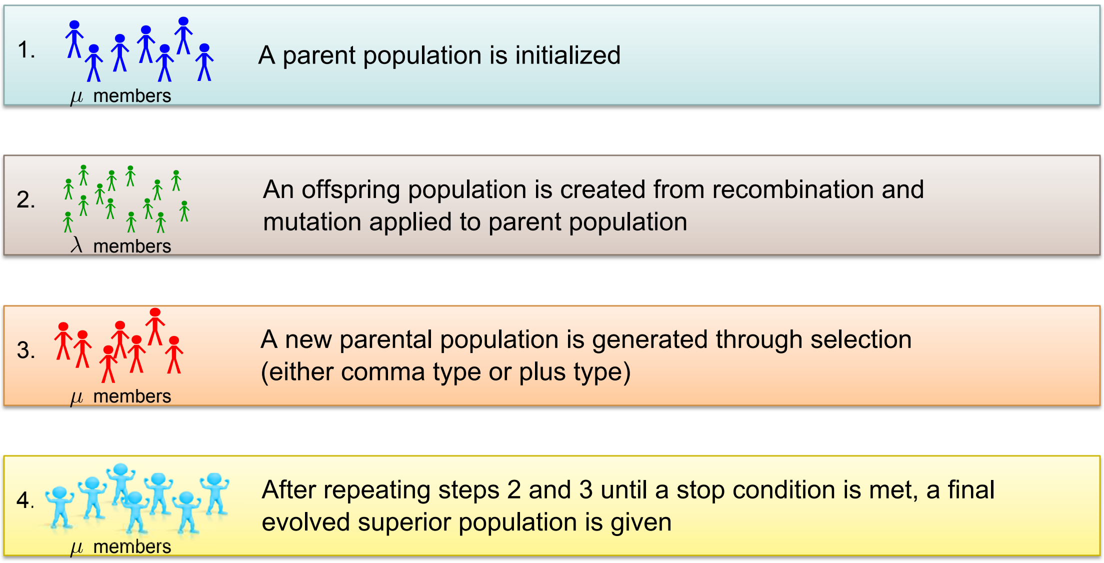


Figure 2.6: Basic evolutionary concept behind Evolutionary Strategies

iteration represents one generation. As time approaches infinity the population will evolve to the most optimal individuals, as shown in Figure 2.6. New generations are born through operators known as recombination and mutation. ES also makes use of the evolutionary idea of *survival of the fittest*, and this is accomplished through the use of a selection operator. The most typical ES technique employs the use of Gaussian distributed random numbers to spawn new members of the population. For CMAES, this distribution can adapt its mean and covariance matrix based on its previous experience, and this adaptation has been claimed to enhance the performance to a faster convergence rate than other nature-inspired optimization techniques.

We begin to shed light on the details of ES by first briefly introducing the terminology often used. Over the course of this text, we have been referring to the parameters inherent to the algorithm as intrinsic parameters. ES further categorizes this into *endogenous parameters* and *exogenous parameters* which are described below.

Endogenous Parameters

The endogenous parameters control statistical properties of the genetic operators (mutation/selection) and can change throughout the optimization run.

Exogenous Parameters

Exogenous parameters are kept constant in a run and they control the size of parent/offspring populations, mixing number, as well as selection type.

One primary difference between the two parameters are that the *endogenous* parameters are often encoded within each of the individuals, while the *exogenous* parameters are encoded into the algorithm as a whole. For CMAES, most of the parameters tend to be exogenous in comparison to other flavors of ES, but these are common ideas within the realm of evolution strategies.

In general, the ES approach has been more mathematically oriented in comparison to other techniques such as PSO, and we will retain the mathematical formalism while still trying to maintain accessibility to the reader. Previously, we had stated that the ES algorithm evolves a population until a final superior generation has been found. At iteration i there are two populations: a parent and an offspring population, to which we will refer by β_p^i and β_o^i, respectively. The parent population is made up of μ individuals $\mathfrak{a}_m^i$ with $m \in 1, 2, \cdots, \mu$. The offspring population is made up of λ individuals $\mathfrak{b}_\ell^i$ where $\ell \in 1, 2, \cdots, \lambda$. Therefore, both populations can be described as the set of its constituents.

$$\beta_p^i = \left\{\mathfrak{a}_1^i, \mathfrak{a}_2^i, \cdots, \mathfrak{a}_\mu^i\right\} \tag{2.6a}$$

$$\beta_o^i = \left\{\mathfrak{b}_1^i, \mathfrak{b}_2^i, \cdots, \mathfrak{b}_\lambda^i\right\} \tag{2.6b}$$

Each individual (both parent and offspring) are represented by their design values $\vec{x}_m$, fitness value $f(\vec{x}_m)$, and their endogenous parameters $\mathfrak{s}_m$. Therefore $\mathfrak{a}_m^i = \{\vec{x}_m^i, f(\vec{x}_m^i), \mathfrak{s}_m^i\}$.

For ES, the three main operations performed on the population guide it towards the superior locations in the solution coordinate system. These three operators include recombination, mutation, and selection, which have similar characteristics to the operators seen in Genetic Algorithms (GA). In the most basic sense, recombination shares the information of a given set of ρ parents (there can be more than two parents), and its ultimate goal is to conserve the good components of the parents. Mutation is the primary source of any variation in the parameters, and changing the scale of mutation either allows for a global search or a more refined search. Lastly, selection directs the population towards the more promising regions in the solution space. Selection implements the *survival-of-the-fittest* by choosing the μ best individuals [80].

There are many different ways to envision the movements of the population as they travel through the solution coordinates. The population centroid $\langle \vec{x} \rangle$ has often been used in the past to describe the whereabouts and general location of the population [80]. CMAES goes a step further and adjusts the population by shifting its mean and adapting the Gaussian covariance matrix throughout the optimization run. At a given iteration i, the location of the mth individual is given by the Gaussian distribution

$$\vec{x}_m^i \sim \mathcal{N}(\langle\vec{x}\rangle^i, (\sigma^i)^2\mathbf{C}^i) \sim \langle\vec{x}\rangle^i + \sigma^i\mathbf{B}^i\mathbf{D}^i\mathcal{N}(0, \mathbf{I}) \tag{2.7}$$

where σ^i is often denoted as the step size, $\mathbf{C}^i \in \mathbb{R}^{N\times N}$ is the covariance matrix of the Gaussian distribution, $\mathbf{B}^i \in \mathbb{R}^{N\times N}$ is the matrix of eigenvectors corresponding to $\mathbf{C}^i$, and $\mathbf{D}^i \in \mathbb{R}^{N\times N}$ is the diagonalized matrix whose elements are the square root of the eigenvalues of $\mathbf{C}^i$ [81]. The corresponding eigendecomposition of the covariance matrix $\mathbf{C}$ is given by

$$\mathbf{C} = \mathbf{B}\mathbf{D}^2\mathbf{B}^T. \tag{2.8}$$

At first glance, this representation may seem very abstract, and some interpretation is required in order to fully understand this representation. In order to simplify the explanation, let us assume that $\mathbf{C} = \mathbf{I}, \forall i$. Now a closer examination of equation 2.7 shows that there are only two variables to control: $\langle\vec{x}\rangle^i$ and σ. The variable $\langle\vec{x}\rangle^i$ controls the center of the Gaussian distribution while σ controls the spread of the distribution.

As a simple illustration of CMAES, we depict the movement of a population as it samples the space in Figure 2.7. We randomly start with an initial $\langle\vec{x}\rangle^0$ and then initialize the other particles based on a given σ^0. The black colored dots represent initialized individuals at $i = 0$, and the lighter colors represent the progression in iterations. The next centroid $\langle\vec{x}\rangle^{i+1}$ is chosen to be equal to the best individual of the offspring population β_o^i, and we arbitrarily set $\sigma^{i+1} = 0.8\sigma^i$. The circles demonstrate the isocontours of the Gaussian distribution at iteration i. This simple algorithm resembles a pattern search algorithm, and of course is not the most suitable optimization algorithm. However, it provides a pictorial understanding on some of these parameters involved. It should be noted that the CMAES algorithm encompasses a much more complicated adaptation process, and it is not only limited to hyperspherical Gaussian

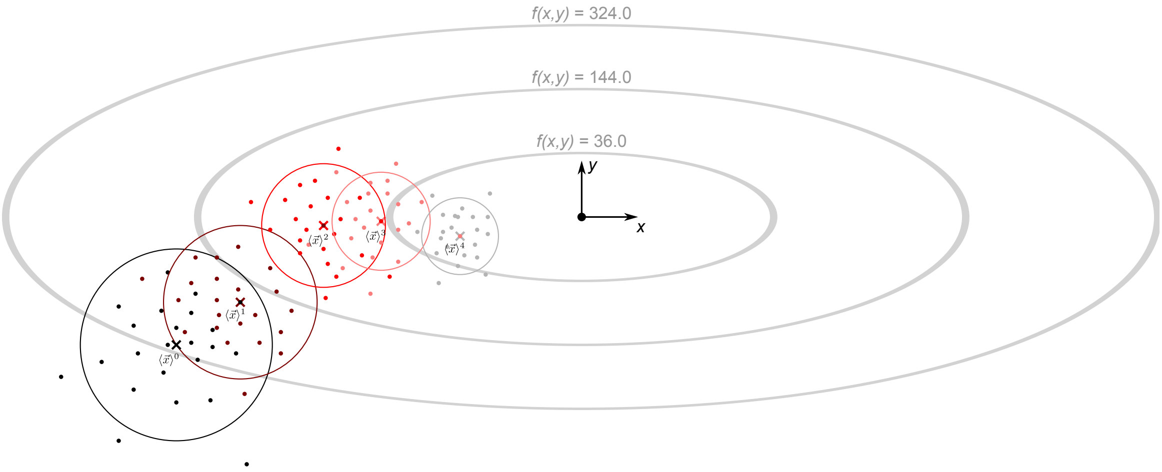


Figure 2.7: Ellipsoidal function $f(x, y) = (x/3)^2 + y^2$ being optimized with a simple best-child evolution strategy. This is a simplified algorithm to explain CMAES.

distributions. In fact, CMAES is able to manipulate the distribution such that the isocontours would form rotated N-dimensional hyperellipses in the solution space in order to accelerate the convergence by using a full covariance matrix $\mathbf{C}$. Now that the concepts, objects, and operations in CMAES have been explained, the focus will shift towards the ES procedure of optimization. We will proceed in the same order as Figure 2.6. We begin with the initialization of the parent population β_p^0. It must be emphasized that the Evolution Strategies technique is an *unbounded* optimization algorithm; it does not require upper and lower bounds on the design parameters. With this in mind, there are two approaches to the initialization given below [82].

1. **Bounded approach**. Many times in electromagnetics we would still prefer to keep the optimization within certain bounds in order to avoid physically unrealizable systems or repeated solutions which can occur for symmetric systems or in periodic fitness functions. Therefore, we initialize the parent population randomly with the distribution at the user's discretion. This distribution might have more weight towards a certain region if some *a priori* knowledge is given about the fitness function.
2. **Unbounded approach**. For more generality, one may avoid a bounded initialization approach, and begin the optimization by providing specified starting point $\vec{x}^0$. The user also provides a $\vec{\Sigma}$ vector which defines the initial variances of the Gaussian distributions. One individual is assigned this position $\vec{x}^0$, and the other $\mu - 1$ individuals are assigned mutated versions of the $\vec{x}^0$ as given by the following.

$$\vec{x}_m^0 = \vec{x}^0 + \sigma \mathbf{B} \mathbf{D} \mathcal{N}(0, \mathbf{I}), \quad \forall m \in 2, 3, \ldots, \mu \tag{2.9}$$

where σ, $\mathbf{B}$, and $\mathbf{D}$ are the same as in equation 2.7 and $\mathcal{N}(0, \mathbf{I})$ is the N dimensional zero-mean unit-variance Gaussian distribution.

The initialization approach that we take in our implementation of CMAES is a hybrid between the two and was recommended in [81, 83]. The hybrid approach starts by generating the initial population centroid $\langle\vec{x}\rangle^0$ with a random position. Since, we do not assume any *a priori* knowledge of the fitness function for our optimization problems in this thesis, we use a uniform distribution for each component of $\langle\vec{x}\rangle^0$ with $\vec{x}_{min}$ and $\vec{x}_{max}$ as our lower and upper bounds. The last part to define is the covariance matrix $\mathbf{C}$ of the Gaussian distribution. Typically, a good initial distribution will include the global extrema within $\pm 3\sigma_i + \langle x\rangle_i^0$ for every dimension, where σ_i is the standard deviation of the initial Gaussian distribution [83]. Therefore, since we assume that the global extrema is located within the hypercube defined by $\vec{x}_{max}$ and $\vec{x}_{min}$, we assign the following parameters [81].

$$\sigma^0 = \max_{i \in \bar{N}} \left(\frac{\vec{x}_{max} - \vec{x}_{min}}{3} \right) \tag{2.10a}$$

$$\mathbf{B} = \mathbf{I} \tag{2.10b}$$

$$\mathbf{D} = \operatorname{diag} \left(\frac{\vec{x}_{max} - \vec{x}_{min}}{\sigma^0} \right) \tag{2.10c}$$

The diag$(\cdot)$ function represents the assignment of the diagonal elements to the vector elements within the parentheses. The distribution is now completely defined, and we can generate the λ offspring for iteration 0 using this distribution. It has also been recommended to use

$$\lambda = 4 + \lfloor 3 \cdot \ln(N) \rfloor \tag{2.11a}$$

$$\mu = \lfloor \lambda/2 \rfloor \tag{2.11b}$$

to start as a minimum population to optimize the fitness function [81, 84]. In cases of extremely multimodal functions, a higher offspring size may be required. Therefore, step 1 in Figure 2.6 has been accomplished. A new parent and a new offspring generation β_o^1 is to be generated. The steps that proceed are repeated until a terminating condition is reached.

Selection is the operator which provides the parent population β_p from the existing populations. It is a deterministic operator, and simply takes the μ best individuals based on their fitness from some specified population. In ES there have been two common types of selection: plus-type which is often represented by $(\mu + \lambda)$-ES and comma-type which is represented by (μ, λ). The $(\mu + \lambda)$-ES applies selection to both the parent population β_p^i and the offspring population β_o^i to form the new parent population β_p^{i+1}. In this algorithm it would be possible for certain individuals to live throughout the entire optimization run without dying, and this has often been referred to as *elitism* [80]. The (μ, λ)-ES only applies selection to the offspring population β_o^i to create the new parent population β_p^{i+1}. Obviously, one stipulation in the comma-selection is that $\lambda > \mu$, and it has been commented that this selection method is less susceptible to local optima while suffering slower convergence [79, 80, 82]. CMAES most often utilizes the (μ, λ) strategy, and it applies both selection and recombination in one operation.

Once the new parent population β_p^{i+1} is created, recombination can be applied on its constituents in order to generate a new offspring population β_o^{i+1}. Recombination takes certain information from the parent population in order to create a new offspring population. For a new offspring individual $\mathfrak{b}_m^{i+1}$, a set of ρ individuals from β_p^{i+1} are chosen to use as parents. The parameter ρ is often denoted when declaring the ES algorithm by $(\mu/\rho + \lambda)$ or $(\mu/\rho, \lambda)$. The recombination operation is then applied to the parent individuals in order to create $\mathfrak{b}_m^{i+1}$. The typical ES programs have either applied *discrete recombination* or *intermediate recombination*. In the discrete recombination, the offspring randomly chooses components from the ρ parents for new design parameter values. In intermediate recombination, each component of the offspring's $\vec{x}$ vector is assigned the arithmetic mean of the ρ parents [79, 80]. However, in CMAES recombination takes place through the assignment of the new population centroid $\langle \vec{x} \rangle^{i+1}$. First the λ offspring of β_o^i are sorted by fitness from lowest $(m = 1)$ to highest $(m = \lambda)$. The the centroid is updated by the assignment

$$\langle \vec{x} \rangle^{i+1} = \sum_{m=1}^{\mu} w_m \vec{x}_m^i \tag{2.12}$$

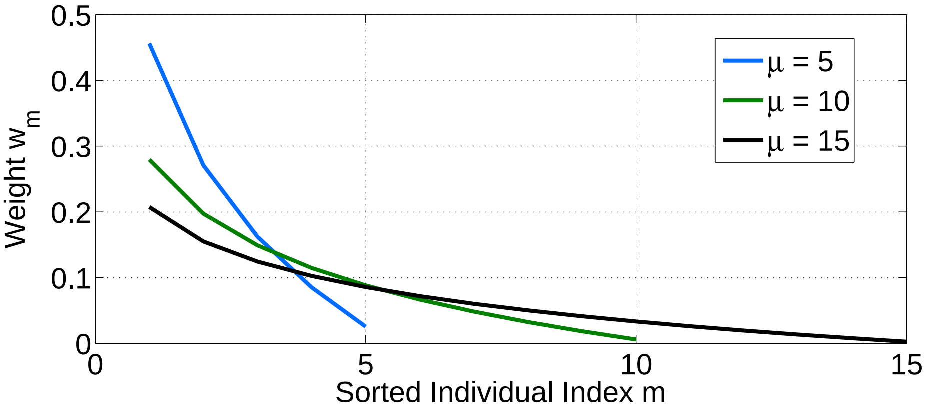


Figure 2.8: Plot of the weights for finding the new population centroid $\langle\vec{x}\rangle^{i+1}$ using equation 2.13

where the weights are assigned the following values.

$$w_m = \frac{\log_2(\mu + 0.5) - \log_2(m)}{\sum_{n=1}^{\mu} (\log_2(\mu + 0.5) - \log_2(j))} \tag{2.13}$$

In our previous simple example, we had found the mean by assigning it the same position as the best offspring location, but these update equations are weighted averages of the μ best performing offspring. It should be noted that the weights are distributed such that $\sum_1^{\mu} w_m = 1$. Figure 2.8 plots some values for the cases where $\mu = 5, 10, 15$ in order to demonstrate some properties of these weights. Our expectation is that the best individuals are weighted most heavily, and indeed Figure 2.8 confirms our expectations. The weight decreases as the index m increases because the fitness worsens as m increases, thereby placing more emphasis on the better individuals. Equation 2.13 applies both selection and recombination to the offspring population β_o^i. Selection is done by the truncation of all other individuals; the summation only includes the first μ individuals and is independent of the others. Since equation 2.13 also takes the weighted average of the individuals, it has been termed as a *weighted intermediate recombination* [83].

At this stage, the new centroid has been accounted for, but the covariance matrix still must be updated. There are a few other terms that must be defined in order to proceed. They are given in the formulas below [81, 84].

$$\mu_{eff} = \left(\frac{\|\vec{w}\|_1}{\|\vec{w}\|_2}\right)^2 = \left(\sum_{m=1}^{\mu} w_m^2\right)^{-1} \tag{2.14a}$$

$$c_\sigma = \frac{\mu_{eff} + 2}{N + \mu_{eff} + 3} \tag{2.14b}$$

$$d_\sigma = 1 + 2\max\left(0, \sqrt{\frac{\mu_{eff} - 1}{N + 1}}\right) + c_\sigma \tag{2.14c}$$

$$c_c = \frac{4}{N + 4} \tag{2.14d}$$

$$c_{cov} = \frac{2}{\mu_{eff}(N + \sqrt{2})^2} + \left(1 - \frac{1}{\mu_{eff}}\right)\min\left(1, \frac{2\mu_{eff} - 1}{(N + 2)^2 + \mu_{eff}}\right) \tag{2.14e}$$

None of these parameters change during the course of the optimization run, and furthermore they only depend on N and μ. The parameters are extremely complicated and are used to enable CMAES for a variety of different applications. The variance effective selection mass μ_{eff} is used to normalize the covariance matrix, and it also represents the effective number of offspring that account for the newly generated mean. The parameters c_σ and d_σ are known as the step-size learning rate and the step-size damping factor, respectively. These two factors are used to control the changes in step size based on a given number of dimensions N and parent population size μ. The next parameter c_c controls the amount of historical information that is retained when adjusting the covariance matrix $\mathbf{C}$. The variable c_{cov} controls the rate of change for the covariance matrix [81, 84].

The next step is to find the next generation's step size σ^{i+1} and covariance matrix $\mathbf{C}^{i+1}$. The first step to find the new step size begins by computing the conjugate evolution path $\vec{p_\sigma}$, which keeps track of the distance traveled by the population centroid.

$$\vec{p}_\sigma^{\,i+1} = (1 - c_\sigma)\vec{p}_\sigma^{\,i} + \sqrt{c_\sigma(2 - c_\sigma)}\frac{\mu_{eff}}{\sigma^g}(\mathbf{C}^i)^{-1/2}\left(\langle\vec{x}\rangle^{i+1} - \langle\vec{x}\rangle^i\right) \tag{2.15a}$$

$$\sigma^{i+1} = \sigma^i \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|\vec{p}_\sigma^{\,i+1}\|}{E\left\{\|\mathcal{N}(0, \mathbf{I})\|\right\}} - 1\right)\right) \tag{2.15b}$$

Since $\mathbf{B}$ and $\mathbf{D}$ are known, the computation of $(\mathbf{C}^i)^{-1/2}$ lends easily to the simplified form

$\mathbf{C}^{-1/2} = \mathbf{B}\mathbf{D}^{-1}\mathbf{B}^T$. Since $\mathbf{D}$ is a diagonal matrix, its inverse is rather simple as given by

$$\mathbf{D}^{-1} = \begin{pmatrix} \frac{1}{d_{11}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{d_{\mathrm{NN}}} \end{pmatrix} \tag{2.16}$$

which is much easier to compute in comparison to the inverse. The last term requiring further discussion in equation 2.15b is $E\left\{\|\mathcal{N}(0,\mathbf{I})\|\right\}$, which is the expected value of the euclidean vector norm of a normally distributed random vector. Its can be calculated by

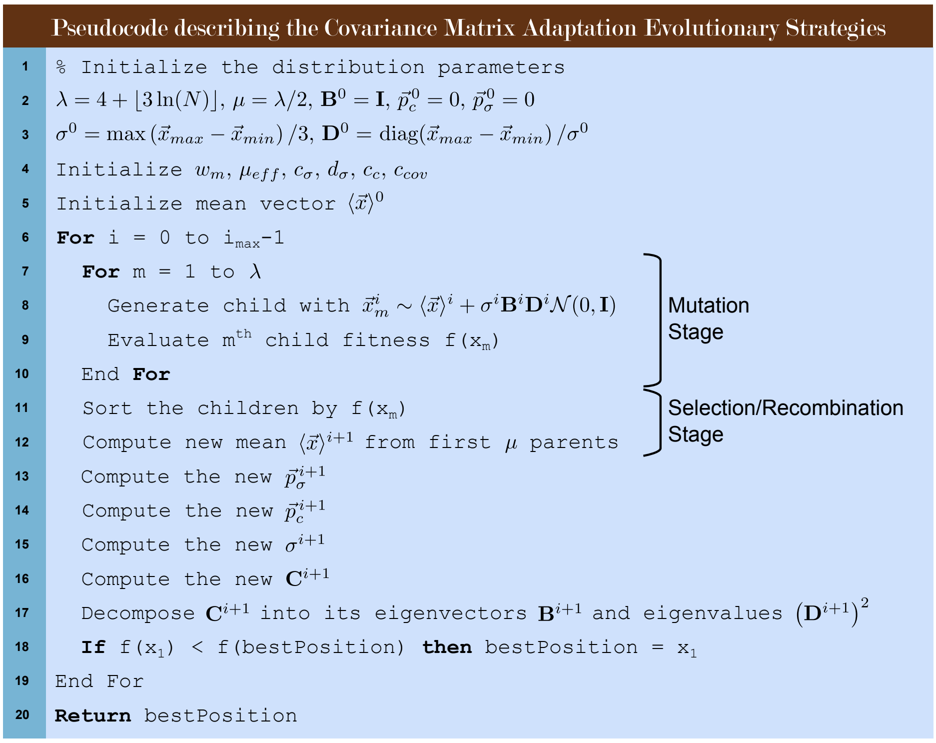
$$E\left\{\|\mathcal{N}(0,\mathbf{I})\|\right\} = \frac{\sqrt{2}\Gamma\left(\frac{N+1}{2}\right)}{\Gamma\left(\frac{N}{2}\right)} \approx \sqrt{N}\left(1 - \frac{1}{4N} + \frac{1}{21N^2}\right) \tag{2.17}$$

This term provides a reference in order to scale the step size σ^{i+1} upwards or downwards. The last component needed is the covariance matrix of the new generation's Gaussian distribution. Computation of the covariance matrix $\mathbf{C}^{i+1}$ follows likewise as shown by

$$\vec{p}_c^{\,i+1} = (1 - c_c)\vec{p}_c^{\,i} + \sqrt{c_c(2 - c_c)}\frac{\mu_{eff}}{\sigma^g}\left(\langle\vec{x}\rangle^{i+1} - \langle\vec{x}\rangle^{i}\right) \tag{2.18a}$$

$$\begin{aligned} \mathbf{C}^{i+1} &= (1 - c_{cov})\mathbf{C}^i + \frac{c_{cov}}{\mu_{eff}}\vec{p}_c^{\,i+1}(\vec{p}_c^{\,i+1})^T \\ &+ \left(1 - \frac{1}{\mu_{eff}}\right)\frac{c_{cov}}{(\sigma^i)^2}\sum_{m=1}^{\mu}\left(\vec{x}_m^{\,i+1} - \langle\vec{x}\rangle^i\right)\left(\vec{x}_m^{\,i+1} - \langle\vec{x}\rangle^i\right)^T \end{aligned} \tag{2.18b}$$

which utilizes the variables defined in equation 2.14. These formulas have been designed such that CMAES will optimize a wide range of problems, and equation 2.18b has two major update terms. The covariance $\mathbf{C}^{i+1}$ inherently is an update from previous history, hence the first term involving the previous iteration's covariance matrix $\mathbf{C}^i$. The second term involves the evolution path vector $\vec{p}_c^{\,i+1}$ and has been termed the *rank one update*. This update elongates the covariance matrix along the path that the mean has traveled. A longer distance traveled by the mean $\langle\vec{x}\rangle^{i+1} - \langle\vec{x}\rangle^{i}$ inherently implies a bigger change in the covariance matrix update. The last term has been denoted as the *rank-μ update*, and this particular term re-orients the Gaussian distribution in order to move toward the direction of the function minimum as well as elongate the ellipse



Pseudocode describing the Covariance Matrix Adaptation Evolutionary Strategies

```
1  % Initialize the distribution parameters
2  $\lambda = 4 + \lfloor 3\ln(N) \rfloor$, $\mu = \lambda/2$, $\mathbf{B}^0 = \mathbf{I}$, $\vec{p}_c^{\,0} = 0$, $\vec{p}_\sigma^{\,0} = 0$
3  $\sigma^0 = \max(\vec{x}_{max} - \vec{x}_{min})/3$, $\mathbf{D}^0 = \mathrm{diag}(\vec{x}_{max} - \vec{x}_{min})/\sigma^0$
4  Initialize $w_m$, $\mu_{eff}$, $c_\sigma$, $d_\sigma$, $c_c$, $c_{cov}$
5  Initialize mean vector $\langle\vec{x}\rangle^0$
6  For i = 0 to $i_{max}$-1
7    For m = 1 to $\lambda$
8      Generate child with $\vec{x}_m^{\,i} \sim \langle\vec{x}\rangle^i + \sigma^i\mathbf{B}^i\mathbf{D}^i\mathcal{N}(0,\mathbf{I})$
9      Evaluate $m^{th}$ child fitness f($x_m$)
10   End For
11   Sort the children by f($x_m$)
12   Compute new mean $\langle\vec{x}\rangle^{i+1}$ from first $\mu$ parents
13   Compute the new $\vec{p}_\sigma^{\,i+1}$
14   Compute the new $\vec{p}_c^{\,i+1}$
15   Compute the new $\sigma^{i+1}$
16   Compute the new $\mathbf{C}^{i+1}$
17   Decompose $\mathbf{C}^{i+1}$ into its eigenvectors $\mathbf{B}^{i+1}$ and eigenvalues $(\mathbf{D}^{i+1})^2$
18   If f($x_1$) < f(bestPosition) then bestPosition = $x_1$
19 End For
20 Return bestPosition
```

Figure 2.9: Pseudocode implementation of the Covariance Matrix Adaptation Evolution Strategy technique which minimizes the fitness function

towards the direction of travel [81, 84]. It should be noted also that the $\vec{x}_m^{\,i+1}$ terms refer to the parents of the new generation β_p^{i+1}, which were the fittest selection from the child generation β_o^i. The offspring mean however corresponds to that of β_o^i, which may be counterintuitive. However it was demonstrated in [81] that this mean $\langle\vec{x}\rangle^i$ produces the correctly directed ellipse in order to predict the location of the function minimum based on the points sampled. In fact, it has also been shown that the covariance matrix roughly approximates the inverse Hessian function for different fitness functions [84, 68]. By using this estimate, the Gaussian distribution is therefore directed towards those points with a predicted zero gradient.

The pseudocode for the CMAES algorithm is shown in Figure 2.9. The previous discussions have already provided some information most of the steps, and this provides an easy-to-follow summary of all the steps. The only minor point that was not discussed was the initialization of the conjugate evolution path and the evolution path. As shown in the pseudocode, these parameters are both initialized to zero by $\vec{p}_\sigma^{\,0} = 0$ and $\vec{p}_c^{\,0} = 0$ on line 2. Line 1 is simply a comment and should be ignored. In this implementation, the best point seen thus far in the

Table 2.2: Recommended values for the CMAES technique

CMAES Parameter	*Recommended Values*
λ	$4 + \lfloor 3 \ln(N) \rfloor$
μ	$\lfloor \lambda/2 \rfloor$
σ^0	$\max(\vec{x}_{max} - \vec{x}_{min})/3$

optimization run is stored in a variable (array) named *bestPosition.* This variable is the final output in this pseudocode implementation as the final values for the design parameters.

In this section, a brief introduction to the concept of Evolutionary Strategies was given, and an implementation of the CMAES version was provided. This algorithm was developed not only to improve convergence over other competing techniques but also to minimize the number of parameters defined by the user. Recommended values for those parameters are given in Table 2.2. As seen from the discussion, this algorithm has been historically more analytical and mathematical in nature in comparison to other techniques such as PSO. It also has been claimed to have improved convergence, and in this thesis we will compare this new technique to verify these claims.

2.3 Applications in Constrained Optimization Problems

Many design problems in engineering often have physical limitations which imply some type of constraints. These design constraints are often due to space or weight limitations as well other various performance issues. Constraints often come in the form of inequalities and constraint equations [85]. Without a proper formulation, the optimization problem can become increasingly difficult due to the limited search space. Therefore these components deserve some attention when working to optimize several classes of antenna optimization problems.

In the beginning of this chapter, the terms *bounded optimization techniques* and *constrained optimization problems* were introduced and delineated. While these terms describe two different pieces of the optimization story, we use these terms *boundary* and *constraint* to describe two different types of inequalities. This dichotomy distinguishes between components that are necessary in *bounded optimization techniques* and those that are applied to all optimization techniques. The inequalities that form the hypercube termed the *solution space* are given as

$$\vec{x}_{max} \leq \vec{x} \leq \vec{x}_{min} \tag{2.19}$$

where $\vec{x}_{max}$ and $\vec{x}_{min}$ form the edges of the solution space. As seen in the previous sections, these boundaries are required in PSO, while in CMAES and ES they are supplementary because they are *unbounded optimization techniques.* An unbounded approach has been applied to many types of problems and provides more generality to the optimization problem. However most antenna design problems have physical upper and lower limits on their dimensions. Therefore these boundaries will be included in the algorithm through various methods in this thesis. Constraints are slightly different inequalities involving more than one design parameter. Using a generalized notation as seen in [85, 86], these constraints can be written as

$$\vec{g}(\vec{x}) \leq 0 \tag{2.20}$$

which captures both types of inequalities through $g_i(\vec{x}) \leq 0$ and $-g_i(\vec{x}) \leq 0$. It is general enough to include equalities $h(\vec{x}) = 0$ as well by defining $g_i(\vec{x}) = h(\vec{x})$ and $g_{i+1}(\vec{x}) = -h(\vec{x})$ to finally use as $g_i(\vec{x}) \leq 0$ and $g_{i+1}(\vec{x}) \leq 0$. To summarize, these *constraints* require attention for both *bounded* and *unbounded* optimization techniques. However, the *boundaries* are only required for *bounded* optimization techniques.

Some terminology has become standard in literature [87, 88, 89, 90] to describe the regions within the solution space $\mathcal{S}$. The *feasible region* $\mathcal{F}$ can be defined as

$$\mathcal{F} = \{\vec{x} \,|\, \vec{g}(\vec{x}) \leq 0, \vec{x}_{min} \leq \vec{x} \leq \vec{x}_{max}\} \tag{2.21}$$

which can be described as the set of points in the solution space which satisfy the constraints, and hence $\mathcal{F} \subseteq \mathcal{S}$. Note also that another region is the infeasible region $\mathcal{I} \subseteq \mathcal{S}$, which can be defined by $\mathcal{I} = \mathcal{S} \cap \overline{\mathcal{F}}$. The region outside of the hypercube will simply be referred to the out-of-bounds region $\overline{\mathcal{S}}$. These regions can be visualized for both two and three dimensional

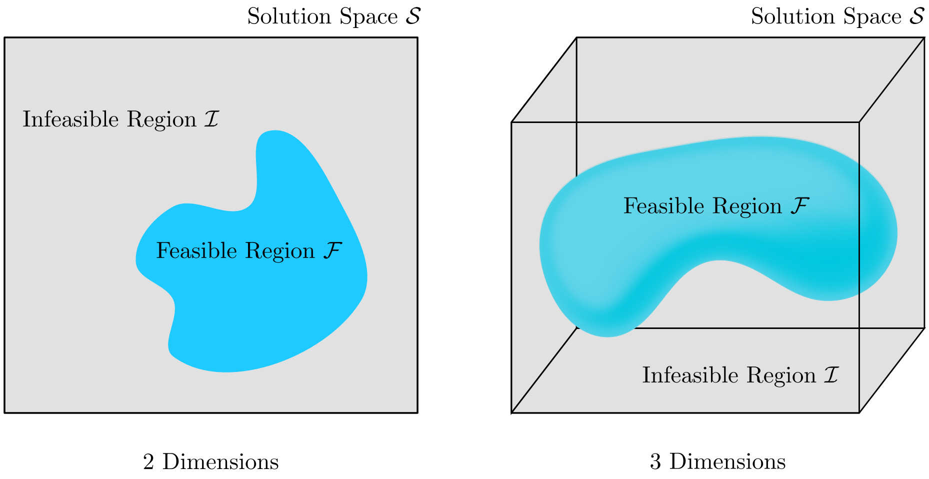


Figure 2.10: Visualization of the Feasible and Infeasible regions

spaces as seen in Figure 2.10.

There are a variety of methods that incorporate the constraints into the optimizer. One technique that has been suggested changes the algorithm's initialization [91] in order to force the initial points to be located within $\mathcal{F}$. This was originally proposed for use in PSO, but this idea could be extended to other algorithms such as GA or CMAES. Others have used used the constraint equations as other fitness parameters in a multi-objective optimization environment. The authors then applied a multi-objective version of PSO to the optimization problem. Another method used for GA as well as other Evolutionary Techniques add a penalty function into the fitness function [87, 88, 86]. This in turn converts the constrained optimization into an unconstrained one. Using this approach, another term $p_c(\vec{x})$ is simply added onto the original fitness function as

$$f(\vec{x}) = f_0(\vec{x}) + p_c(\vec{x}) \tag{2.22}$$

where $f_0(\vec{x})$ is the term which describes the fitness of the antenna alone, i.e. the original fitness function. For many of the optimization runs used in this thesis, the constraint penalty function

$$p_c(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{F} \\ 10^{20} & \text{if } \vec{x} \notin \mathcal{F} \end{cases} \tag{2.23}$$

is used, which simply increases the fitness to a large number if outside the feasible region. This may not always be the best penalty function because it does not provide any information as to the location of the feasible region. Others actually implement the constraint equations $g_i(\vec{x})$ into $p_c(\vec{x})$ [89], and this has been preferred because it can help guide the optimizer towards $\mathcal{F}$. We use the stepped penalty approach for ease of implementation, and it fits naturally into PSO since this has the same appearance as the invisible condition. It should also be pointed out that one advantage of penalty functions is that they can be applied to every technique, and therefore they form a widely usable approach in handling constraints. The stepped constraint penalty function was applied to the optimization problems involving either PSO or CMAES in this thesis.

CMAES and other unbounded techniques also have to incorporate the design boundaries $\vec{x}_{max}$ and $\vec{x}_{min}$ into their algorithm. Whenever running CMAES for optimization problems we incorporate another boundary penalty function $p_b(\vec{x})$. The newly expanded fitness function $f(\vec{x})$ would then become

$$f(\vec{x}) = f_0(\vec{x}) + p_c(\vec{x}) + p_b(\vec{x}) \tag{2.24}$$

where $p_b(\vec{x})$ incorporates the boundaries of the design. A recommended boundary penalty function has been given by

$$p_b(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{S} \\ \|\vec{x} - \vec{x}_{ctr}\| & \text{if } \vec{x} \notin \mathcal{S} \end{cases} \tag{2.25}$$

where $\vec{x}_{ctr}$ is the center of the hypercube given by $\vec{x}_{ctr} = (\vec{x}_{min} + \vec{x}_{max})/2$. By adding this penalty function, the CMAES algorithm naturally is drawn to the solution space $\mathcal{S}$. Without this added penalty function, the population of individuals is free to roam at any point in space, and this can be undesirable. Again it is emphasized that this does not have to be done for PSO because they are bounded algorithms and incorporate the design limits $\vec{x}_{min}$ and $\vec{x}_{max}$ using its boundary conditions. Therefore $p_b(\vec{x}) = 0$ for the PSO technique.

When constraints are introduced into an optimization, it becomes increasingly difficult for optimizer to find the feasible region. In fact, the difficulty increases as the number of dimensions increases. Since the search space is limited by the constraints, the probability that the next test point will fall in $\mathcal{S}$ is dependent on the number of dimensions N. This can be demonstrated with a simple probability exercise. If we assume the distribution of the next test point is uniform throughout the solution space, and then the probability that the next test point will fall in the feasible space is the ratio of the volume of $\mathcal{F}$ to the volume of $\mathcal{S}$. If we further assume for simplicity that the region $\mathcal{F}$ is a hypercube with sides $\Delta s_i, \forall i \in 1, \ldots, N$ then the probability is simplified to the ratio of each side of the two hypercubes as

$$P\{\vec{x} \in \mathcal{F}\} = \frac{\int_{\mathcal{F}} \cdots \int_{\mathcal{F}} dx_1 \cdots dx_N}{\int_{\mathcal{S}} \cdots \int_{\mathcal{S}} dx_1 \cdots dx_N} = \frac{\Delta s_1 \cdots \Delta s_N}{\Delta x_1 \cdots \Delta x_N} \tag{2.26}$$

where $\Delta x_i = x_{max,i} - x_{min,i}$. We observe that the ratio $\Delta s_i / \Delta x_i = p_i \leq 1$, due to $\mathcal{F} \subseteq \mathcal{S}$. This implies that for greater dimensions, the probability either decreases or stays equal due to the product of another since $p \leq 1$. In general the distribution of the next test point is not uniform, but this helps visualize the difficulty that the constraints place on the optimizer. From this we can see the importance and the challenge that a constrained optimization problem places on the designer and therefore it must be considered in the algorithm in order to guarantee good convergence.

2.4 Convergence Analysis using Mathematical Functions

Convergence is always an important issue when discussing optimization methods, and one of the theoretical advantages of the classical optimization techniques is that one can prove their convergence. By analytically demonstrating their convergence, one can also compare how fast they converge towards the optima. However, analytical proofs on the rate of convergence are not typically available with the stochastic optimization algorithms, and convergence towards the global optimum in some cases is impossible to prove. Some have claimed to prove that certain techniques, such as Evolution Strategies, have a given convergence

$$P\left\{\lim_{i\to\infty} f(\vec{x}^i) = f^*\right\} = 1 \tag{2.27}$$

where f^* is the global optima value for a given fitness function [79]. However, this does not describe the speed at which an algorithm will find the global optima. Even a completely random search of the search space will eventually find the global optimum after ∞ iterations. Other researchers have demonstrated parameters related to convergence such as the probability of a successful mutation analytically for specific optimization problems [79]. Again this does not prove its convergence for all problems. Therefore this does not provide much worth for convergence purposes other than some insight on the best values for the intrinsic parameters of a given algorithm.

However, the lack of these proofs does not necessarily take away from their value in global optimization problems, and their usefulness has been proven in a wide range of applications and research projects. Indeed, the nature-inspired algorithms are not always guaranteed global convergence, but researchers have observed that these techniques demonstrate good global convergence on the average case. It is obvious that the pure random search would eventually find the global optima, and these techniques lie somewhere in between a pure random search and their more analytical gradient-based counterparts. They simply exploit the history of points with good fitness and adapt the next testing points' distribution in order to emphasize the areas with good history.

Therefore many researchers have resorted to comparison of these techniques by applying them to several different types of mathematical functions. Some are unimodal and ill-conditioned while others are highly multimodal. It is necessary to compare a wide variety of different functions in order to test their performance. In this section we wanted to introduce this concept as well as the typical curves seen in these optimizations. For example, the two dimensional Schwefel

Table 2.3: Intrinsic Parameters Used to Optimize the 2D Schwefel Function

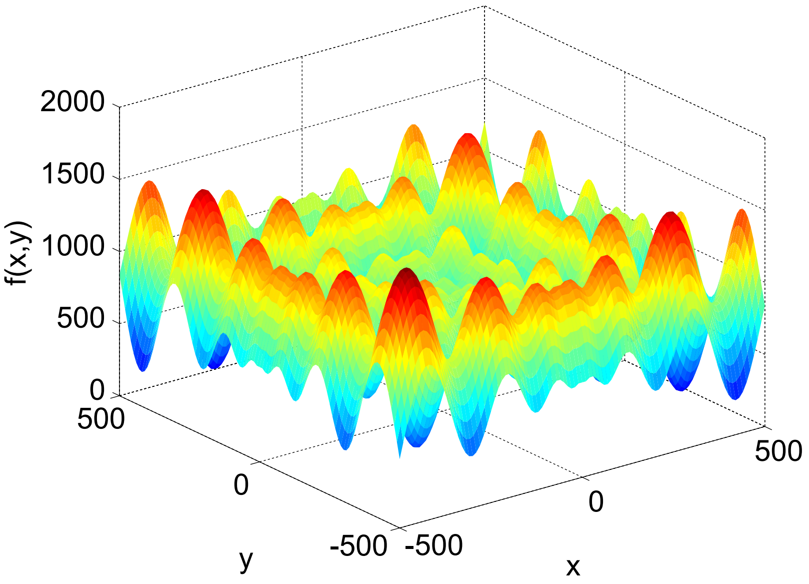
PSO Parameter	*Values Used*
c_1	2.0
c_2	2.0
Swarm Size	4
Δt	1.0
Max Iterations (i_{max})	2000
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

function is given in Figure 2.11. This function can be defined as

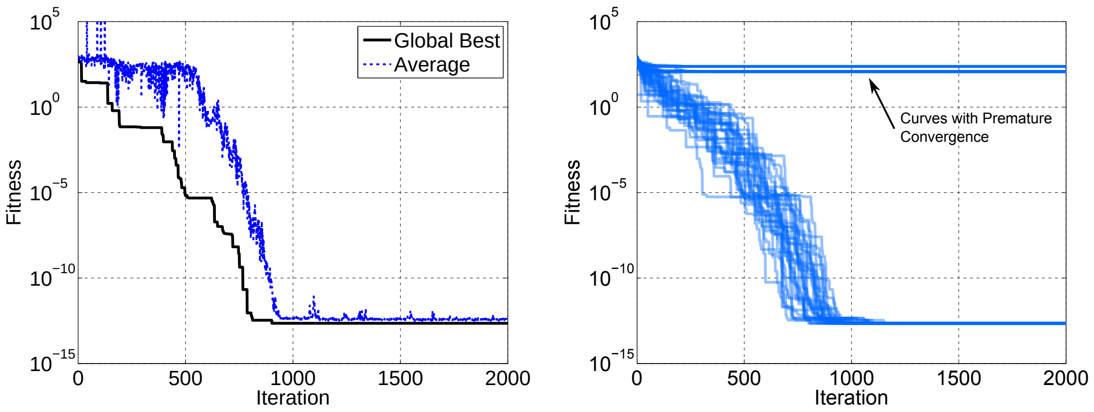
$$f_{schwefel}(\vec{x}) = 418.9828872724339 \cdot N - \sum_{i=1}^{N} x_i \sin(\sqrt{|x_i|}) \tag{2.28}$$

which has a global minimum at $x_i = 420.96874636, \forall i \in 1, \ldots, N$ [92]. This optimization is typically quite difficult for the PSO algorithm, which can fall prey to many of the Schwefel function's local minima. An optimization run using PSO was applied to the two-dimensional Schwefel function, and the results of the optimization run are also shown in 2.11. The solution space boundaries are $[-500, 500]^n$, and the intrinsic parameters for the PSO algorithm used in this run are shown in Table 2.3. We applied more iterations than the recommended values given in Table 2.1 because the Schwefel function is relatively fast to compute. In electromagnetics problems, this is not usually the case, and therefore less iterations are recommended (usually around 500) in order to obtain a solution within a reasonable amount of time. There are two plots that show the results of the optimization run. In Figure 2.11b, two curves are given: the average fitness and the global best fitness for one run. This is a typical plot seen for PSO, and the global best fitness is nothing more than a plot of the fitness evaluation at gBest $f(\vec{x}_g^i)$ for iteration i. The average fitness is simply the average fitness of every particle (or individual) at iteration i given by

$$f_{avg} = \frac{1}{M} \sum_{m=1}^{M} f(\vec{x}_m^i) \tag{2.29}$$



(a) Topology of Schwefel Function

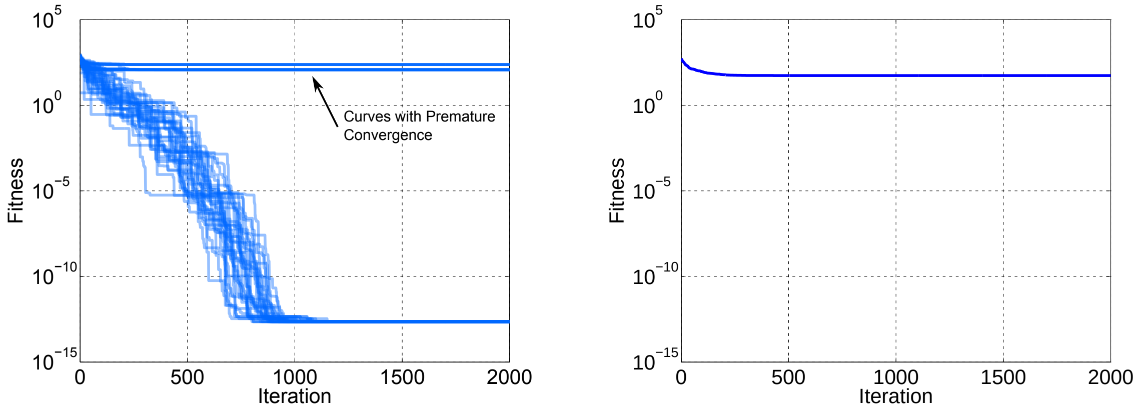


(b) 1 Run **(c)** 50 Runs

Figure 2.11: Application of PSO on a 2-dimensional Schwefel function

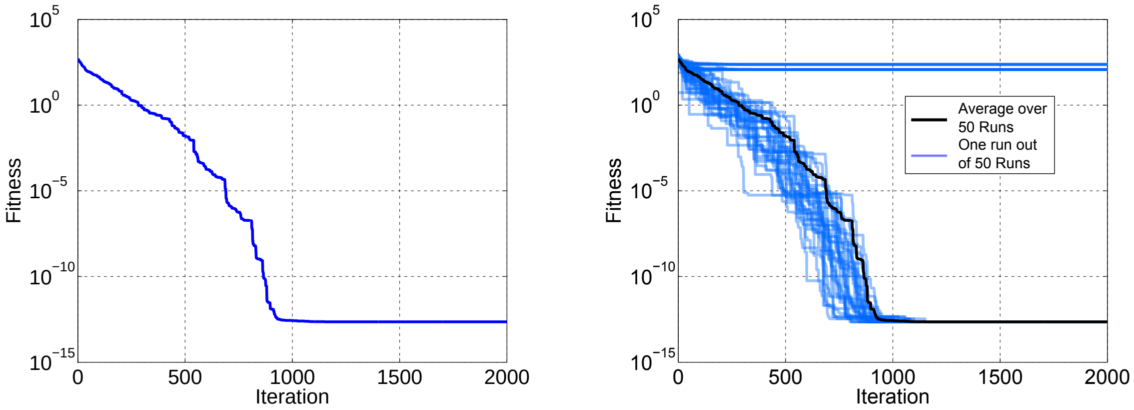
where M is either equal to the Swarm Size (PSO) or λ (CMAES). Now the data seen in Figure 2.11b are only for 1 run, while the Figure 2.11c shows the global best fitness for all 50 runs in one plot. In order to prove that an algorithm is robust, it is often required that one demonstrate its effectiveness for several independent runs. In many cases people provide a 50-run average in order to demonstrate the algorithm's performance, but this may not be effective because some runs might have converged upon a local optima prematurely as seen in Figure 2.11c. The values between the premature runs and the optimized runs are different by several orders of magnitude, and thus any averaging would force the averaged value to be equal to the premature values. Therefore this is not always the best way to depict this information and should be avoided unless no premature runs exist. This is why the transparent plot shown in Figure 2.11c provides a better picture; one can see the results for all runs in one plot. The darker lines occur when more than one curve sits atop one another. For this optimization, 17 runs converged prematurely, while the other 33 runs found the global optimum at $x_i^* = 420.96874636, \forall i \in 1, \ldots, N$. It should be noted that this function should have a value of zero at $\vec{x}^*$, but the value calculated at $\vec{x}^*$ is roughly 4.55×10^{-13} due to some truncation errors. Lastly, we show several different methods for calculating the average convergence over 50 runs in Figure 2.12. The problematic method is shown in Figure 2.12b, where the premature curves take over the converged runs. Another proposed way to plot the average convergence is by excluding the premature curves from the average as shown in Figure 2.12c. This provides a good picture that demonstrates the average curve of a PSO run for the Schwefel function and is compared to the 50 run plot in Figure 2.12d.

In summary, convergence is a difficult thing to prove for every problem in the world of stochastic global optimizers. Therefore comparison between different techniques is made by examining their convergence for a library of functions. The Schwefel function was provided as an example, and several representative curves were explained. These curves will be quite commonly used throughout this thesis to discuss the convergence of a particular run, and some explanation was needed in order to proceed. While global convergence may not be guaranteed for every run in these algorithms, they provide good convergence on the average case scenario.



(a) All 50 Runs Displayed

(b) Averaging with premature convergence curves included in the average



(c) Averaging without the premature convergence curves

(d) Comparison of the 50 runs with the averaging in (c)

Figure 2.12: Comparison of Different Averaging Procedures

2.5 Implementation

The algorithms for both PSO and CMAES are quite simple to finally implement into a program, and several program interfaces were designed in order for the algorithms to communicate with the electromagnetic solvers. The numerical engines that were used throughout this work include HFSS and IE3D. The algorithms for both PSO and CMAES were implemented in Matlab due to its ease in implementation as well as versatility. The baseline CMAES code was also provided by Prof. A. Hoorfar [93] and was further edited. Some of the optimization runs for PSO and HFSS used a VB-scripts implementation, and more on this will be discussed for those particular optimization runs. These codes were run on a server equipped with two quad-core Intel Xeon 2.5 GHz Processors with 32 GB of RAM. Most of the optimization runs in this thesis used a serial computation configuration, where each test point was calculated one-at-a-time.

These algorithms also lend to easy parallelization of the code in order to drastically reduce computation time. By assigning k nodes to evaluate the fitness function, the algorithm experiences a near-linear increase in speed. Clearly the situation where every particle (or children for CMAES) has a designated node represents the fastest and most efficient possible implementation of the nodes, but this is not the only configuration to implement the parallel solution. For this algorithm, there is a small amount of sequential code (roughly 0.1% to 1%), and therefore one can predict the process acceleration by Amdahls Law, which is given by

$$A\left(n_p\right) = \frac{n_p}{1 + (n_p - 1)f} \tag{2.30}$$

where A represents the acceleration (or speedup) of the program, n_p is the number of nodes (or processors) used, and f is the sequential fraction of the code. Since f is approximately zero in all practical applications, one can see that a linear increase in acceleration can occur by using parallelized coding. One can also implement this algorithm as a multi-threaded program on one computer, but the speedup is not linear in comparison to using independent processors for the computation. Our code also incorporated multi-threaded capabilities for the later runs, and only some of the optimization runs use this for program acceleration.

The program implemented for CMAES and PSO has many added functionalities. First, it can work with any provided external fitness function. In order to run a non-Matlab function, a Matlab function interface must be created in order to output the correct fitness value. The functionality in Matlab streamlines the process so that little effort must be used to create these interfaces. Once a fitness function has been developed, the program starts with the default recommended values as given in the previous tables 2.1 and 2.2. If other values for these parameters are desired then the user can change the first few lines of code in order to implement those changes. Once completed, the user can run the full global optimization on the desired fitness function and find a solution.

CHAPTER 3

Comparison Between PSO and CMAES

With the proliferation of evolutionary computing, it becomes difficult to choose between different algorithms for a particular optimization problem. It becomes even more difficult in scenarios where no *a priori* knowledge on the fitness function topology is given, which happens to be the case in most electromagnetics applications. Many of the algorithms have their own inherent advantages, and it remains difficult to provide any concrete mathematical evidence of convergence. As discussed in Section 2.4, the primary way that researchers make comparisons is by conducting optimization tests on a library of mathematical functions.

This chapter will be devoted to a comparison between CMAES and PSO for several different problems involving analytical mathematical functions. This provides a relatively faster methodology for comparison since mathematical functions typically take a relatively short time to evaluate. The first section will focus on a comparison of typical mathematical functions common to the evolutionary computation field. The last section will compare abilities of each algorithm in antenna array problems, which have been a standard electromagnetics optimization problem in literature.

3.1 Mathematical Function Comparison

In order to make a full comparison between the two algorithms, many different types of functions should be considered. In order to provide general observations on these algorithms, one must test several different types of functions. In this way one can conclude whether a certain algorithm has certain deficiencies for a certain type of problem. Below is a list qualities that we are considering in our tests.

Table 3.1: Benchmark Testing Functions

Name	***Function***	***Solution Space***
Sphere	$f_{sph}(\vec{x}) = \|\vec{x}\|^2$	$[-100, 100]^n$
Rosenbrock	$f_{ros}(\vec{x}) = \sum_{n=2}^{N} \left[(1 - x_{i-1})^2 + 100(x_i - x_{i-1}^2)^2\right]$	$[-10, 10]^n$
Cigar	$f_{cig}(\vec{x}) = x_1^2 + \sum_{n=2}^{N}(1000x_i)^2$	$[-5, 5]^n$
Rastrigin	$f_{ras}(\vec{x}) = 10N + \sum_{n=1}^{N}(x_n^2 - 10\cos(2\pi x_n))$	$[-5, 5]^n$
Griewank	$f_{gr}(\vec{x}) = \frac{1}{4000}\sum_{n=1}^{N} x_n^2 - \prod_{n=1}^{N}\cos(x_n/\sqrt{n}) + 1$	$[-600, 600]^n$
Schwefel[1]	$f_{sch}(\vec{x}) = c_{sch} \cdot N - \sum_{n=1}^{N} x_n \sin(\sqrt{\lvert x_n\rvert})$	$[-500, 500]^n$

- Multimodal
- Unimodal
- High dimensionality
- Low dimensionality
- Ill-conditioned
- Well-conditioned

There are many other facets of the fitness functions that could also be considered, and this list does not include all characteristics for brevity. Table 3.1 shows the test functions to be considered in this section. In order to test the effect of dimensionality on the algorithms, $N = 50$ and $N = 5$ is chosen for each function. Some of the test functions have been known to be easier to optimize for higher dimensions, such as the Griewank function. This may seem counter-intuitive, but a demonstration was provided in [67].

The sphere function and the cigar function are both unimodal functions, and the sphere is quite simple to optimize. The difficulty with the cigar function is that it forms a narrow ridge problem that is ill-conditioned [94]. The rest of the functions listed in Table 3.1 are

---

[1] The value for $c_{sch} = 418.9828872724339$

Table 3.2: PSO Intrinsic Parameters for $N = 50$

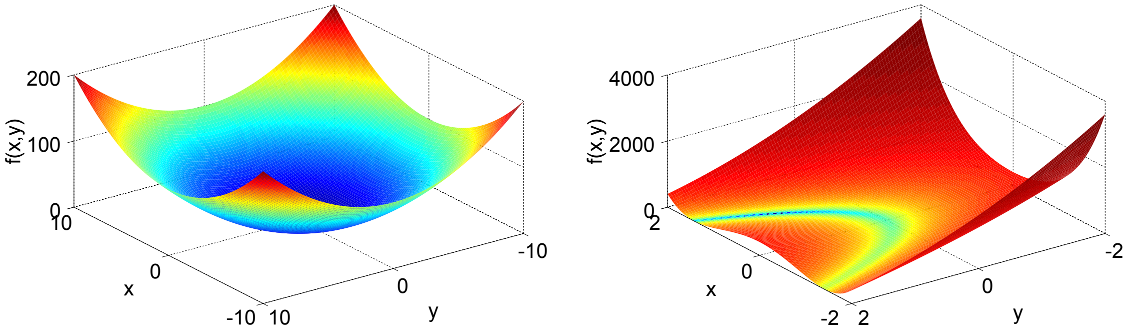
PSO Parameter	*Values Used*
c_1	2.0
c_2	2.0
Swarm Size	100
Δt	1.0
Max Iterations (i_{max})	2000
Boundary Condition	Invisible
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

Table 3.3: CMAES Intrinsic Parameters for $N = 50$

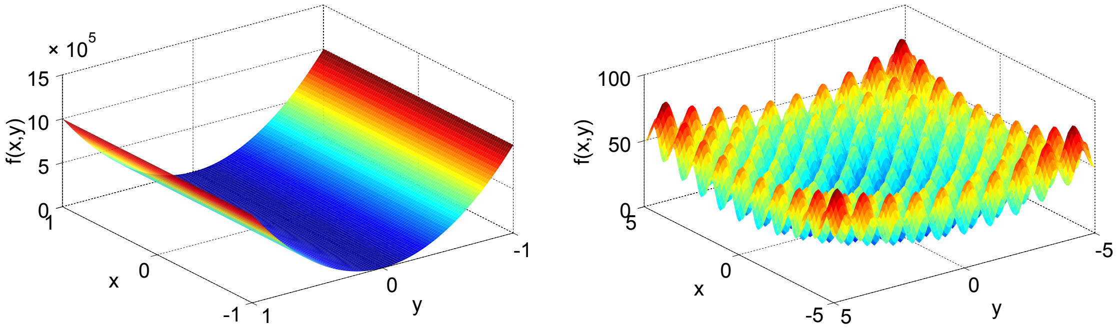
CMAES Parameter	*Values Used*
λ	15
μ	7
σ^0	$\max(\vec{x}_{max} - \vec{x}_{min})/3$

all multimodal functions. The Rosenbrock function is typically known as a moderately ill-conditioned unimodal function for two dimensions, but for higher dimensions there can exist more than one extremum [66]. In summary, we hope to achieve all the characteristics listed in order to make a proper comparison between PSO and CMAES.

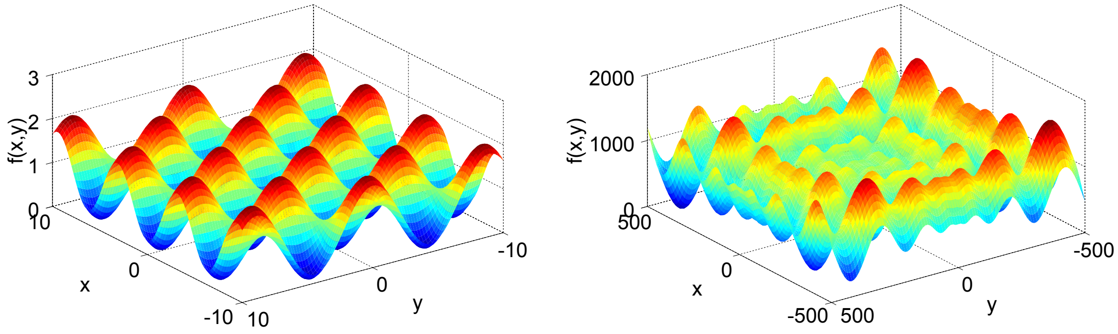
Figure 3.1 shows each function for two dimensions, and this can provide a better understanding of the fitness function topology of each function. Most of the functions have a global minimum at $\vec{x} = 0$, except for the Rosenbrock and the Schwefel functions, which have a global optimum at $\vec{x} = [1, \ldots, 1]$ and $\vec{x} = [420.96874636, \ldots, 420.96874636]$, respectively. Note that these problems represent unconstrained optimization problems. No constraints on the optimizer have been imposed and therefore it is free to search the whole solution space $\mathcal{S}$. In order to make a fair comparison between PSO and CMAES, a boundary penalty function $p_b(\vec{x})$ was applied for the CMAES runs. This ensures that CMAES only searches the specified solution space $\mathcal{S}$ as opposed to an infinite space.



(a) Sphere Function **(b)** Rosenbrock Function



(c) Cigar Function **(d)** Rastrigin Function



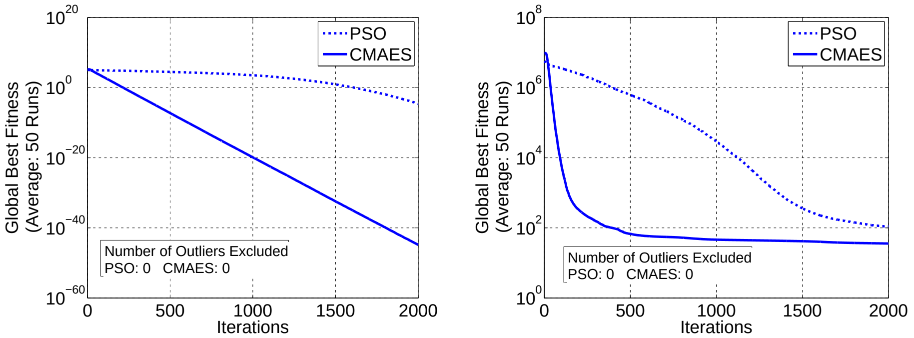
(e) Griewank Function **(f)** Schwefel Function

Figure 3.1: Two dimensional versions of the six testing functions

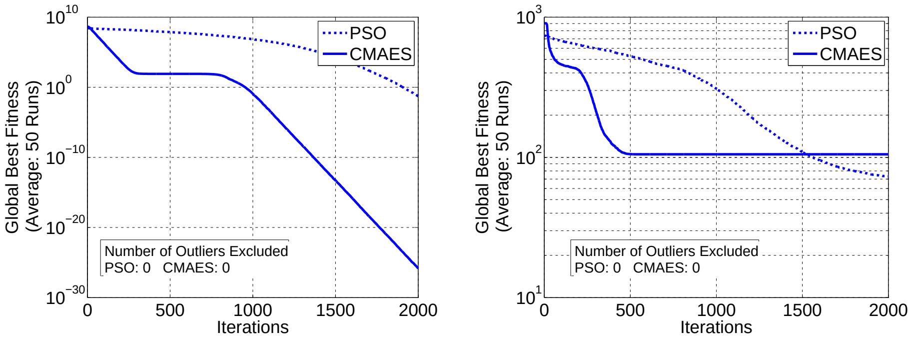
The first case optimized using PSO and CMAES uses $N = 50$ for each test function. The parameters for each algorithm are shown in Tables 3.2 and 3.3. The averaged global best results are provided in Figures 3.2 and Figures 3.3. The first figure shows the results versus iteration while the second shows the comparison versus function evaluation. When making these comparisons, one must also consider the convergence versus function evaluations. In electromagnetics this is more important than the number of iterations because the function evaluation is computationally costly. This is also important in this comparison because a different number of evaluations was used between the two algorithms. For PSO, 100 function evaluations happen at each iteration (unless particles fall outside the boundaries), whereas only 15 function evaluations are performed per iteration for CMAES. Therefore both types of plots are used in order to compare these two algorithms fairly. For the PSO optimizations, the number of function evaluations totals 200,000, while for CMAES the total number of function evaluations is 30,000. The last note to make is that the stopping criterion was a maximum number of iterations, which was set at $i_{max} = 2000$. These things should be taken into consideration for this comparison.

A few points should be understood when making this comparison. First, our interest is primarily to compare these two functions for cases with a relatively small number of evaluations. This means both that the search population sizes as well as the total number of iterations will be small. A minimal number of particles (individuals) were used for each algorithm, and the number of iterations was set to 2000 as a compromise between a thorough search versus a realistic number of function evaluations for electromagnetic problems. Other researchers have even used up to 10^4 individuals and $500 \cdot N$ iterations for different test problems [92]. Increasing these better guarantee convergence for a given problem, but this thesis is focused on optimization in a time and resource limited setting since our interest is for electromagnetics.

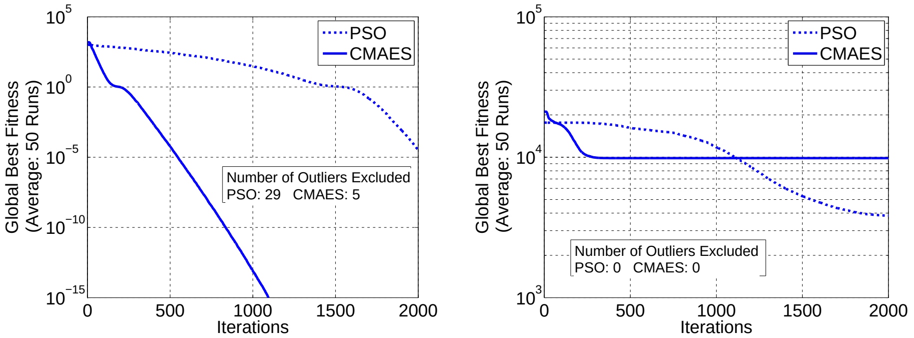
The choice to use a smaller number of evaluations can force nearly all runs to converge prematurely, especially for a large number of dimensions. Therefore most of the optimization runs did not converge upon the global optimum, but rather on some local optimum. While this draws away from guaranteeing the global optimality of the design, the designers are still happy if the final design meets the desired specifications. As discussed in Section 2.4, any outliers which have significantly higher fitness (at least two orders of magnitude higher) are excluded from the



(a) Sphere Function **(b)** Rosenbrock Function

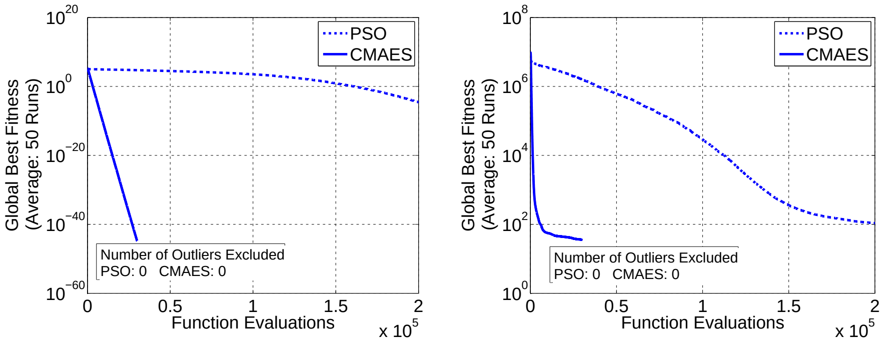


(c) Cigar Function **(d)** Rastrigin Function

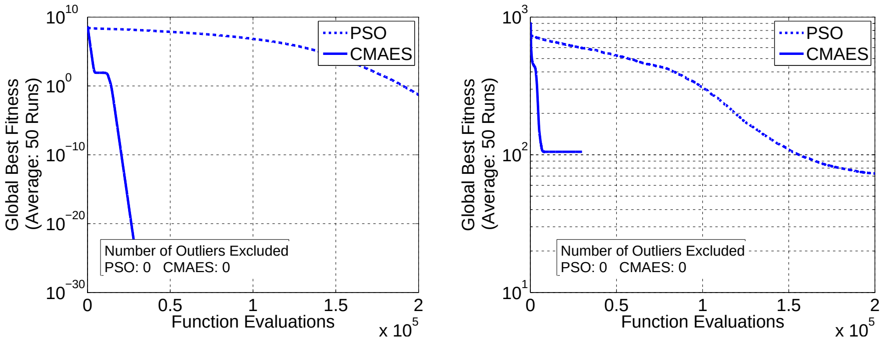


(e) Griewank Function **(f)** Schwefel Function

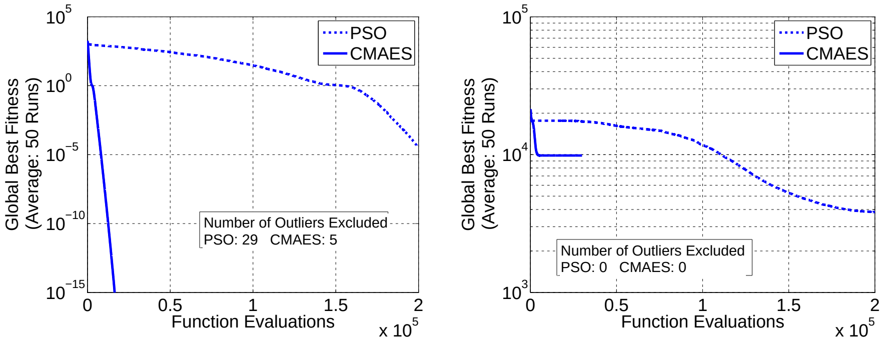
Figure 3.2: Optimization results when $N = 50$ versus *iterations* for the mathematical test functions



(a) Sphere Function **(b)** Rosenbrock Function



(c) Cigar Function **(d)** Rastrigin Function



(e) Griewank Function **(f)** Schwefel Function

Figure 3.3: Optimization results when $N = 50$ versus *function evaluations* for the mathematical test functions

average, and the number of runs excluded are written in the plot for both algorithms.

The CMAES algorithm appears to demonstrate excellent performance and rapidly converges upon the optimal point of the sphere function. PSO also converges upon the point, but has some difficulty in fine tuning the design parameters. In other words, PSO lacks somewhat in the local optimization stage in comparison to the CMAES algorithm. However, The plot scale can be somewhat deceiving, and it should be noted that the PSO algorithm still performed well for this particular problem. On the average it was able to achieve a value near 10^{-7}, which is a significant improvement from its initialization point which had a fitness near 10^{8}.

Both the sphere and cigar function represent unimodal functions, and similar performance is seen in both problems. Similarly to the sphere function, CMAES quickly converged upon the global maximum within a small number of iterations. PSO also converged upon the global optimum and was slower than CMAES to reach a low value. Again, this demonstrates PSO's lack of ability to fine tune its parameters in the local optimization stage. The Rosenbrock function is a multimodal function for higher dimensions, but it is only moderately multimodal in that the number of local optimal are small in comparison to other functions [66]. With this in mind, it makes sense to see comparable performance for the Rosenbrock case as well. CMAES proves to converge extremely quickly upon the global optima, and this type of moderately ill-conditioned problem is one of its specialties. PSO, however, took a longer amount of function evaluations to achieve the same function value as CMAES, but it was still able to achieve significantly better fitness (in comparison to its initialization) regardless of the function's known difficulties.

Overall, we see that CMAES appears to rapidly decrease the fitness function, but for multimodal functions it quickly falls into a local optimum. PSO is observed to make slower progression but is able to overcome some of the local optima and get a lower fitness overall. The Schwefel function is a difficult multimodal optimization problem to solve for high dimensions with limited resources. It has been shown that the CMAES algorithm can optimize this function for higher dimensions with 95% probability of finding the global extremum if a much larger population size is used [92]. CMAES may be able to better overcome some of the local optima by increasing the population sizes to a comparable size used by PSO.

Table 3.4: PSO Intrinsic Parameters for $N = 5$

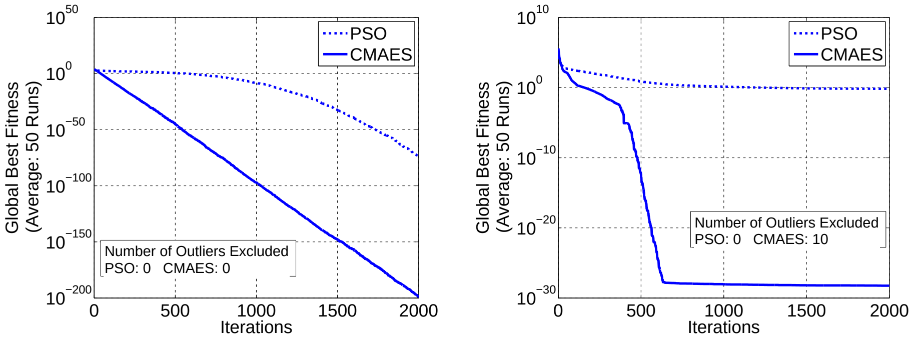
PSO Parameter	*Values Used*
c_1	2.0
c_2	2.0
Swarm Size	10
Δt	1.0
Max Iterations (i_{max})	2000
Boundary Condition	Invisible
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

Table 3.5: CMAES Intrinsic Parameters for $N = 5$

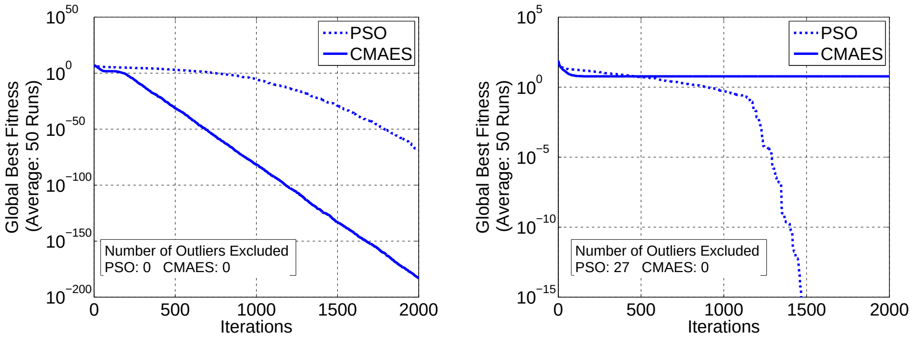
CMAES Parameter	*Values Used*
λ	8
μ	4
σ^0	$\max(\vec{x}_{max} - \vec{x}_{min})/3$

The next comparison to be investigated is the case of 5 dimensions ($N = 5$). We use the intrinsic parameters listed in Tables 3.4 and 3.5, which are similar to the ones previously, except for the population sizes. The same termination criterion was used as well, stopping the runs at 2000 iterations. As seen in the two tables, the population sizes are smaller in comparison to the 50 dimensional case. This is because the choice of swarm size (λ for CMAES) is linearly (logarithmically) dependent on the number of dimensions N.

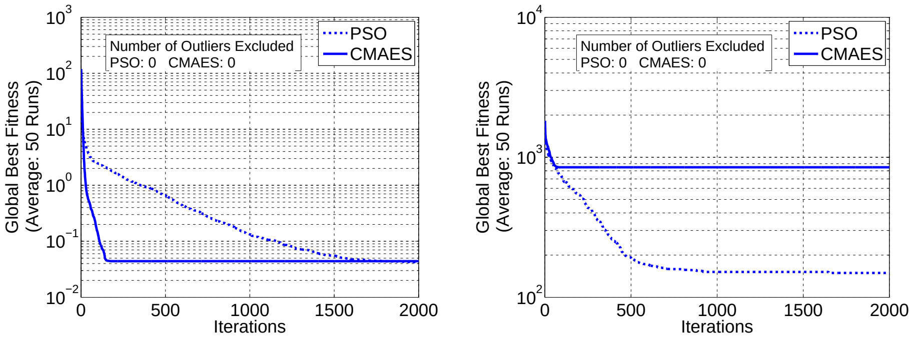
Similar features are observed in the lower dimensional problems when compared to those of higher dimensions. Both PSO and CMAES converged to a very small number for the sphere and cigar function. Again, PSO was slower in finding the global minimum and then fine tuning the parameters. Better final fitness values are seen in comparison to the same functions at higher dimensions. This is expected since the optimizers have an easier time searching a lower number of dimensions. In the Rosenbrock function, PSO does not find the global optimum, whereas CMAES gets close to the global optimum. There were a few times where CMAES would get stuck in a local optimum, and these were excluded from the average data. One interesting feature seen in the Rosenbrock optimization is the threshold that CMAES seems to run into



(a) Sphere Function **(b)** Rosenbrock Function

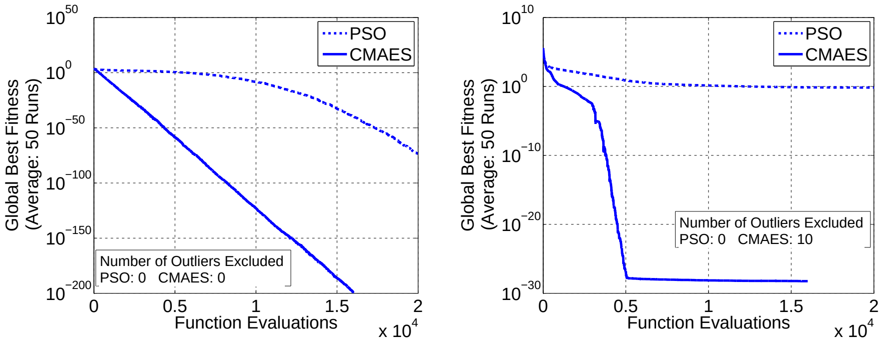


(c) Cigar Function **(d)** Rastrigin Function

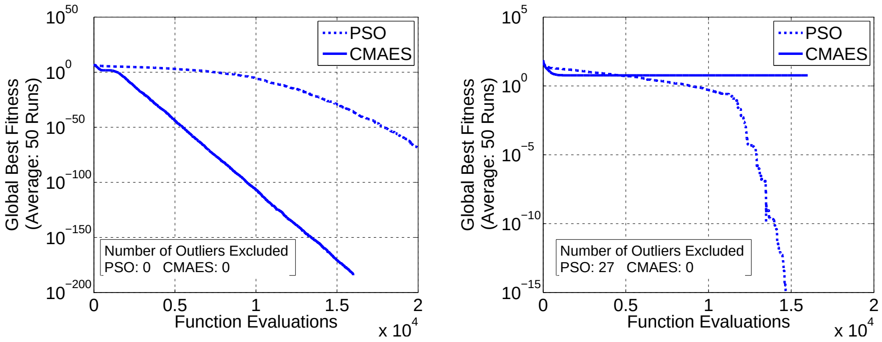


(e) Griewank Function **(f)** Schwefel Function

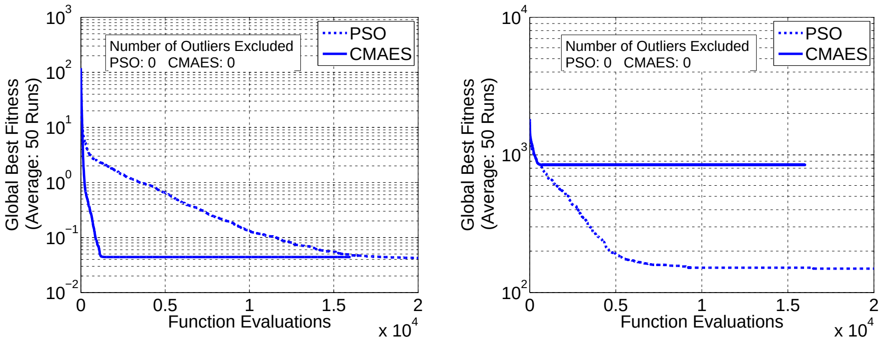
Figure 3.4: Optimization results when $N = 5$ versus *iterations* for the mathematical test functions



(a) Sphere Function **(b)** Rosenbrock Function



(c) Cigar Function **(d)** Rastrigin Function



(e) Griewank Function **(f)** Schwefel Function

Figure 3.5: Optimization results when $N = 5$ versus *function evaluations* for the mathematical test functions

once it hits a fitness value around 10^{-30}. This may be due to possible numerical errors that cause the optimizer to get stuck as opposed to continuing towards zero.

For the multimodal functions, the results were not as clear in comparison to the higher dimensions. These are still quite difficult topologies to navigate at lower dimensions. In the Rastrigin function, PSO seems to dominate over CMAES in performance, but it is also important to remember that the number of outliers excluded from the average is also quite large. Therefore the PSO algorithm has less chance of achieving the results that it accomplished in the other runs. CMAES did not perform as well as expected for the Rastrigin function, and all results were similar to those shown in Figure 3.5d. A similar performance is experienced for the Griewank and Schwefel function; CMAES seems to have some issues with getting stuck at local optima for the multimodal problems. For these last two problems, PSO is able to reduce the function value but also gets stuck in a local optima similar to CMAES. In order to overcome these problems one might use a larger population. In [92], an offspring population size of roughly $\lambda = 400$ was implemented to ensure a 95% probability that the CMAES would converge for the Griewank function of 5 dimensions. This is generally the same for PSO as well, where increasing the swarm size helps overcome local optima.

Overall, good performance has been demonstrated for both algorithms, although a few functions proved challenging enough to force them to a premature convergence. The goal was to simulate and compare their performance in a time and resource limited setting. In a real-world electromagnetics application, the fitness function evaluation can take as long 30 minutes, and the overall optimization time can accrue quite rapidly with those computational speeds. Therefore we used the lowest recommended population sizes and a low number of iterations in order to minimize the total number of function evaluations. In general, CMAES was observed to achieve a more rapid convergence in comparison to PSO, however it was also prone to fall prey to local optima. While PSO did converge at a slower rate, there were a few problems where it overcame the local optima and found the global optima or it found a local optima closer to the global one. It is interesting to see a similar performance between PSO and CMAES even though the PSO algorithm is simplistic in comparison to CMAES.

3.2 Nonuniform Antenna Array Optimization

Antenna array design has been a standard problem seen in the electromagnetics community, and it happens to be a perfect test problem for optimization algorithms because array problems are real-world applications and their performance can often be computed quickly. Their rapid calculation owes to the fact that an array of antennas can be viewed as a superposition of isotropic point sources, and the interference created between the radiated waves can manipulate the radiation pattern into a desired shape. With the assumption that an array of antennas acts as a superposition of point sources and the assumption that the antenna elements have exactly the same radiation, the array radiation pattern can be calculated by

$$\vec{E}_{tot}\left(\theta,\phi\right)=\vec{E}_{\ell}\left(\theta,\phi\right)\cdot AF\left(\theta,\phi\right) \tag{3.1}$$

where $\vec{E}_{tot}\left(\theta,\phi\right)$ is the total far-field electric field radiation in the (θ,ϕ) direction, $\vec{E}_{\ell}\left(\theta,\phi\right)$ is the far-field electric field radiation for a single element, and $AF\left(\theta,\phi\right)$ is the array factor. For an antenna array, the only factor that will change is the array factor, and this happens to be easily computable using mathematical formulas [15].

There are many types of problems for arrays which have primarily been solved by global optimization algorithms, and a common type is aperiodic array design. Aperiodic arrays have been previously employed to minimize the sidelobe level and even place nulls in certain specified locations as an alternative to controlling the amplitude and phase of each element [95]. The element location is much easier to control tolerances and to predict in comparison to controlling each element's phase and amplitude, which can be different than expected due to unforeseen coupling. In general there are two methods of achieving aperiodicity: nonuniform arrays and thinned arrays. In the first, the number of elements is kept constant and the inner elements are moved about in order to achieve the desired goal. The outer elements are kept stationary in order to maintain the same directivity regardless of the inner element position. This would be a real-valued optimization where the design variables are continuous real numbers. The second type is achieved by placing M elements uniformly throughout the space and turning each element

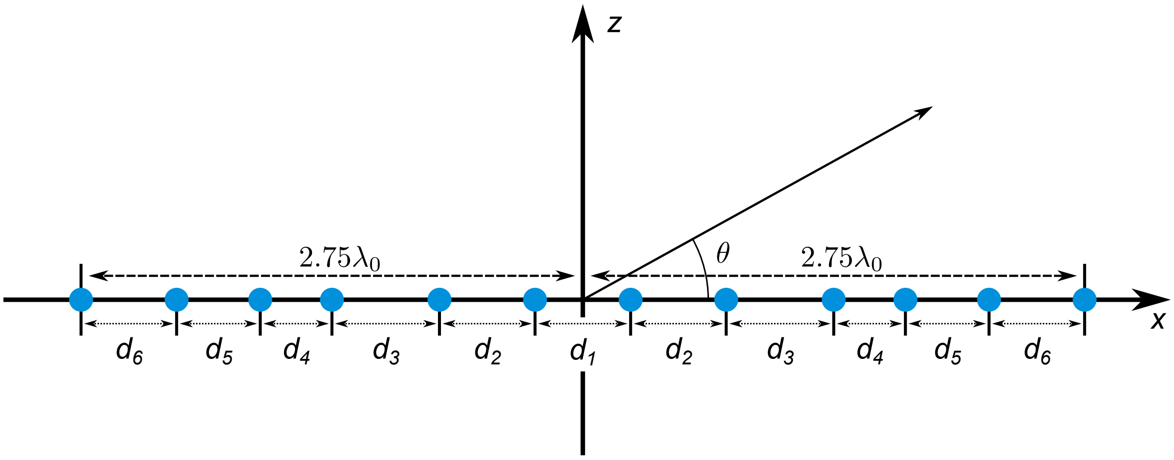


Figure 3.6: Geometry used for the symmetrical nonuniform antenna array to be optimized with PSO and CMAES. The outer elements are located at $x = \pm 2.75\lambda_0$ and the other element positions are to be optimized

ON or OFF. The ON state implies that the element is included in the array, and the OFF state implies the contrary. The thinned array approach is a discrete (or binary) optimization.

Since the focus remains on real-valued algorithms, the nonuniform array approach will be taken. Both PSO and CMAES will be applied to a symmetric 12 element array as seen in Figure 3.6. This array problem is similar to that in [75], except that a 12 element array is optimized over a larger overall distance. For this optimization, the goal is to reduce the sidelobe level (SLL) for the given SLL region, i.e. outside of the main lobe region. Since the array is symmetric, there are only five design variables to optimize, and the outer elements are positioned at $\pm 2.75\lambda_0$ so that the total array size will remain at $5.5\lambda_0$. In order to avoid a large number of constraints inequalities, the distance between each elements are optimized, which is an alternative to the approach used in [75]. The solution space is limited to

$$d_i \in (.25\lambda_0, 2.75\lambda_0), \quad i \in \{1, 2, 3, 4, 5\} \tag{3.2}$$

where d_i is the distance in wavelengths between adjacents nodes. Since patch antennas are typically around $0.5\lambda_0/\sqrt{\varepsilon_r}$, the lower limit of $0.25\lambda_0$ was employed to guaratee a sufficient distance between elements while allowing the search space to be extensive.

A constraint must be enforced in order to ensure that new elements are located outside of

Table 3.6: PSO Intrinsic Parameters for the Nonuniform Array Optimization

PSO Parameter	*Values Used*
c_1	2.0
c_2	2.0
Swarm Size	10
Δt	1.0
Max Iterations (i_{max})	500
Boundary Condition	Invisible
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

the maximum dimensions given for the array. Therefore, the only constraint needed in this formulation is given by

$$\frac{d_1}{2} + \sum_{i=2}^{5} d_i \leq 2.75\lambda_0 \tag{3.3}$$

which ensures that none of the elements fall outside of the designated space $|x| \leq 2.75\lambda_0$. This provides for a simpler and more compact formulation. These account for all the geometrical constraints in the problem, and the intrinsic parameters used in PSO and CMAES are shown in Tables 3.6 and 3.7, where the typical values for electromagnetics problems are used.

The fitness function must now be developed. The primary goal is to lower the SLL of the array by changing the interelement spacings. This provides for a relatively simple fitness function given by

$$f(\vec{x}) = \max\left(20\log|AF(\theta)|\right), \quad \forall\theta \text{ in sidelobe region} \tag{3.4}$$

where the array factor AF for a symmetric 12 element array is given by

$$\begin{aligned} AF(\theta) &= \sum_{i=1}^{6} \cos\left(2\pi x_i \cos(\theta)\right) \\ &= \sum_{i=1}^{5} \cos\left(2\pi x_i \cos(\theta)\right) + \cos\left(2\pi \cdot 2.75\cos(\theta)\right). \end{aligned} \tag{3.5}$$

Table 3.7: CMAES Intrinsic Parameters for the Nonuniform Array Optimization

CMAES Parameter	*Values Used*
λ	8
μ	4
σ^0	$\max(\vec{x}_{max} - \vec{x}_{min})/3$

The last equality in equation 3.5 was made due to the outer elements being set at a distance of $\pm 2.75\lambda_0$. With the fitness function defined in equation 3.4, the optimizer will be minimizing the maximum SLL found in the sidelobe region. The maximum SLL represents the worst case scenario, and this often is termed a *minimax* optimization which guarantees the SLL output throughout the entire sidelobe region. It is also important to note the definition of *sidelobe region.* In this problem, the sidelobe region is defined as the points not including the main lobe. This can be more formally defined as $\mathcal{R} = \{\theta : 0 \leq \theta \leq \theta_n \cup (180^\circ - \theta_n) \leq \theta \leq 180^\circ\}$, where the angle θ_n is defined as the first null found by decreasing from 90°, which is the location of the main beam. The first null seen below 90° and above 90° will be the same distance from 90° since the array is symmetric. These points are found by searching for the closest zero-derivative points on the plot.

In order to make a proper comparison, two penalty functions must be incorporated for the CMAES algorithm. This is to ensure that CMAES searches the same feasbile space $\mathcal{F}$ as the PSO algorithm. Otherwise, it may be possible that CMAES will test designs that are to be avoided. The penalty functions implemented into the CMAES algorithm are listed below.

$$p_b(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{S} \\ 10^5 + 1000\|\vec{x} - \frac{\vec{x}_{max}+\vec{x}_{min}}{2}\|^2 & \text{if } \vec{x} \notin \mathcal{S} \end{cases} \tag{3.6a}$$

$$p_c(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{F} \\ 10^5 + 100\left(\frac{d_1}{2} + \sum_{i=2}^{5} d_i - 2.75\right) & \text{if } \vec{x} \notin \mathcal{F} \end{cases} \tag{3.6b}$$

The boundary penalty function $p_b(\vec{x})$ first adds a high penalty 10^5 in order to encourage CMAES to search in other areas. It also adds the sphere function which is centered at the solution space center. Without this term, CMAES will keep expanding its variances until the optimization finishes. This has been confirmed experimentally, and therefore the last term has been added onto the penalty function as recommended by [83]. This helps guide the optimizer towards the search solution space $\mathcal{S}$. The constraint penalty function $p_c(\vec{x})$ adds a penalty 10^5. The other term in $p_c(\vec{x})$ is simply $100g_1(\vec{x})$, where $g_1(\vec{x})$ is the constraint function given in equation 3.3. Similarly to the parabolic term in equation 3.6a, this term provides some guidance towards the feasible space $\mathcal{F}$ as recommended by [85]. Without this term, the optimizer has difficulty finding the feasible solution space since the fitness function is a constant within the space $\vec{x} \in \mathcal{S} \cap \vec{x} \notin \mathcal{F}$. Just to reemphasize, the final fitness function for the CMAES optimization run was given by

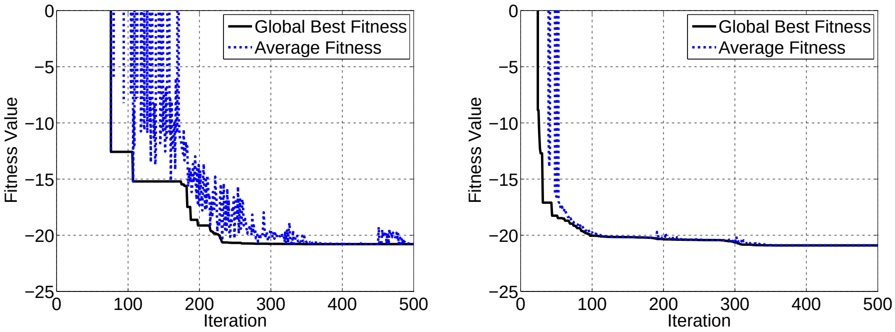
$$f_C(\vec{x}) = f(\vec{x}) + p_b(\vec{x}) + p_c(\vec{x}) \tag{3.7}$$

where $f_C(\vec{x})$ is the fitness function evaluated by the CMAES optimizer. This function is only used in the CMAES optimizer and not the PSO optimizer.

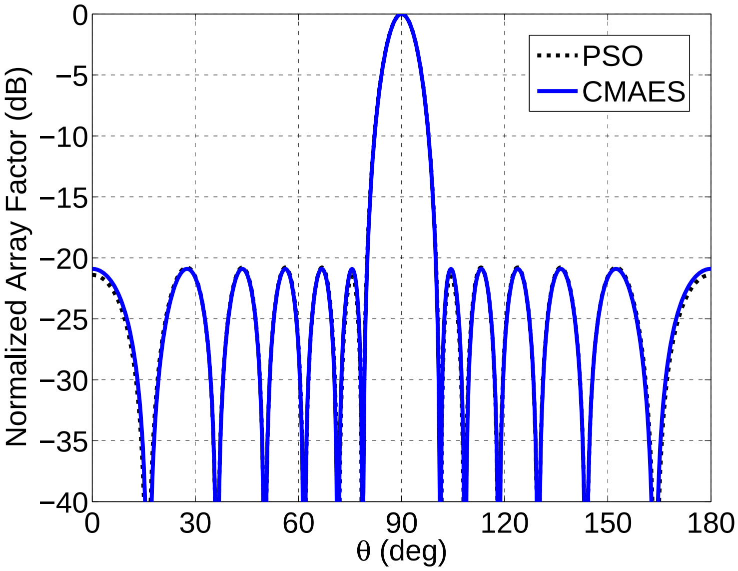
The two algorithms were applied to the optimization problem, and the convergence results are shown in Figure 3.7a and 3.7b. In 3.7c, the final array pattern is shown using the final CMAES design parameters given in Table 3.8. There is almost no noticeable difference between the two patterns, and this is expected since the final design values are so close. These final designs provided excellent results and also compared well with the results shown in [75]. The maximum SLL seen in both designs are -20.79 dB in the PSO design and -20.9 dB in the CMAES design. A slight decrease from the results in [75] are to be expected since more elements are being used. The final design locations are depicted in Figure 3.8, where a uniform array as a reference to compare to the optimized design. Only the CMAES design is shown in this figure because the designs would be right on top of each other. Since this roughly represents both final designs it has been termed the optimized array. From this figure and Table 3.8 it can be observed that the interelement spacing gradually increases for elements farther away from the array center. This is similar to the idea of weighting the center element amplitudes higher than the outer elements, which is often used in Chebyshev arrays in order to minimize the SLL to a certain level for periodically spaced elements. Indeed, the SLL is distributed uniformly throughout the region,

Table 3.8: Final Design Values provided by CMAES and PSO in the Nonuniform array optimization

	d_1	d_2	d_3	d_4	d_5
PSO	0.37477	0.40296	0.41518	0.46320	0.58083
CMAES	0.36921	0.41676	0.41091	0.46920	0.57328



(a) PSO Convergence **(b)** CMAES Convergence



(c) Comparison of the array patterns for both designs

Figure 3.7: Comparison of optimization results between PSO and CMAES for the nonuniform array with 12 elements

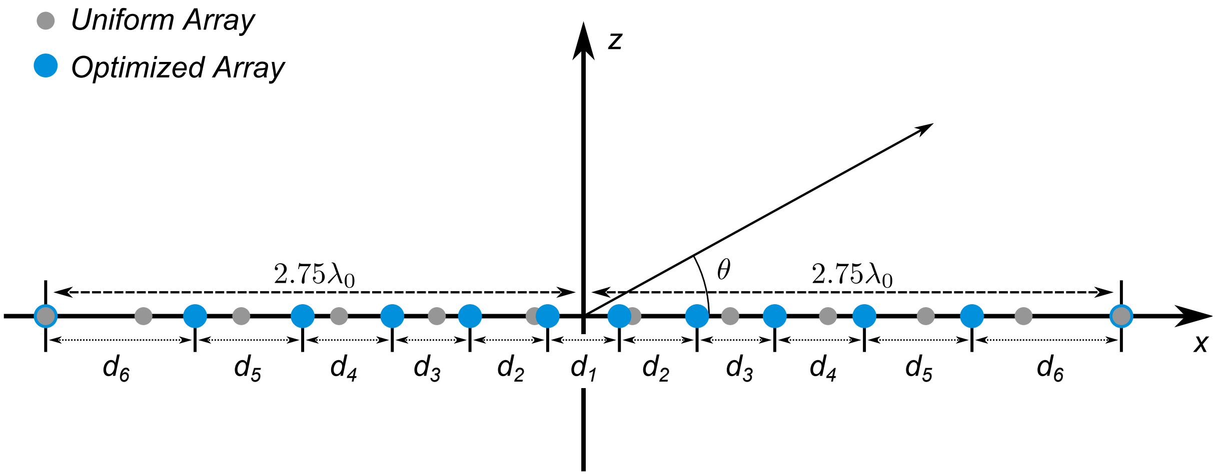


Figure 3.8: Final design geometry provided from CMAES. The resulting design from PSO is almost identical to the one depicted. A 12 element uniform array with $\lambda_0/2$ element spacing is also shown in gray as a reference.

which another characteristic similar to the Chebyshev arrays [15].

Overall, it appears that both CMAES and PSO performed well and produced nearly identical designs. It becomes a little more difficult to make any conclusive claims since each technique found the feasible space $\mathcal{F}$ at different iterations. CMAES may have found this region by the help of the penalty functions. However, CMAES still demonstrates faster convergence upon the same optimum when using the the first iteration which test solutions in $\mathcal{F}$. It rapidly minimized the fitness, whereas PSO took some time to fine-tune the search space.

CHAPTER 4

Optimization of Polarimetric Radar Systems

Weather radar is the secret weapon in the current meteorologist's arsenal and is used to make more confident predictions about the next weather events. There are many different types of weather radar systems, each having their own unique set of advantages and abilities. In particular, the dual polarization weather radar has stirred recent interest due to its added capabilities. By adding another linear polarization, more features of precipitation and the weather can be detected. With these additional benefits, meteorologists can make better predictions on the amounts of precipitation. This can also help provide hydrologists the information to make accurate estimates in stream flow forecasts and for river flooding in emergency weather situations as well as water management. Dual polarization radar has also been known to provide better discrimination between precipitation types, such as hail and rain [96, 97]. These new potential improvements have generated new attraction, and currently a new dual polarization weather radar system is being built to upgrade the current systems in the U.S. [98].

A basic weather radar radiates electromagnetic pulses and measures the backscattered fields from the precipitation as shown in Figure 4.1. Polarimetric radars are more general because they obtain backscattered information for two different polarizations. In one system, each wave is generated at separate times and measured at separate times, which is similar to the depiction in Figure 4.1. Other systems can operate by transmitting both polarizations simultaneously and receiving both backscattered signals simultaneously [96, 99]. These have often been termed alternate and simultaneous transmit modes [99], and each have their distinct advantages and disadvantages.

The backscattered fields from the precipitation are often characterized by the backscattering

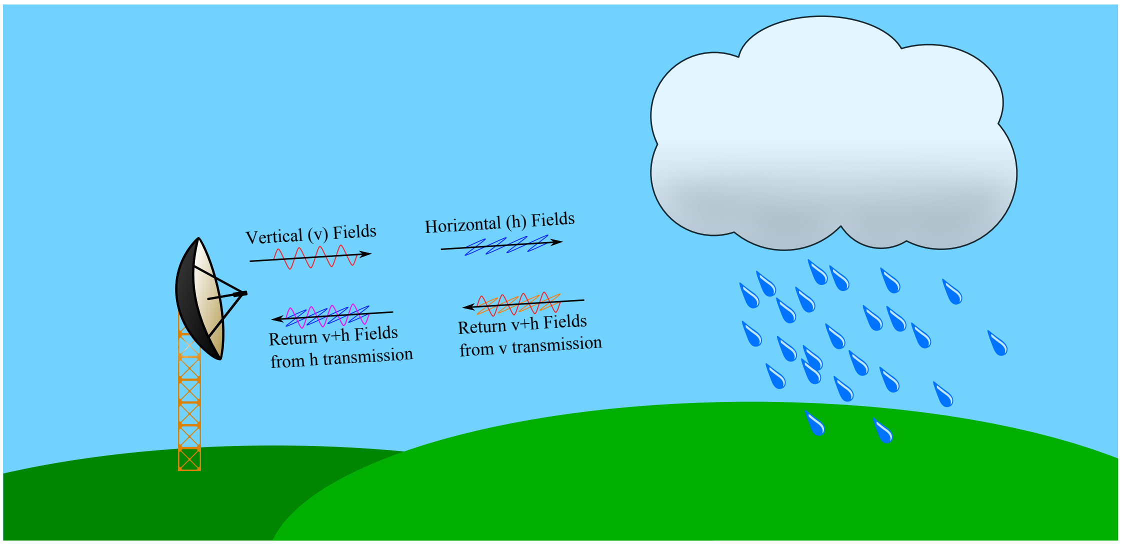


Figure 4.1: Illustration of the basic polarimetric weather radar used to measure properties of the precipitation

matrix **S**, where

$$\mathbf{S} = \begin{pmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{pmatrix} \tag{4.1}$$

and the first and second subscript denotes the transmitted and received polarization, respectively. Each of the matrix elements are complex numbers, and the phase and magnitude both have important implications on the measurements. The magnitude is often used as a measure of the radar cross sectional area and is highly correlated to the size of rain droplets. The phase can provide information about distance traveled as well as some characteristics on the rain rate. Some of the features that are detectable from dual polarization weather radars using this scattering matrix are given in the following list.

Vertical Reflectivity

This is the vertically polarized backscattered power observed when the vertical polarization is transmitted. It is often denoted by $Z_v = \langle |S_{vv}|^2 \rangle$.

Horizontal Reflectivity

This is the horizontally polarized backscattered power observed when the horizontal polarization is transmitted. It is often denoted by $Z_h = \langle |S_{hh}|^2 \rangle$.

Differential Reflectivity

This is the ratio of the horizontal backscattered power to the vertical backscattered power.

This can be used to describe the elongation of the raindrops in the horizontal direction, which directly relates to the raindrops' size [100, 101]. It is often denoted by $Z_{dr} = Z_h/Z_v$.

Copolar Correlation Coefficient

This is the measure of the correlation between the horizontal and vertical backscattered fields. This parameter can be used to identify hail and mixed precipitation scenarios [102]. It is typically denoted by $\rho_{hv} = |\langle S_{hh}^* S_{vv} \rangle| / \sqrt{Z_h Z_v}$.

Cross-Polar Reflectivity

This is the horizontally polarized backscattered power observed when a vertically polarized wave is transmitted. Likewise, it is also the vertically polarized backscattered power when a horizontally polarized wave is transmitted. This is often denoted by $Z_x = \langle |S_{vh}|^2 \rangle$.

Each of these parameters provide some unique information about the weather, and there has been a widespread effort to characterize weather conditions based on these parameters. Out of this research came several models (or functions) based on previous evidence that can predict weather conditions such as rain rate among other things. While there are a few other measure parameters used to describe the current weather status, these are the primary ones that are of importance to the antenna designers.

In this chapter, the focus is to investigate the best configuration of a possible array element which would be implemented in a cylindrical array. This array configuration has been proposed as a possible design solution for a new weather radar system to be implemented in the near future [103]. There are several parameters of interest to the weather radar designers, and are often termed the bias weighting factors (BWF's). These are to be minimized in order to have better weather prediction [104]. It can be shown that these parameters relate highly to the cross polarization levels in the antenna, and it is not yet known whether decreasing the cross polarization levels for specific cuts in the radiation pattern gives rise to the minimum of the BWF's. As will be shown in the next section, the BWF's depend on both the amplitude and phase of the far-field radiation, and therefore these parameters become unclear as to their best value. As a first investigation, this will test whether an optimization of the cross polarization levels is equivalent to the minimization of the BWF's. Another aspect of this chapter is to test the effectiveness of CMAES in real-world antenna applications, and both CMAES and PSO are utilized as the optimization algorithm for these particular optimizations.

4.1 Antennas for Polarimetric Radar Systems

When implementing these radar systems in practice, errors can occur in many different areas. The primary assumption made when taking these radar measurements is that no coupling exists between each polarization within the antenna and RF system. If a vertically polarized wave was incident upon the antenna, then the expectation is that only the corresponding port would receive the signal. However, cross-polarized radiation and coupling between the two ports are challenges that must be overcome in order to ensure good polarization isolation. In a typical compact antenna system with dual polarization, two lines (which could be either microstrip lines or coaxial feeds) connect to the two different polarization ports. Obviously, if the two lines have strong coupling then some of the signal for a particular polarization will appear at the other port, thus creating a false presence of the other polarization. This type of coupling can be difficult to overcome in antenna designs where the lines feeding the antenna for two different polarizations are in close proximity.

While transmission line coupling is one challenge that must be overcome in the system design, another important component is the radiation pattern, which can create coupling between both polarizations even if the transmission lines are completely isolated. The terms copolar and crosspolar fields refer to the desired and undesired polarized fields, respectively, and there are a few different ways to define the polarization vector over the spherical space (ϕ, θ) [105]. There are several different definitions of cross polarization in the literature. For example, one choice of reference polarization is given by Ludwig's 2nd cross polarization definition [106], and the copolar and crosspolar fields are shown in Figure 4.2. If the main lobe of the radiation pattern faces the X direction and the antenna has Y directed currents, the copolar reference unit vector would be $\hat{\phi}$, whereas the cross polarization reference unit vector would be $\hat{\theta}$ for this particular coordinate system. These vectors provide a reference direction to define the copol and crosspol fields for a given direction (ϕ, θ).

The presence of cross polarized radiation can cause errors in the different weather measurement parameters defined above, and one critical parameter that can be affected is the differential reflectivity, Z_{dr}. Since the antenna patterns are reciprocal, i.e. the receive and transmit patterns

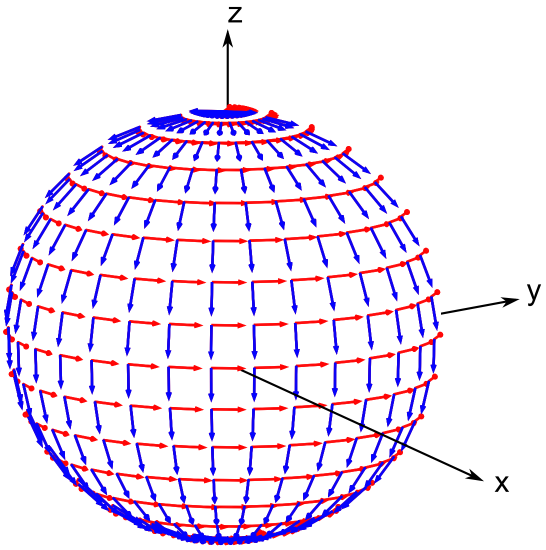


Figure 4.2: Ludwig's 2nd definition of copolarized (red vectors) and cross polarized (blue vectors) radiation from an antenna having a main lobe in the X direction with Y-directed currents

are the same, any crosspolar radiation observed in the radiation pattern implies that leakage will come from the undesired polarization. Investigations have been made in order to quantify the Z_{dr} errors associated with the isolation between the associated ports as well as the radiation pattern. Any sort of bias present in Z_{dr} can cause errors in the rain rate calculation. Therefore, the Z_{dr} bias, often denoted by δZ_{dr}, is desired to be minimized in order to alleviate any possible errors [104]. In order to minimize this bias the following bias weighting factors (BWF)[1] must be minimized [106].

$$W_{hv} = \frac{\int_0^{\pi} \int_{-\pi}^{\pi} \left(F_{hh}(\theta,\phi)\right)^2 F_{hh}^{*}(\theta,\phi) F_{hv}^{*}(\theta,\phi) \sin\theta d\phi d\theta}{\int_0^{\pi} \int_{-\pi}^{\pi} |F_{hh}(\theta,\phi)|^4 \sin\theta d\phi d\theta} \tag{4.2a}$$

$$W_4 = \frac{\int_0^{\pi} \int_{-\pi}^{\pi} |F_{hh}(\theta,\phi)|^2 |F_{hv}(\theta,\phi)|^2 \sin\theta d\phi d\theta}{\int_0^{\pi} \int_{-\pi}^{\pi} |F_{hh}(\theta,\phi)|^4 \sin\theta d\phi d\theta} \tag{4.2b}$$

The bias weighting factor W_{hv} is related to δZ_{dr} for an antenna with a cross polarization lobe

---

[1] The original formulation for W_{hv} presented in [104] has a slightly different form than presented in this work. However the authors had assumed that the copolar radiation would radiate with uniform phase at all the points on the spherical space (ϕ, θ). In practice, this assumption does not hold true and therefore this equation was altered per communication with Dr. Guifu Zhang.

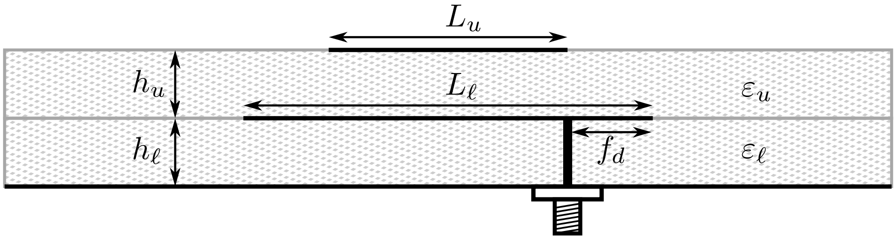


Figure 4.3: Stacked patch antenna configuration for possible use in weather radar systems

that is coaxial to the copolar fields. This type of pattern is often termed a single lobe crosspolar pattern. The bias weighting factor W_4 is related to the bias in differential reflectivity due to an antenna with four lobes in its crosspolar radiation pattern, and it is assumed that these lobes are asymmetric in phase [104]. The terms $F_{hh}(\theta, \phi)$ and $F_{hv}(\theta, \phi)$ correspond to the copolar and crosspolar radiation patterns, respectively. These formulas are also independent of the radiation pattern normalization, and therefore a normalized pattern can be applied with the same results as any other scaled pattern. The copolar and crosspolar radiation terms are denoted by $E_{co}(\theta, \phi)$ and $E_x(\theta, \phi)$, respectively.

Clearly, the equations shown in equation 4.2 are dependent upon the cross polarized fields, and BWF reduction can be achieved by reducing the total cross polarized fields seen throughout the entire spherical space. Minimizing the overall cross polarized fields throughout the entire radiation sphere is quite difficult and non-intuitive. In general, antenna engineers have much more experience working with the radiation patterns and minimizing the cross polarization in the principle cuts in the radiation pattern, i.e. $\phi = 0, 45°, 90°, 135°$. However, it is not clear whether an optimization of the fields in these cuts would minimize the BWF's in general, especially since the phase of the far-field pattern is also included in the calculation. Therefore, two separate optimizations can be compared to determine the differences in optimizing the BWF's versus optimizing the cross polarized radiation in the principle cuts as a first investigation. If the two optimizations have comparable results, then this validates the principle cut method of optimizing the fields in the principle planes to achieve a minimum for W_4 and W_{hv}.

The stacked patch antenna shown in Figure 4.3 is chosen as an example antenna due to its possibility for low cross polarized radiation and wide bandwidth [107]. These particular antennas are quite suitable for use as elements in the new cylindrical phased arrays proposed for use in

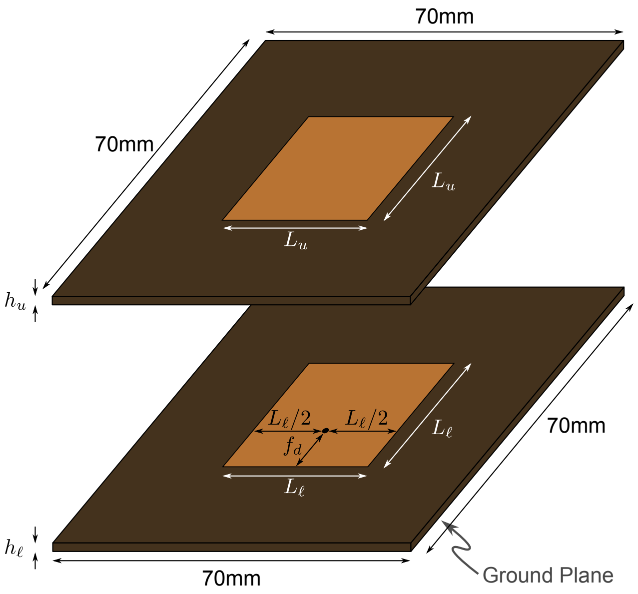


Figure 4.4: Exploded view of the stacked patch used in the optimization

weather radars [103]. Note that this particular example does not achieve dual polarization as needed in polarimetric radar, but the goal of this optimization is to minimize the crosspolar fields, which does not require that the antenna under investigation be dual polarized. An exploded view of the stacked patch configuration showing the layers is given in Figure 4.4. The upper and lower substrate are chosen to be Rogers Duroid 5880 since it is commercially inexpensive, and this substrate has a permittivity of $\varepsilon_r = 2.2$. The highest available height for this substrate was 125 mil (or 3.175mm), and this height was chosen for maximum bandwidth. In order to make the cross polarization performance equivalent for both polarizations, the patches are chosen to be made square, hence both sides being equal to L_u and L_ℓ. As shown in both Figures 4.3 and 4.4, a coaxial probe is connected to the middle layer a distance f_d away from the edge of the lower patch. This leaves only the lower patch length L_ℓ, the upper patch length L_u, and the feed distance from the edge f_d for the optimization.

The organization of the optimization runs is given as follows. First, both PSO and CMAES will be applied to the stacked patch separately in order to minimize the cross polarized fields in the main cuts as well as ensuring good impedance matching. After the optimization, the radiation patterns and finally the BWF's will be computed for the two optimized designs. Next, PSO and CMAES will be applied to the same stacked patch antenna using a different fitness function which incorporates the BWF's, W_{hv} and W_4. Comparisons will be made about both algorithms' convergence, and the final design performance from both algorithms will be compared. Once all these runs have been demonstrated, a final comparison will shed light onto possible differences between these two different approaches to optimize the BWF's. The last section of this chapter will also shed light on the meaning of these integrals and suggest recommendations to antenna engineers for future designs in weather radar applications.

4.2 Optimization of Cross Polarized Fields

4.2.1 Optimization Problem Development

The first optimization applies both PSO and CMAES to minimize the cross polarized fields in the principle pattern cuts, as well as minimize the impedance matching of the antenna over the required band, 2.7-3.0 GHz. The coordinate system must be properly defined in order to have any discussion on the crosspolar fields, which is shown in Figure 4.5. In order to distinguish this optimization from the other, this optimization will be denoted as the cross polarization fields optimization (XPFO). This optimization will utilize the more common 3rd definition of cross polarization by Ludwig [105], whose reference vectors are given as

$$\hat{i}_{co} = \sin(\phi)\hat{\theta} + \cos(\phi)\hat{\phi} \tag{4.3a}$$

$$\hat{i}_{x} = \cos(\phi)\hat{\theta} - \sin(\phi)\hat{\phi} \tag{4.3b}$$

where $\hat{i}_{co}$ and $\hat{i}_x$ are the unit vectors pointing in the copolar and the crosspolar field directions, respectively. Note that these reference vectors are given for an antenna with y-directed currents, as shown in Figure 4.5. The principle cuts in the radiation pattern often include the $\phi = 0°, 45°, 90°$, and 135° cuts. The $\phi = 0°$ and the $\phi = 90°$ cuts are considered the H-plane and the E-plane, respectively, for this geometry and coordinate system. The cross polarization level

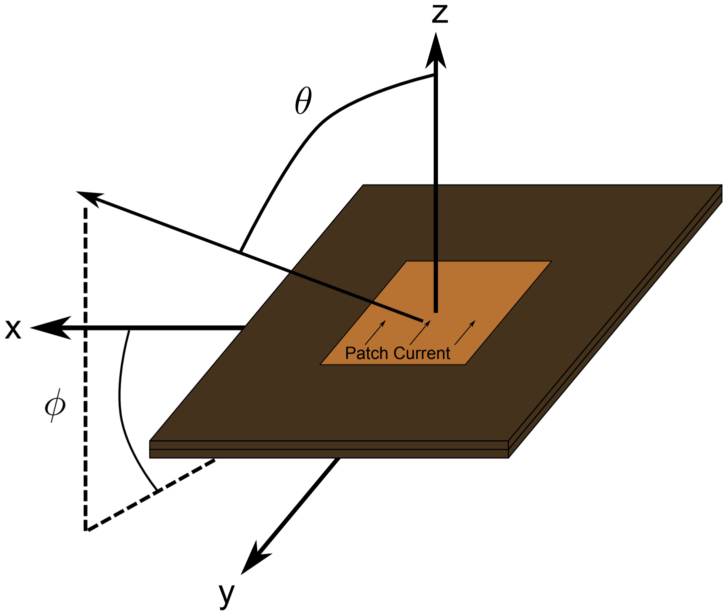


Figure 4.5: Coordinate system used for measuring the crosspolar fields in the cross polarized field optimization (XPFO)

(xpol level) for a given cut ϕ_0 is often given as

$$X_p(\phi_0) = \max_{\theta\in[-\pi,\pi]} \left(E_x^{dB}(\theta,\phi_0)\right) - \max_{\theta,\phi} \left(E_{co}^{dB}(\theta,\phi)\right) \tag{4.4}$$

where $X_p(\phi_0)$ is the xpol level for the ϕ_0 cut, E_x^{dB} is the cross polarized field in dB, and E_{co}^{dB} is the copolar fields in dB. The angle θ is applied over the range $[-\pi,\pi]$, which is typically not the range set for θ in the spherical coordinate system. This is purely a shorthand notation for denoting a full circular cut in the spherical space with the more common notation (ϕ,θ) and $(\phi+\pi,\theta), \forall\theta \in [0,\pi]$. These patterns are calculated for only one specific frequency, and in order to represent the fields over the desired bandwidth (2.7 - 3.0 GHz), these fields are computed for only 2.85 GHz.

Both the xpol levels in the principle cuts and the impedance matching are the primary concerns for the final antenna design. In order to incorporate these features into the optimization, they must be included in the fitness function, which can be written as

$$f(\vec{x}) = \beta \max_{f\in F} \left(S_{11}^{dB}\right) + \alpha_1 X_p(0^\circ) + 2\alpha_2 X_p(45^\circ) + \alpha_3 X_p(90^\circ) \tag{4.5}$$

Table 4.1: Average Values for Design Objectives

Parameter	Average Value
$\max(S_{11})$	-10 dB
$X_p(0°)$	-15 dB
$X_p(45°)$	-12.5 dB
$X_p(90°)$	-25 dB

where S_{11}^{dB} represents the reflection coefficient in dB, and $f \in F$ represents a frequency f within the set $F = \{2.7 \text{ GHz - } 3.0 \text{ GHz}\}$. The cut $\phi = 135°$ was not included in the fitness function, because $X_p(45°) = X_p(135°)$ due to the symmetry of the antenna. Therefore the weight has been doubled in equation 4.7. The next step is to find the best weighting coefficients α_i. To accomplish this, one can use either *a priori* knowledge of the average values observed or assume no prior knowledge and weight the coefficients by their values desired (e.g. $S_{11} = -10$ dB). Since stacked patch antennas are well documented in literature, some *a priori* knowledge was used to determine these coefficients. Other researchers within the UCLA Antenna Research, Analysis, and Measurement (ARAM) lab had also conducted prior investigations with stacked patch antennas and provided recommendations as to their average value observed [108], which are shown in Table 4.1. Since the crosspolar terms are somewhat dependent, the coefficients are set to $\alpha_1 = \alpha_2 = \alpha_3 = 1$. The concern when weighting the objectives are that one extremely different parameter may have a very low/high value in comparison to the others. This also comes from the argument that there are only two objectives: lower xpol level in principle cuts and lower the impedance matching. The summation of the xpol levels represents a scaled averaging of the four cuts. Therefore, one can equate the averages such that these two objectives have equal weight in the fitness function. This can be done by writing

$$\beta\langle\max_{f\in F}\left(S_{11}^{dB}\right)\rangle = \alpha_1\langle X_p(0°)\rangle + 2\alpha_2\langle X_p(45°)\rangle + \alpha_3\langle X_p(90°)\rangle \tag{4.6}$$

where the $\langle\cdot\rangle$ implies the average value of the parameter within. With the values given in Table

4.1 and the previously assigned weights α_i, one can find the following relation.

$$-10\beta = -25 - 2(12.5) - 15 \rightarrow \beta \approx 6 \tag{4.7}$$

Therefore this weighted averaging provides the final weights needed in the fitness function. Now the fitness function can be written in its entirety.

$$f(\vec{x}) = 6 \max_{f \in F} \left(S_{11}^{dB} \right) + X_p(0°) + 2X_p(45°) + X_p(90°) \tag{4.8}$$

This fitness function is computed using full wave electromagnetic simulation tools, and HFSS was chosen in particular to simulate this antenna. This allows a full characterization of the antenna performance, whose results should be comparable to the results observed in a practical implementation.

With the fitness function now defined, the only components left to define are the solution space and any constraints required. Due to the design simplicity, the solution space is relatively easy to define, and the upper and lower bounds are

$$L_u = [10\text{mm}, 60\text{mm}] \tag{4.9a}$$

$$L_\ell = [10\text{mm}, 60\text{mm}] \tag{4.9b}$$

$$f_d = [1\text{mm}, 30\text{mm}] \tag{4.9c}$$

where the upper limits for each patch length were chosen to be slightly smaller than the size of the Duroid substrate. The lower limits were chosen as a compromise between a limited and wide solution space. The lower limits for f_d was set to 1mm in order to avoid a feed that crossed the patch edge due to the finite radius of the feed. Symmetry of the patch was exploited in order to limit the solution space in terms of the upper limit of f_d, and therefore the limit was set to half of the largest length possible of the patch. This does not always guarantee that the feed will be

Table 4.2: PSO Intrinsic Parameters for XPFO

PSO Parameter	*Values Used*
c_1	2.0
c_2	2.0
Swarm Size	6
Δt	1.0
Max Iterations (i_{max})	500
Boundary Condition	Invisible
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

Table 4.3: CMAES Intrinsic Parameters for XPFO

CMAES Parameter	*Values Used*
λ	7
μ	3
σ^0	$\max(\vec{x}_{max} - \vec{x}_{min})/3$

connected to the lower patch, and therefore an additional constraint

$$f_d \leq \frac{L_\ell}{2} \tag{4.10}$$

must be enforced in order to keep the feed within the the patch as well as avoid similar solutions coming from symmetry. This concludes the optimization formulation and all necessary aspects have been defined.

4.2.2 PSO and CMAES Results and Comparison

Both PSO and CMAES are now applied to the optimization problem described in the previous section. The original RPSO algorithm was used with the intrinsic algorithm parameters shown in Table 4.2. All the recommended values given in Section 2.1.1 were used for this run. The CMAES algorithm described in Section 2.2 was also applied to this optimization problem for the sake of comparison in a real-world application utilizing a full-wave electromagnetic solver. This algorithm also utilized the recommended parameters given in Section 2.2 and are shown in

Table 4.3. In order to apply the CMAES algorithm to the stacked patch, penalty functions also had to be provided. The functions

$$p_c(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{F} \\ 10^5 & \text{if } \vec{x} \notin \mathcal{F} \end{cases} \tag{4.11a}$$

$$p_b(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{S} \\ 10^5 + 100\|\vec{x} - \frac{\vec{x}_{max}+\vec{x}_{min}}{2}\| & \text{if } \vec{x} \notin \mathcal{S} \end{cases} \tag{4.11b}$$

were applied in order to force the optimizer towards the feasible region $\mathcal{F}$. Note that this constraint penalty function does not incorporate any of the constraints as functions. This is due to the fact that this is a rather simple constrained optimization problem with a minimal number of constraints. Therefore a complicated constraint penalty function was unnecessary. The CMAES algorithm had also been tested without the penalty functions, and the algorithm failed to find $\mathcal{F}$ for every run tested. This goes to show that the penalty functions are a requirement in the CMAES algorithm. The final fitness functions for the CMAES and PSO algorithm are given by

$$f_C(\vec{x}) = f(\vec{x}) + p_c(\vec{x}) + p_b(\vec{x}) \tag{4.12a}$$

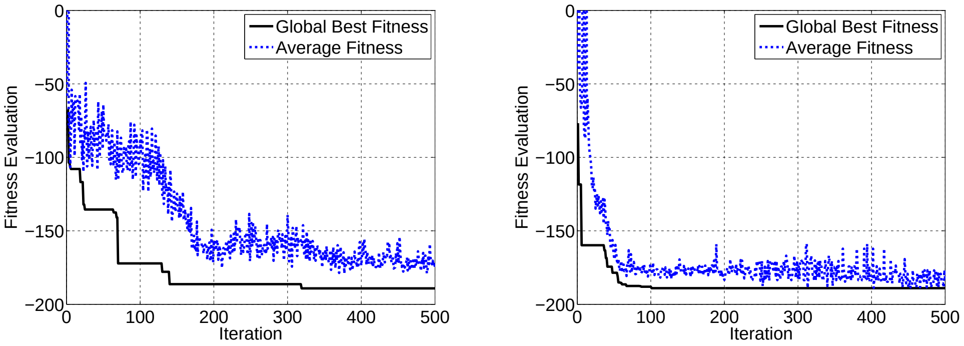
$$f_P(\vec{x}) = f(\vec{x}) + p_c(\vec{x}) \tag{4.12b}$$

where $f_C(\vec{x})$ is the fitness function optimized by CMAES and the function $f_P(\vec{x})$ is the fitness function optimized by the PSO algorithm. Note that the constraint penalty equation had to be enforced in order to ensure that the PSO algorithm would avoid the infeasible region $\mathcal{I}$. Now the PSO algorithm only requires the use of the constraint penalty function because it accounts for the solution space boundaries by its boundary conditions. The CMAES algorithm however must incorporate the boundary penalty function because it is an unbounded optimization algorithm, as discussed in Section 2.3. The functions $f_C(\vec{x})$ and $f_P(\vec{x})$ represent the implemented fitness function optimized by the PSO and CMAES algorithms, respectively.

Both algorithms demonstrated good results in their separate optimizations. The convergence plots for both the PSO and the CMAES runs are shown in Figure 4.6. Both algorithms showed

Table 4.4: Final Design Values provided by CMAES and PSO in the XPFO

	L_ℓ (mm)	L_u (mm)	f_d (mm)
PSO	31.812	30.121	6.085
CMAES	31.877	29.965	5.977



(a) PSO Convergence **(b)** CMAES Convergence

Figure 4.6: Convergence for the XPFO problem

rapid convergence upon the global optima, although the CMAES algorithm found its final design slightly faster than PSO around the 100th iteration. PSO found a good design (not the final) by the 140th iteration which demonstrated comparable performance to the final design. The final design values given by both optimization runs are given in Table 4.4. Their affinity is fairly remarkable, and the fact that two completely different algorithms found a similar solution reinforces the global optimality of the design.

The antenna performance should also be shown in order to make the final comparison between the PSO and the CMAES optimization runs. The final fitness values were approximately -189.2 and -189.0 for the PSO and the CMAES algorithms, respectively. Therefore, the fitness for only the PSO design will be demonstrated due to their similarities. The impedance matching is shown for the final PSO design in Figure 4.7. The shaded region highlights the frequency band of interest, namely 2.7-3.0 GHz. The S_{11} response of the final design demonstrates satisfactory performance ($S_{11} \leq -10$ dB) in the region of interest. In a stacked patch antenna, the wideband performance is achieved by the addition of another coupled patch above the primary driven

Table 4.5: Pattern Characteristics of the XPFO Final Design

Parameter	*Observed value*
Directivity	7.97 dB
Max Xpol	-16.8 dB
W_{hv}	-33.39 dB
W_4	-19.75 dB

patch, and the dual resonance feature in the S_{11} response is expected. By introducing this coupled patch, one can increase the overall impedance bandwidth.

The other results of interest are the radiation patterns, which are plotted in Figure 4.8. The patterns observed are fairly typical of rectangular patch antennas, with most of the energy propagating towards the +z direction. Symmetrical patterns are also observed for $\phi = 0$, and the $\phi = 45°$ and the $\phi = 135°$ patterns are the same due to the patch symmetry. Negligible xpol levels are observed for the E-plane ($\phi = 90°$) which is common because of typical geometrical symmetry seen in stacked patch designs. The maximum xpol level is observed in the $\phi = 45°, 135°$ planes, which was observed to be roughly -16.8 dB lower than the maximum directivity in the boresight direction ($\theta = 0, \phi = 0$). Other various radiation characteristics of the final design are listed in Table 4.5. The BWF's were also calculated for the final antenna design given, and their values are shown in Table 4.5. Note that the BWF's were computed using Ludwig's second definition of cross polarized radiation, as shown in Figure 4.2. The coordinate system shown in Figure 4.9 was used to compute them rather than the coordinate system shown in Figure 4.5. The values were computed using equations 4.2a and 4.2b and were converted into dB using $10 \log |W_{hv}|$ and $10 \log |W_4|$.

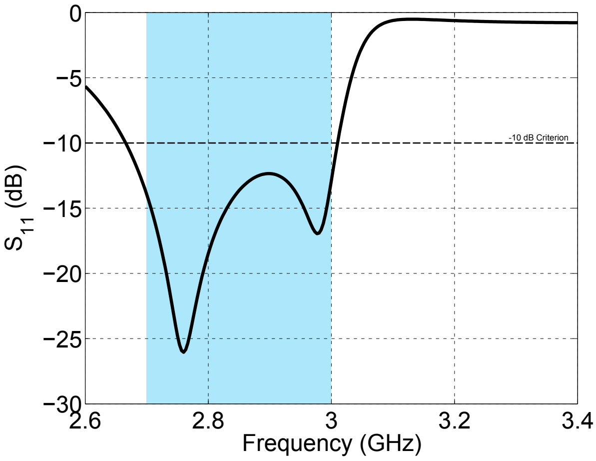
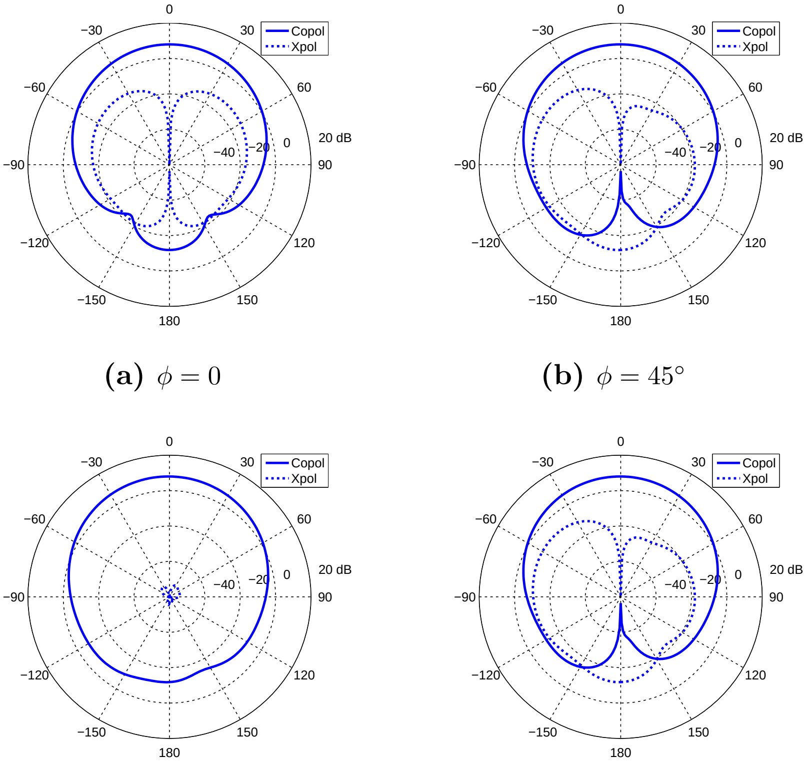


Figure 4.7: S_{11} response for the final PSO design (similar characteristics are observed with the CMAES design)



(a) $\phi = 0$

(b) $\phi = 45°$

(c) $\phi = 90°$

(d) $\phi = 135°$

Figure 4.8: Radiation patterns of the final optimized stacked patch antenna design in the principle planes

4.3 Direct Optimization of Bias Weighting Factors

4.3.1 Optimization Problem Development

While the optimization of the cross polarization can lead to good designs providing a lower W_{hv} and W_4 in general, it is unclear whether these final designs provide the lowest values possible. Therefore, a comparison is made by optimizing the same stacked patch antenna as the previous optimizations with a different fitness function that focuses on the BWF's. This optimization is denoted as the Bias Weighting Factor Optimization (BWFO). The optimization has many identical characteristics as the previous optimization, and therefore it would be unnecessary to re-explain each similar characteristic. The boundary conditions and the constraints are kept the same and are listed below for convenience.

$$L_u = [10\text{mm}, 60\text{mm}] \tag{4.13a}$$

$$L_\ell = [10\text{mm}, 60\text{mm}] \tag{4.13b}$$

$$f_d = [1\text{mm}, 30\text{mm}] \tag{4.13c}$$

$$f_d \leq \frac{L_\ell}{2} \tag{4.14}$$

The primary difference is in the fitness function, which must incorporate the BWF's in some manner in order for the optimizer to distinguish designs with better BWF performance. The impedance matching must also be included in order to ensure good port performance. This can be accomplished by incorporating both of these objectives in a weighted fitness function having multiple objectives. Using the weighted objective approach, one can write the fitness function as

$$f(\vec{x}) = \beta \max_{f \in F} \left(S_{11}^{dB} \right) + \alpha_1 10 \log |W_{hv}| + \alpha_2 10 \log |W_4| \tag{4.15}$$

where S_{11}^{dB} is in dB form while the BWF's, W_{hv} and W_4, are on a magnitude scale. Note that

$10\log(\cdot)$ is applied to the BWF's since they represent a power-like value [106]. It should also be noted that a stacked patch antenna is often classified as an antenna having a four lobe crosspolar pattern, which would imply that only W_4 is applicable to this antenna since theoretically $W_{hv} \to 0$ for antennas with these asymmetries. However, this is not the case with practical realizations of a stacked patch antenna, and even simulations will reflect an asymmetric crosspolar pattern which will have some finite value for W_{hv}. Both parameters are equally important as an objective of the optimization. Therefore, both parameters should be used to quantify the cross polarization performance of the antenna in relation to the differential reflectivity bias δZ_{dr}, and both are incorporated into the fitness function as shown in equation 4.15.

The next item to be addressed is the proper weighting of each objective in order that the optimizer minimize each one equally. A similar approach to the average weight equalization shown in the previous section is used to weight each of these parameters. The average values expected of each of these parameters for typical designs are shown in Table 4.6 [108]. These average values provide some rough estimate on the numerical values to be expected with each parameter. These values may be somewhat different than the actual observation from the optimization, but this provides a systematic weighting procedure based on a small amount of prior experience. With these average values observed, the weights can be found by

Table 4.6: Average Values for Design Objectives in the BWFO runs

Parameter	Average Value
$\max(S_{11})$	-10 dB
$10\log\lvert W_{hv}\rvert$	-30 dB
$10\log\lvert W_4\rvert$	-20 dB

$$\beta\langle \max_{f\in F}\left(S_{11}^{dB}\right)\rangle = \alpha_1\langle 10\log|W_{hv}|\rangle = \alpha_2\langle 10\log|W_4|\rangle \tag{4.16}$$

where the average values $\langle\cdot\rangle$ are equated for each parameter. There are an infinite number of ways to solve these equations, and the value β is arbitrarily set to $\beta = 1$. With this, the other two parameters are found to be $\alpha_1 = 1/3$ and $\alpha_2 = 1/2$. This now allows each parameter to be

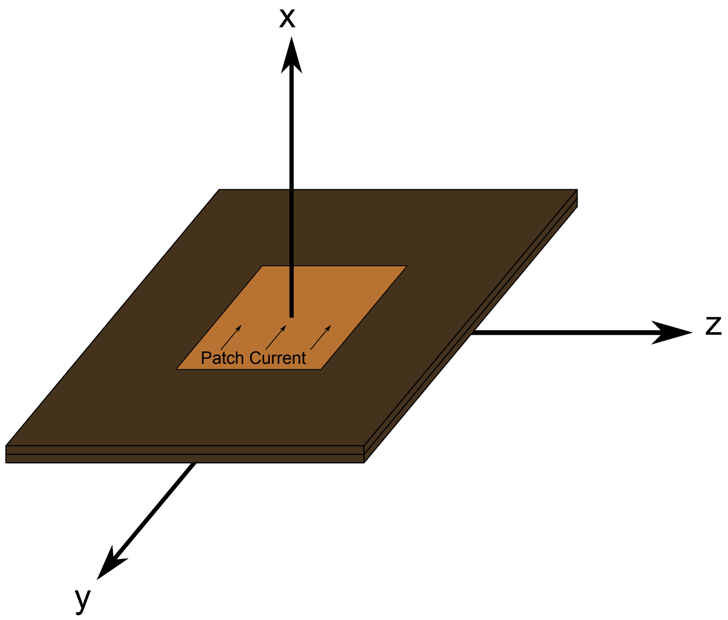


Figure 4.9: Coordinate system used for determining the crosspolar fields in the bias weighting factor optimization (BWFO)

equally weighted in the fitness function, which can be fully written as

$$f(\vec{x}) = \max_{f \in F}\left(S_{11}^{dB}\right) + \frac{1}{3}10\log|W_{hv}| + \frac{1}{2}10\log|W_4| \tag{4.17}$$

Every design to be tested is simulated within a full electromagnetic solver, HFSS, and the radiated field data is extracted in order to finally evaluate the integrals seen in equations 4.2a and 4.2b. For these parameters, Ludwig's second definition of cross polarization was used to define the copolar and crosspolar reference vectors. This however requires a slightly different coordinate system, which is shown in Figure 4.9. The choice of coordinate system is somewhat arbitrary, and simplifications can be made if a proper choice is made. By using the coordinate system in Figure 4.9, the reference vectors can be easily set to

$$\hat{i}_{co} = \hat{\phi} \tag{4.18a}$$

$$\hat{i}_{x} = \hat{\theta} \tag{4.18b}$$

using Ludwig's 2nd definition of cross polarization [105]. One last item that must be addressed is the penalty function used in the CMAES optimization. These are also identical to the previous optimization and can be written as

$$p_c(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{F} \\ 10^5 & \text{if } \vec{x} \notin \mathcal{F} \end{cases} \tag{4.19a}$$

$$p_b(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{S} \\ 10^5 + 100\|\vec{x} - \frac{\vec{x}_{max}+\vec{x}_{min}}{2}\| & \text{if } \vec{x} \notin \mathcal{S} \end{cases} \tag{4.19b}$$

where the final fitness function for CMAES is given as

$$f_C(\vec{x}) = f(\vec{x}) + p_c(\vec{x}) + p_b(\vec{x}) \tag{4.20}$$

where the function $f_C(\vec{x})$ represents the fitness function which is optimized by the CMAES algorithm. While the CMAES algorithm optimizes the given fitness function $f_C(\vec{x})$, the PSO algorithm optimizes the function

$$f_P(\vec{x}) = f(\vec{x}) + p_c(\vec{x}) \tag{4.21}$$

in a separate optimization, where $f_P(\vec{x})$ is the implemented fitness function optimized by PSO. As discussed in Section 2.3, the PSO algorithm can exclude the boundary penalty function $p_b(\vec{x})$ because it already incorporates the boundary conditions. With this all necessary aspects have been properly formulated and the two algorithms can be applied to the optimization problem.

4.3.2 PSO and CMAES Results and Comparison

The fitness function shown in the previous section incorporates the BWF parameters which are non-intuitive parameters that may have multiple local optima and whose global optimum is not easily described. Therefore, this represents a difficult optimization in which the fitness can be highly nonlinear and multimodal. The standard PSO technique was applied to the BWF optimization problem with no further modifications to the algorithm described in Section 2.1.1. The parameters for the PSO algorithm used in the BWFO problem are shown in Table 4.7. The

Table 4.7: PSO Intrinsic Parameters for BWFO

PSO Parameter	*Values Used*
c_1	2.0
c_2	2.0
Swarm Size	6
Δt	1.0
Max Iterations (i_{max})	500
Boundary Condition	Invisible
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

Table 4.8: CMAES Intrinsic Parameters for BWFO

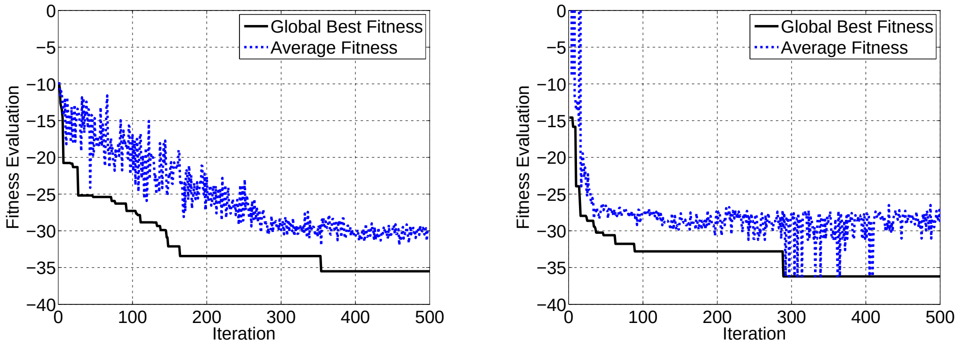
CMAES Parameter	*Values Used*
λ	7
μ	3
σ^0	$\max(\vec{x}_{max} - \vec{x}_{min})/3$

CMAES algorithm discussed in Section 2.2 was also applied to this problem for comparative purposes, and its parameters used for this problem are shown in Table 4.8.

PSO and CMAES demonstrated good convergence for the fitness function given in the previous section. Their convergence results are shown in Figure 4.10. Both algorithms started with roughly the same fitness value, and both were able to bring the fitness down to approximately $f(\vec{x}) = -35$, which is a considerable improvement over the initial points tested. The CMAES algorithm again seems to optimize the function slightly faster than PSO, but some interesting features can also be observed from these convergence plots. First, a considerable drop in fitness is observed late in the optimization run. This can either imply that the fitness function is ill-conditioned and the algorithm is converging slowly towards the optimum, or it can imply that the fitness function is sensitive to small changes in the design variables. It could also imply that the fitness function is discontinuous, but this is not the case since each parameter used in the fitness function and their linear combinations behave fairly well with respect to the design variables. It turns out that one of the BWF's is quite sensitive to any small changes in the design variables, and this can cause sudden changes in fitness seen in the convergence plots. More on

Table 4.9: Final Design Values provided by CMAES and PSO in the BWFO

	L_ℓ (mm)	L_u (mm)	f_d (mm)
PSO	31.885	30.056	5.630
CMAES	31.994	30.072	3.687



(a) PSO Convergence **(b)** CMAES Convergence

Figure 4.10: Convergence for the BWFO problem

this will be explained in the following section. This also explains why the average fitness jumps up and down for the CMAES optimization. The neighborhood which has similar fitness to the global best seen in the CMAES optimization is small, therefore the average fitness jumps up and down because the algorithm can generate points outside of the neighborhood.

The final design values are shown in Table 4.9. Both have fairly comparable values, and strong similarity can be seen for the lengths of the two patches. This is expected because the resonances of the patch are strongly correlated to the patch lengths, and the resonances must be properly placed in order to obtain the wideband response required for this application. The feed distance from the edge shows some discrepancies between the two final designs and most likely is the cause of the difference of $f(\vec{x}_{cmaes}) = -36.21$ versus $f(\vec{x}_{pso}) = -35.50$ for the CMAES and PSO final designs, respectively. Since there are fairly substantial differences between both designs, the design performance of both final designs will be shown for completeness.

The final design performance is summarized in Figures 4.11 and 4.12. The S_{11} response for the final PSO and CMAES designs are plotted in Figure 4.11, and the shaded region represents

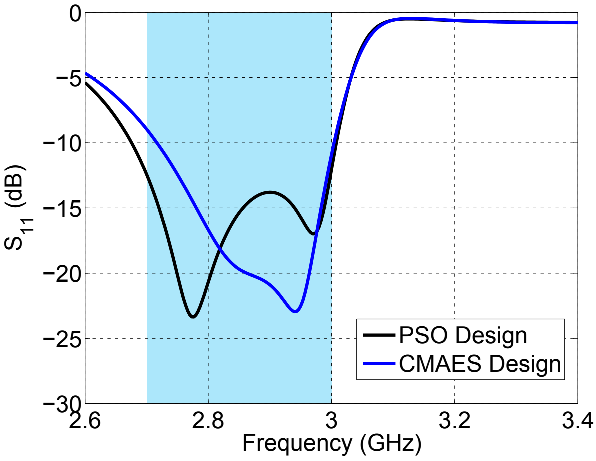
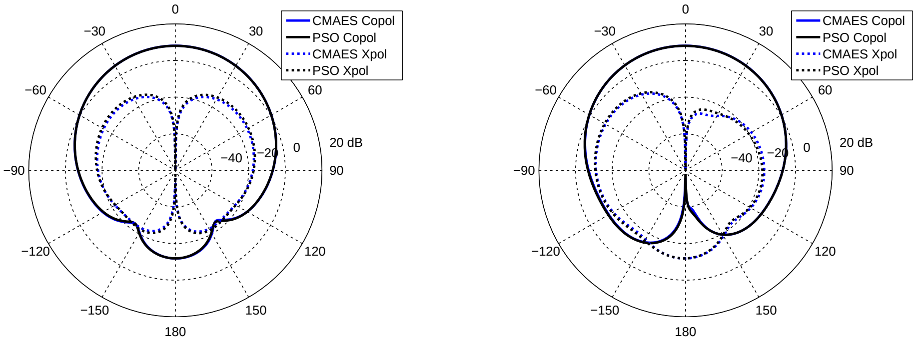
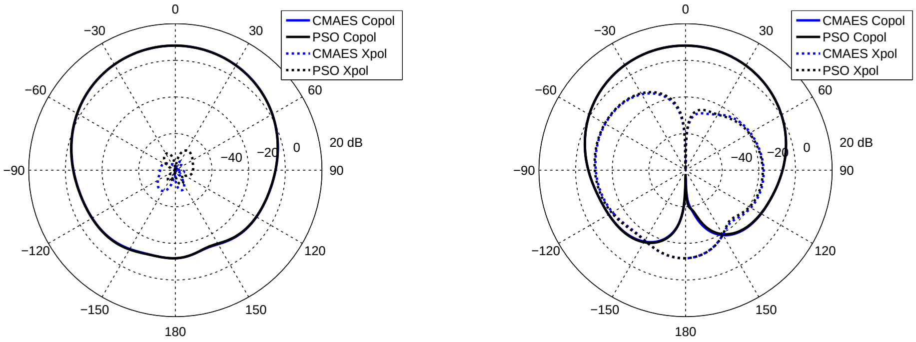


Figure 4.11: S_{11} response for the final PSO and CMAES designs



(a) $\phi = 0$ **(b)** $\phi = 45°$



(c) $\phi = 90°$ **(d)** $\phi = 135°$

Figure 4.12: Radiation patterns of the final optimized stacked patch antenna designs in the principle planes

Table 4.10: Pattern Characteristics of the BWFO Final Designs

Parameter	*PSO value*	*CMAES value*
Directivity	8.03 dB	8.02 dB
Max Xpol	-17.03 dB	-17.6 dB
W_{hv}	-40.49 dB	-51.25 dB
W_4	-19.90 dB	-20.33 dB

the frequency band of interest, 2.7-3.0 GHz. Some fairly significant differences are observed between the two designs in the curves, and this can be attributed to the change in feed probe location in the CMAES design. Even slight shifts in the probe can alter the impedance match for a given frequency. The CMAES design is closer to the probe edge than the PSO final design, which can indicate that a higher input impedance is observed at the antenna input port due to the currents approaching zero near the edge. The reason that CMAES pinpointed this final design as opposed to the design found by PSO is that the W_{hv} parameter was significantly larger in magnitude compared to the final PSO design. These values for W_{hv} and W_4 are shown in Table 4.10 along with other pattern characteristics. As can be seen from Table 4.10, $W_{hv} = -51.25$ dB for the CMAES in comparison to $W_{hv} = -40.49$ dB for the PSO final design. This decrease in W_{hv} outweighed the advantage of the lower S_{11} observed by the PSO design. This demonstrates one of the disadvantages of a single valued weighted-objective fitness function, where one objective can outweigh another important objective. This is also why the weights can be critical in distinguishing some designs from others. One way to avoid this is through the implementation of $\max(W_{hv}, -35)$ in the fitness function in order to cap the value at a certain minimum number. However, this can further complicate the fitness function development.

Overall, similar characteristics between the two final designs are still observed in the radiation patterns shown in Figure 4.12. As shown in the table, the directivities are nearly identical. The copolar radiation patterns also show strong similarities. These radiation patterns in Figure 4.12 use Ludwig's 3rd definition of cross polarization, and the coordinate system shown in Figure 4.5

is used to create the radiation pattern cuts. There are some slight differences observed in the crosspolar radiation including their maximum values. The difference between the maximum xpol levels account for the differences in the BWF's, where a lower W_4 and W_{hv} are observed for the CMAES design. While lower xpol levels can decrease these parameters in general, the phase of the fields can also be critical in their reduction, and therefore this is not the only cause for their decrease.

As demonstrated from the results, both algorithms performed well in this optimization run. Both PSO and CMAES found a good solution that either satisfied or nearly satisfied the criterion. CMAES found a design that did not meet the -10 dB criterion for impedance matching within the desired frequency band, however it did discover a design with better overall fitness that provided a better W_{hv} and W_4. This may have been due to CMAES better ability to finely search the local space once the optimization shifts towards a local one. All in all, similar results are demonstrated by both algorithms on the optimization of the BWF's.

4.4 Some Final Comparisons and Discussion

In the previous sections, some comparison was made between the two different algorithms, PSO and CMAES. However a side by side comparison between the two fitness function approaches has not been fully shown. This final comparison provides some insight as to the best approach to optimize and design antennas for weather radar applications in the future with regards to minimizing their error. This error comes from an imperfect isolation between the two polarizations which is quantified as δZ_{dr}, which is the error in the differential reflectivity Z_{dr}. The antenna performance is related to the δZ_{dr} by the bias weighting factors

$$W_{hv} = \frac{\int_0^\pi \int_{-\pi}^\pi \left(F_{hh}(\theta,\phi)\right)^2 F_{hh}^*(\theta,\phi) F_{hv}^*(\theta,\phi) \sin\theta d\phi d\theta}{\int_0^\pi \int_{-\pi}^\pi \left|F_{hh}(\theta,\phi)\right|^4 \sin\theta d\phi d\theta} \tag{4.22a}$$

$$W_4 = \frac{\int_0^\pi \int_{-\pi}^\pi \left|F_{hh}(\theta,\phi)\right|^2 \left|F_{hv}(\theta,\phi)\right|^2 \sin\theta d\phi d\theta}{\int_0^\pi \int_{-\pi}^\pi \left|F_{hh}(\theta,\phi)\right|^4 \sin\theta d\phi d\theta} \tag{4.22b}$$

Table 4.11: Final Design Values provided by PSO in the Optimization Runs

	L_ℓ (mm)	L_u (mm)	f_d (mm)
XPFO	31.812	30.121	6.085
BWFO	31.885	30.056	5.630

Table 4.12: Pattern Characteristics of the BWFO Final Designs

Parameter	*BWFO*	*XPFO*
Directivity	8.03 dB	7.97 dB
Max Xpol	-17.03 dB	-16.8 dB
W_{hv}	-40.49 dB	-33.39 dB
W_4	-19.90 dB	-19.75 dB

where the functions $F_{hh}(\theta, \phi)$ and the $F_{hv}(\theta, \phi)$ represent the copolar and crosspolar fields radiated by the antenna. These formulas can be fairly complex in nature, and it is difficult to pinpoint the best way to minimize these parameters. Our approach was to optimize the antenna with two different objectives in two separate optimizations as a first investigation.

The first optimization was designed to minimize the crosspolar radiation in the principle cuts. This mimics an antenna designer trying to minimize the fields in these cuts, which is typically done by antenna designers when minimizing the cross polarization. The second optimization was designed to directly minimize the BWF's, and the BWF's were incorporated into the fitness function. For simplicity sake, only the final designs given by PSO will be compared. Their final design variable results are shown in Table 4.11. A comparison of their S_{11} response and their radiation patterns are shown in Figures 4.13 and 4.14.

The final design values in Table 4.11 show a fairly strong agreement, although a slight change in the feed distance f_d is observed. Therefore similar input impedance performance is expected as demonstrated in Figure 4.13. The shaded region shows the region of frequency operation

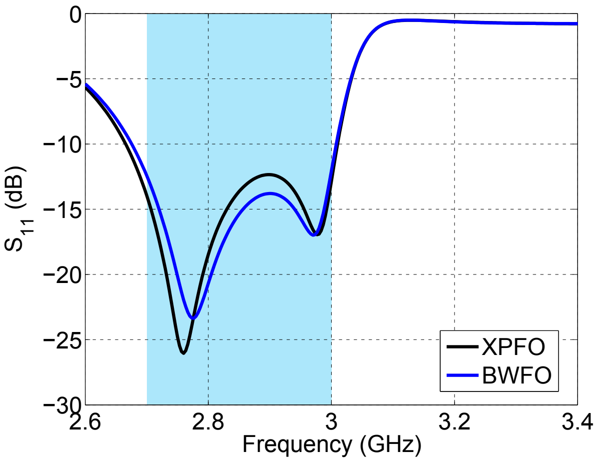
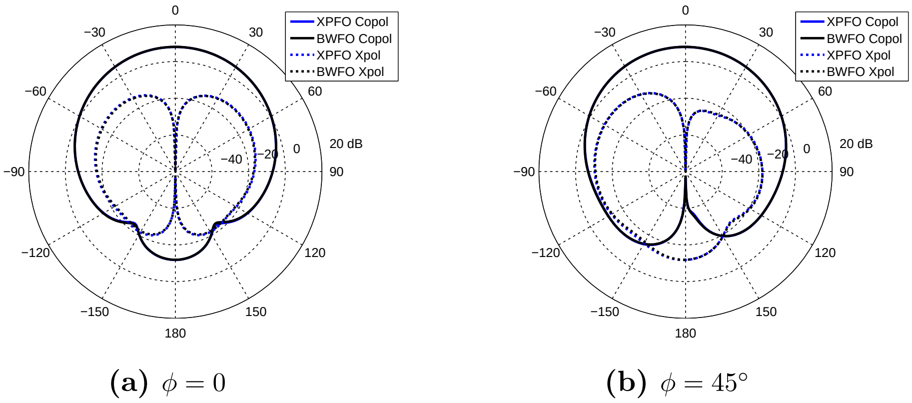
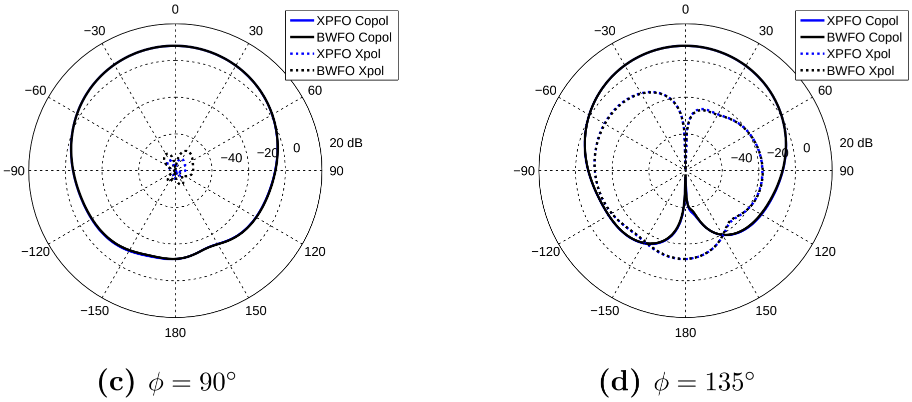


Figure 4.13: S_{11} response for the final PSO and CMAES designs



(a) $\phi = 0$

(b) $\phi = 45°$



(c) $\phi = 90°$

(d) $\phi = 135°$

Figure 4.14: Comparison of the radiation patterns at 2.85 GHz for the final optimized stacked patch antenna designs of the two different optimization approaches

intended for the radar system, which is located at 2.7 - 3.0 GHz. Both optimization runs show satisfactory S_{11} (≤ -10 dB) in this region of interest. The radiation patterns are plotted in Figure 4.14 using Ludwig's 3rd definition of cross polarization as well as the coordinate system shown in Figure 4.5. Both the copolar and the crosspolar fields show very similar radiation patterns, which is somewhat expected since their final design values are so similar. This affinity leads to a fairly comparable W_4 for their final design, as shown in Table 4.12. However their W_{hv} shows some significant discrepancies, therefore a closer investigation into these integrals may be needed in order to explain these similarities and differences.

One can start by rewriting the terms in the integrals for the BWF's in a more understandable manner. The W_4 integral remains the same as given by

$$W_4 = \frac{\displaystyle\int_0^{\pi}\int_{-\pi}^{\pi} |F_{hh}(\theta,\phi)|^2 \, |F_{hv}(\theta,\phi)|^2 \sin\theta d\phi d\theta}{\displaystyle\int_0^{\pi}\int_{-\pi}^{\pi} |F_{hh}(\theta,\phi)|^4 \sin\theta d\phi d\theta} \tag{4.23}$$

and for W_{hv}, the integral can be written as

$$\begin{aligned} W_{hv} &= \frac{\displaystyle\int_0^{\pi}\int_{-\pi}^{\pi} (F_{hh}(\theta,\phi))^2 \, F_{hh}^*(\theta,\phi) F_{hv}^*(\theta,\phi) \sin\theta d\phi d\theta}{\displaystyle\int_0^{\pi}\int_{-\pi}^{\pi} |F_{hh}(\theta,\phi)|^4 \sin\theta d\phi d\theta} \\ &= \frac{\displaystyle\int_0^{\pi}\int_{-\pi}^{\pi} |F_{hh}(\theta,\phi)|^3 \, |F_{hv}(\theta,\phi)| \, e^{j\Phi_{hh}(\theta,\phi) - j\Phi_{hv}(\theta,\phi)} \sin\theta d\phi d\theta}{\displaystyle\int_0^{\pi}\int_{-\pi}^{\pi} |F_{hh}(\theta,\phi)|^4 \sin\theta d\phi d\theta} \end{aligned} \tag{4.24}$$

where $\Phi_{hh}(\theta,\phi)$ and $\Phi_{hv}(\theta,\phi)$ are the phase of the copolar and crosspolar radiation patterns for a given point (θ,ϕ) in the spherical space. With the terms described under the integral by their phase and magnitude, one can begin to explain the observations made in the previous values for the BWF's. We note that the W_4 did not change very drastically, and this is expected because the terms under the integral are only dependent on the field magnitude. As shown in Figure 4.14, the magnitudes between the two radiation patterns are quite similar, and therefore no large change should be expected in the integral. For the W_{hv} integral, the phase comes into the picture, and a slight change in phase can account for big changes in W_{hv}, even though the magnitudes are relatively similar. The integral can be further expanded into the numerical summation

$$\begin{aligned}W_{hv} &= \frac{\displaystyle\int_0^{\pi}\int_{-\pi}^{\pi} |F_{hh}(\theta,\phi)|^3 \, |F_{hv}(\theta,\phi)| \, e^{j\Phi_{hh}(\theta,\phi)-j\Phi_{hv}(\theta,\phi)} \sin\theta d\phi d\theta}{\displaystyle\int_0^{\pi}\int_{-\pi}^{\pi} |F_{hh}(\theta,\phi)|^4 \sin\theta d\phi d\theta} \\ &\approx \frac{\sum_i^M \sum_j^N |F_{hh}(\theta,\phi)|^3 \, |F_{hv}(\theta,\phi)| \, e^{j\Phi_{hh}(\theta,\phi)-j\Phi_{hv}(\theta,\phi)} \sin\theta}{\sum_i^M \sum_j^N |F_{hh}(\theta,\phi)|^4 \sin\theta}\end{aligned} \tag{4.25}$$

assuming that the same $\Delta\phi$ and $\Delta\theta$ was used for both integrals in the numerator and denominator. Now in order to find the integral that accounts for the large change, both integrals are recalculated for the two different cases. For the XPF optimization, the magnitude of the integral can be calculated by

$$\left|\sum_i^M \sum_j^N |F_{hh}(\theta,\phi)|^3 \, |F_{hv}(\theta,\phi)| \, e^{j\Phi_{hh}(\theta,\phi)-j\Phi_{hv}(\theta,\phi)} \sin\theta\right| \approx 26.76 \tag{4.26a}$$

$$\sum_i^M \sum_j^N |F_{hh}(\theta,\phi)|^4 \sin\theta \approx 5.00 \times 10^4 \tag{4.26b}$$

where (θ,ϕ) are sampled over the entire spherical space. For the BWF optimization, the integral can be calculated by

$$\left|\sum_i^M \sum_j^N |F_{hh}(\theta,\phi)|^3 \, |F_{hv}(\theta,\phi)| \, e^{j\Phi_{hh}(\theta,\phi)-j\Phi_{hv}(\theta,\phi)} \sin\theta\right| \approx 4.44 \tag{4.27a}$$

$$\sum_i^M \sum_j^N |F_{hh}(\theta,\phi)|^4 \sin\theta \approx 4.98 \times 10^4 \tag{4.27b}$$

One can see from the following two calculations that the integral in the denominator is nearly identical between the two optimized designs. This is expected because only the magnitude is used, and the copolar magnitudes were nearly identical between the two design radiation patterns. The integral in the numerator, however, shows an increase by almost a full magnitude for the XPFO case. Therefore, this integral accounts for the drastic change in the W_{hv}, and explains why such a high discrepancy between the designs exists. The phase must also be taken into account whenever considering these integrals, and any slight changes in the phase can contribute to a large increase in W_{hv}.

This investigation has shed some insight on the behavior of the two bias weighting factors, W_4 and W_{hv}. Both of these terms help predict the error that will be encountered for the designed antenna. Among the two factors, W_4 appears to be more stable, meaning that any slight deviations in the fields will not cause large changes in the end result. The factor W_{hv} appears to be highly dependent on both the phase and magnitude of the fields. This implies that there may be a minimum value for this parameter where its value does not have any meaning. For instance, if a W_{hv} of -50 dB is predicted by simulation, there is an extremely low probability that a measured result would show a similar value. This is due to the fact that both simulation and measurement are prone to errors in both the magnitude and phase. If only a slight variation in the phase (or in the magnitude) causes such as a large change, then it is reasonable to expect that a large change will occur between the simulation and measurement due to possible variations observed between the two results. Ultimately both parameters are important in the final prediction of the weather radar performance, and caution must be exercised when predicting the antenna performance based on the W_{hv} parameter for very small values.

As a final note, it should be seen that both optimizations produced fairly similar results, and this implies that an optimization of the crosspolar fields seems to place the design in the right direction for best BWF performance. While this may be the case for small aperture antennas, such as patch antennas, more investigation must be conducted for antennas with a larger aperture. These antennas can have much more variation in their crosspolar and copolar fields, and therefore an optimization of the crosspolar fields in their principle cuts may not be sufficient to guarantee low values for W_{hv} and W_4. However, for smaller aperture antennas with slowly changing field magnitudes, the crosspolar field optimization is the most natural approach, and this procedure can be used as a first order technique for minimizing the BWF's.

CHAPTER 5

Smooth Contour Septums for High Power Microwave Systems

Circular polarized (CP) radiation can be highly desirable for many antenna applications. A circularly polarized wave can be decomposed into two different linear polarizations with a 90° phase shift, and this alleviates losses due to improper antenna orientation and fading when using one CP antenna with another linear antenna. If one linear polarization experiences more fading, the system can still rely on the link provided by the other linear polarization. Some common examples of systems providing coverage with CP antennas are GPS [109, 110] and handset antennas for communications [111]. Circular polarized horn antenna systems are another class of CP antennas which are widely used as feeds in reflector antenna systems. Their applications include satellite communications, deep-space telemetry, and radio astronomy. The feed antennas are often the most restricting component of the whole reflector antenna system, and proper design can lead to good performance. Therefore they are both an important component and a difficult one to design.

The metallic horn antenna is fed by a metallic waveguide whose cross-section is usually circular or rectangular, having the appearance of a circular or rectangular pipe. In their most typical form, these horn antennas produce a linear polarization, and therefore some system complexity must be added in order to create a CP horn antenna. Several techniques that are used in the industry include an orthogonal mode transducer (OMT) which is a three port device capable of controlling the TE_{10} and the TE_{01} modes independently in a connected square waveguide. This system can suffer from design complexity due to the size of the OMT and the connected microwave network [112, 113]. Another technique that has gained popularity in

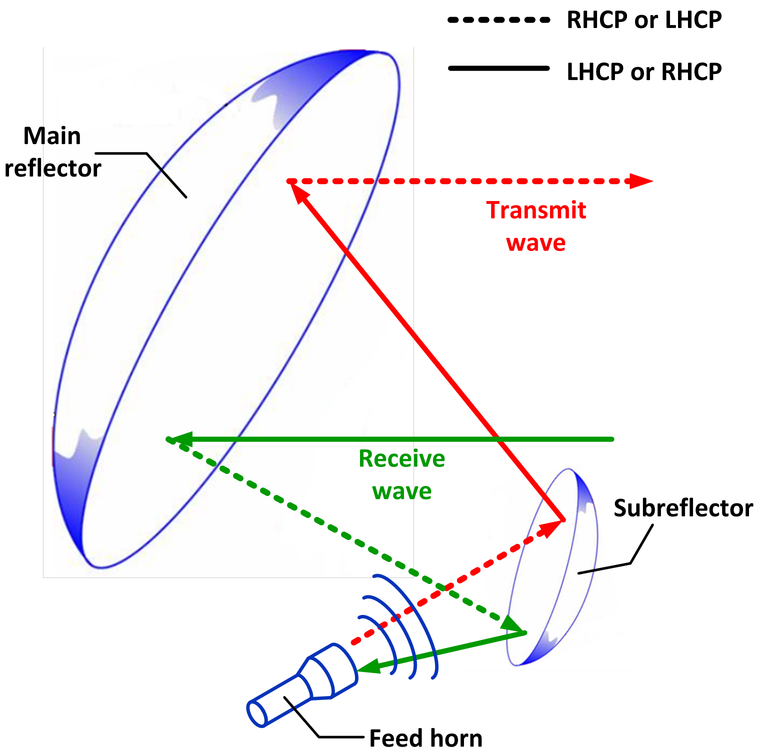


Figure 5.1: Reflector antenna system using circular polarization duplex

industry is the use of a septum, as shown in Figure 5.2, which creates a circularly polarized wave from a linearly polarized (LP) wave. Depending on the orientation of the septum, one port will produce right-handed circular polarization (RHCP) at the output port if excited, and left-handed circular polarization (LHCP) will be generated when the other port is excited. These microwave devices provide a compact design solution for use in two way communication systems. Many systems can be polarization-duplexed, e.g. the transmission are sent as RHCP waves and the received signal is LHCP waves [114] as shown in Figure 5.1. One of the first septum designs had used a slanted septum in order to achieve a CP wave [115]. The stepped septum as seen in Figure 5.2 was later introduced [116] and is currently the standard for most horn antennas. Excellent circularity, impedance matching, and isolation can be achieved with these types of structures [116, 117].

Our goal in this chapter is to optimize the septum structure for good impedance matching and axial ratio with Particle Swarm Optimization in order to explore more possible capabilities with the septum. We then focus on smoothed septums which eliminate the sharp corners of the

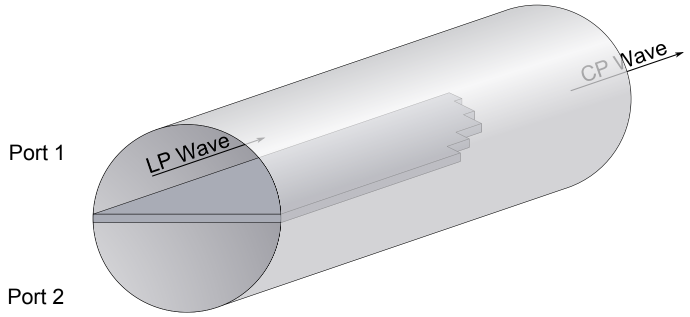


Figure 5.2: Septum polarizer within a circular waveguide cross section

step septum. These sharp corners inevitably cause air breakdown for high power applications and are preferred to be avoided. The smoothing is accomplished through the use of the sigmoid function, which is also formally introduced. The sigmoid septum is then optimized using PSO for good impedance matching and axial ratio over narrow and wide bandwidths.

5.1 Septum Design and Current Issues

The stepped septum design as shown in Figure 5.2 has seen widespread use in many different applications, and it is most commonly placed in designs using square waveguide [116] or circular waveguide [117]. In order to obtain a CP wave, one only excites fields in the upper or lower portion of the waveguide. This can be accomplished using waveguide to coaxial adapters, which are not shown in Figure 5.2 but are still a critical component in the design. The modes excited in either the top or bottom portion of the waveguide can be decomposed into an even and odd mode. In [116] this even-mode/odd-mode analysis was shown, and it provides an understanding on how each mode (TE_{10} and TE_{01}) is generated by the excitation of port 1 or 2. The even mode excitation is responsible for the generation of the TE_{10} mode for a rectangular waveguide, while the odd mode excitation creates the TE_{01} mode in the rectangular waveguide. A graphic visualization of these even/odd modes for the rectangular waveguide was presented using HFSS [118]. It was demonstrated using HFSS that an even mode would produce a vertically polarized

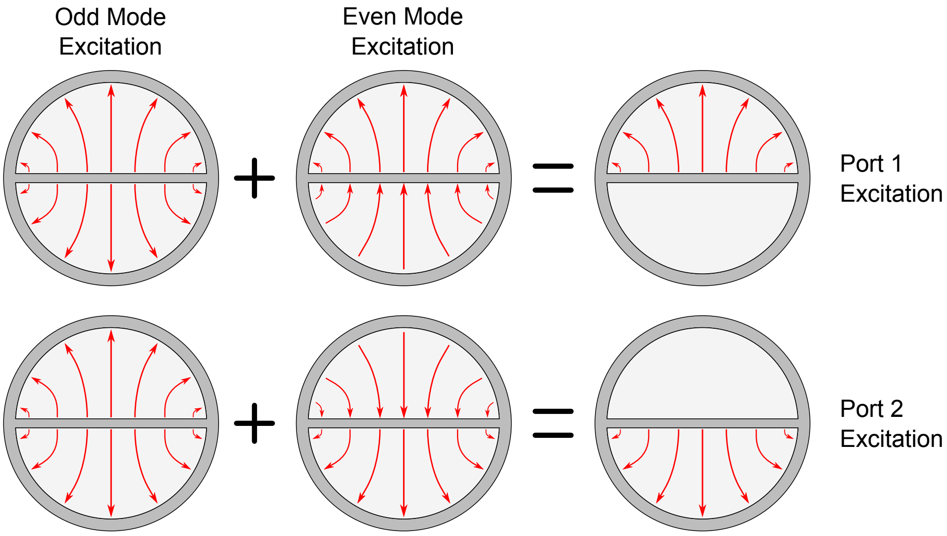


Figure 5.3: Decomposition of the port 1 and port 2 excitation into the even and odd modes, which are responsible for generating the TE_{11} as well as the rotated TE'_{11}, respectively

wave while the odd mode would create a horizontally polarized field. For a circular waveguide the mode distribution will be different than the rectangular waveguide. In this type of cross section, the dominant mode is the TE_{11} mode, and this mode can exist with any arbitrary rotation since the waveguide is circularly symmetric [63]. Just as with the rectangular waveguide, the circular waveguide septum has its even and odd modes as shown in Figure 5.3, which generate a TE_{11} mode whose vectors are normal to the septum and a rotated TE_{11} mode whose vectors are parallel to the septum, respectively. We will designated this rotated TE_{11} mode as TE'_{11} for dinstinction later on.

The key with septums is that the steps (or contour) are designed such that both modes have equal magnitudes and are 90° out of phase. With these characteristics a circularly polarized wave is created, and the septum can be easily connected to a horn antenna in order to provide CP radiation. However, the design of the step or the contour is not as straightforward. In [116] a stepped septum design with good impedance matching performance was found through *trial-and-error*, and a dielectric loading block was placed in the waveguide in order to compensate for non-orthogonal phase differences over a wide bandwidth. Albertsen and Skov-Madsen analyzed compact septums with the Wiener-Hopf technique combined with Galerkin's method [119]. In order to achieve an orthogonal phase difference, they introduced notches into the plate rather than using a stepped approach, which allowed a more compact design. However, this design only showed good results for narrowband applications. Others have investigated thick stepped septums and stepped-thickness stepped septums using Evolution Strategies [120]. In their analysis they had used a mode-matching technique orginally proposed in [121], which splits the septum into sections of ridge waveguide and applies mode-matching to each section. With their analysis and investigation, they proposed some guidelines for the design of wideband septum designs. Dielectric septums have also been recently introduced as another alternative to the conventional metallic septum [122]. A few other notable projects using septums to achieve good axial ratio performance can be found in [114, 123, 117, 124].

While these designs have shown excellent results, one main disadvantage in their design is the use of sharp corners for the stepped septums. Since a sharp corner (and also discontinuities in curvature) can give rise to field singularities, the power handling capabilities of these devices can be severly limited. Much of the research in High Power Microwave (HPM) systems has focused on other design aspects instead of the septum. As a first step to the investigation, a smooth contour septum using the sigmoid function is proposed for possible use. The septum will be optimized for impedance matching and axial ratio performance in order to determine whether a smooth contour is possible for septum applications. Our investigation also wanted to determine if there are other features that can be exploited from the curved septums in comparison to the traditional stepped septums.

The septum design under investigation is shown in Figure 5.4, which is a scaled version of the septum design seen in [117] for operation in C-band at 5.8 GHz. The first section (not shown) is a waveguide to coaxial adapter used to connect a 50 Ohm coaxial line to the waveguide. The next section is the septum, which is either a stepped septum or a sigmoid septum. The last section is the circular horn used to radiate the CP fields. Now, there are several ways to approach the design problem. One can use a modular approach where the waveguide to coaxial adapter, septum, and horn antenna are designed separately, and this can allow for a faster optimization since each piece can be designed separately. However, electromagnetic coupling between each component is not incorporated in the approach, and these effects may take away from good performance in the final design. The other approach would be to include the entire antenna

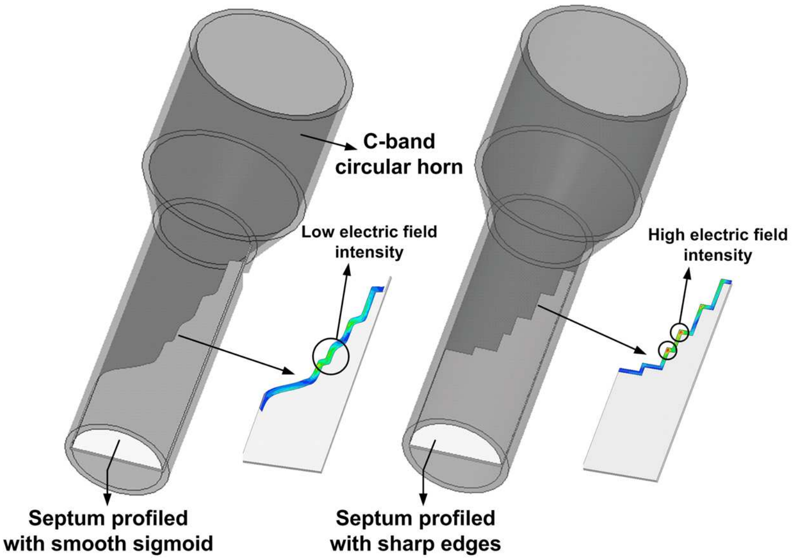


Figure 5.4: Septum and circular horn design; the design to the left uses the sigmoid function and the one to the right uses a stepped septum

structure in order to account for any coupling between elements. While this may take a longer amount of time to optimize, the final design will be ready immediately after the optimization is finished.

We shall proceed by introducing the sigmoid function and some its various properties of interest. Once the sigmoid contour is described, then the optimization problem will be properly formulated for both the stepped septum and the sigmoid septum. The final designs provided by the optimization will then be compared in order to make some conclusions about possible properties of the smoothed septum versus the stepped septum.

5.2 Sigmoid Function and Its Properties

The sigmoid function is well-known in the statistical and solid state physics research areas by the Fermi-Dirac distribution [125]. However, our use of the function is simply to define the contour of the septum. One version of the sigmoid function can be defined by

$$S(x) = \frac{H}{1 + \exp\left(\frac{4(L-x)}{HC}\right)} \tag{5.1}$$

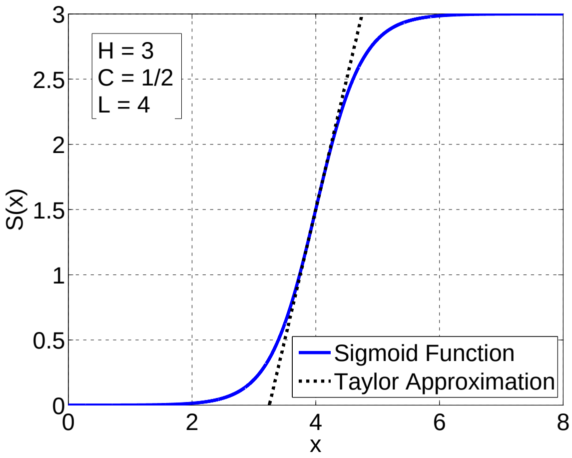


Figure 5.5: Sigmoid function and its Taylor expansion about the point $x = L$

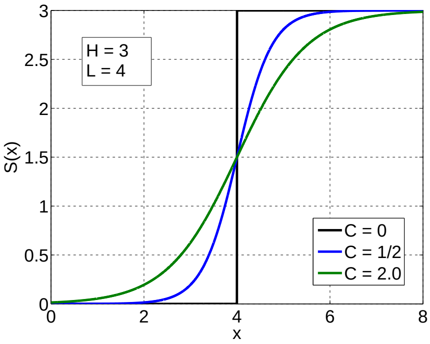
where H defines the height of the sigmoid, L defines the position of the sigmoid midpoint where $S(L) = H/2$, and $1/C$ defines the slope of the sigmoid at the midpoint. Note that this is not the most standard definition of the sigmoid function, but this representation was chosen in order to control the parameters such as the slope, height, and position of the sigmoid more effectively. The first-order Taylor expansion of $S(x)$ about the point $x = L$ is given by the following.

$$S'(x) = \frac{4\exp\left(\frac{4(L-x)}{HC}\right)}{C\left(1+\exp\left(\frac{4(L-x)}{HC}\right)\right)^2} = \frac{\operatorname{sech}^2\left(\frac{2(L-x)}{HC}\right)}{C} \tag{5.2a}$$

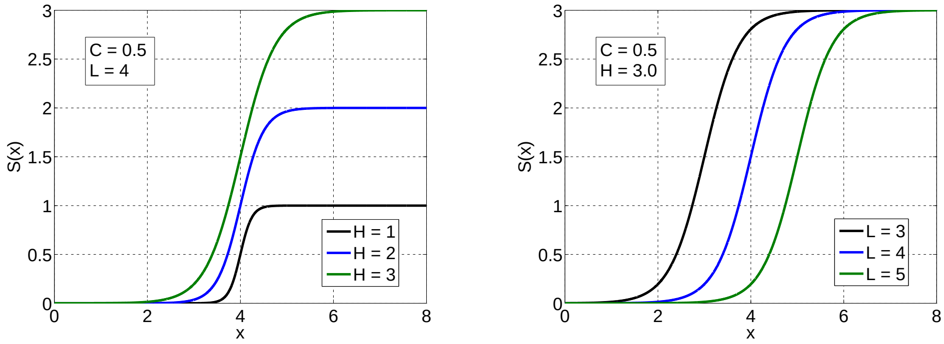
$$S(x) \approx S(L) + S'(L)(x-L) = \frac{H}{2} + \frac{1}{C}(x-L) \tag{5.2b}$$

Clearly, the parameter $1/C$ controls the slope independently and L controls the midpoint location of the sigmoid, as depicted in Figure 5.5. Another consequence from this equation is that H is the height of the sigmoid. This may not be as intuitive from the other parameters, but since the point $x = L$ is the location where the sigmoid function reaches half its height, i.e. $S(L) = 0.5S(x)|_{x\to\infty} = H/2$, the height is then defined by H. Overall, this choice of representation allows a clear visualization of the sigmoid function and which parameters control these aspects. Several different sigmoid curves were plotted in Figure 5.6 in order to demonstrate how each of these parameters shapes the sigmoid function.

The sigmoid function is also able to realize the step function by setting $C = 0$, which makes



(a) Varying the slope $1/C$



(b) Varying the height H

(c) Varying the midpoint location L

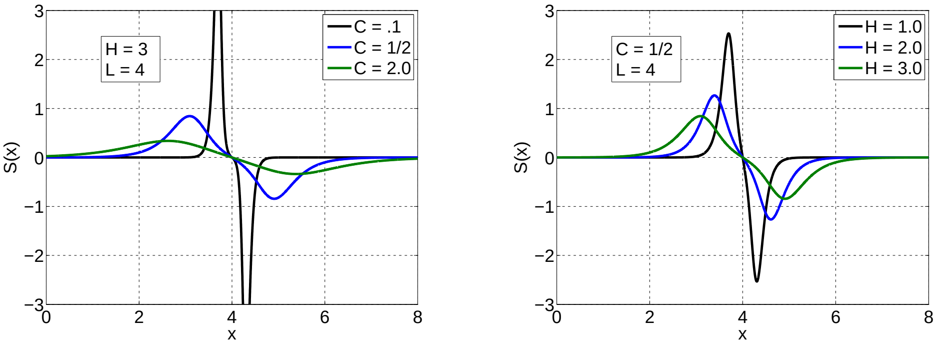
Figure 5.6: The effect of each parameter on the sigmoid function

the slope infinite. This is an important aspect of the sigmoid function which is exploited in the optimization. It is not well-known whether a smooth contour may have better performance than the stepped septums, and the optimizer will demonstrate which curve is superior by the result of the optimization. If the stepped septum truly has better performance than a smoothed septum, then a stepped septum should result from the optimization. If the opposite occurs then new curves may need to be investigated in the septum design area for possible design improvements.

Another feature of interest is the curvature of the sigmoid function. The curvature of a planar curve within the x and y coordinates can be computed using

$$\kappa(t) = \frac{|\vec{r}' \times \vec{r}''|}{|\vec{r}'|^3} \tag{5.3}$$

where $\vec{r}$ is the parameterized coordinate vector $(x(t), y(t), z(t))$ describing the sigmoid function. In this particular example, $\vec{r} = (t, S(t), 0)$, since the curve is planar and does not change position



(a) Varying the slope $1/C$ **(b)** Varying the height H

Figure 5.7: The effect of each parameter on the curvature of the sigmoid function

in z. It can be shown that this leads to

$$\kappa(t) = \frac{S''(t)}{(1+S'(t)^2)^{3/2}} = \frac{4C\tanh\left(\frac{2(L-t)}{CH}\right)\operatorname{sech}^2\left(\frac{2(L-t)}{CH}\right)}{H\left(C^2+\operatorname{sech}^4\left(\frac{2(L-t)}{CH}\right)\right)^{3/2}} \tag{5.4}$$

where t is the parameter of each component and κ is the curvature of the function. In high power applications, a sharp change or discontinuity in the curvature are the primary cause in breakdown and high electric field concentrations. This is avoided by the sigmoid function, where the curvature is continuous for all x. It should be noted that a high curvature is still achievable by the sigmoid function, allowing the optimizer to choose the curvature that provides the best design. As shown in Figure 5.6 and 5.7, this occurs as C approaches zero, where the step function is achieved with the sigmoid function. The continuity of κ versus x is shown in Figure 5.7 as well as the effect of the height and slope on $\kappa(t)$. It is observed that the curvature increases as both C and H decreases. These curves are expected from our previous observations. As the sigmoid function approaches the step function ($C \to 0$), the curve gradually becomes sharper. Once the step function is reached, both the derivative and second derivative have singularities at $x = L$, which typically implies that the curvature is discontinuous or even infinite. The relationship between the curvature and the height H is somewhat counterintuitive, but the quickest way to understand this is through the original sigmoid function in equation 5.1. As $H \to 0$, the function also becomes more like a step function, even though its height tends to zero. The H

in the exponential function forces this transition as H approaches zero. This is a compromise in our choice of representation, and this is not as critical as the proper control of the sigmoid function in order to simplify the constraints in the optimization as seen in the next section.

There are plenty of different shapes that are sigmoidal in nature, e.g. the arctangent function, however the choice to use the sigmoid function was not arbitrary by any means. The sigmoid function is convenient due to its exponential decay properties: the function approaches a constant value relatively quickly outside of the region $L \pm 0.75CH$. At these points the function is within 5% of the final value which it approaches at $\pm\infty$, i.e. $S(L + 0.75CH) = 0.953H$ and $S(L - 0.75CH) = 0.474H$. This is useful because no other changes must be made in order to force the contour to exist in a given region.

In order to create a similar contour to the stepped septum, a summation of sigmoids are used in order to create the sigmoid contour on the septum. The summation

$$s_c(x) = \sum_{i=1}^{P} S_i(x) = \sum_{i=1}^{P} \frac{H_i}{1 + \exp\left(\frac{4(L_i - x)}{H_i C_i}\right)} \tag{5.5}$$

will be used to define this contour, where $s_c(x)$ is the septum contour, P is the number of sigmoid steps, and the ith sigmoid $S_i(x)$ with its parameters L_i, H_i, and C_i. Each sigmoid function $S_i(x)$ can change its parameters to create a general shaped contour which can still achieve the original stepped septum shape by setting $C_i = 0$, $\forall i \in 1, \ldots, P$. A general four-sigmoid curve is shown in Figure 5.8.

5.3 Formulation of the Optimization Problem

5.3.1 Stepped Septum Optimization

As usual, the optimization must be formally defined in order to proceed forward with the design. This optimization focuses specifically on the septum design. Therefore the only design parameters that will be provided to the optimizer will be the coefficients shown in the previous section. A top view of the overall horn and waveguide geometry to be used in the simulations is shown in

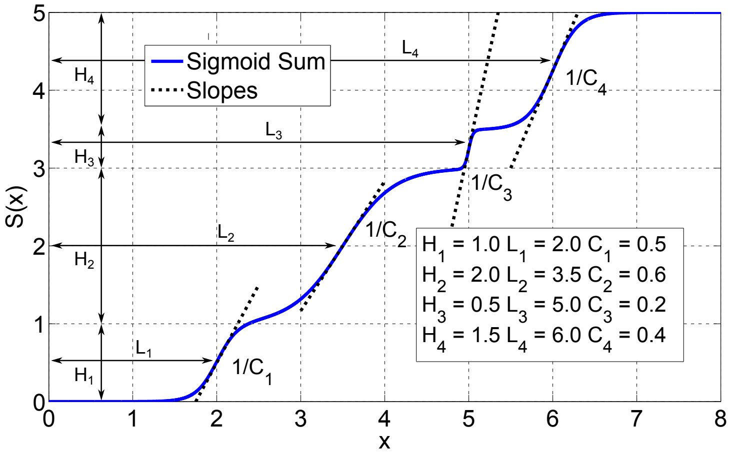


Figure 5.8: Sigmoid summation to represent the septum contour

Figure 5.9, and the values for each dimension are provided in Table 5.1. These values were found based on a frequency scaled version of the horn seen in [117], which was originally designed to operate at 1.296 GHz. The design values in the table are given both in wavelengths[1] λ_0 as well as in cm. Both the sigmoid and the stepped septums are traced in order to show their placement within the waveguide in Figure 5.9. These figures demonstrate a possible design configuration and are not the final optimized septum designs. The final optimized designs may have a different appearance and location.

We first apply optimization to the stepped septum, which is attained by setting $C_i = 0, \forall i \in \ 1, \ldots, P$, assuming that there are P steps. A five step septum design is provided for visualization in Figure 5.10. For this particular configuration, the number of steps P, is determined by the number of parameters needed to completely specify the step dimensions. As seen in Figure 5.10, there are five length (L) parameters and five height (H) parameters. Therefore this is termed the 5-step septum, and this vocabulary will be used throughout the rest of this chapter. The L, C, and H parameters represent the design variables of this problem, and now the solution space must be defined. We define the solution space with the following assumptions. The first is that the H parameters must satisfy the constraint $\sum H_i = H$, as seen in Figure 5.10 where $H = W_1 - T = 4.04$cm. Equality constraints are very difficult if not impossible to meet when generating a random vector for $H_1, \ldots, H_P$. Therefore we remove one of the variables from

---

[1] At 5.8 GHz, the free space wavelength $\lambda_0 = 5.17241$cm

Table 5.1: Design Specifications for both the Stepped and Sigmoid Septum designs in terms of free space wavelength λ_0 at 5.8 GHz

Parameter	Length (λ_0)	Length (cm)	Parameter	Length (λ_0)	Length (cm)
V_1	$2.4581\lambda_0$	12.714	W_1	$0.7949\lambda_0$	4.112
V_2	$0.5223\lambda_0$	2.702	W_2	$1.3167\lambda_0$	6.811
V_3	$1.3167\lambda_0$	6.811	W_p	$0.2665\lambda_0$	1.378
T	$0.0614\lambda_0$	0.318			

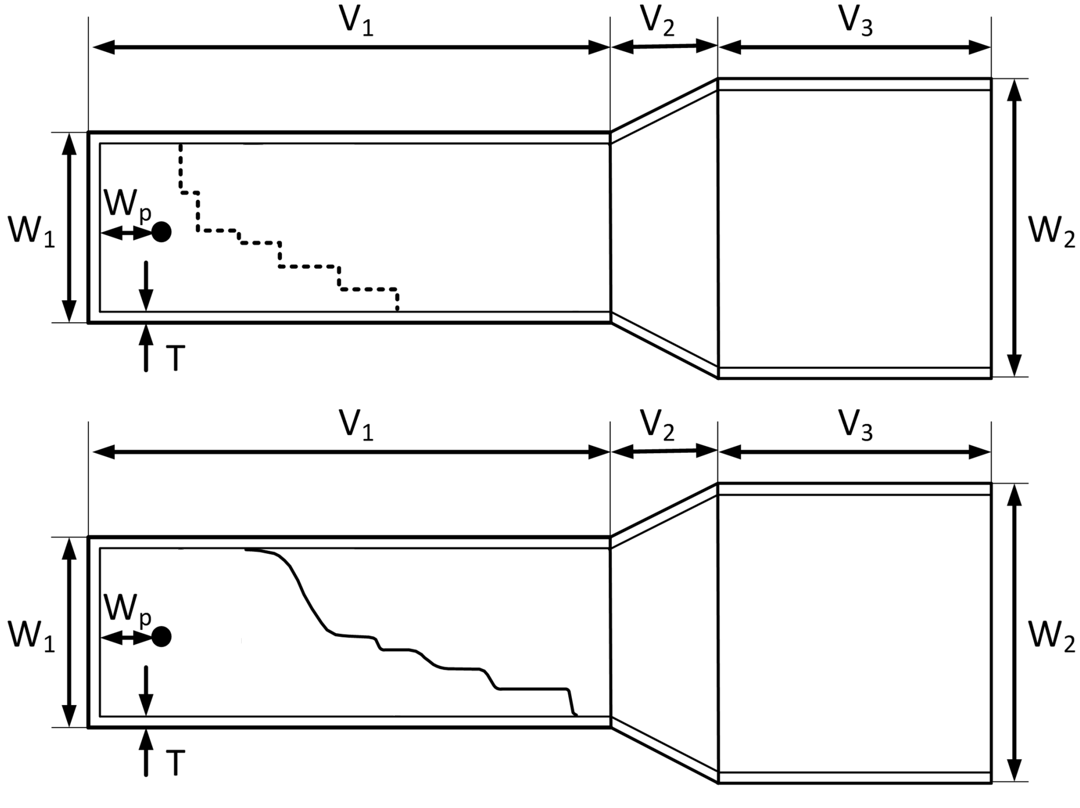


Figure 5.9: Top view of the stepped (dashed) and sigmoid (solid) septums

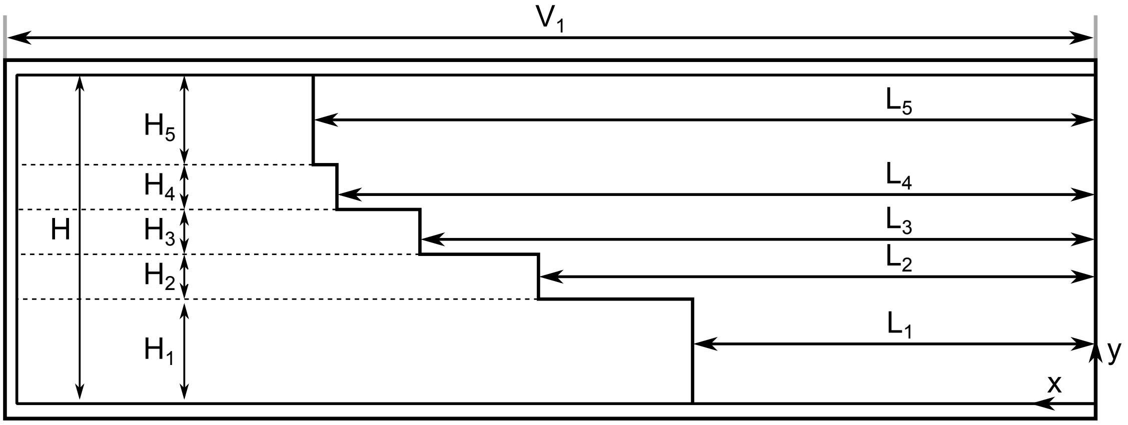


Figure 5.10: Stepped septum design for the 5-step case

the optimization by defining

$$H_P = H - \sum_{i=1}^{P-1} H_i \tag{5.6}$$

which will always satisfy the equality constraint. From this there will only be $P-1$ of the height variables in the optimization.

During the preliminary investigation of the septum it was discovered that, for a reasonably sized septum, the overall placement of the septum within the waveguide did not drastically affect its overall performance. Once the overall shape of the septum was known, a shift by ΔL to each L parameter did not cause any adverse effects. Therefore, in order to eliminate the multi-optima scenario, the parameter L_1 was set to the lowest possible value before any effects were noticed when testing another design given in [117]. This also allows the septum to have the widest design range possible. The value was set to

$$L_1 = 1.793\text{cm} \tag{5.7}$$

and the only parameters left to find are $L_2, \ldots, L_P$. For the stepped septum designs, the number of dimensions in the solution space are $N = 2(P-1)$. For the stepped septum, the design variables will be denoted as $\vec{L} = [L_2, \ldots, L_P]$ and $\vec{H} = [H_1, \ldots, H_{P-1}]$, and the optimization design variable vector is $\vec{x} = [\vec{L}, \vec{H}]$.

The solution space can now be defined by the following

$$H_i = [0, 4.04\text{cm}], \quad \forall i \in 1, \ldots, P-1 \tag{5.8}$$

$$L_i = [1.793\text{cm}, 9.973\text{cm}], \quad \forall i \in 2, \ldots, P \tag{5.9}$$

where the upper limit for each height variable H_i is given by the total waveguide height $H = 4.04\text{cm}$. The lower limit for the length parameters L_i is given by the value set for L_1, since none of the lengths should be at a position lower than the first step. The upper limit for the length parameters L_i comes from a limit on the closeness to the probe feed. This limit minimizes any possible designs which may depend heavily on the probe geometry and its exact location. These limits force the optimizer to test designs which only occupy the specified maximum design area as shown in Figure 5.11. The feasible space $\mathcal{F}$ must also be defined in order to fully characterize the solution space to be tested. The constraints present in this optimization are

$$H_P = H - \sum_{i=1}^{P-1} H_i \leq 4.04\text{cm} \tag{5.10}$$

$$H_P = H - \sum_{i=1}^{P-1} H_i \geq 0 \tag{5.11}$$

$$L_i \geq L_{i-1}, \quad \forall i \in 2, \ldots, P \tag{5.12}$$

where the first two constraints ensure that the last step is also within the same value range as the other heights, i.e. $0 \leq H_P \leq 4.04\text{cm}$, and the last $P-1$ constraints ensures non-repeated designs. Depending on the configuration of the septum design within the commercial solver, the last conditions are also necessary to ensure non-overlapping steps.

Since this optimization problems represents a highly dimensional optimization problem with complex constraints, a constraint penalty function may not be the most efficient solution. Therefore, we approached this problem by initializing all test solutions within the feasible space $\mathcal{F}$. This was done in the usual sense by initializing each particle with a uniform distribution within the solution space $\mathcal{S}$. If a particle's location fell outside of $\mathcal{F}$ then the particle would be reinitialized using the same uniform distribution. While this may take a while longer at the beginning,

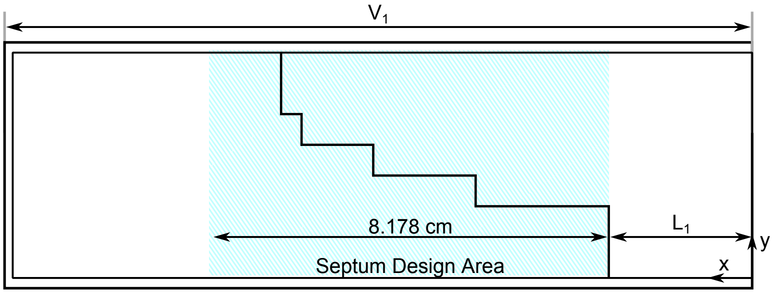
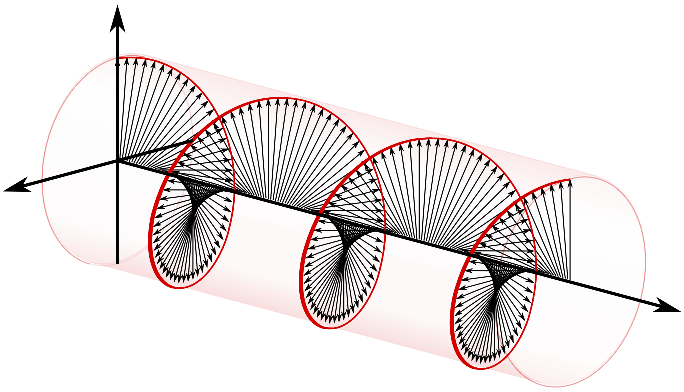


Figure 5.11: Maximum space that could be possibly occupied by the septum design (using the max/min value for L_i)

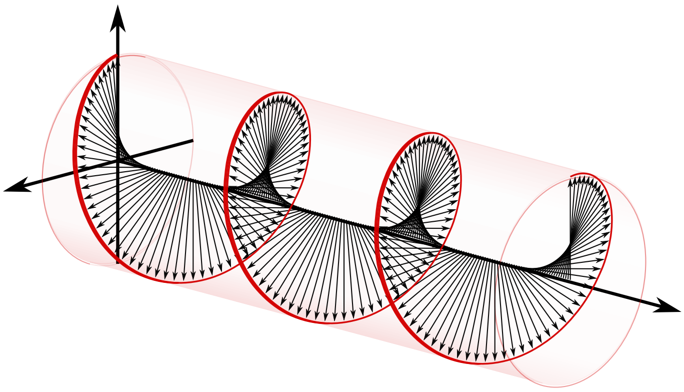
this is a straightforward approach to ensure that all particles begin in $\mathcal{F}$. The approach for this initialization is documented in [91].

The primary goal in this design is to minimize the impedance matching S_{11} and the overall axial ratio (AR) of the radiated fields from the horn. The axial ratio is a figure of merit when discussing the circularity of circularly polarized fields. For a pure CP wave, the tip of the electric field vector will trace out a circle both as time and space changes, as shown in Figure 5.12. For an elliptically polarized wave (an impure CP wave), the tip of the electric field vector traces out an ellipse, which have two axes of different lengths. The axial ratio quantifies the circularity of the wave by the ratio of the major to minor axis. Therefore, the AR varies from $1 < \text{AR} < \infty$, with $\text{AR} = 1$ being the best value, which represents a pure CP wave whose major and minor axes are equal. Many times the dB scale is used by computing $\text{AR}_{dB} = 20\log(AR)$, and an EM wave is classified as a CP wave if $\text{AR}_{dB} < 3$ dB. With the knowledge of the field vector magnitude and phase, the AR is fairly straightforward to compute with the formulas provided in [12].

The only difficulty is that the AR is quite time-consuming to compute versus frequency using many of the commercial electromagnetic solvers available, and therefore some assumptions must be made in order to reduce the computational burden. One assumption that can be made is that the axial ratio does not increase very drastically over a wide angular range. For horn/septum combinations, this is a fairly reasonable assumption and does not pose any problems. This assumption allows us to use the AR at the boresight direction (the direction normal to the horn antenna aperture) as a representation of the overall radiated fields. Therefore the fitness function



(a) Left-handed CP



(b) Right-handed CP

Figure 5.12: Visualization of a circularly polarized wave (Adapted from the Wikimedia Commons files given in [126, 127])

chosen for this optimization is given by

$$f(\vec{x}) = \alpha_1 \max_{f_i \in F}(\text{VSWR}|_{f=f_i}) + \alpha_2 \max_{f_i \in F}(\text{AR}|_{f=f_i}) \tag{5.13}$$

where $\vec{x} = [\vec{H}, \vec{L}]$, $f_i \in F$ is one testing frequency point f_i within the set of frequencies F, the VSWR is the voltage standing wave ratio, and AR is the axial ratio in a magnitude scale. The set of frequencies used for this stepped septum optimization is

$$F = \{5.075, 5.317, 5.558, 5.800, 6.042, 6.283, 6.525\} \tag{5.14}$$

The VSWR is another port parameter that describes the impedance matching on a different scale. The VSWR is defined by the equation

$$\mathrm{VSWR} = \frac{1+|\Gamma|}{1-|\Gamma|} \tag{5.15}$$

where Γ is the reflection coefficient in a magnitude scale. The VSWR was used in this fitness function because its scale is $1 \leq \mathrm{VSWR} \leq \infty$, which is comparable to the AR scale. As can be seen from the function, the VSWR decreases as the reflection coefficient magnitude $|\Gamma| = |S_{11}|$ decreases. In Section 1.3, it was shown that a good value for the reflection coefficient was $|\Gamma| = -10$ dB. A close value for VSWR that corresponds to this upper bound is VSWR = 2. For the AR, 3 dB is often used as a benchmark, and in a magnitude scale that translates to AR = $\sqrt{2}$.

This represents a multiobjective function which is wrapped into one fitness function, where one objective is to minimize the reflection coefficient and the other is to minimize the AR. No *a priori* knowledge is known about the fitness function, and recommendations for setting the weights α_i for each objective in the fitness function $f(\vec{x})$ becomes unclear. One possible approach to find the best choice of weights can be done by equating their threshold values. Since the theshold values were 2 and $\sqrt{2}$ for VSWR and AR, respectively, the weights assignment results from

$$\alpha_1 \mathrm{VSWR}_t = 2\alpha_1 = \alpha_2 \mathrm{AR}_t = \sqrt{2}\alpha_2 \rightarrow \frac{\alpha_2}{\alpha_1} = \sqrt{2} \tag{5.16}$$

where VSWR_t and AR_t are the threshold values for the VSWR and AR, respectively. The above equation implies that $\alpha_2 = \sqrt{2}$ if we make the arbitrary assignment that $\alpha_1 = 1$. With these choices the fitness function can be written as

$$f(\vec{x}) = \max_{f_i \in F}(\mathrm{VSWR}|_{f=f_i}) + \sqrt{2}\max_{f_i \in F}(\mathrm{AR}|_{f=f_i}) \tag{5.17}$$

This weighting scheme may not always be the most optimal, but it represents a systematic and intuitive approach to generating a set of weights for a weighted multiobjective to single objective mapping scheme without any *a priori* knowledge of the design. The goal in this optimization is to minimize this given function.

One point to note is the use of the max function in equation 5.13. This is often known as a *minimax* optimization problem, where the goal of the optimizer is to minimize the maximum value [75]. By using this representation, the worst case scenario in the band of interest is being optimized. These are extremely suitable for problems where the user has some understanding about the maximum bandwidth of the structure. However, this function is not a good representation in problems that explore the limits of devices. For instance, if the bandwidth was set incredibly high (200% of the center frequency), the optimization would return meaningless designs because the worst case scenario will never be optimized. In these cases it would be advisable to start an optimization using other classes of fitness functions. Other options include the sum function, where the VSWR/AR at multiple frequencies are added together. This however loses the guarantee of good performance throughout the frequency band of interest, since this function represents the average value of the points scaled by the number of points used. Optimizing the average does in general provide some useful information, but it is usually encouraged to use the *minimax* approach. We use this approach since we have a rough idea on the bandwidth of the structure.

This covers the required aspects in the optimization of the stepped septum design. The algorithm used for this optimization problem is RPSO, with a slight modification to the initialization of the particles as documented in [91]. The initialization algorithm forces the particles to start their search within the feasible solution space $\mathcal{F}$. In order to provide many degrees of freedom, the number of steps is chosen to be $P = 6$, which makes this a six-step septum optimization. The parameters for the PSO algorithm can be found in Table 5.2, which were used for this optimization problem.

The convergence of the six-step septum optimization is shown in Figure 5.13. With the given fitness function from equation 5.13, the best possible value could occur when VSWR $= 1$ and AR $= 1$. Using equation 5.13, this would produce a value at $f(\vec{x}_{best}) = 1 + \sqrt{2} \approx 2.414$. The best value seen at iteration 500 is 2.858, which is close to this value. At this iteration, the fitness function value can be broken down into VSWR $=$ 1.085 and AR $=$ 1.254 at boresight.

Table 5.2: PSO Intrinsic Parameters for the Stepped Septum Optimization using Six Steps

PSO Parameter	*Values Used*
c_1	2.0
c_2	2.0
Swarm Size	19
Δt	1.0
Max Iterations (i_{max})	500
Boundary Condition	Invisible
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

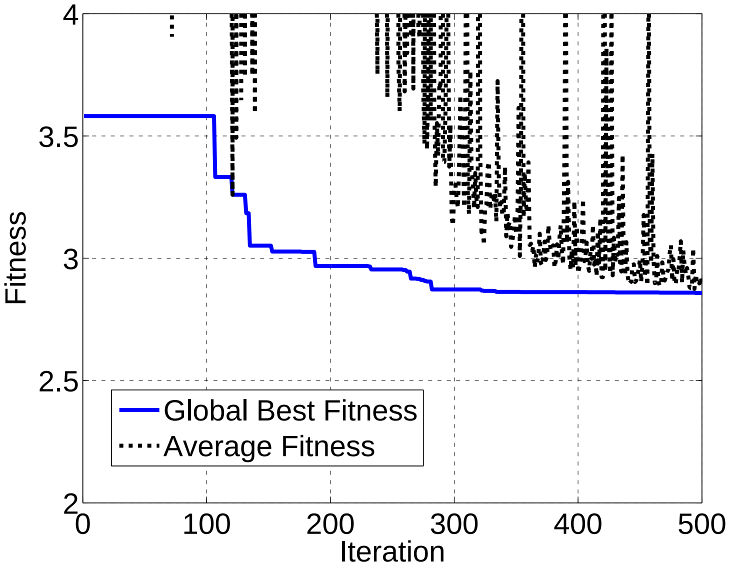


Figure 5.13: Convergence of the six-stepped septum

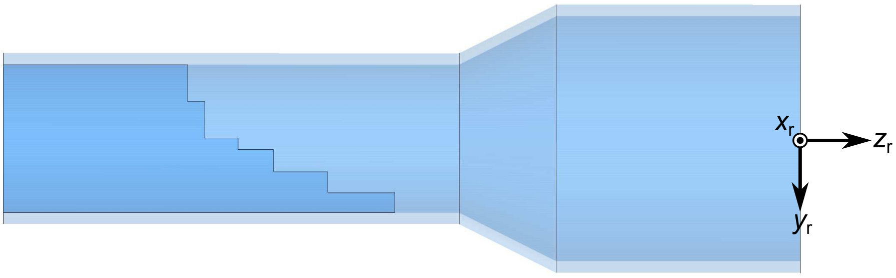


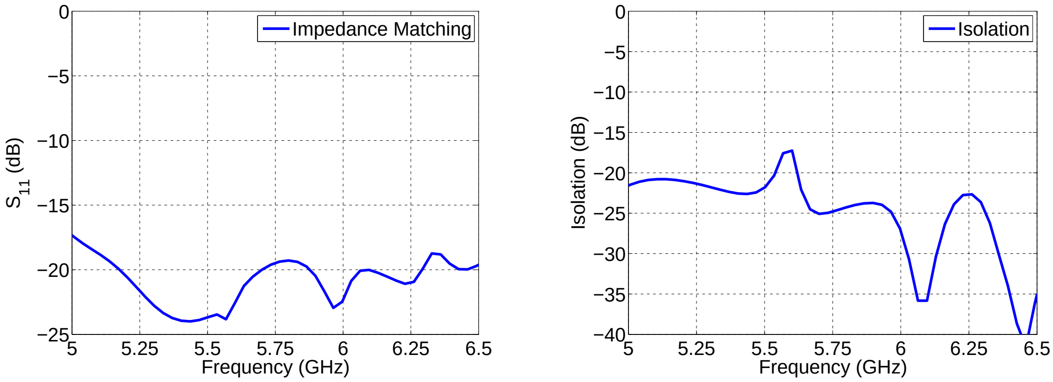
Figure 5.14: Scaled drawing of the final optimized six-step septum design

Table 5.3: Final Design Values from the Stepped Septum Optimization

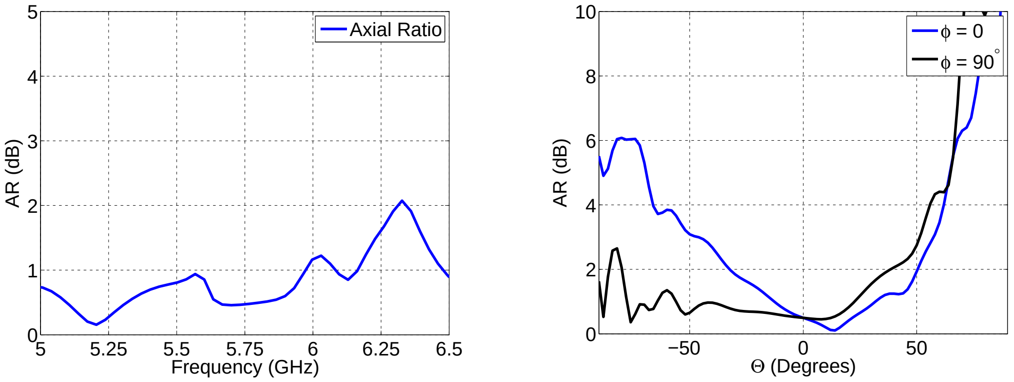
Parameter	Value (cm)	Parameter	Value (cm)
H_1	0.555	L_1	1.793
H_2	0.582	L_2	3.662
H_3	0.612	L_3	5.172
H_4	0.322	L_4	6.160
H_5	1.015	L_5	7.096
H_6	1.026	L_6	7.567

This translates into a value of $S_{11} = -27.79$ dB and AR $= 1.966$ dB. Note that the average fitness settles down to a value near the global optimum, which can demonstrate that the design maintains good performance despite any small changes to the parameters. It is also interesting to observe the spikes in the average fitness until the end of the optimization. This often happens for complex constrained optimization problems, where the constraints form a difficult landscape to navigate for the particles. If the particles are outside the boundary then they are assigned a high fitness, e.g. 10^{20}, and this explains why the spikes can rise to a very high number and then fall back down to the region plotted. The final values for the six step septum are shown in Table 5.3, and a scaled drawing of the final design is shown in Figure 5.14.

With the design provided from the optimization, the ultimate interest is in the overall antenna performance. The design shown in Table 5.3 was simulated, and the antenna performance is depicted in Figure 5.15. The simulated design has the geometry as shown in Figures 5.9 and 5.14, and the probe feed was used to simulate this final design performance shown in Figure 5.15 in order to demonstrate a practical implementation. The impedance matching is shown in Figure 5.15, and good impedance matching can be seen over the whole plot. Less than -15 dB is observed over the entire region, which is quite exceptional for many applications. Some applications may require an impedance matching less -20 dB for high performance systems, and a different fitness function with a smaller bandwidth may be applied in order to achieve this. The isolation also demonstrates good performance with levels less than -15 dB for the frequencies shown, even though isolation was not incorporated into fitness function. For communication



(a) Impedance Matching **(b)** Isolation



(c) Axial Ratio vs. Frequency **(d)** Axial Ratio vs. Angle at 5.8 GHz

Figure 5.15: Performance of the optimized stepped septum design

systems using the CP duplex configuration, this isolation may be further optimized by incorporating it into the fitness function. The last and possibly one of the most important performance aspects is the axial ratio, which is plotted in Figure 5.15c and 5.15d. The plot shown in Figure 5.15c depicts the AR performance in the boresight direction versus frequency. Good AR performance is often classified by $AR < 3$ dB, and this is also observed over the frequencies shown in the plot. Some applications may require better circularity with $AR < 0.5$ dB, and this can be achieved by optimizing over a narrow bandwidth. Overall, good circularity is observed over a wide bandwidth (roughly 27.5%). The AR is also plotted versus angle for 5.8 GHz in Figure 5.15d, and good performance can be observed over a wide angular region. The coordinate system used for plotting the AR versus angle is shown in Figure 5.14 as the coordinate system (x_r, y_r, z_r). The angles ϕ and θ are defined in the usual manner for spherical coordinate sys-

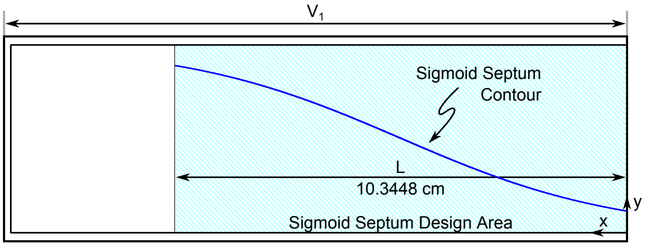


Figure 5.16: Scaled drawing of an oversized sigmoid function within a waveguide for $H_1 = H$, $L_1 = L/2$, and $C_1 = 10$

tems. The performance versus angle confirms the original assumption made that optimizing the boresight ($\theta = 0$) axial ratio would optimize the overall performance for a wide angular region.

5.3.2 Sigmoid Septum Optimization

Much of the formulation of the sigmoid septum optimization is identical to the stepped septum, and we will simply restate the similar qualities and provide an explanation for any new features. Since $C_i \neq 0$, many of the assumptions made in the prior formulation are no longer valid. When the contour is more sigmoidal, it has been observed that the location of the sigmoid septum will have some effects on the overall performance, especially since the sigmoidal curves are longer in length in comparison to the stepped septums in general. This led us to include L_1 in the optimization. With this the $\vec{L}$ vector is defined by $\vec{L} = [L_1, \ldots, L_P]$. The H vector remains the same since there are no changes made in the height designation.

The only parameter that remains to be discussed is the $\vec{C}$ vector, which provides the slope for each sigmoid function $S_i(x)$. In order to allow the optimizer to achieve the step function (if it has better fitness than the stepped function), the value of $C_i = 0$ must be attainable, and therefore this will be the lower limit on the slope constants. The best suggestion for the upper limit on C_i is not as clear as the lower limit, and therefore we make a recommendation based on some properties of the sigmoid function. Obviously, an upper limit of $C_i = \infty$ is not very practical, as it produces a simple constant valued function at 1/2. Even relatively low values of C_i can cause problems in the septum context. For example, take a sigmoid whose height is H (the total height of the waveguide), whose position is in the middle of the waveguide, and whose

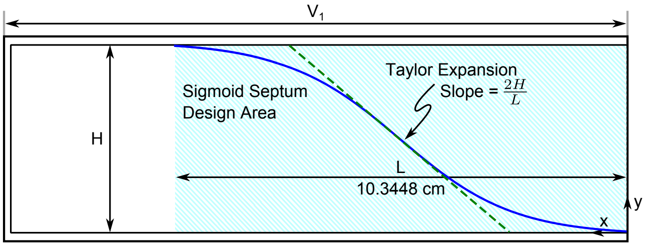


Figure 5.17: Scaled drawing of the worst case sigmoid allowable within the septum design area where $H_1 = H$, $L_1 = L/2$, and $C_1 = L/2H$

slope $C_i = 10$. This sigmoid is shown within the waveguide in Figure 5.16. We can see that the sigmoid is oversized for these dimensions, and it does not form a smooth connection to the waveguide wall on both the right and left sides. These types of designs are undesirable, but this aids in developing an upper bound for the constants C_i.

In order to limit the optimizer from exploring undesirable designs of this nature, an upper limit should be place on the slope constants for the sigmoidal functions $S_i(x)$. The approach we take in deriving this upper limit uses the slopes of the Taylor series expansion. The worst case scenario happens when the height of one sigmoid is equal to its maximum possible value $H_i = H_{max} = H$. This case requires a longer distance to approach a given threshold Δf which tells the commercial electromagnetic solvers that the interface is flush with the waveguide wall. In order to guarantee that a centered sigmoid would be flush to the waveguide walls, the length was set to half of the septum design area ($L/2$). This would allow the sigmoid function to reduce to $S(0) = 0.02H$ at the right and $S(L) = 0.98H$ on the left, which can be easily smoothened out to make the contour flush with the waveguide. Therefore, the upper limit on C_i was given as

$$C_i \leq \frac{L}{2H} \tag{5.18}$$

which is depicted in Figure 5.17. It is observed that the sigmoid is not guaranteed to always be flush with the waveguide walls, and the upper limit also does account for the cases where the sigmoid is located near the edge. In this case there still might be an undesirable discontinuity at the ends of the sigmoid septum design area. These cases are avoided using other explicit constraints designed to ensure a certain tolerance at the edges of the septum design area. At this point we have fully defined the solution space $\mathcal{S}$, and we have most of the components in the optimization problem.

This optimization uses the same fitness function in order to minimize the impedance matching and axial ratio of the overall design. The sigmoidal septum contour optimization problem can then be formulated as

$$\begin{aligned}
&\text{Minimize } f(\vec{x}), \quad \vec{x} \in \mathcal{S} \\
&\text{Given } H_P \leq 4.04\text{cm} && (5.19)\\
&\qquad H_P \geq 0 && (5.20)\\
&\qquad L_i \geq L_{i-1}, \quad \forall i \in 2, \ldots, P && (5.21)\\
&\qquad S(0) < 0.005\text{cm} && (5.22)\\
&S(L) - H > 0.005\text{cm} && (5.23)
\end{aligned}$$

where the first three constraints have been discussed in the stepped septum section and the last two constraints ensure that the tolerance between the waveguide wall and the septum edges form a complete divider between the bottom and top half of the septum. Otherwise gaps and discontinuities will ensue in the design, which are undesirable. To reiterate, the solution space $\mathcal{S}$ for this problem is given by

$$H_i = [0, 4.04\text{cm}], \quad \forall i \in 1, \ldots, P-1 \tag{5.24}$$

$$L_i = [0, 10.3448\text{cm}], \quad \forall i \in 1, \ldots, P \tag{5.25}$$

$$C_i = [0, 1.25], \quad \forall i \in 1, \ldots, P \tag{5.26}$$

where the upper limit on the C_i is given by $L/2H \approx 1.25$. The fitness function given in equation 5.19 is the same fitness function seen in equation 5.13, where the impedance matching and the axial ratio are being minimized over a similar band to the stepped septum. Again, the PSO algorithm is applied to this optimization problem, and its intrinsic parameters used in the run

Table 5.4: Final Design Values from the Stepped Septum Optimization

Parameter	Value (cm)	Parameter	Value (cm)	Parameter	Value
H_1	0.631	L_1	0.916	C_1	0.1044
H_2	0.546	L_2	3.139	C_2	0.5295
H_3	0.448	L_3	4.798	C_3	1.1451
H_4	0.343	L_4	6.074	C_4	0.4353
H_5	0.103	L_5	7.558	C_5	1.0937
H_6	1.969	L_6	7.919	C_6	0.4681

are shown in Table 5.5. This feasible space $\mathcal{F}$ is even more difficult to initialize and navigate the particles than the stepped septum optimization. Therefore the initialization scheme used in [91] was implemented in this optimization. The convergence results are shown in Figure 5.18. For the six sigmoid septum, there are 17 design variables to be optimized, a swarm size of 33 particles was chosen for the optimization. This represents a highly dimensional optimization problem which can take a long time to reach convergence due to the large swarm size. Since each iteration takes quite some time to evaluate, a termination criterion was implemented which stopped the optimizer after stagnation was observed. The plotted convergence shows the first 400 iterations of this run. Note also that the average convergence did not meet the global best curve because the feasible solution space $\mathcal{F}$ was difficult to stay within. The particles were assigned a high fitness value if they were outside of the feasible region $\mathcal{F}$. Another reason why the average fitness curve did not meet the global best curve is due to the commonly high fitness value of several outliers in the swarm. Even though these outliers would be located within $\mathcal{F}$, their fitness would be extremely high because the VSWR and AR can take on very high values (up to ∞), which would skew the average fitness away from the global best. This is another cause that did not allow the average fitness to settle towards the global best at the 400th iteration.

The final optimized design is shown in Figure 5.19. It is interesting to see that a mixture of both sharp and smooth sigmoids are present on the optimized contour. This may suggest that sharp edges are necessary to meet impedance matching and axial ratio requirements. Lastly, the six-sigmoid optimization converted the six-sigmoid into a five-sigmoid septum, as can be seen

Table 5.5: PSO Intrinsic Parameters for the Sigmoid Septum Optimization using 6 Sigmoids

PSO Parameter	*Values Used*
c_1	2.0
c_2	2.0
Swarm Size	33
Δt	1.0
Max Iterations (i_{max})	500
Boundary Condition	Invisible
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

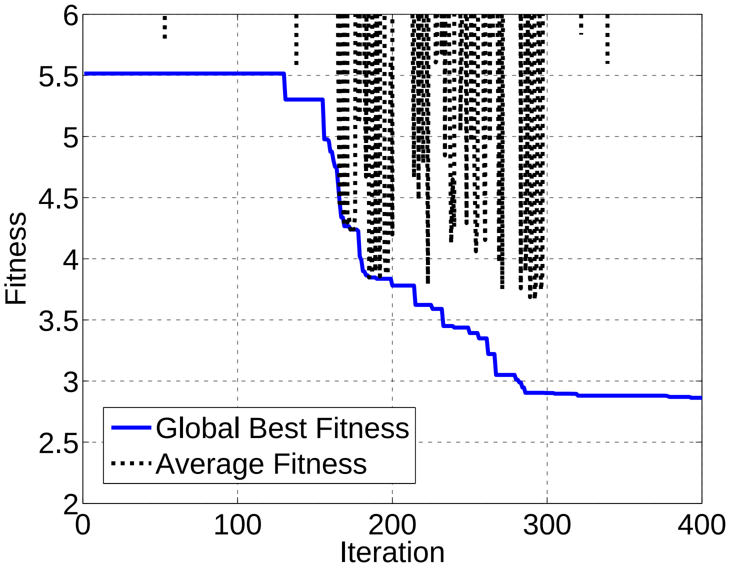


Figure 5.18: Convergence of the six-sigmoid septum

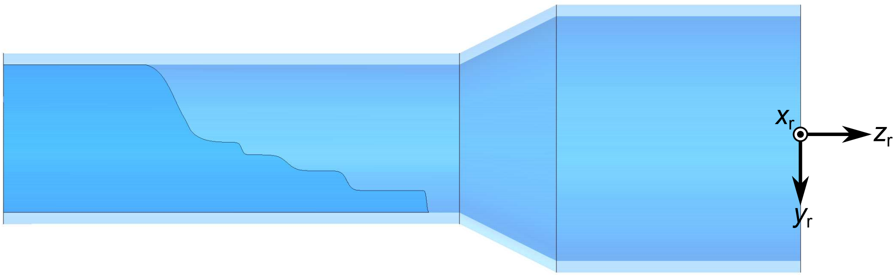
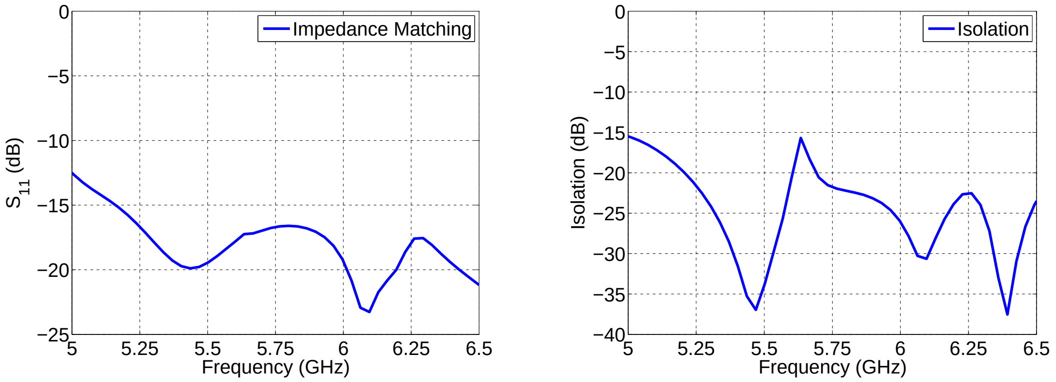
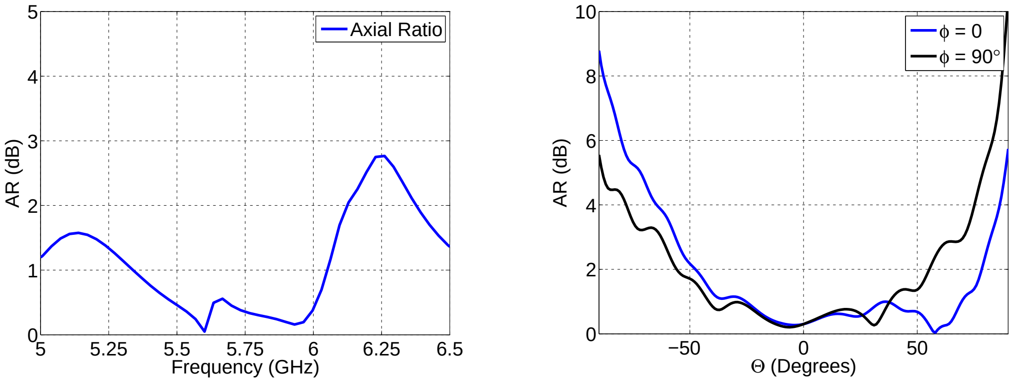


Figure 5.19: Scaled drawing of the final optimized six-sigmoid septum design



(a) Impedance Matching

(b) Isolation



(c) Axial Ratio vs. Frequency

(d) Axial Ratio vs. Angle at 5.8 GHz

Figure 5.20: Performance of the optimized sigmoid septum design

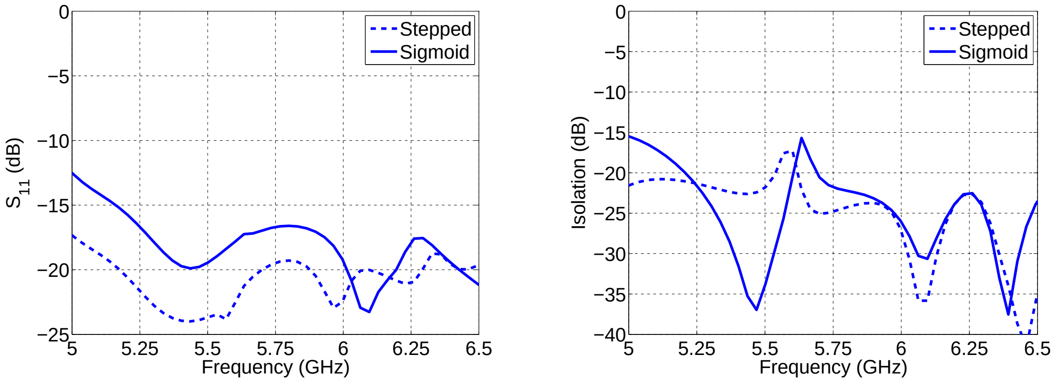
by counting the number of sigmoids in Figure 5.19. It was able to achieve this by placing both sigmoids close to one another, and the leftmost sigmoid in Figure 5.19 is actually two sigmoids placed close together. Overall, there are many interesting features seen in the results of this optimization.

The antenna performance for the final sigmoid septum shown in Table 5.4 is given in Figure 5.20. The sigmoid septum design provides a good impedance match as shown in Figure 5.20a, with $S_{11} < -15$ dB achieved for most of the band. The final design also incorporates good isolation performance with the isolation less than -15 dB over the entire plot. Again, if better performance is required then a smaller bandwidth should be chosen to optimize. The AR for the sigmoid design is observed to be less than 3 dB over the entire band shown. Note that the sigmoid septum can even provide less than 0.5 dB AR for the region 5.5-6.0 GHz, which is roughly 8% bandwidth at 5.75 GHz. This is a decently wide bandwidth with excellent performance for CP applications. The angular data for the AR performance is also shown for 5.8 GHz in Figure 5.20d, where a fairly large angular region within ±60° demonstrates good circularity. The two curves correspond to the principal cuts using the (x_r, y_r, z_r) coordinate system shown in Figure 5.19.

In this section, a sigmoid septum was introduced for possible use in high power microwave systems, and PSO was applied to discover whether good performance could be realized with such a design. It also helped in answering the question on whether sharp edges are required for good septum performance. The sigmoid function is able to realize the step function, and PSO had the option to use a stepped design versus a smooth curve design. The final design shown in Figure 5.19 showed both smooth and sharp curves. Since the sigmoid design shows good performance with a mixture of sharp edges and smooth edges, one can conclude that sharp edges are not mandatory for good performance. However, more investigation is needed in order to find whether there exist any distinct advantages of the smooth surfaces in terms of antenna performance.

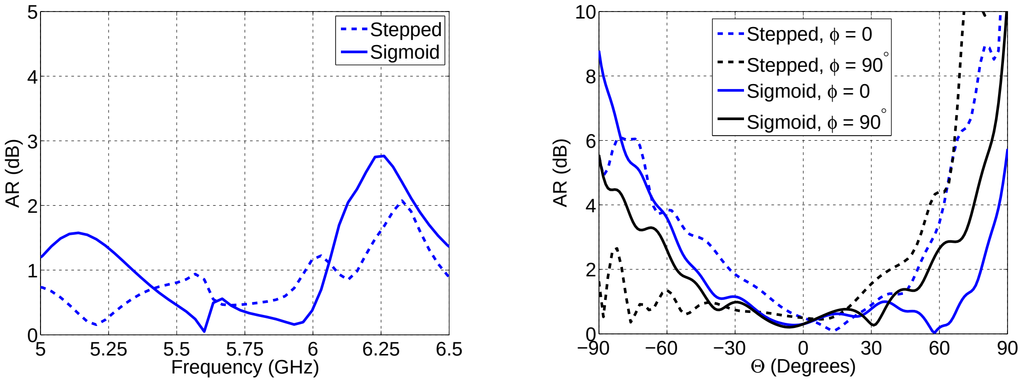
5.4 Final Design Comparison

As a final investigation, the two final designs from the optimizations are compared in terms of their respective antenna performance. The subfigures in Figure 5.21 provide a side by side comparison of the performance of both designs. First, it can be noted that the sigmoid septum has a slightly increased impedance matching in comparison to the stepped septum. It is possible that either the smoothed edges or the slanted contours are responsible in this increase, and a further investigation into the sigmoid septum is needed in order to determine the reasons why these are experienced. The isolation performance is quite comparable between both designs, and it is interesting to observe many similarities that both curves share. As seen in Figure 5.21, the sigmoid septum provides better axial ratio than the stepped septum for the given region near 5.75 GHz. However, the sigmoid septum has a large spike in AR for higher frequencies which may be undesirable for applications requiring a broad bandwidth with a high degree of circularity.



(a) Impedance Matching

(b) Isolation



(c) Axial Ratio vs. Frequency

(d) Axial Ratio vs. Angle at 5.8 GHz

Figure 5.21: Comparison of the performance of the optimized stepped septum design and the optimized sigmoid septum design

Both designs show a similar angular performance, although the stepped septum has a smaller angular region where $AR < 3$ dB. This may be due to the fact that the AR towards boresight ($\theta = 0$) is larger than the sigmoid septum. Overall, the performance of both the sigmoid and the stepped septum designs are quite comparable with some slight differences in the S_{11} and the AR.

The septum is able to provide a high quality of circularity with a simple passive structure that is fairly easy to fabricate. The stepped septum and the sigmoid septum were both introduced as possible candidates for the creation of CP waves from a linear excitation in feed horn applications. To the author's best knowledge, this is the first time that a stepped septum has been optimized within a full circular horn apparatus using PSO, and a final design has been obtained directly from the optimization without the need for any further design. The sigmoid septum was investigated as a new septum candidate that might be used in high power applications. The sigmoid septum was also optimized with PSO in a similar manner as the stepped septum, although the sigmoid septum posed a much more difficult optimization problem with higher dimensionality. This preliminary investigation focused on testing whether the sigmoid septum could provide comparable or better performance in comparison to the stepped septum. As shown in the previous discussion, the final sigmoid septum offers a comparable antenna performance, however there are no overwhelming antenna performance advantages of the sigmoid septum over the stepped septum. Since the sigmoid septum may be more difficult to fabricate, it might not be the most suitable for the typical wireless communications application. However, more investigation is required in order to determine whether better designs might exist using this platform for different applications (and hence different fitness functions). Further investigation will also be taken in quantifying its power handling advantage over the stepped septums, which is one distinct advantage of the sigmoid septum.

CHAPTER 6

MEMS Reconfigurable E-Shaped Patch Antennas

The wireless communication environment often suffers from adverse effects such as multipath, interference, and fading. Each of these effects can be detrimental on the overall system and are difficult to combat. One methodology that has become popular among the wireless community is the use of multiple antennas in order to increase the so-called antenna diversity. By introducing multiple antennas with different characteristics, the probability that one may work properly is increased in a general sense. Of course, the inner workings of these antenna systems are much more complex than this simple analogy. These multiple antenna systems may give rise to a larger system altogether, which can be difficult to manage. The added size of the antenna system can also be quite cumbersome, and therefore this might not always be the best approach to alleviating some of these issues.

Although the multiple antenna technique has found widespread acceptance throughout many research communities, new methodologies have begun to gain popularity within the antenna engineering community. Another technique to overcome some of the difficulties encountered in wireless systems is through the use of reconfigurable antennas. The definition of a reconfigurable antenna is given from the following quote from J. T. Bernhard.

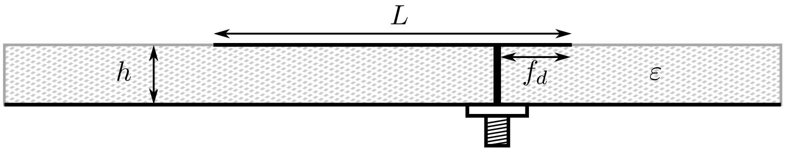
> *Reconfigurability, when used in the context of antennas, is the capacity to change an individual radiator's fundamental operating characteristics through electrical, mechanical, or other means. Thus, under this definition, the traditional phasing of signals between elements in an array to achieve beam forming and beam steering does not make the antenna 'reconfigurable' because the antenna's basic operating characteristics remain unchanged in this case*
>
> J. T. Bernhard, *Reconfigurable Antennas [128]*

These reconfigurable antennas can often be placed in three primary categories: frequency reconfigurable, polarization reconfigurable, and radiation pattern reconfigurable [128]. Reconfigurable antennas can overcome challenges inherent in multiple antenna scenarios by replacing all antennas with different functionalities with one antenna having the same functionality. The use of reconfigurable antennas alleviates the antenna's isolation and space requirements. While diversity is not maintained, these reconfigurable antennas can alter their characteristics in order to make the system more robust. For instance, a frequency reconfigurable antenna would be able to combat frequency selective fading [129] by changing its operational frequency. It might also reduce the interference from a given angle by shifting a null in the radiation pattern towards the direction of the interferer. There exist a wide variety of applications where reconfigurable antennas might be used to make a system more robust.

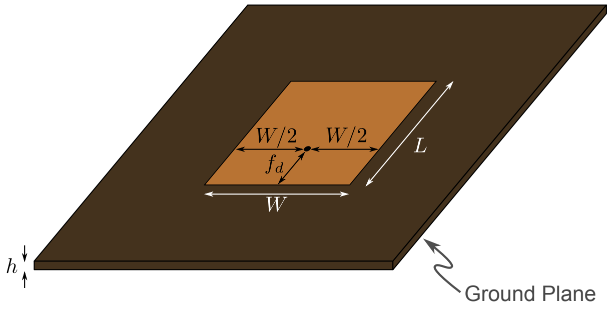
This chapter will examine some possible wideband reconfigurable antennas using a special patch antenna commonly known as the E-shaped patch antenna [130]. An introduction to the E-shaped patch antenna will be given as well as its performance in comparison to some of its competitors. Next, a MEMS switch model and its implementation will be discussed in order to add reconfigurability to the E-shaped patch antenna. Then two possible reconfigurable E-shaped patch antennas will be discussed, namely a polarization reconfigurable E-shaped patch antenna and a frequency reconfigurable E-shaped patch antenna. Lastly, the bias network for the switches must be implemented into the E-shaped design, and the chapter will be concluded by demonstrating the bias networks for a full implementation of the reconfigurable E-shaped patch antenna.

6.1 Introduction to E-Shaped Patch Antennas

Patch antennas have a widespread use due to their ease of implementation and manufacturing. Their low profile allows them to be versatile for use in a wide range of applications. Their design also fits into a planar microstrip topology, which makes it straightforward to integrate into a microstrip circuit. In the array context, simple feeding structures can be easily created using microstrip lines to feed each array element. Some common applications for their use can be found in mobile handsets, GPS receivers, conformal antennas, as well as low profile, high gain antenna



(a) Side View



(b) Perspective View

Figure 6.1: Demonstration of the simplicity of a coaxial probe-fed rectangular patch

arrays. A common rectangular patch antenna fed with a coaxial probe is shown in Figure 6.1 to demonstrate their simplicity and low profile.

Some disadvantages of patch antennas include a low bandwidth and high cross polarization. However, these can be alleviated with the proper redirection of the patch currents as well as other clever techniques. The bandwidth can discussed in the context of the typical input circuit representation of the patch antenna [131] shown in Figure 6.2, where L_p is the probe feed inductance and R, L, and C form the tank circuit representing the cavity model of the patch antenna. In general, the bandwidth of patch antennas can be increased by increasing the substrate height, but eventually a point will be reached where the probe inductance L_p can cause an impedance mismatch for the circuit. There are several techniques to avoid or compensate for this inductance. A stacked patch antenna configuration [132], which was discussed in Chapter 4, avoids this by using a driven patch and a coupled patch, thereby using a shorter probe feed. By shortening the probe feed, a smaller inductance is seen at the input port. The upper patch couples to the driven patch in an effort to provide another resonant frequency, which can extend the bandwidth. While this design avoids the probe inductance by using another layer, a

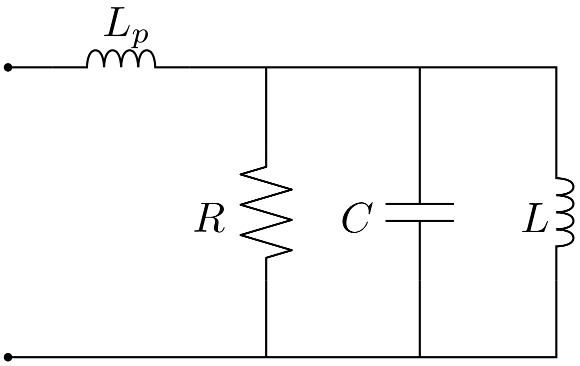


Figure 6.2: Circuit model of the patch input port

multilayered antenna can have its own challenges in manufacturing and alignment.

Another technique to overcome the probe inductance is through the use of an L-shaped probe [133]. The L-shaped probe-fed antenna antenna utilizes an L-shaped probe that is capacitively coupled to the patch and overcomes the probe inductance by introducing some series capacitance seen at the input port. This design is able to get roughly 27% to 35% bandwidth with fairly high gain, which is comparable to other antennas in its class. However, this patch antenna can be somewhat tricky to fabricate and requires either a three dimensional manufacturing approach or a multilayered approach in order to construct it. This increases the cost and complexity for manufacturing, making it a less attractive candidate. This design also does not pose any straightforward possibilities for implementation as a simple reconfigurable antenna.

Another clever technique that has gained wide attention in the antenna community is the use of slots to extend the impedance bandwidth. The most popular are the U-slotted antenna [134, 135] and the E-shaped patch antenna [130, 136]. The inclusion of these slots can provide some compensation for the probe inductance in order to improve matching. These slots also introduce another mode by redirecting the currents to realize a new lower resonant frequency. Both of these effects combined can provide an incredibly wide bandwidth for a patch antenna. Recently, a thorough investigation meticulously compared the performance of both the E-shaped patch and the U-slotted patch in terms of their bandwidth, cross polarization, and overall size [137]. With this study and others, it has been shown that approximately 24% up to 45%

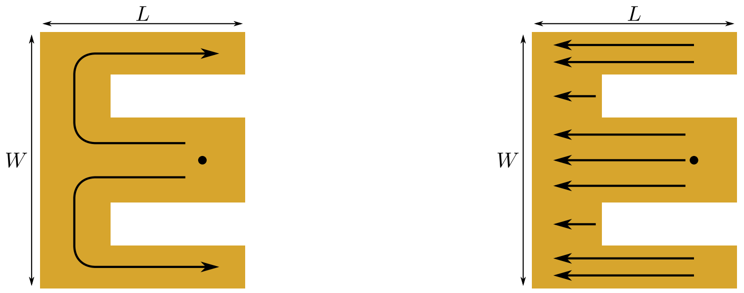
bandwidth can be achieved with the U-slotted patch antenna. The E-shaped patch antenna is able to achieve a comparable bandwidth with reports showing anywhere from 25% to 34%. These antennas are also easy to manufacture due to their single layer and single feed designs. In comparison to other alternatives, these two antenna types can be easily designed and built for many applications. Both topologies are drawn in Figure 6.3, where the U-slot and the E-shape can be clearly seen from the patch antenna shapes.

(a) U-Slotted Patch **(b)** E-Shaped Patch

Figure 6.3: Top view of the U-slotted and the E-shaped patch antenna topologies

The principle behind their wideband nature is similar, where the currents are manipulated in such a way that two different frequencies resonate within the device. For the upper frequency mode, also called the patch mode, the resonant frequency is guided by the length of the patch. The currents for this mode on the E-shaped patch are shown in Figure 6.4b. This is the typical current distribution on a rectangular patch without the slots. The slots introduce another mode, the E-shaped mode, which forces the currents to meander around the slots. This current distribution is depicted in Figure 6.4a, where the resonant length is guided by the patch length and the dimensions of the slot. Both the E-shaped and U-slot antenna operate with the same principle of adding a new resonant mode.

This investigation will examine possible implementations of these slotted antennas as reconfigurable antennas. Reconfigurability may be possible by changing the slot dimensions, since these dimensions can have an appreciable effect on the resonant frequency. As will be discussed in Sections 6.3 and 6.4, reconfigurability can be achieved by changing the effective dimensions of the slots. This can be accomplished by shorting the slot with a switch. Both Figure 6.3 and



(a) Lower Frequency Mode **(b)** Higher Frequency Mode

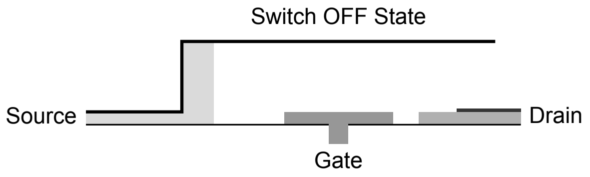
Figure 6.4: Depiction of the possible resonant modes that can occur in the E-shaped patch antenna

6.4 show that the E-shaped patch has clear access to both of the slots. This allows for a simple implementation of the bias network required to control the switches. The U-slotted patch has slots which are confined within the patch and can make it difficult to connect any sort of bias line on the same layer as the patch antenna. Therefore, the E-shaped patch antenna provides a more convenient design for reconfigurability, and only this antenna will be further investigated.

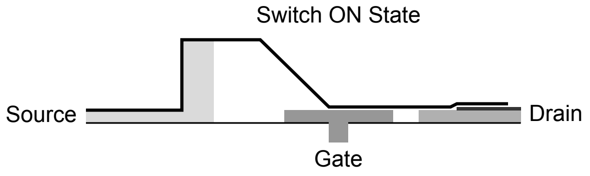
6.2 Development of a Simple MEMS Model

The basic E-shaped patch antenna topology has been shown, and the dimensions will be provided in Sections 6.3 and 6.4. The next step is to characterize the switch which enables the antenna reconfigurability. Currently, the most popular devices in microwave switching technology include varactors, PIN diodes, and MEMS (Microelectromechanical Systems) switches, and these devices are frequently implemented in reconfigurable antennas. Among these three different types, MEMS switches have gained more popularity among the antenna and microwave communities due to their low insertion loss, good impedance matching, and high isolation. MEMS are also potentially able to avoid nonlinearities and intermodulation effects in comparison to the other switches [138]. Therefore, MEMS switches were chosen for this design in order to minimize these losses, which can decrease the antenna's dielectric-conductor efficiency discussed in Section 1.3.

A MEMS switch is a three-terminal device which can act as a switch operating from DC



(a) OFF State



(b) ON State

Figure 6.5: Simple illustration of a MEMS switch

to microwave frequencies (up to 20GHz). The three terminals are often denoted as the Source, Gate, and Drain. Once a voltage (higher than the threshold voltage) is applied between the gate and the source, the static electrical force can bring down the cantilever beam to connect the Source and Drain. This represents the ON state for the MEMS switch. The threshold voltage for typical commercial MEMS switches is usually around 90 volts. When the voltage is well below the threshold then high isolation between the Source and Drain can be achieved, which represents the OFF state. The MEMS switch that was used in the reconfigurable E-shaped patch antenna design was a Radant MEMS *RMSW100HP* switch [139].

With the type of switch chosen, the next task is to chose a method to incorporate its presence in the antenna for optimization purposes. In the optimization run, the antenna simulation can be critical to obtaining a working design in a reasonable amount of time. One should always be cognizant on the simulation time for a particular structure. If one simulation takes up to 1 hour, then the total optimization may take several thousand hours, which is quite unreasonable in a practical setting. Therefore, one must consider any alternatives which provide enough complexity to properly model the switch but are computationally inexpensive. There are several options when modeling the MEMS switches, which can be categorized into the following list.

Ideal Switch Model

The ideal switch model assumes that the MEMS switch acts as a simple ON/OFF switch, with no internal effects accounted for. With this model, the OFF case is represented by an open circuit, and the ON case is represented by a short circuit (or a PEC connection).

Simple Circuit Model

With the circuit model of the MEMS switches, the OFF case is simulated with a capacitive connection between the two nodes. The ON case is then simulated as a resistive connection between the two terminals due to the contact resistance of the cantilever beam. More complicated circuit models exist to account for other features of the MEMS switch, but these can be difficult to model within electromagnetics simulators.

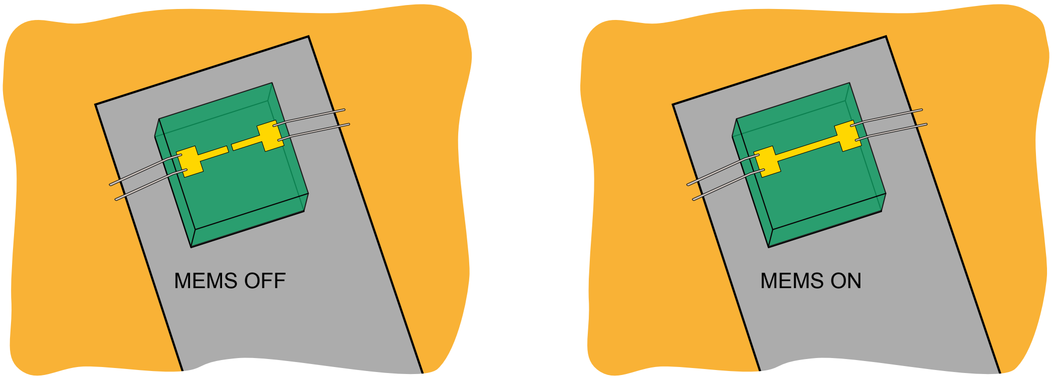
MEMS Full Model

One can simulate the full electromagnetic model of the MEMS switch by incorporating all of its geometrical features. This includes a silicon substrate as well as some representative transmission lines. In the OFF position, the transmission lines form an open connection. In the ON position the transmission lines are connected metallically. This full model is based on the model shown in [140], where parametric studies were used to match the simulation with measurements.

Each of these models are depicted in Figure 6.6. The level of complexity in the MEMS model increases from the top to the bottom. The ideal switch represents the easiest implementation of the MEMS switch and would most likely have the fastest simulation time. The circuit model involves slightly more complexity in its representation, but it does not represent a drastic increase in computational time. The only disadvantage of this model is that the internal parameters C and R are not easily measured, and any shift in these values might cause drastic changes in performance. This depends on the input port's sensitivity to any changes in capacitance or resistance which comes internal to the MEMS switch. The last possible implementation of the MEMS switch for simulation purposes is to use the full MEMS model, shown in Figure 6.6c. The geometry of the MEMS switch is given in [140], and this design has demonstrated good agreement with measurements in the past. One major hindrance in analyzing a MEMS reconfigurable antenna with the full MEMS switch model is that each simulation requires nearly 1-2 hours to provide all necessary information needed to evaluate its fitness. This is unacceptable for optimizations which typically analyze over 4000 simulations, and this precludes any use of the full MEMS model for optimization purposes. Therefore a decision must be made between the ideal switch model and the simple circuit model.

(a) Ideal Switch Model

(b) Simple Circuit Model



(c) Full MEMS Model

Figure 6.6: Several possible implementations of the MEMS model within a full wave electromagnetics simulator

An investigation was conducted on the accuracy of the ideal switch model in comparison to a measurement with MEMS switches. More discrepancies are expected for the OFF case in comparison to the ON case. This owes to the low internal resistance values typically seen for the MEMS switches, which is on the order of $R = 0.7\Omega$ [141]. Since the resistance is fairly close to the ideal case of a short circuit, the discrepancies between the ideal switch and the actual MEMS model should be minimal. However, the OFF state provides more opportunities for disagreement, and therefore only the OFF state will be considered in this investigation to properly model the MEMS switches.

Several design optimizations were conducted to test whether a simple ideal switch model provided the necessary amount of information to properly model all characteristics of the MEMS switch in the OFF state. The details of the optimizations will be excluded in order to keep the focus on the MEMS switch model. A preliminary optimization which used a similar strategy described in Section 6.4 implemented the ideal switch model shown in Figure 6.6a into the frequency reconfigurable E-shaped patch antenna. Once the design was optimized to provide a dual frequency band performance, a measurement of the ideal case was used to verify the ideal switch model worked using ideal switch measurements. Ideal switches were implemented by applying copper tape to effectively short the bars in the E-shaped slot for the ON state. The OFF state was measured using an ideal switch with an open circuit, as shown in Figure 6.8b. The S_{11} response is shown in Figure 6.7 which compares the simulation and measurement. Both demonstrate good agreement, which implies that the optimization found a good design using ideal switches. This is not necessarily the case when MEMS switches are applied to the same design.

The next step aims to determine whether any discrepancies exist between the ideal switch model and the measured MEMS switch performance for the OFF case. In order to make this comparison, MEMS switches are then implemented into the same design seen in Figures 6.7 and 6.8. The MEMS switches were glued to the pads and then the source and drain terminals were wirebonded to the E-shaped patch structure as shown in Figure 6.10. Since bias lines had not been developed at those stages, the gate was not wirebonded, and this should not cause any detrimental effects because MEMS provide a high isolation between the gate and the

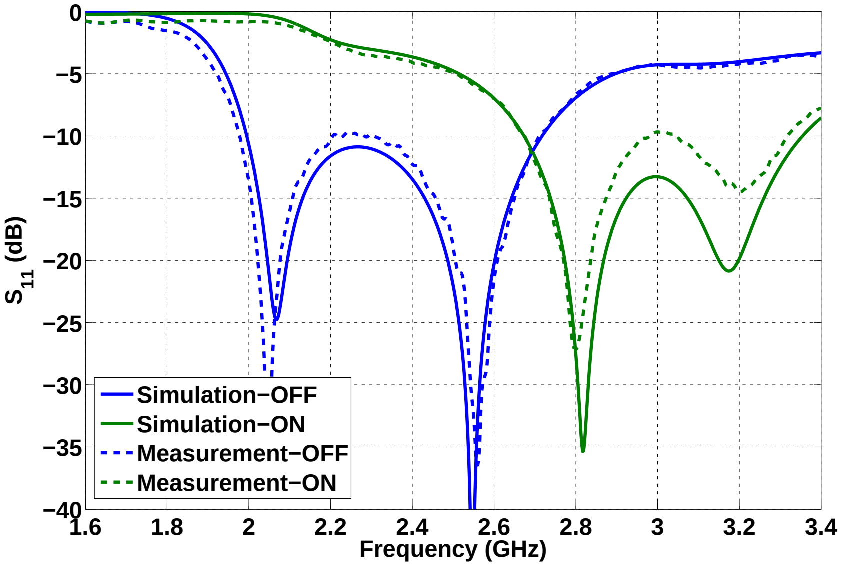
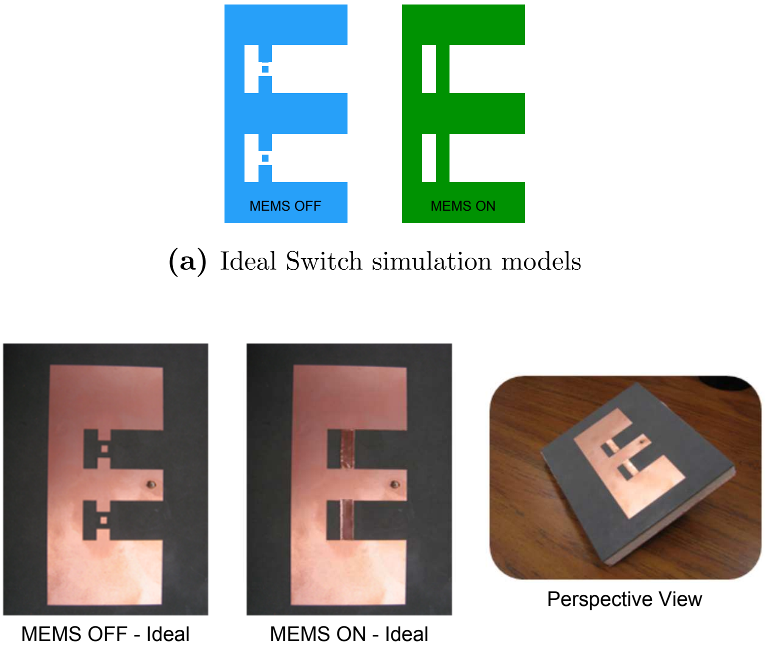


Figure 6.7: S_{11} comparison of the ideal case between simulation and measurement



(a) Ideal Switch simulation models

(b) Ideal Switch fabricated designs

Figure 6.8: Fabricated antennas and simulation topology to measure the ideal switch case

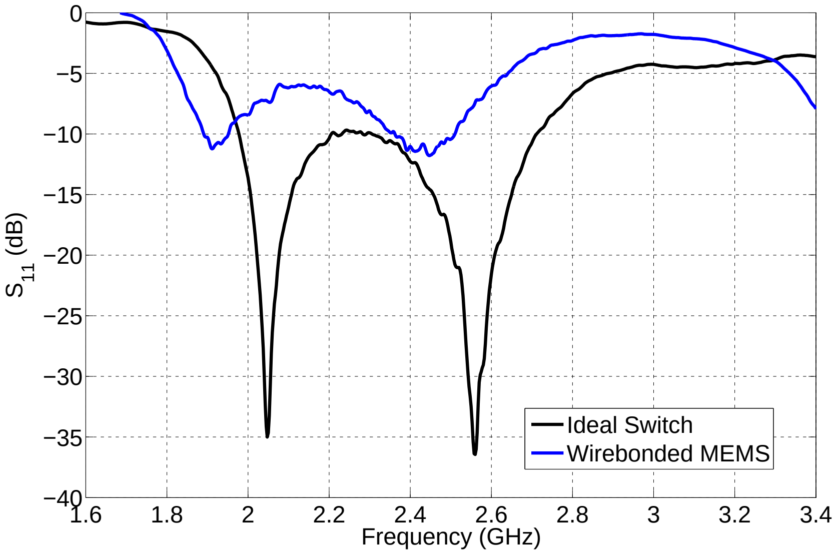


Figure 6.9: S_{11} comparison of the ideal switch measurement versus the wirebonded MEMS implementation

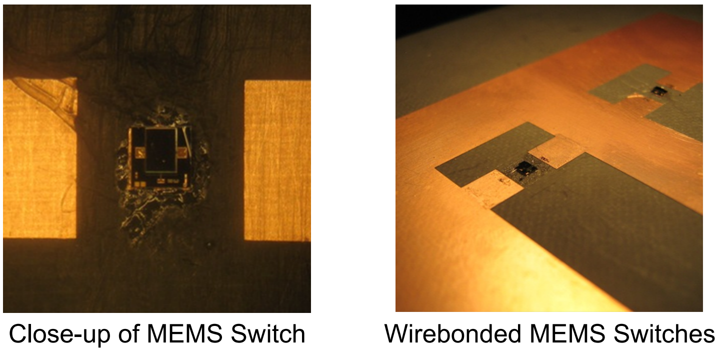

Close-up of MEMS Switch
Wirebonded MEMS Switches

Figure 6.10: Implementation of the MEMS switches into the E-shaped patch antenna optimized with the ideal switch model

drain/source. The S_{11} was then measured with the wirebonded MEMS implementation shown in Figure 6.10, and the performance is shown in comparison to the ideal switch case with no MEMS switches in Figure 6.9. From this figure, the performance of the design with the MEMS switches has drastic changes in comparison to that of the ideal switch case.

Obviously, this poor agreement between the two models indicates that the ideal switch model would be inadequate for optimization purposes. In order to arrive at a good practical design, it is imperative that the model for a reconfigurable antenna using MEMS switches predicts its performance to an acceptable degree. These results prompted further investigation in order to account for these disagreements and ultimately provide a better model. We approached this by simulating slightly more complicated circuit models to test whether they provide better agreement with the measurement. A simple model for the OFF case includes both an inductor to account for the wirebonds and a capacitor to account for any parasitic coupling between the drain and the source [141]. Since HFSS, the electromagnetic software used for simulation purposes, was able to model wirebonds, the model that was tested used both wirebonds and a capacitive impedance to model the internal capacitance of the MEMS switch. This is shown in Figure 6.12b, where the wirebonds connect to an impedance surface replicating a capacitor. This model showed excellent agreement with the simulation using the full MEMS model and the measurement of the final design with MEMS switches. The comparison of the measurement with these two simulations are shown in Figure 6.11, where most of the features are demonstrated by both the full MEMS model and the circuit model.

The results from these measurements brought about several observations for modeling the MEMS switches. First, the internal capacitance can alter the performance quite considerably, even with a small capacitance of 80 pF. This makes sense when comparing the impedance of this capacitor at these frequencies, which can be given by $|Z_C| = (\omega C)^{-1} \approx 800\Omega$ at 2.5 GHz, which may not be best represented as an open circuit. From this investigation, it was concluded that this internal capacitance was the primary cause for the major discrepancies in the OFF state. Second, any simulations must account for this capacitance in order to provide good agreement between simulation and measurement. Both the circuit model and the full MEMS model account for this internal capacitance in their models while the ideal switch model does not. Lastly, the

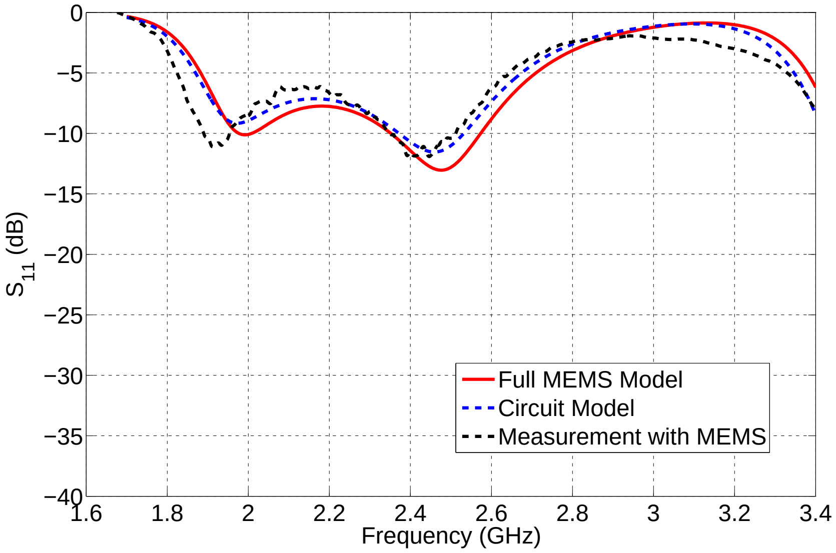
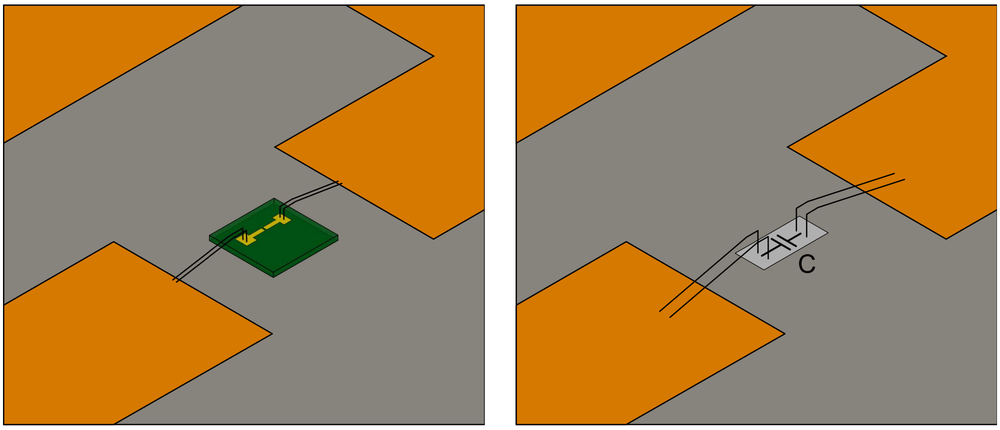


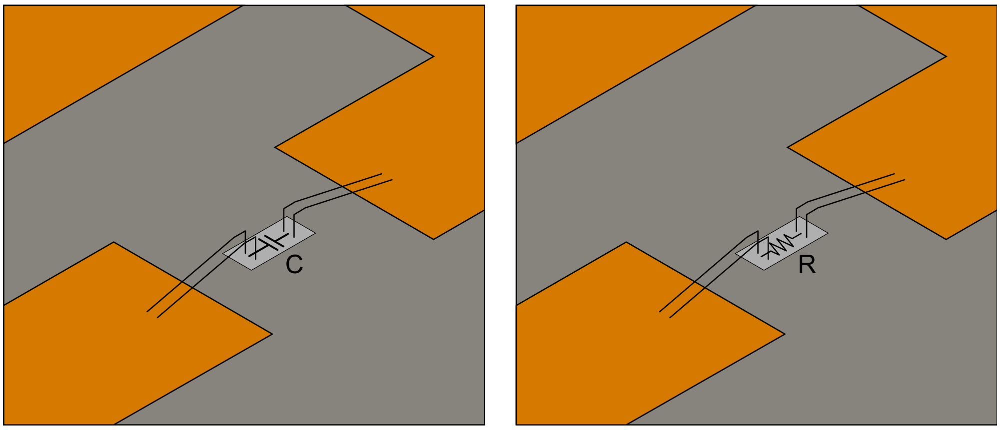
Figure 6.11: Comparison of the S_{11} from the wirebonded MEMS measurement with the full MEMS model and the circuit model using $C = 80\text{pF}$



(a) Full MEMS model

(b) Circuit model using wirebonds and a capacitor

Figure 6.12: Models for comparison to the measured wirebonded MEMS performance



(a) OFF State, $C = 80\text{pF}$ **(b)** ON State, $R = 0.7\Omega$

Figure 6.13: Circuit model used in the optimization for the reconfigurable E-shaped patch antenna

simple circuit model provides good agreement even though it represents a first order approach to simulating the MEMS switch. While more research needs to be accomplished in order to better characterize the internal effects of the MEMS switch, this model seems to provide the best of all worlds. It accurately predicts the performance of the antenna design while also being computationally efficient. This can make it more attractive for design purposes than the full MEMS model.

Overall, the experiments and results provide good insight into the best approach for simulating the reconfigurable E-shaped patch antenna using MEMS switches in the optimization. The final circuit design used in the optimization are shown in 6.13, where the ON state is represented by a small resistance and the OFF state is represented by a capacitor. With this investigation, the simulated performance of the reconfigurable E-shaped patch antenna should be well predicted by the simulator in the optimization process. In Sections 6.3 and 6.4, there will be two reconfigurable antennas that will be developed. The first is a RHCP/LHCP switchable antenna whose polarization can be altered. The next design shows the development of a frequency reconfigurable E-shaped patch antenna. Both of these designs will be optimized by the PSO algorithm.

6.3 Circular Polarization Reconfigurable E-Shaped Patch

With the proliferation of cognitive radio and wireless communications, there is a growing need to develop wideband circularly polarized antennas. Within the broad scope of wireless communications, circularly polarized systems can offer many advantages over linearly polarized systems. It has been shown in previous work that circularly polarized wireless systems gain improvement by mitigating multipath signals [142, 143]. These multipath signals change from right-handed to left-handed polarization (or vice-versa) after a reflection from a surface, and with this new polarization these signals are nearly unabsorbed by the receiving antenna. Circular polarization also adds polarization flexibility for receiving antennas which may be either linearly polarized or circularly polarized. While this may reduce polarization efficiency in the case of linear polarization, one can guarantee good signal strength with a low dependence on antenna orientation. The CP design strategy has been used in many existing systems including GPS and RFID.

One might be able to maximize polarization efficiency through the use of a circularly polarized receiving antenna, but if the wrong polarization is used then there is a risk for further reduced signal strength. System designers will typically prefer to avoid this due to a higher probability for poor signal quality. However, this problem can be completely removed with the development of a polarization reconfigurable antenna which can switch between RHCP and LHCP. With this added flexibility, one can ensure that the system can maintain a good signal quality in many various scenarios. This antenna is also useful in non line of sight (non-LOS) applications where the reflected wave may have an orthogonal polarization. By switching between CP polarizations, one can guarantee good polarization efficiency in the adverse environments seen in wireless systems.

6.3.1 Concept Design

Wideband circularly polarized patch antennas are often difficult to achieve, and for many designs the axial ratio (AR)-impedance bandwidth is only a few percent. There exist several techniques for single layer single feed designs achieving CP. One method involves placing the probe feed along the diagonal of a nearly square patch [131] or placing tabs on an elliptical patch [15].

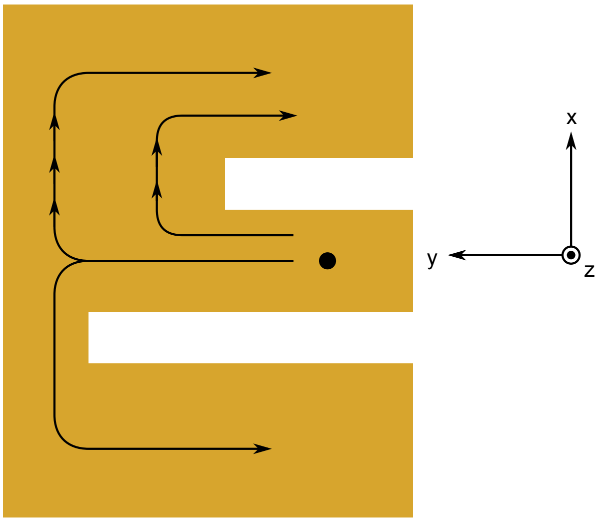


Figure 6.14: Asymmetric slots in the E-shaped patch antenna creates a wideband CP antenna

Other techniques cut opposite corners of a square patch and place the feed similarly to the linearly polarized case [15]. However, these techniques often produce a bandwidth less than 1%. Multifeed and multilayer configurations also exist, but these designs require more complicated fabrication techniques [15] and are still limited by the impedance bandwidth. Newly proposed techniques using slotted patches antennas are a promising alternative, and wide bandwidths have been recently obtained [144, 145]. Recent research in microstrip patch antenna technology has shown that wideband circular polarization is achievable with the E-shaped patch antenna platform. This can be accomplished by using asymmetric slots in the E-shape, as shown in Figure 6.14. In the standard E-shaped patch antenna design, the patch has primarily y-directed currents, which implies a linearly polarized antenna. As discussed in [145], these asymmetric slots allow the creation of x-directed currents of nearly equal magnitude to the y-directed currents. With the proper geometry, one can obtain x-directed currents with both equal magnitude and quadrature phase. However, finding the best geometry is quite challenging and non-intuitive.

Clearly, creating an asymmetry in the slots is required in order to obtain good AR and impedance matching performance. This design, however, only creates one polarization, which is either RHCP or LHCP. Theoretically, reconfigurability can be achieved if the slot asymmetry could be flipped dynamically. In other words, one could achieve polarization reconfigurability (RHCP/LHCP) by changing the effective slot dimensions in real time. Our proposed design

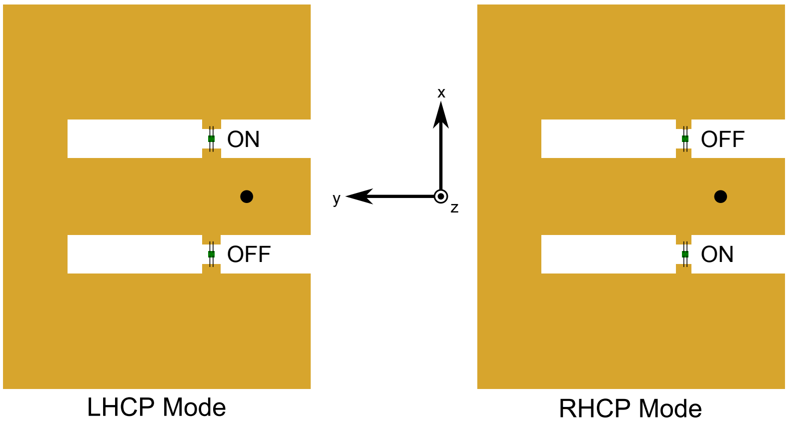
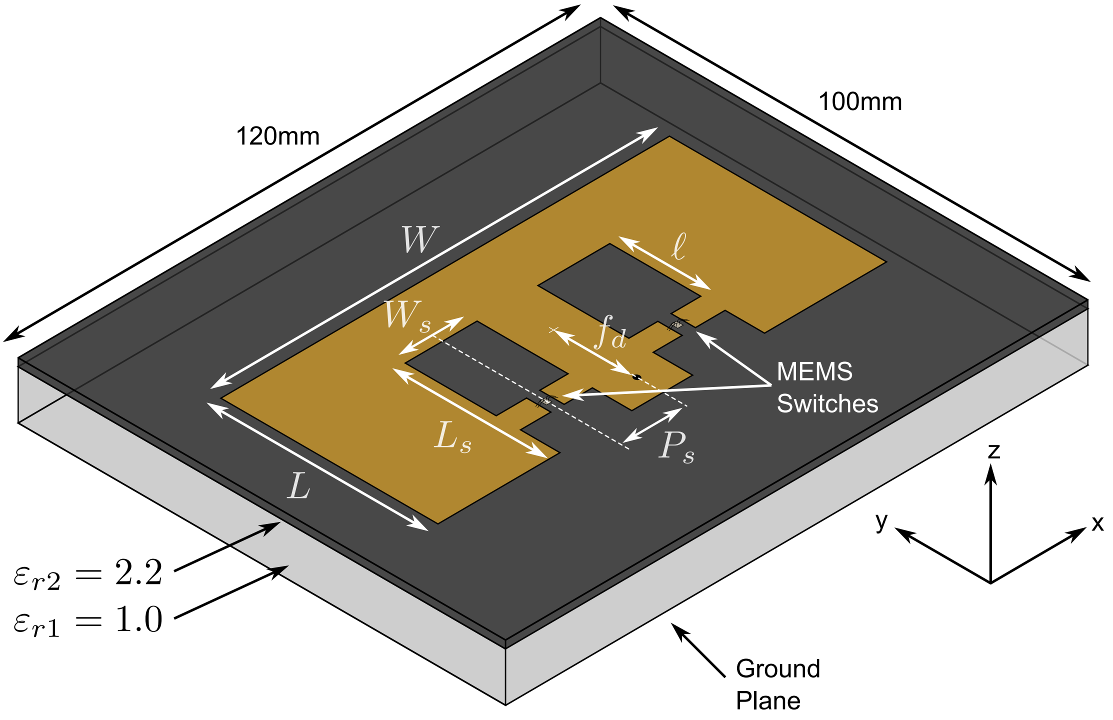


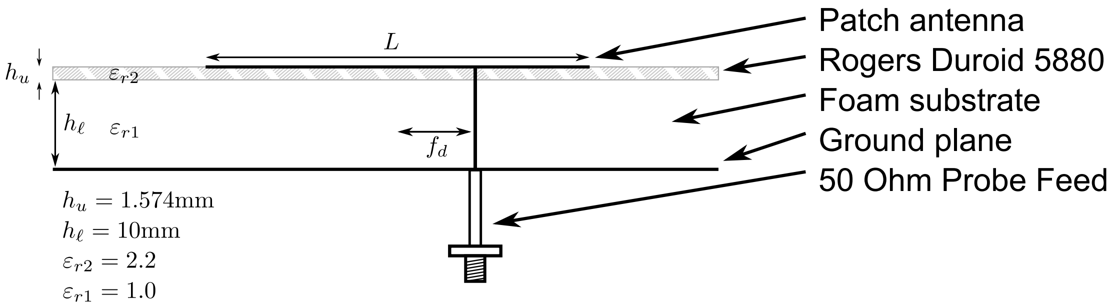
Figure 6.15: LHCP/RHCP reconfigurability implemented into the E-shaped patch design

attempts to effectively change the slot dimensions by placing MEMS switches within the slots as shown in Figure 6.15. The polarization reconfigurable E-shaped antenna would be able to achieve the LHCP radiation with the top switch on and bottom switch off. The RHCP mode could also be excited by switching the slot asymmetry, i.e. reversing the polarity of the switches. Ultimately, this design combines the wideband nature of the E-shaped patch antenna with reconfigurability in order to provide a highly versatile CP patch antenna.

The final implementation of the circularly polarized E-shaped patch antenna is shown in Figure 6.16. The MEMS switches are also included in the figure, however the bias lines are not included in the optimization procedure. The reconfigurable E-shaped patch was implemented in a multilayer fashion in order to facilitate simple and accurate fabrication with photolithography. This methodology was employed in order to retain the accuracy of photolithography (used on the Rogers Duroid layer) while increasing bandwidth by using an effective substrate permittivity of $\varepsilon_r \approx 1$. We can assume that the patch antenna will act similar to an air filled substrate, since most of the volume is occupied by the foam substrate which has a relative permittivity close to unity. The geometry used in this design is similar to the E-shaped patch antenna seen in [59], where a wideband and dual-band E-shaped patch are optimized. As seen in Figure 6.16, there are seven variables to optimize, which can be a difficult optimization problem especially when the fitness function is highly multimodal and nonlinear. The next section will discuss the formulation of the optimization problem and then conclude with some results.



(a) Perspective View with Design Variables. Note that the ground plane will have dimensions 200mm x 200mm.



(b) Side View

Figure 6.16: Reconfigurable E-shaped patch implementation for polarization reconfigurability using MEMS switches

6.3.2 Applying Particle Swarm Optimization

With the antenna design and the design variables clearly defined, we proceed to define the components needed in the optimization algorithm. The design in Figure 6.16 is similar to the design in [59], and good results were demonstrated for both designs optimized. Similar boundary conditions are applied to the reconfigurable CP design, where the boundaries are given by

$$L \in (30, 96) \text{ mm} \tag{6.1a}$$

$$W \in (30, 96) \text{ mm} \tag{6.1b}$$

$$L_s \in (0, 96) \text{ mm} \tag{6.1c}$$

$$P_s \in (0, 48) \text{ mm} \tag{6.1d}$$

$$W_s \in (0, 48) \text{ mm} \tag{6.1e}$$

$$f_d \in (-48, 48) \text{ mm} \tag{6.1f}$$

$$\ell \in (0, 96) \text{ mm} \tag{6.1g}$$

In these boundaries, L is the length of the patch, W is the width of the patch, L_s is the slot length, W_s is the slot width, P_s is the slot position, f_d is the x position of the feed, and ℓ is the position of the MEMS bars. This design allows a wide variation of designs and does not restrict other possible designs in any way. The constraints must also be included in the optimization and are shown below.

$$L - L_s > 5\text{mm} \tag{6.2a}$$

$$P_s - \frac{W_s}{2} > 2.5\text{mm} \tag{6.2b}$$

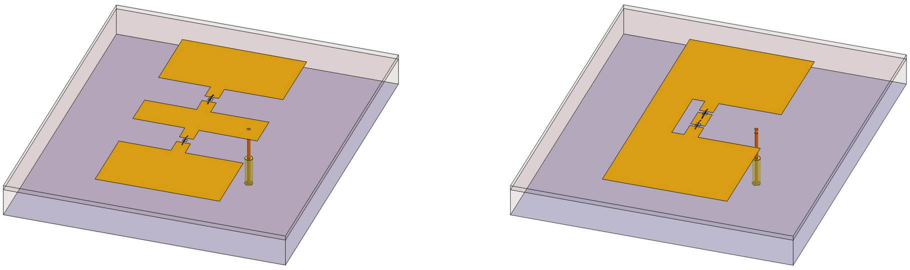
$$\frac{W}{2} - \left(P_s + \frac{W_s}{2}\right) > 5\text{mm} \tag{6.2c}$$

$$|f_d| < \frac{L}{2} \tag{6.2d}$$

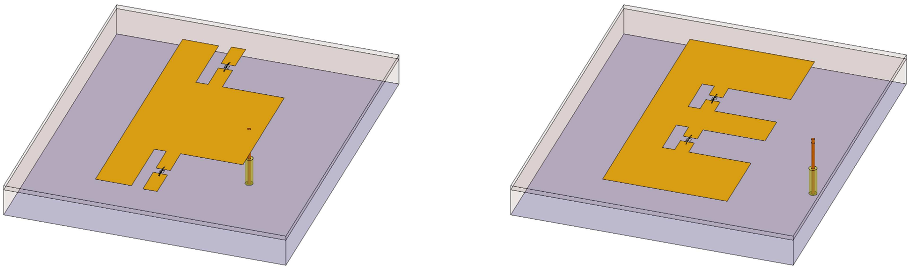
$$L_s - \ell > 5\text{mm} \tag{6.2e}$$

These equations are formulated to avoid designs which do not maintain the E-shaped design or are not of any interest. The constraints also avoid designs that may pose any challenges in fabrication. For example, in equation 6.2a, setting the inequality to be greater than 5mm forces the section of the E-shaped patch above the slots to have width greater than 5mm. By doing this, a design with extremely thin bars is avoided which may cause increased sensitivity to any fabrication errors. The second constraint equation circumvents any possible designs which remove the inner section to which the probe feed is connected. The third constraint does not allow the slots to eliminate the outer bars by forcing the slots to be within the E-shaped patch. The fourth equation confines the probe feed location within the E-shaped patch dimensions. The fifth equation ensures that the MEMS bars are always attached to the patch. Figure 6.17 depicts each constraint and demonstrates an example of a violation for each equation. If all equations in 6.2a-e were written in the form $g_i(\vec{x}) \leq 0$, then there would be six constraints equations total, since equation 6.2d can be split into two inequalities. This is not only a highly dimensional optimization problem, but it is also a highly constrained optimization. Both of these properties make this optimization more difficult.

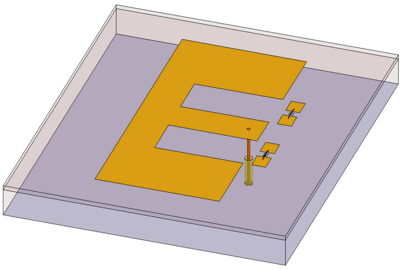
With the geometrical aspects of the optimization defined, the only part left to define is the fitness function for the CP reconfigurable E-shaped patch antenna. This design aims to achieve good impedance matching and good circularly polarized radiation. Similar to the septum optimization in Chapter 5, we assume that the antenna will maintain good AR performance over a wide angular range and therefore will use only the boresight AR (the direction normal to the patch antenna) to measure the antenna's circularity. Since research on this particular antenna is fairly recent, the upper limit on the AR-S_{11} bandwidth is not yet known. This bandwidth is defined as the region in which the antenna demonstrates good performance with $S_{11} < -10$ dB and $AR < 3$ dB. Therefore, the fitness function should include three primary terms: S_{11}, AR, and AR-S_{11} bandwidth. Since the bandwidth is currently unknown, this makes it difficult to predict the frequency band to determine the worst S_{11} performance. This bandwidth approach allows the use of one static evaluation frequency to be included in the fitness function. The frequency was chosen as 2.4 GHz, which is a common frequency band used for WLAN and other various wireless applications. A single objective fitness function incorporating each of these objectives



(a) $L - L_s \ngtr 5\text{mm}$ **(b)** $P_s - \frac{W_s}{2} \ngtr 2.5\text{mm}$



(c) $\frac{W}{2} - \left(P_s + \frac{W_s}{2}\right) \ngtr 5\text{mm}$ **(d)** $|f_d| \nless \frac{L}{2}$



(e) $L_s - \ell \ngtr 5\text{mm}$

Figure 6.17: Possible designs which violate a constraint. It is desired to avoid these designs due to possible fabrication issues or to avoid distortions to the E-shaped design

can be written as

$$f(\vec{x}) = \alpha_1 \; S_{11}|_{f=2.4\text{GHz}} + \alpha_2 \; AR|_{f=2.4\text{GHz}} - \alpha_3 BW_{\text{MHz}} \tag{6.3}$$

where $S_{11}|_{f=2.4\text{GHz}}$ is the S_{11} at 2.4 GHz in a magnitude scale, $AR|_{f=2.4\text{GHz}}$ is the AR at 2.4 GHz in a magnitude scale, and BW_{MHz} is the AR-S_{11} bandwidth in MHz. Since the optimizer has been written for minimization problems, the negative sign for the BW_{MHz} term allows the minimization algorithm to maximum this term by increasing its magnitude.

Of course, the S_{11} and AR are not only evaluated for this particular frequency. In order to find the AR-S_{11} bandwidth, both the S_{11} and AR must be evaluated at multiple frequencies in order to determine the upper and lower frequencies which determine its bandwidth according to $S_{11} < -10$ dB and $AR < 3$ dB. Including the bandwidth is a different approach from using the max function seen in Chapter 5 for the sigmoid septum. Using this type of fitness function allows the optimizer to explore new designs without restricting its focus to one frequency band. If the fitness function used a max function similar to the fitness functions in Chapter 5, then there may be no guarantee that the best design would have the most bandwidth. The fitness function in equation 6.3 hopes to maximize this bandwidth.

Next, the weights α_i must be determined in order to balance the objectives equally. The average values expected for each objective in the design are shown in Table 6.1. These values come from some *a priori* knowledge of the E-shaped patch designs. An average value of $S_{11} = 0.3162 = -10dB$ is written since 2.4 GHz will not necessarily be guaranteed to have the best performance over the entire bandwidth. Using this value expects the worst possible case scenario where $S_{11} = -10$ dB at 2.4 GHz. The same argument is also given for the axial ratio, where the worst case scenario is assumed by using $AR = 1.414 = 3$ dB. The bandwidth assumes an average value of 50 MHz, which represents a 2% bandwidth at 2.4 GHz. While this bandwidth may be small in comparison to the results reported in [145], a smaller bandwidth may be seen on average. It is fairly reasonable to assume that the CP E-shaped would be able to achieve this bandwidth since other single layer single feed CP designs are able to achieve a similar bandwidth. Using these average values, each of the weights can be found by roughly equating the average

Table 6.1: Average Values for Design Objectives

Parameter	Average Value
$S_{11}\|_{f=2.4\text{GHz}}$	0.3162
$AR\|_{f=2.4\text{GHz}}$	1.414
BW_{MHz}	50 MHz

value of each parameter. This leads to

$$\alpha_1 \langle S_{11}|_{f=2.4\text{GHz}} \rangle = \alpha_2 \langle AR|_{f=2.4\text{GHz}} \rangle \tag{6.4}$$

where $\langle \cdot \rangle$ represents the average value. Setting $\alpha_2 = 1$ leads to finding $\alpha_1 \approx 4$. The bandwidth term represents a peculiarity in the averaging scheme due to the negative sign. Another aspect that is different about the BW term is that it is set to zero when $S_{11}|_{f=2.4\text{GHz}} > 0.3162$ and $AR|_{f=2.4\text{GHz}} > 1.414$, and equating the average may not be the most appropriate method for determining the weight. More emphasis should be placed on the BW term, since this term is of most importance. Therefore a larger average value of the term $\alpha_3 BW_{\text{MHz}}$ must be used while still keeping the term at roughly the same magnitude at the first two terms. Therefore a value of $\alpha_3 = 1/10$ is proposed for use in the fitness function in order to maintain this balance. With the weights now chosen, the final implemented fitness function can be written as

$$f(\vec{x}) = 4\, S_{11}|_{f=2.4\text{GHz}} + AR|_{f=2.4\text{GHz}} - \frac{1}{10} BW_{\text{MHz}} \tag{6.5}$$

which is evaluated for one state of the switches. Only one simulation is required since both states will have similar S_{11} and AR performance due to the geometrical symmetry. The switch states create the asymmetry required to generate CP waves, and reversing these states provides a design with roughly (if not exactly) the same performance for one state versus another state. Since constraints were also involved with this optimization, a penalty function $p_c(\vec{x})$ must be

Table 6.2: PSO Intrinsic Parameters for the RHCP/LHCP Reconfigurable E-shaped patch antenna design optimization

PSO Parameter	*Values Used*
c_1	2.0
c_2	2.0
Swarm Size	16
Δt	1.0
Max Iterations (i_{max})	500
Boundary Condition	Invisible
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

introduced. For this particular problem, a stepped value penalty function written as

$$p_c(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{F} \\ 10^5 & \text{if } \vec{x} \notin \mathcal{F} \end{cases} \tag{6.6}$$

was included within the PSO optimization implementation, and the final fitness function optimized by PSO was given as

$$f_P(\vec{x}) = f(\vec{x}) + p_c(\vec{x}) \tag{6.7}$$

where $f_P(\vec{x})$ is the fitness function implemented for PSO to optimize. The PSO algorithm was then applied to the model shown in Figure 6.16 using the final implementation of the simple MEMS circuit model shown in Figure 6.13. The PSO intrinsic parameters used for the CP E-shaped patch optimization are shown in Table 6.2. The only modification made to the original algorithm discussed in Section 2.1.1 was that a seed was applied to this optimization run by setting one of the particle's initial position to the seed location. This seed was taken from a previous optimization run using the ideal switch model. This previous run on the CP E-shaped patch demonstrated good performance, and their results are reported in [146, 147]. The final design reported in [146] was used as the seed.

6.3.3 Results and Measurements

Among the optimizations conducted in this thesis, this was among one of the lengthiest runs in terms of time. This is due to the lengthy computational time required for calculating the axial ratio versus frequency. In order to provide some acceleration in computing equation 6.7 , the AR was only computed for 2.4 GHz unless both bandwidth conditions ($S_{11} < 0.3162$ and $AR < 1.414$) were met. The AR and S_{11} were calculated versus frequency once the S_{11} and the AR were below these thresholds. The algorithm would then find the bandwidth by performing a step search for the frequency where these thresholds were crossed. Once the thresholds were crossed, the frequencies would be calculated by interpolation. One simulation requiring a bandwidth simulation could take up to 30 minutes, and this led to extremely lengthy optimization runs. The convergence is shown in Figure 6.18 for this optimization run. The entire time required for this optimization was nearly 4 weeks, which indeed is quite a long time for one optimization run. This lengthy amount of time necessitates the use of a good optimization tool such as PSO.

Several features can be observed from this plot. First, only 115 simulations were used in the optimization. This is because a special termination condition had been applied in the case that an extraordinary amount of bandwidth was obtained. The termination condition would stop the optimizer if the simulated bandwidth was above 400 MHz in order to avoid any higher order modes within the E-shaped patch from being excited. These higher order modes can have a tendency to degrade the performance at non-boresight angles. The optimization would continue to run for a few more iterations to search the space for points with better impedance matching or AR. This optimization stopped at roughly 110 iterations.

Next, it can be seen that the global fitness is fairly constant for the first 100 iterations, and this is due to the challenge that a constrained optimization space poses for navigation purposes. The average fitness is in agreement with this postulation, where an average fitness higher than the upper limit of the plot represents an iteration where all particles were outside of $\mathcal{F}$. The average fitness curve shows that the swarm was frequently outside of the constrained space. Hence, it was difficult to maintain a position within $\mathcal{F}$ due to its complex boundaries. The average fitness begins to approach the global best fitness towards the end of the run, which indicates that most

Table 6.3: Final Design Values from the CP E-shaped Patch Optimization

Parameter	Value (mm)
L	44.955
W	89.029
L_s	35.990
W_s	9.251
P_s	7.466
f_d	10.458
ℓ	15.059

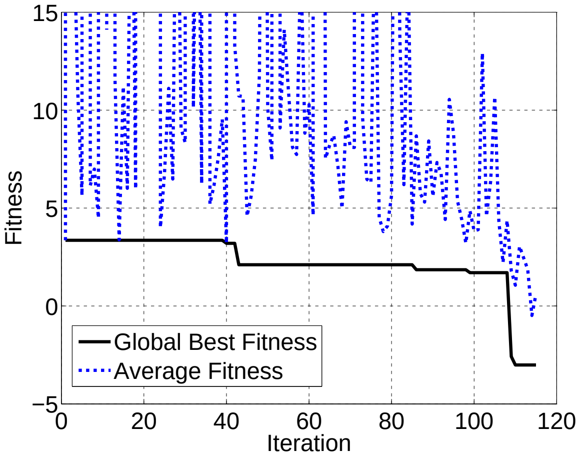


Figure 6.18: Convergence of the CP E-shaped patch design optimization

of the particles were finding points within $\mathcal{F}$. It also demonstrates that most of the particles were near the neighborhood of the global best point observed.

While there may be other existing designs with possibly higher bandwidths, it was decided to investigate this design further before continuing the optimization. The final bandwidth provided by the simulation was nearly 510 MHz, which is roughly a 21% AR-S_{11} bandwidth. A bandwidth this large certainly was not anticipated, and further investigation was conducted on this particular design. This is quite a considerable bandwidth improvement in comparison to the reported AR-S_{11} bandwidth of 9.27% given in [145] and is considered very wideband for a single layer single feed CP patch antenna. The final design values for the CP reconfigurable E-shaped patch antenna are shown in Table 6.3.

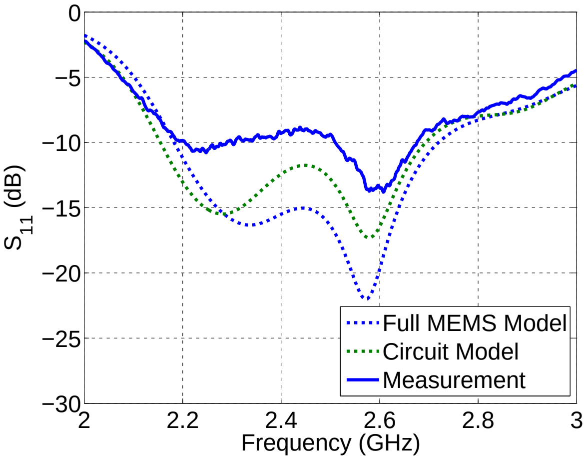
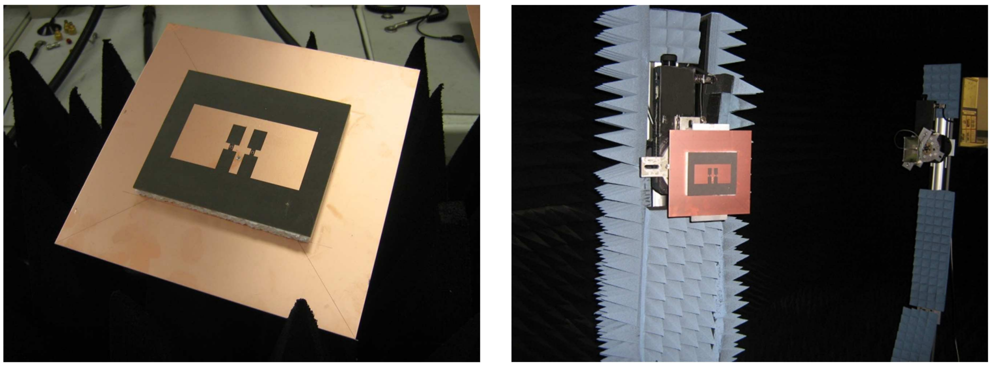


Figure 6.19: Comparison of the impedance matching performance for the CP E-shaped patch antenna between the circuit simulation model, full MEMS simulation model, and a measurement of a fabricated CP E-shaped patch antenna with wirebonded MEMS shown in Figure 6.20



(a) S_{11} measurement

(b) Pattern measurement

Figure 6.20: Fabricated CP E-shaped patch antenna with wirebonded MEMS switches for port and radiation pattern measurements

These design values compared similarly to those reported in [145], which had a height of 10mm and operated in the 2.4-2.5 GHz frequency range. The most notable difference is that the patch width W was considerably wider by roughly 13mm. The slot dimensions are also somewhat different, and these dimensions are often the most critical in terms of performance. Therefore, the final design brought by PSO may have given the improvement over the design presented in [145]. It is also possible that this design experiences better bandwidth performance due to the slight modification to the patch antenna topology in [145], which used unequal length slots. However, in this design a set of bars force the slots to be unequal, and this may have performance advantages. The design shown in Table 6.3 was simulated using both the circuit model and the full MEMS model, and the S_{11} performance is shown in Figure 6.19. The experimental configuration for the S_{11} measurement is depicted in Figure 6.20a. The circuit simulation predict good performance from 2.15-2.69 GHz, which is roughly 540 MHz bandwidth. However the measurements shows some slight disagreement with respect to the circuit model and the full MEMS model simulations. This can be attributed to the approximate nature to these models as well as fabrication errors which are common to multilayer structures. Each of these errors can potentially increase the impedance mismatch observed. Meanwhile, the measured design still exhibits adequate performance that is still acceptable in many cases.

For the CP reconfigurable E-shaped patch antenna, the radiation pattern measurement is just as critical as the impedance matching. These quantify the circularity of the radiation by means of the boresight AR, which is plotted in Figure 6.21. The measurement is compared to the circuit model and full MEMS model simulations, and fairly good agreement is observed between the curves. According to the simulation, the model predicts a circularly polarized wave (AR < 3 dB) for the frequencies 2.18-2.70 GHz using the circuit model. The two bandwidths would therefore intersect in the frequency region 2.18-2.69 GHz, which provides nearly 510 MHz AR-S_{11} bandwidth as discussed previously. However, the measurement exhibits some slight disagreement from the simulated results, and the AR band is given by 2.150-2.525 GHz, which gives a 375 MHz (15.6%) bandwidth. Some of the discrepancies can be expected when comparing results from an approximate MEMS versus an actual measurement. Other errors can be attributed to possible scattered waves and misalignment within the anechoic chamber used for measurement.

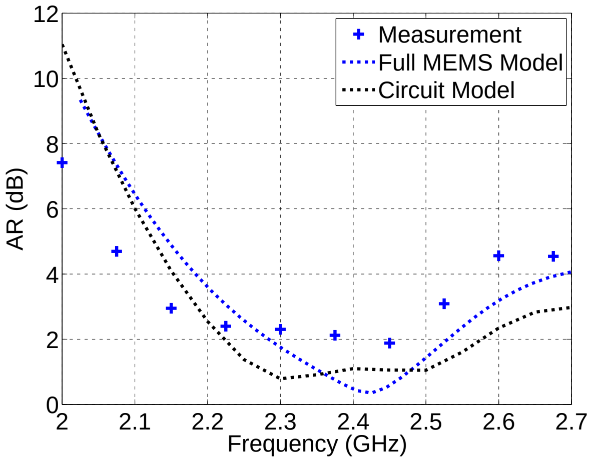
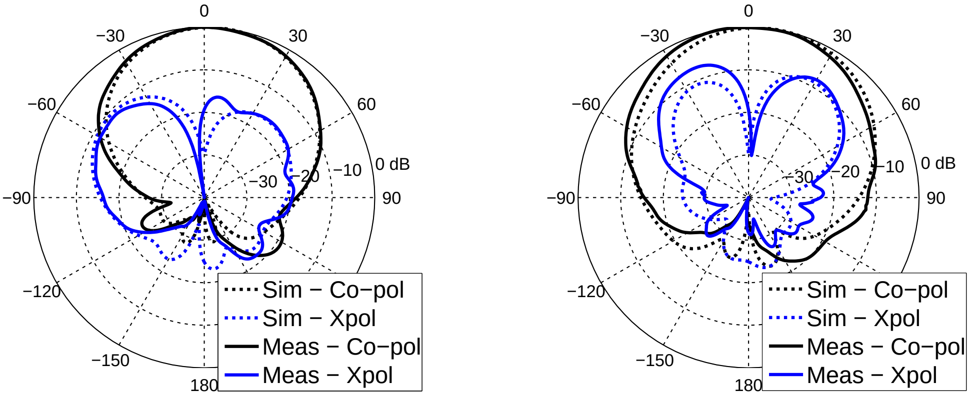


Figure 6.21: Comparison of the axial ratio for the CP E-shaped patch antenna between the circuit simulation model, full MEMS simulation model, and a wirebonded MEMS measurement shown in Figure 6.20b



(a) E-plane ($\phi = 90°$)

(b) H-plane ($\phi = 0°$)

Figure 6.22: Principal patterns for the fabricated RHCP/LHCP reconfigurable E-shaped patch antenna with wirebonded MEMS switches at 2.45 GHz. Its directivity is $D_0 = 8.34$ dB.

The radiation patterns for the CP E-shaped patch are given in Figure 6.22, and only the principal planes (E-plane and H-plane) are shown for brevity. The radiation patterns remain fairly constant versus frequency in the given bandwidth, and therefore allows this concise representation of the antenna radiation performance. Note how the nulls in the cross polarization pattern are not aligned between the simulation and measurement. This could indicate possible misalignments in the chamber which account for any increases observed in the AR measurement. Overall, the radiation patterns show a fairly wide beamwidth as expected in CP patches, and the radiation patterns are fairly typical of CP patches which exhibit no symmetry.

In summary, the CP E-Shaped patch antenna concept was introduced, and an implementation of reconfigurability has been built into the antenna to allow it to switch from RHCP to LHCP. The PSO algorithm was applied to extend the AR-S_{11} bandwidth further than previously reported bandwidths. The final design has been demonstrated and verified through measurements and simulations. The only piece that must be implemented to finalize this design are the bias networks to control the MEMS switches, which are discussed in Section 6.5.

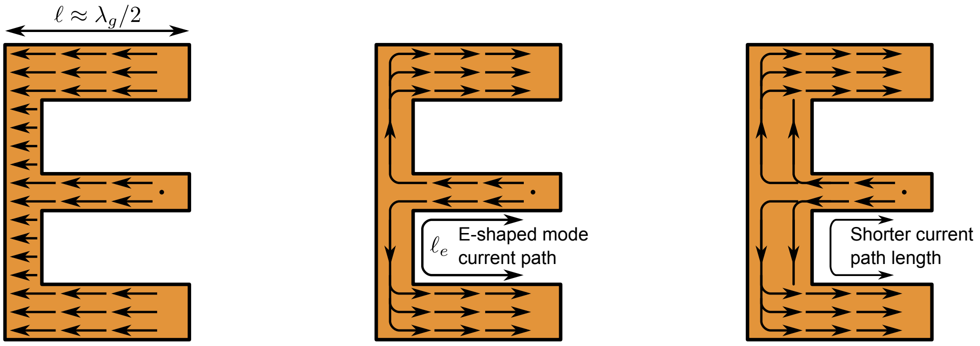
6.4 Frequency Reconfigurable E-Shaped Patch

Versatility and adaptability in wireless systems has been a research topic that has stirred interest in multiple scientific circles, including the antenna systems community. Frequency reconfigurability is the ability to alter the antenna's frequency of operation, where the antenna is either impedance matched or has a certain required gain. One of the primary uses of this can be for the replacement of large multi-antenna systems where each antenna has its own independent use. With the use of frequency reconfigurable antennas, one can replace the complex multi-antenna systems with just one antenna, where each frequency bandwidth can be achieved by a certain switch state. This eliminates the requirement to design high isolation antenna systems, which can be a challenging design problem. It also helps to meet size and weight requirements, since only one antenna may be used to represent multiple antennas. Frequency reconfigurable antennas may also become an important research subject for use in possible wireless systems such as cognitive radio. With the fairly recent proposal of cognitive radio, adaptive and intelligent system research has become a popular research subject. A large amount of research has been spent in solving many problems dealing with the communications aspects of a cognitive radio, such as infrastructure, free spectrum detection, resource allocation, and interference minimization [148, 149]. However there has only been a limited amount of research in the antenna community to provide design solutions in the cognitive radio context [150, 151, 152]. Overall, there exist many system applications for frequency reconfigurable antennas to provide extra system versatility or reduce the system size.

Frequency reconfigurable antennas are typically implemented by shifting the resonance of the antenna, where the resonant frequency is often defined by a particular length of the antenna. This resonance can be shifted in a variety of different ways. One method could be to alter the resonant length of the antenna via the use of switches [153, 154]. Mechanically changing the antenna structure has also been implemented using actuation to alter a patch antenna's structure [155, 156]. Another technique used to alter the resonant frequency can be accomplished by altering the reactive loading upon an antenna. A slot-loaded patch antenna is a great example, and there have been a variety of different patch antenna structures which have had celebrated success in providing frequency reconfigurability [157, 158].

This frequency reconfigurable E-shaped patch antenna design can be classified in the third category of slot-loaded reconfigurable patch antennas. Frequency reconfigurability is achieved by changing the effective dimensions of the slots which load the patch to create the E-shaped patch. To the author's best knowledge, there has only been one other proposed reconfigurable E-shaped patch antenna design given in [159]. This antenna design utilizes PIN diodes in order to achieve two E-shaped states, where the resonant frequency is shifted by reconfiguring to a larger E-shaped design. However, this design requires many more switches (roughly 17 switches) than the proposed design in this section, which only requires two switches.

The next sections will provide some background on the frequency reconfigurable (FR) E-shaped patch antenna. PSO will be applied to achieve a wideband reconfigurability for this design using a similar approach as the CP reconfigurable E-shaped patch antenna. The final results of the optimization and measurements will then be provided to conclude this section.



(a) Patch mode **(b)** E-Shaped mode **(c)** New design with higher resonance frequency

Figure 6.23: Depiction of the slot modes and the effect of changing slot dimensions on the resonant frequency

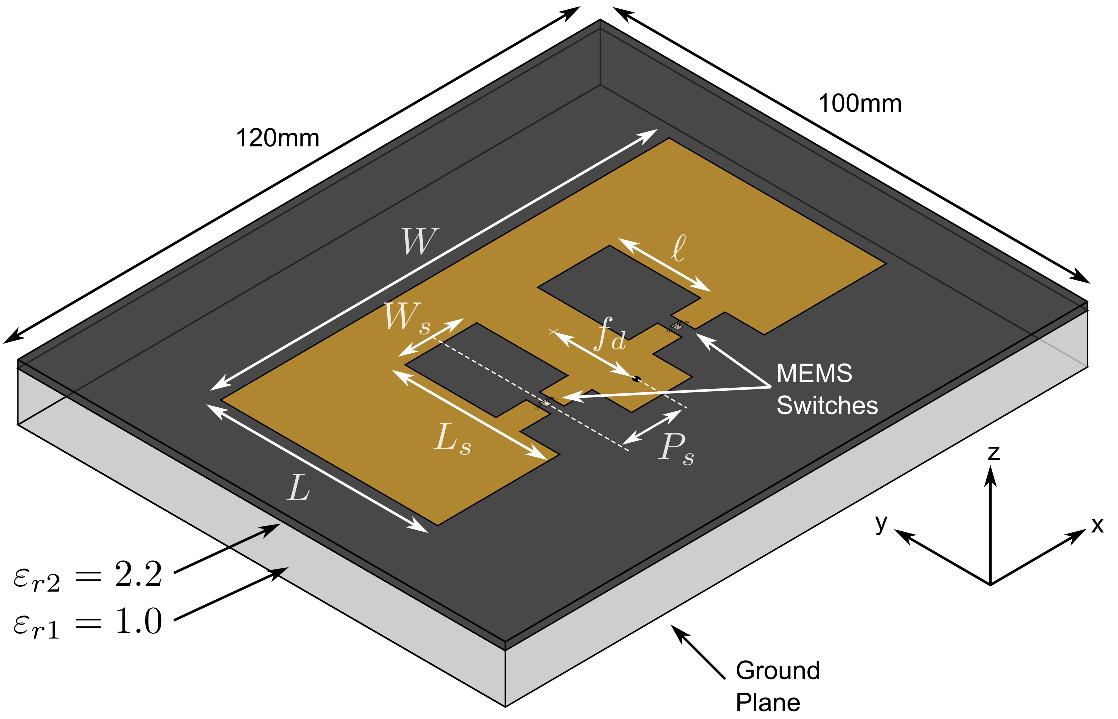
6.4.1 Physical Principles

As discussed in Section 6.1, the E-shaped patch antenna has two primary modes of operation: the patch mode and the E-shaped mode. The patch mode resonance can be predicted by the patch length ℓ depicted in Figure 6.23, which at resonance is roughly half the guided wavelength λ_g within the substrate. The E-shaped mode has currents that meander around the slots, and its resonance is roughly guided by the length that would be travelled by the current. These are only rough insights on the resonance condition within the E-shaped patch, and their exist much more complicated models that better predict the resonant frequencies of this design. However, these first order models provide insight into possible implementations of frequency reconfigurability into the E-shaped patch antenna design.

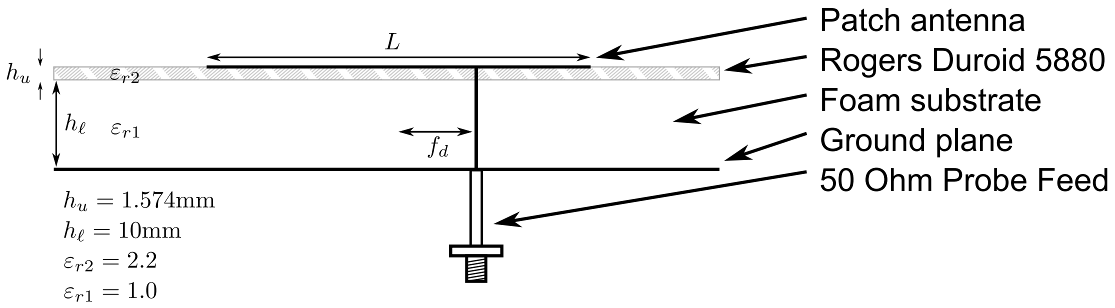
A closer look at the E-shaped patch mode shows that the resonant frequency is highly dependent on the slot dimensions. A wider and longer slot would extend the overall path length traveled by the current, which would lower the resonant frequency of the E-shaped patch mode. A shorter or thinner slot would accomplish the opposite; it would increase the resonant frequency of this mode. Another E-shaped design is shown in Figure 6.23c in order to demonstrate this concept. When comparing the two designs in Figure 6.23b and 6.23c, one can see that the overall length traveled by the currents is shorter for the Figure 6.23c, which would imply that the resonant frequency would be higher for this design.

These observations lead to an intriguing question on possible reconfigurability. What if the slot dimensions could be dynamically changed in real-time? By changing the slot dimensions, it can be seen that the resonant frequency of the E-shaped mode would be shifted. In particular, if the slots were shortened using switches, then one could theoretically shift this resonant frequency in order to obtain frequency reconfigurability. Therefore, a frequency reconfigurable E-shaped patch antenna design is proposed based on the concept of dynamically changing the slot dimensions. This can be realized by placing two switches within the slots. The OFF case represents the normal E-shaped patch case, where the original performance would be observed. The ON case represents the new case where both slots are effectively shortened by turning the switches on. This is the premise of the FR E-shaped patch antenna design, and good performance must be obtained for the two different states in terms of impedance matching.

The FR E-shaped patch concept was implemented on a similar platform used for the CP reconfigurable E-shaped patch. The schematic is shown in Figure 6.24, where the same multilayer structure is given. The primary difference between this design versus the one shown in Figure 6.16 is that the ground plane is the same size as the duroid layer. This provides a more compact design and can be done since the pattern measurements are not as sensitive with linear polarization in comparison to the CP design. The MEMS switches are also implemented in the same fashion as in the CP reconfigurable case. However, the MEMS switches are either both turned OFF or both turned ON, and they create a symmetric antenna structure. The MEMS switches in the CP antenna are opposite, where one is turned off and the other is on in order to create the asymmetry needed for creating CP radiation. This design shown is also the design that is used in conjunction with the MEMS circuit model for optimization purposes. Since the design is symmetric, one can also apply a symmetry boundary condition in order to reduce the computational burden of the simulation.



(a) Perspective View with Design Variables. Note that the ground plane will have dimensions 120mm x 100mm.



(b) Side View

Figure 6.24: Reconfigurable E-shaped patch implementation for frequency reconfigurability using MEMS switches

6.4.2 Optimization for Wideband Designs

The optimization of the FR E-shaped patch design has many similar characteristics to the optimization applied to the CP reconfigurable E-shaped patch design. Therefore a summary of the similar aspects will be recapped briefly and the equations rewritten for clarity. First, the boundary conditions for the FR E-shaped design are identical to those of the CP design and are shown in the following list.

$$L \in (30, 96) \text{ mm} \tag{6.8a}$$

$$W \in (30, 96) \text{ mm} \tag{6.8b}$$

$$L_s \in (0, 96) \text{ mm} \tag{6.8c}$$

$$P_s \in (0, 48) \text{ mm} \tag{6.8d}$$

$$W_s \in (0, 48) \text{ mm} \tag{6.8e}$$

$$f_d \in (-48, 48) \text{ mm} \tag{6.8f}$$

$$\ell \in (0, 96) \text{ mm} \tag{6.8g}$$

Each variable represents the same aspect as in the CP design and are given the same range of values. The design constraints are also identical, as seen by

$$L - L_s > 5\text{mm} \tag{6.9a}$$

$$P_s - \frac{W_s}{2} > 2.5\text{mm} \tag{6.9b}$$

$$\frac{W}{2} - \left(P_s + \frac{W_s}{2}\right) > 5\text{mm} \tag{6.9c}$$

$$|f_d| < \frac{L}{2} \tag{6.9d}$$

$$L_s - \ell > 5\text{mm} \tag{6.9e}$$

These constraints are provided to avoid the same problematic designs as shown in Figure 6.17 shown in Section 6.3.2. The equations 6.8 and 6.9 define the solution space $\mathcal{S}$ and the feasible space $\mathcal{F}$, and therefore account for all geometrical aspects of the optimization process.

The following step aims to define the fitness function for this optimization, which will be different than the CP design optimization. In this frequency reconfigurable design, both possible states must be optimized rather than just one. Therefore, two simulations are required in order to evaluate a given set of geometrical parameters. One simulation outputs the performance for the OFF state (where both switches are off), and the other outputs the performance for the ON state. Assuming that the directivity is fairly constant in the band of interest, one can solely optimize the impedance matching in order to provide better connectivity on the average case scenario. The fitness function should then distinguish designs having a good impedance match over the band of interest from those which do not. A methodology that has shown good results is the *minimax* approach as shown in Chapter 5. This is incorporated in the fitness function

$$f(\vec{x}) = \alpha_1 \max_{f \in F_{\text{off}}} S_{11,\text{off}} + \alpha_2 \max_{f \in F_{\text{on}}} S_{11,\text{on}} + \alpha_3 \left| \max_{f \in F_{\text{off}}} S_{11,\text{off}} - \max_{f \in F_{\text{on}}} S_{11,\text{on}} \right| \tag{6.10}$$

where the term $\max S_{11,\text{off}}$ represents the worst case S_{11} in dB for the band F_{off} when the MEMS switches are in the OFF case. The term $\max S_{11,\text{on}}$ represents the worst case S_{11} in dB for the band F_{on} when the MEMS switches are in the ON state. After some investigation, the bands

$$F_{\text{off}} = (2.0, 2.6) \text{ GHz} \tag{6.11a}$$

$$F_{\text{on}} = (2.6, 3.2) \text{ GHz} \tag{6.11b}$$

were chosen for the optimization due to the antenna geometry and the possible bandwidth expected from this design. The optimization will attempt to minimize this fitness function, and therefore it will minimize the sum of the worst S_{11} values for the two states. However, these two terms alone do not best represent the optimization problem, for there may be cases where the fitness function would output small values for designs that are not necessarily better than others. One must ensure that both objectives are optimized rather than just one alone. Using the shorthand $S_{11,\text{off}}^{\max} = \max_{f \in F_{\text{off}}} S_{11,\text{off}}$ and $S_{11,\text{on}}^{\max} = \max_{f \in F_{\text{on}}} S_{11,\text{on}}$, consider the following two cases assuming that $\alpha_i = 1$:

$$\begin{pmatrix} S_{11,\text{off}}^{\max} = -20 \text{ dB} \\ S_{11,\text{on}}^{\max} = -20 \text{ dB} \end{pmatrix} \quad \rightarrow \quad S_{11,\text{off}}^{\max} + S_{11,\text{on}}^{\max} = -40 \tag{6.12a}$$

$$\begin{pmatrix} S_{11,\text{off}}^{\max} = -40 \text{ dB} \\ S_{11,\text{on}}^{\max} = 0 \text{ dB} \end{pmatrix} \quad \rightarrow \quad S_{11,\text{off}}^{\max} + S_{11,\text{on}}^{\max} = -40 \tag{6.12b}$$

In the first case shown by equation 6.12a, both states have the worst $S_{11} = -20$ dB, which implies that good performance is observed for both bands. The second case in equation 6.12b shows the worst $S_{11} = -40$ dB for the OFF state, but the ON state has the worst $S_{11} = 0$ dB, which implies that not all frequencies satisfies the criterion $S_{11} < -10$ dB for the ON state. Yet, the addition of the two terms is equal for both cases as shown in equation 6.12. This implies that both designs have equal fitness and are therefore equivalent in performance. This statement is false however since the second case has an $S_{11} > -10$ dB. Therefore an additional term in the fitness function is needed. The term $|S_{11,\text{off}}^{\max} - S_{11,\text{on}}^{\max}|$ hopes to negate this effect and properly distinguish good designs from poor ones. These two cases are added again using this extra term and assuming that $\alpha_i = 1$.

$$\begin{pmatrix} S_{11,\text{off}}^{\max} = -20 \text{ dB} \\ S_{11,\text{on}}^{\max} = -20 \text{ dB} \end{pmatrix} \quad \rightarrow \quad S_{11,\text{off}}^{\max} + S_{11,\text{on}}^{\max} + |S_{11,\text{off}}^{\max} - S_{11,\text{on}}^{\max}| = -40 \tag{6.13a}$$

$$\begin{pmatrix} S_{11,\text{off}}^{\max} = -40 \text{ dB} \\ S_{11,\text{on}}^{\max} = 0 \text{ dB} \end{pmatrix} \quad \rightarrow \quad S_{11,\text{off}}^{\max} + S_{11,\text{on}}^{\max} + |S_{11,\text{off}}^{\max} - S_{11,\text{on}}^{\max}| = 0 \tag{6.13b}$$

Again, the optimization algorithm attempts to minimize the fitness function, and therefore it will try to make this absolute subtraction term have a value closer to zero. Hence, this term compares the relative difference between both bands and forces them to be similar. By doing so, it alleviates the problem shown in equation 6.12.

The weights in the fitness function must be set in order to proceed with the optimization. Luckily, each of the S_{11} terms will have roughly the same average value and therefore the weights $\alpha_1 = \alpha_2 = 1$ can be chosen. Since, the absolute subtraction term will have an average of zero,

Table 6.4: PSO Intrinsic Parameters for the Frequency Reconfigurable E-shaped patch antenna design optimization

PSO Parameter	*Values Used*
c_1	2.0
c_2	2.0
Swarm Size	14
Δt	1.0
Max Iterations (i_{max})	500
Boundary Condition	Invisible
w^k	$0.9 - 0.5\left(\frac{i}{i_{max}}\right)$
$\vec{v}_{max}$	$\frac{1}{2}\left(\vec{x}_{max} - \vec{x}_{min}\right)$

the averaging technique does not provide a clear suggestion for the value of α_3. Therefore, the value is arbitrarily set to $\alpha_3 = 1$ since the magnitude will be roughly the same when the two bands have some differences. With these choices, the fitness function can be finally written as

$$f(\vec{x}) = \max_{f \in F_{\text{off}}} S_{11,\text{off}} + \max_{f \in F_{\text{on}}} S_{11,\text{on}} + \left| \max_{f \in F_{\text{off}}} S_{11,\text{off}} - \max_{f \in F_{\text{on}}} S_{11,\text{on}} \right| \tag{6.14}$$

Lastly, the same penalty function for the CP reconfigurable optimization run was utilized in this run, and the function was written as

$$p_c(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{F} \\ 10^5 & \text{if } \vec{x} \notin \mathcal{F} \end{cases} \tag{6.15}$$

and was implemented into the fitness function by

$$f_P(\vec{x}) = f(\vec{x}) + p_c(\vec{x}) \tag{6.16}$$

where $f_P(\vec{x})$ represents the function which PSO optimizes.

All required aspects for this optimization have been defined, and PSO was applied in order to minimize the given fitness function with the intrinsic parameters shown in Table 6.4. This optimization used the model provided in Figure 6.24 in conjunction with the simple MEMS circuit model in order to finally simulate the performance of the FR E-shaped patch antenna. Two simulations are computed in order to simulate the ON and OFF cases. However, the simulations are not very computationally expensive since only the port parameters are used in the fitness function.

6.4.3 Final Designs

The optimization run was quite successful in finding a design that provided the performance desired in the bands of interest. The fitness function had a well defined frequency band, and therefore the computation of the fitness function was quite simple. First, the simulation model was simulated and all frequency data obtained for the two switch states and frequency bands. Then a max function was implemented by comparing all S_{11} values within the frequency band and finding the maximum value. Once the maximum is found for each band/state, the values are then substituted into equation 6.16. Since half-symmetry was exploited, the average time was reduced to a minimal 2-3 minutes, leading to a function evaluation time of approximately 5 minutes. If every design provided by PSO was simulated, then this would result in roughly 24 days for a full optimization of 500 iterations. However, not all designs are simulated since there can be many that are out of the boundaries or that do not satisfy the constraints. The total time for this particular optimization run was roughly 23 days.

The convergence plot is given in Figure 6.25, where both the average and global best fitness are plotted. PSO was somewhat slow in finding a design that met the -10 dB S_{11} criterion, but it did find a design near 350 iterations. This is the point where the global best fitness crosses the -20 mark, which implies that both states have a maximum $S_{11} < -10$ dB, which meets the desired criterion for those bands. Several other features to this optimization can be observed from this convergence plot. First, the given feasible space $\mathcal{F}$ is quite tricky to navigate as evidenced by the sharp spikes seen in the beginning of the optimization. Further into the optimization all of the particles make it into the feasible region, where they navigated the space to find the global best minimum. We also see that the average fitness approaches the value of the global best near the end of the optimization. This is often the point where the inertial weight has decreased to

Table 6.5: Final Design Values from the FR E-shaped Patch Optimization

Parameter	Value (mm)
L	44.299
W	95.870
L_s	28.580
W_s	11.436
P_s	13.100
f_d	18.664
ℓ	4.882

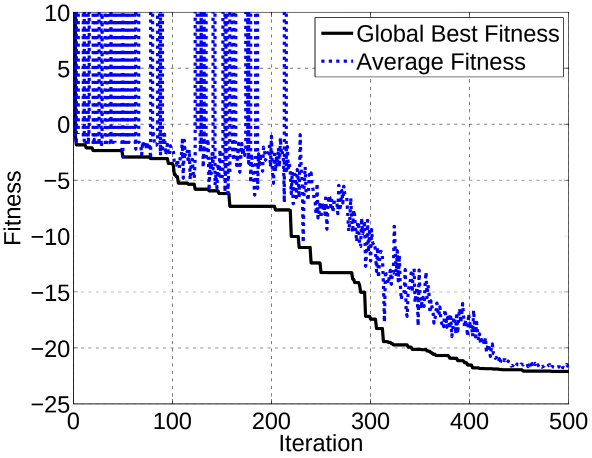


Figure 6.25: Convergence of the frequency reconfigurable E-shaped patch design optimization

a point where most (if not all) of the particles are searching the local neighborhood around the gBest point. This can therefore imply that the design can maintain good performance even with some slight variations in the design variables, since the average fitness is similar to the global best fitness.

Even though this constrained optimization was quite difficult, PSO found a design that met the required design criterion. The final design values are provided in Table 6.5. From this table, one can see that a wide patch antenna must be used in order to achieve a FR E-shaped patch antenna. This width W is slightly less than one wavelength at 2.4 GHz, which can be considered somewhat large for a patch antenna. The length is still near our expectations and is quite similar to the one found in the CP optimization run, as shown in Table 6.3. The slot dimensions are fairly common for the wideband E-shaped patch designs. One major difference between this design and the CP reconfigurable design is the switch bar placement within the slot. For the CP reconfigurable design, the bars are placed nearly halfway into the slots, leading to a value nearly at $\ell \approx L_s/2$. For this design, the bars are placed very close to the bars that form the E-shape, which implies a small value of ℓ, as shown in Table 6.5.

The design shown in Table 6.5 was both simulated and measured with a fabricated design and wirebonded MEMS switches. The S_{11} performance for both states is plotted in Figure 6.26 and compared between the circuit simulation model, the full MEMS simulation model, and the measurement. The blue curves correspond to the OFF state, and the green curves correspond to the ON state. The measured prototype is shown in Figure 6.27a, and a close-up view of the implemented MEMS switch is shown in Figure 6.27b. It should be noted that in the measurement, two prototypes were made to measure each state of the antenna design, since no bias lines were present to switch the MEMS switches. Radant MEMS offers a Thru-Line version of their Radant MEMS *RMSW100HP* switch [139], and this was utilized to measure the ON state of the antenna. The MEMS switch shown in the close-up figure is the OFF state switch, which is represented by the standard Radant MEMS *RMSW100HP* switch. Good agreement is observed by all curves on the plot, especially for the OFF case. In the ON case there is slightly more disagreement, and the measurement slightly goes over the -10 dB mark. Nevertheless, this demonstrates the concept of the frequency reconfigurable E-shaped patch antenna.

The radiation patterns of the FR E-shaped patch antenna were also measured and are plotted in Figure 6.28. Note that the frequency was 2.55 GHz, which implies that the OFF state must be used for the measurement. Good agreement is also seen between the measurement and the simulation, which uses the circuit model for the radiation patterns. In the E-plane, a slightly larger amount of cross polarization was detected, and this may be due to possible diffractions or misalignments within the UCLA anechoic chamber. This is also noted for the boresight direction in the H-plane plot, where a high cross polarization is observed rather than a null. However, these patterns in general show fairly good agreement between simulation and measurement and overall the patterns match the standard characteristics of a typical rectangular patch antenna.

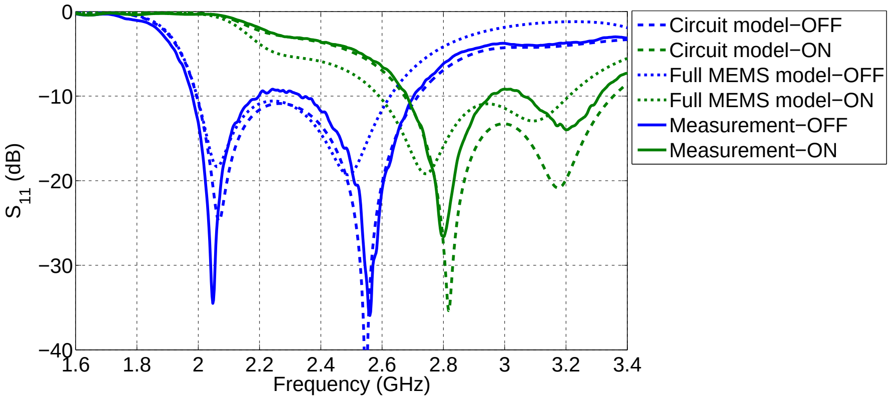
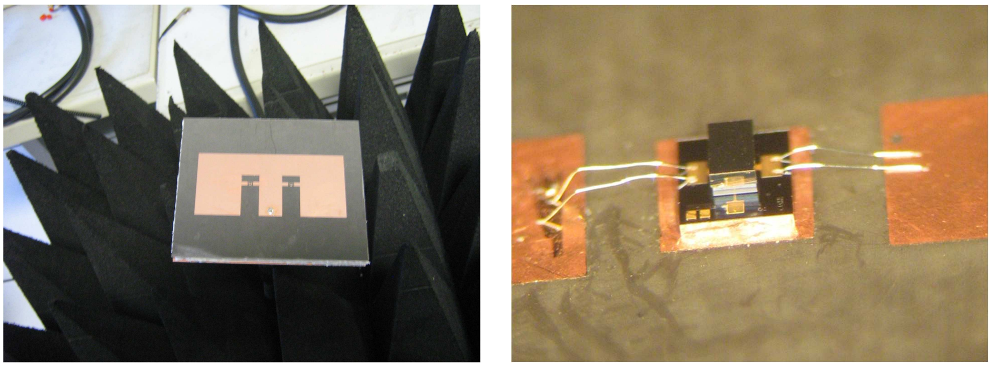
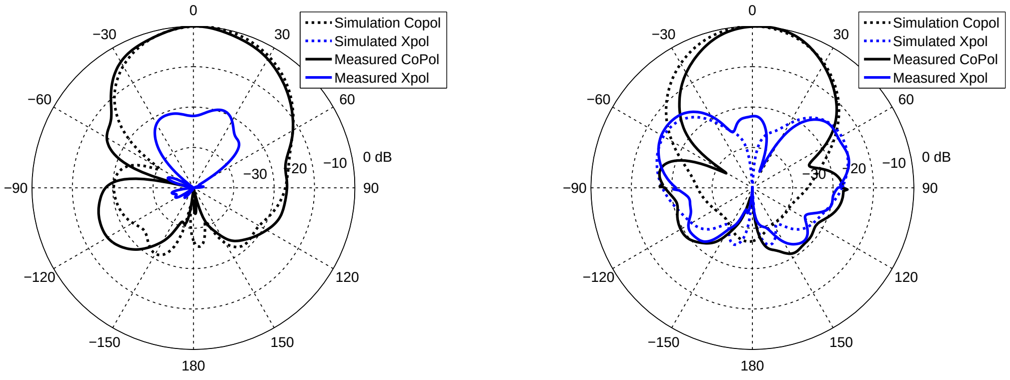


Figure 6.26: Comparison of the impedance matching performance for the frequency reconfigurable E-shaped patch antenna between the circuit simulation model, full MEMS simulation model, and a fabricated E-shaped patch antenna with wirebonded MEMS shown in Figure 6.27



(a) S_{11} measurement **(b)** Close up of MEMS switch

Figure 6.27: Fabricated frequency reconfigurable E-shaped patch antenna with wirebonded MEMS switches



(a) E-plane ($\phi = 90°$) **(b)** H-plane ($\phi = 0°$)

Figure 6.28: Simulated (circuit model) and measured frequency reconfigurable E-shaped patch antenna patterns at 2.55 GHz in the principle planes. The directivity $D_0 = 10.57$ dB for this frequency.

In summary, the results of the optimization are very pleasing, and the optimizer provided a final design which fit the criterion in the bands provided. Much of the agreement between the simulation and the measurement can be attributed to the circuit model. These final implementations demonstrate the use and power of the simple circuit model in the effort to optimize these MEMS reconfigurable E-shaped patches. Much of the work was spent in finally achieving good agreement between the simulation model and measurement. The final design presented a marvelous boost in effective usable bandwidth for the FR E-shaped patch antenna. E-shaped patches typically demonstrate roughly 31% bandwidth, as discussed previously in Section 6.1, and the final FR design showed roughly 1.25 GHz bandwidth, which is roughly 48% bandwidth. This nearly doubles the previously achievable bandwidth while utilizing nearly the same volume, and this can be considered a difficult achievement which can have many different applications.

6.5 Bias Network Development

As with any electrically controlled switch, a bias line to control the switch must be implemented in order for the reconfigurable antenna design to have any practical use. The design and implementation of the bias lines is typically done as an afterthought due to their limited interaction with the RF performance. However, if stronger interactions are anticipated, then many clever techniques can be used in order to connect the bias lines for the proper biasing of the antenna and the RF switch. For example, the use of a quarter wavelength stub was used in [158] in order to short the patch antenna to DC ground while not affecting the performance of the patch antenna. An integrated bias network was implemented into the patch antenna using capacitively coupled lines in order to provide access to the patch slot while allowing the RF current to travel on its normal path. Unfortunately, this methodology is not applicable to the E-shaped patch antenna due to its wideband characteristics.

We will use the following approach in order to develop a bias network design solution. First, we used the ideal switch model in order to test any possible designs using measurements. This avoided the costly and wasteful use of MEMS switches. The ideal switch E-shaped model was discussed in Section 6.2 and assumes that the MEMS switches operate as either ideal shorts or ideal open circuits (which in not the case in reality). Nonetheless, this design still operates rather well when operating without MEMS switches wirebonded, which is why it provides a great deal of information with regard to the implementation of the bias network. Another important point is that only the results for the OFF case are presented in this section for brevity. The ON case can be achieved either by fabricating the ideal switch design or by placing copper tape across the bars in order to provide a short circuit. In this manner, achieving a good design for the OFF case will also imply achieving a good design for the ON case.

As a first implementation, metal lines are placed into the ideal switch design, where the metal lines would have clear access to bias the MEMS switches if placed on the pads. The S_{11} performance comparison is shown in Figure 6.29, where three cases are investigated. They include a simulation without any sort of bias lines included, a simulation with metal lines 45mm long included, and a measurement with the same metal lines. Each of these cases are depicted in the same order in Figure 6.30. The S_{11} response observed for the simulation without any bias lines is expected, where good $S_{11} < -10$ dB performance is achieved throughout the 2.0-2.7 GHz frequency band. However, both the simulation and the measurement which incorporate the metal bias lines show a drastic performance change in the design, and this is unacceptable for a final design implementation. The impedance matching performance is almost completely ruined for the band of interest, and therefore this design solution could not be used for the final

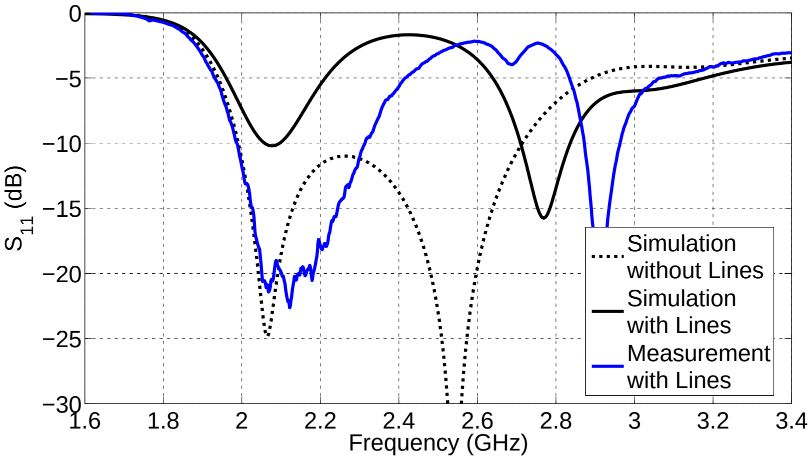
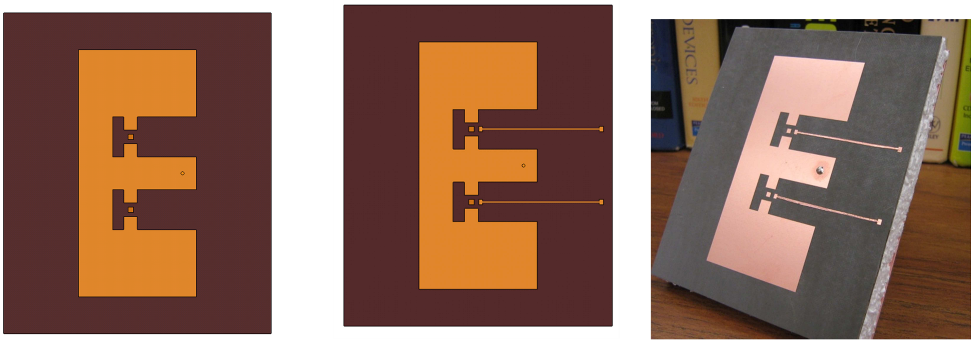


Figure 6.29: Comparison of the impedance matching performance for the simulation with no bias lines, the simulation with metallic lines included, and the measurement with metallic lines included as depicted in Figure 6.30



(a) Simulation without Lines

(b) Simulation with Lines

(c) Measurement with Lines

Figure 6.30: Simulated and fabricated models to test metallic bias networks

implementation.

This certainly was not the expectation for this design, and it forced a further examination into possible reasons to explain this phenomenon and pinpoint the situation to avoid. One immediate observation is that the length of these given lines provides a resonant frequency (half wavelength) near 3 GHz, which may induce some strong coupling between the line and the E-shape patch. Although the resonance is not exactly at a frequency within the band, the near resonance effects may still be substantial, and therefore more experiments must be conducted. One method to investigate this observation further is a test of the dependence of the S_{11} performance on the line length. Therefore, a line with variable length was placed in the HFSS simulation in order to simulate these effects. The line length was varied from 5mm to 45mm in intervals of 10mm, and Figure 6.32 depicts the simulation model. The S_{11} performance for each length of line is provided in Figure 6.31. The original E-shaped simulation model without the line corresponds to the case where $d = 0$mm.

The first and most obvious observation from this plot is that the S_{11} is indeed dependent on the length of the line. It seems that the impedance matching is relatively unaffected by the presence of the line for lengths lower than or equal to 25mm. Above these values, the S_{11} becomes notably different, and therefore any lengths above 25mm are not recommended in order to maintain the performance of the E-shaped antenna. This also shows that the lines under 25mm are relatively invisible to the RF waves, possibly due to the non-resonance condition. However, this does not improve the situation because a connection is needed in order to provide a bias to the MEMS switch. Since it may be possible that the coupling between the bias lines and the patch is strongest when the lines are in resonance, the bias line investigation shifts its focus towards finding methods for breaking a line resonance at the RF frequencies but maintaining a DC connection. The space is also limited due to the slot geometry, and therefore passive tuning stubs and low pass filter design are out of the question.

One might be able to break the near-resonance condition by adding loss to the bias line network. The next part of this investigation looked into the possibility for implementing resistors or inductors into the design. This could be accomplished by splitting the bias line into small

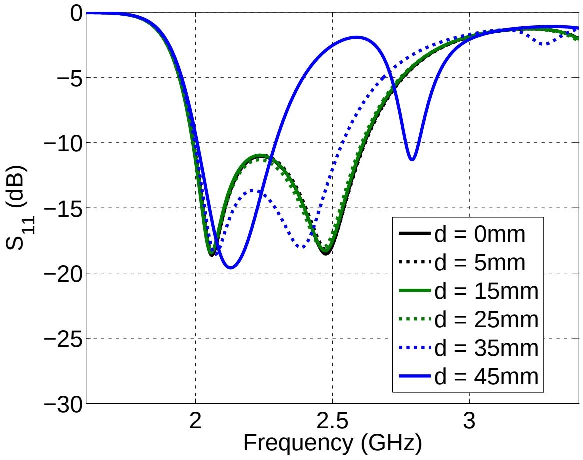


Figure 6.31: Comparison of the effect of line length, shown in Figure 6.32, on S_{11} performance in order to determine whether a resonant line causes the impedance mismatch

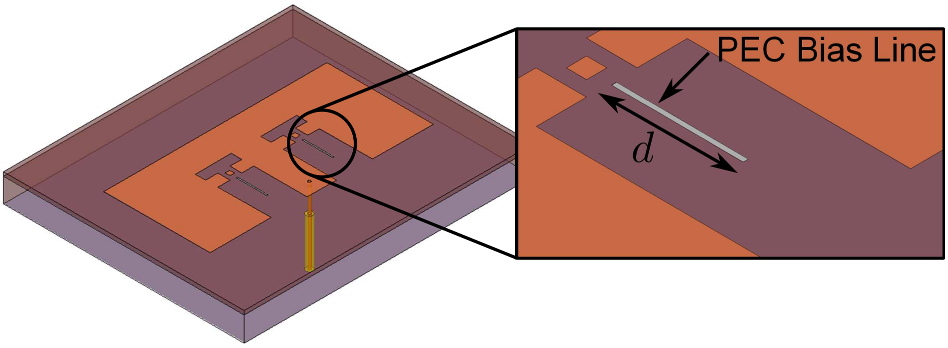
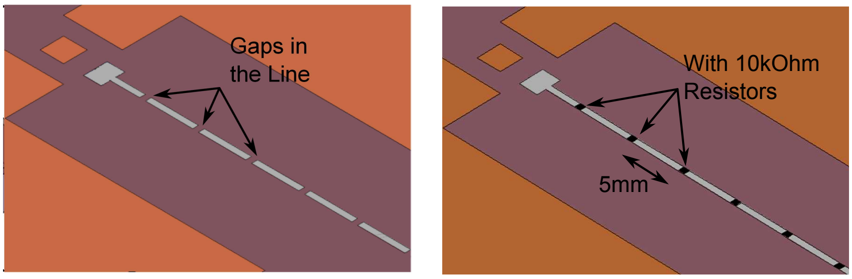


Figure 6.32: Simulation model used to determine the dependence of S_{11} on the length, d, of the bias line

sections with gaps between each section. The gap scenario represents the ultimate scenario with infinite resistance, however it is also unknown whether a structure with multiple lines of length d would cause problems. It was demonstrated that a 5mm line would not cause any problems in Figure 6.31, but this does not imply that an array of these objects within the slot would not have adverse effects. Therefore the lines with gaps and the case without lines are compared in Figure 6.34. The bias network with multiple sections of bias line is shown in Figure 6.33a.

There seems to be little interaction between the patch and the bias line with gaps, which implies that the near-resonance condition has been broken for this bias network. Overall, the performance between the case without any lines and the case incorporating a line with gaps is nearly indistinguishable. Since this design was successful in breaking the near-resonance condition, another design with perfect 10 kΩ resistors are placed in the gaps in order to provide the DC connectivity to the MEMS switches. Perfect resistors represent the ideal case where only a resistance is seen across the terminals of a resistors. However, this may not be the case since a surface-mount technology (SMT) resistor is used for a practical implementation. The SMT resistor chips are most often modeled with a series inductance and a terminal to terminal capacitance. This model shows that the impedance of the resistor degrades after a certain frequency and either increases or decreases indefinitely [160]. The case with 10 kΩ resistors is then simulated and the results are also shown in Figure 6.34, and an illustration of the resistors in place is shown on Figure 6.33b. Again, perfect agreement is observed between the resistor bias line and the case without lines. This is good news, since it demonstrates that the resonance condition can be broken while still providing a DC connection to the MEMS switches. Even though 10-100 kΩ is a large amount of DC resistance, the MEMS switches typically are actuated by voltage rather than by current. In other words, the input impedance of the MEMS switch is typically on the order of 100-1000 MΩ. Therefore, a line impedance of 80 kΩ is negligible in comparison, and only a small portion of the voltage will be lost to the line. It should be noted that one may be able to accomplish the same task using inductors, which would have very little DC loss. The only issue encountered with this approach was that the required inductance needed for good performance was at least 15 μH, which is not commercially available with a self-resonant frequency (SRF) higher than 3 GHz. Therefore, an inductor design would not be



(a) Bias line with gaps (open circuit)

(b) Bias line with perfect 10kΩ resistors

Figure 6.33: Simulation models to compare a bias line with gaps (open circuit) and 10 kΩ resistors filling the gaps

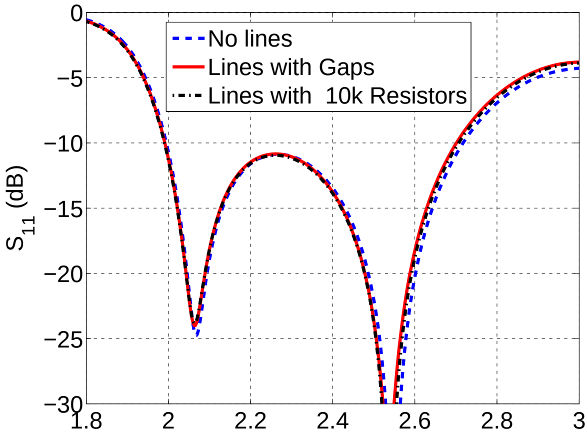


Figure 6.34: Comparison of the S_{11} performance for the simulation with no bias lines, the simulation with metallic lines with gaps shown in Figure 6.33a, and the simulation with resistors shown in 6.33b

possible unless commercial technology changes in the near future.

The bias network design which allowed a DC connection and broke the near-resonance condition using resistors was prototyped to ensure that the design was plausible and maintained the same performance in comparison to the design without bias lines. The prototype is shown in Figure 6.35, where 8-10 kΩ resistors are implemented via solder for each bias line. The connection was tested using an ohmmeter, and a reading of 80 kΩ was measured, which implied that the connection between each resistor was solid. A measurement of the S_{11} for several different cases ensued, and the cases compared in Figure 6.36 include simulations and measurements for the bias line with gaps case and the bias line with resistors case. Note that the E-shaped patch was measured before the resistors were soldered onto the bias lines, and therefore there were two measurements: a design with bias lines having gaps and a design with resistors. There are some discrepancies observed between the simulation and measurement for the bias line with gaps and for the bias line with resistors. However, the best plots to compare are the two simulation scenarios and the two measurement scenarios. Clearly, adding the resistors in the simulation and in the measurement design showed no significant changes in their S_{11} plot. With both the simulation and the measurement results, one can conclude that the resistor bias line is a possible choice of bias network which can be fully implemented into the E-shaped patch design platform with good performance.

Another approach that one may take is the use of highly resistive lines, which we will denote as lossy bias lines for clarity. These are quite popular and have been implemented in systems where micro-fabrication facilities are available. Typically, the designs reported in literature have used TiW lines [161] or even Silicon Chrome (SiCr) lines [162], both of which have a high DC resistance. In the case of the SiCr lines, the total impedance of the line was 500 kΩ. Yet, these micro-fabrication techniques were not available to our lab at the time, and other avenues were explored in order to implement this technique. After some research, a conductive ink was found to provide the required DC resistance for our application, and it would be fairly easy to implement onto the E-shaped patch using a brush. The conductive paint was a PELCO Isopropanol Based Graphite Based Paint distributed by Ted Pella, Inc [163]. This paint provides approximately 2400 Ω/sq for a thickness of 1 mil (25 μm), which is the approximate thickness on average for

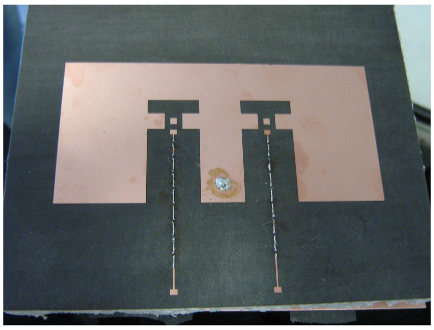


Figure 6.35: Fabricated E-shaped patch antenna with metallic bias lines connected by 8 SMT 10 kΩ resistors, where the total line resistance should be roughly 80 kΩ

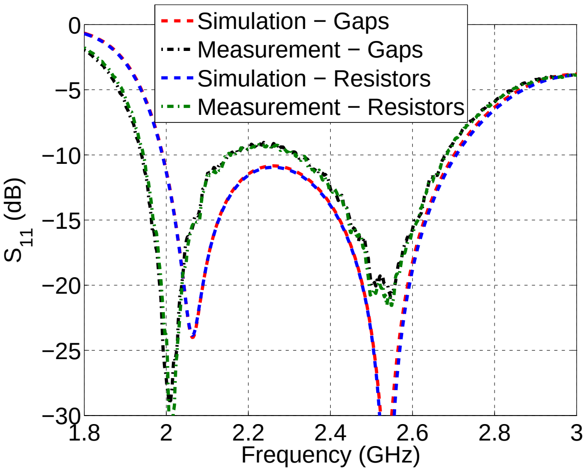


Figure 6.36: Comparison of the impedance matching performance simulation and measurement results for 2 cases: the E-shaped patch design with metallic bias lines having gaps and the same design with 10 kΩ resistors filling gaps

a brush coated application of the paint. This resistance nicely fit with our requirement for high resistance and was also fairly inexpensive. Using supplies around the lab, stencils were created in order to apply the paint via brush onto the Rogers Duroid substrate.

As a test of these lossy lines, an experiment using the same ideal switch E-shaped patch design was fabricated. The design without the lossy bias lines is shown in Figure 6.38b, where four pads are included to for use with the bias line. Before applying the conductive ink, this design was measured as a reference measurement without the lossy lines. Next, the lossy lines were applied to the design using the stencils to trace out the lines. After applying, the ink must cure for 30 minutes in order fully bond to the substrate and to dry completely. The final absolute resistance was roughly 200 kΩ for both lines, which provides the necessary resistance to appear invisible to the RF but allows a DC connection. The final design with the conductive ink is shown in Figure 6.38c. The S_{11} measurements and simulations are shown in Figure 6.37. The simulation and measurement without lines show some discrepancy, but this is somewhat expected due to errors in fabrication common to multilayer structures. However, the general trend shows good agreement between the two. Even further, very good agreement between the measurement without lines and the measurement with the lossy lines is observed. This demonstrates that the lossy lines have a minimal effect on the design in comparison to its performance without the bias lines, and therefore are a viable option for bias lines. This may even be a preferred option due to its ease in application. The resistor bias line tends to be quite messy and challenging to solder SMT chips at extremely small sizes.

In summary, the bias lines represent another challenge to the E-shaped patch design that is typically taken for granted in the reconfigurable antenna research area. The length of the bias lines was the primary factor which degraded the E-shaped performance with the presence of metallic bias lines. Therefore other clever techniques had to be sought. Two viable options were investigated: a resistor bias line structure and a lossy bias line network. Both provide good performance, but the lossy line approach may be superior because it is less challenging to fabricate in comparison to the resistor networks. Overall, these results demonstrate the implementation of the bias lines into the design and ultimately this demonstrates the full practical implementation of both reconfigurable E-shaped patch antenna designs.

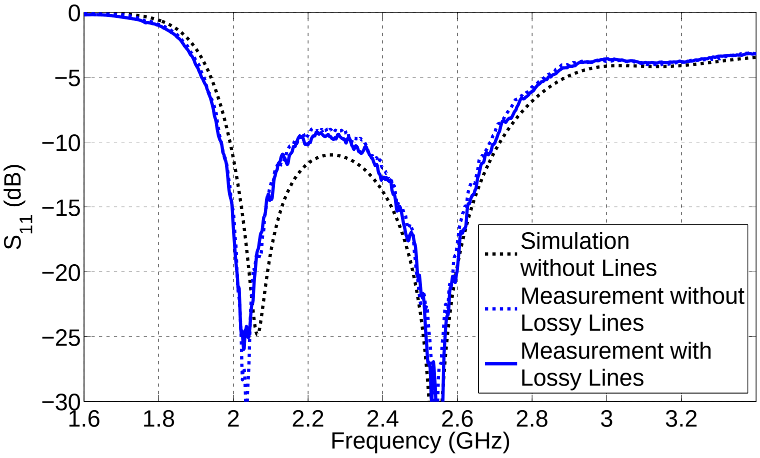
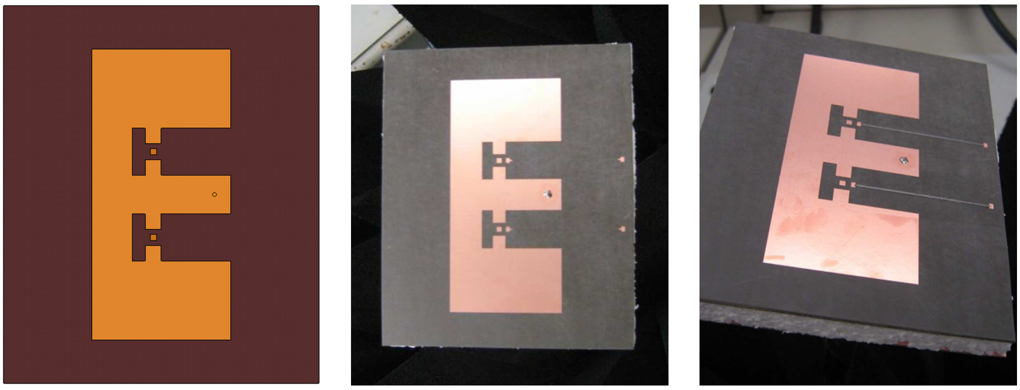


Figure 6.37: Comparison of the impedance matching performance for the simulation with no bias lines (Figure 6.38a), the measurement without bias lines (Figure 6.38b), and the measurement with lossy lines included (Figure 6.38c)



(a) Simulation without Lines **(b)** Measurement without Lines **(c)** Measurement with Lossy Lines

Figure 6.38: Simulated and fabricated models to test lossy bias line networks

6.6 Further Discussion

The E-shaped patch antenna has been a widely celebrated patch antenna design which can provide a wideband performance with a simple methodology. In this chapter, two new reconfigurable E-shaped patch antenna designs have been presented: a RHCP/LHCP reconfigurable E-shaped patch and a frequency reconfigurable E-shaped patch antenna. PSO was applied to both designs in order to optimize performance according to the particular antenna. The final designs, simulations, and measurements proved that PSO did an excellent job in finding good design solutions. The final CP E-shaped patch antenna provided a AR-S_{11} bandwidth of roughly 15.6%, which is quite high for a CP patch antenna in its class. The FR E-shaped patch antenna was able to nearly double its bandwidth from 31% to 48%, by adding the new state which altered the dimensions of the slot. Overall, both designs provided good results.

One consideration that is hoped to achieve in the future is the implementation of every component discussed in this chapter into one fully functional E-shaped patch design. This design would have the ability to switch between linear polarization to circular polarization, and it could switch from one frequency to another. In other words, it would have two switches where all four states were usable states in order to take full advantage of all reconfigurability currently available to the E-shaped patch design. This may turn out to be more difficult than it looks because the two designs shown in this chapter had inherently different characteristics. As noted previously, the MEMS bar position within the slots was one of the key dimensions that was different between the two. For the CP reconfigurable antenna, the bars were placed at roughly $\ell \approx L_s/2$. However, the bars on the FR patch antenna were closer to the other bars which form the E-shape. This may pose problems in trying to find a compromise between each design in order to realize the full picture, but it is still hopeful that this type of design could be achieved.

CHAPTER 7

Conclusions

Indeed, nature-inspired optimization techniques provide a unique tool that can find good designs where other algorithms often fail. They are still newly developed algorithms, and much more research has been going on to improve the various techniques. A discussion was given to provide some background on the role of optimization within the discipline of electromagnetics and antenna engineering. The parameters of interest for optimization of antenna designs were then described based on the idea of improving the antenna link for the average case scenario. In this work, the Particle Swarm Optimization (PSO) and the Covariance Matrix Adaptation Evolution Strategy (CMAES) algorithms were introduced as possible candidates for use in antenna design optimization, and definitions were provided for every aspect of the optimization formulation.

The two algorithms were compared in terms of convergence, efficiency, and robustness. They were first compared by the optimization of several mathematical fitness functions. The goal of this optimization was to provide some insight on their convergence in resource limited settings for different types functions which have qualities often seen in electromagnetics applications. Both techniques showed good convergence which provided significantly better fitness. In general, CMAES showed faster convergence than PSO for many of the functions, and required a significantly less number of function evaluations for the optimization due to a smaller population size. There were also some functions that were difficult to optimize, and the multimodal problems often posed a challenge to both algorithms. There were a few cases where PSO demonstrated better robustness by achieving better fitness value at the end of the run.

The next comparison was conducted on a standard array problem to minimize the sidelobe level using a non-uniform array technique. Both techniques found nearly identical design solutions which provided a minimal SLL in the sidelobe region. Both were able to converge upon a solution within the given amount of iterations, however CMAES approached the final solution faster than PSO for this particular problem, which took some time to fine-tune the design parameters.

A weather radar antenna was optimized to improve its dual polarization performance for the context of dual polarized weather radar systems, which has been the forefront of research in atmospheric radar technology. This provided a simple but challenging real-world optimization problem to compare PSO and CMAES. Two different optimization formulations were discussed in the context of providing equivalent optimization results. Both PSO and CMAES demonstrated good results for both optimization problems, however the CMAES algorithm had some trouble with the minimization of the S_{11} for the specified frequency band for one run. The difference in the optimization of the cross polarization in the principle planes ($\phi = 0, 45, 90, 135^\circ$ planes) versus the direct optimization of the bias weighting factors was also discussed. From the results it was concluded that for small antennas a minimization of the crosspolar fields in the principle planes is fairly equivalent to the direct optimization of the bias weighting factors.

The next two chapters then provided two other antenna designs which were optimized by PSO. In Chapter 5, both a stepped septum and a sigmoid septum were introduced for use within a circular waveguide horn. The role of these septums was to produce a circular polarization from a linear one in order to radiate highly circular waves. The optimization was designed to minimize the impedance matching and axial ratio over a specified band. These designs proved to be a challenging optimization problem in which PSO was able to reduce the fitness. Both designs were evaluated via simulations and showed similar performance, although there were areas where each design performed better than the other.

The last chapter investigated possible reconfigurable antenna designs for E-shaped patch antennas for use in wireless applications or even cognitive radio. The E-shaped patch antenna concept was introduced, and then a simple MEMS circuit model was provided for rapid optimization. First, a polarization reconfigurable (RHCP/LHCP) E-shaped patch antenna was introduced, and its design was optimized using PSO. Next, a frequency reconfigurable E-shaped patch antenna design was optimized also using PSO. Good design performance was realized by both designs. The bias network design proved to be another challenging problem due to resonant coupling, however a final design solution was provided in order to practically implement either reconfigurable E-shaped patch design.

This work was yet another demonstration of the wide applicability of nature-inspired optimization techniques to a practical discipline interested in obtaining a final design solution. Even though the fitness functions are often ill-conditioned, multimodal, and nonlinear, the algorithms demonstrated fast convergence for the problems discussed in resource limited settings. Both techniques (PSO and CMAES) have also been discussed in terms of their applicability to antenna design problems, and each have their advantages. Ultimately, it is important to keep current with the newly proposed algorithms within the evolutionary computation field, and collaboration between researchers in microwave engineering as well as evolutionary computation is highly encouraged in order to incorporate the needs of practical design optimization problems.

REFERENCES

[1] S. Parkvall and D. Astely, “The evolution of LTE towards IMT-Advanced,” *Journal of Communications*, vol. 4, no. 3, 2009.

[2] T.-H. Lee and T. Choi, “Self-powered wireless communication platform for disaster relief,” in *2011 13th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 1 –3, September 2011.

[3] UNICEF, “Bee system prototypes represent the future of connectivity in emergencies.” `http://www.unicef.org/emerg/index_45259.html`, August 2008.

[4] H. Hariyanto, H. Santoso, and A. Widiawan, “Emergency broadband access network using low altitude platform,” in *2009 International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME)*, pp. 1 –6, November 2009.

[5] S. Haykin, *Communication Systems.* John Wiley and Sons, 2001.

[6] NASA, “CloudSat Mission, Overview, and Instrumentation.” `http://cloudsat.atmos.colostate.edu/`, January 2012.

[7] NASA, “GRACE Overview.” `http://www.csr.utexas.edu/grace/overview.html`, August 2010.

[8] O. B. Andersen, S. I. Seneviratne, J. Hinderer, and P. Viterbo, “GRACE-derived terrestrial water storage depletion associated with the 2003 European heat wave,” *Geophysical Research Letters*, vol. 32, September 2005.

[9] NASA, “Aquarius Mission.” `http://aquarius.nasa.gov`, March 2012.

[10] NASA, “Soil Moisture Active Passive Mission.” `http://smap.jpl.nasa.gov/instrument/`, March 2012.

[11] NASA, “Gravity Recovery and Interior Laboratory Mission.” `http://solarsystem.nasa.gov/grail`, October 2011.

[12] C. Balanis, *Advanced Engineering Electromagnetics.* Wiley, 1989.

[13] R. W. P. King, *Theory of Linear Antennas.* Cambridge: Harvard University Press, 1956.

[14] S. A. Schelkunoff and H. T. Friis, *Antennas: Theory and Practice.* John Wiley and Sons, 1952.

[15] C. A. Balanis, *Antenna Theory: Analysis and Design.* Wiley-Interscience, 3 ed., 2005.

[16] R. Harrington, *Field Computation by Moment Methods.* Wiley-IEEE Press, 1993.

[17] W. C. Gibson, *The Method of Moments in Electromagnetics.* Chapman and Hall, CRC, 2008.

[18] M. Ney, “Method of Moments as Applied to Electromagnetic Problems,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 33, pp. 972 – 980, October 1985.

[19] J. Volakis, A. Chatterjee, and L. Kempel, *Finite Element Method Electromagnetics: Antennas, Microwave Circuits, and Scattering Applications.* IEEE-Wiley, 1998.

[20] P. P. Silvester and G. Pelosi, *Finite Elements For Wave Electromagnetics.* IEEE Press, 1994.

[21] T. Itoh, *Numerical Techniques for Microwave and Millimeter-Wave Passive Structures.* Wiley, 1989.

[22] K. Kunz and R. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics.* CRC Press, 1993.

[23] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method.* Artech House Publishers, 3 ed., 2005.

[24] W. Yu, R. Mittra, T. Su, Y. Liu, and X. Yang, *Parallel Finite-Difference Time Domain Method.* Artech House Publishers, 2006.

[25] S. Mikki and A. Kishk, “Theory and Applications of Infinitesimal Dipole Models for Computational Electromagnetics,” *IEEE Transactions on Antennas and Propagation*, vol. 55, pp. 1325 –1337, May 2007.

[26] S. Karimkashi, A. Kishk, and G. Zhang, “Application of Invasive Weed Optimization in designing aperiodic antenna arrays using dipole models,” in *2011 IEEE International Symposium on Antennas and Propagation*, pp. 2437 –2440, July 2011.

[27] S. Karimkashi, A. Kishk, and D. Kajfez, “Antenna Array Optimization Using Dipole Models for MIMO Applications,” *IEEE Transactions on Antennas and Propagation*, vol. 59, pp. 3112 –3116, August 2011.

[28] Y. Rahmat-Samii, J. M. Kovitz, and H. Rajagopalan, “Nature Inspired Optimization Techniques in Communication Antenna Designs,” *Proceedings of the IEEE*, July 2012. In print for the July 2012 Issue.

[29] K. Parsopoulos and M. Vrahatis, “Particle swarm optimization method for constrained optimization problems,” in *Proceedings of the Euro-International Symposium on Computational Intelligence*, 2002.

[30] J. Bandler, “Optimization Methods for Computer-Aided Design,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 17, pp. 533 – 552, August 1969.

[31] C. Charalambous, “A Unified Review of Optimization,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 22, pp. 289 – 300, March 1974.

[32] D. Cheng, "Optimization techniques for antenna arrays," *Proceedings of the IEEE*, vol. 59, pp. 1664 – 1674, December 1971.

[33] W. E. Smith, H. H. Barrett, and R. G. Paxman, "Reconstruction of objects from coded images by simulated annealing," *Optics Letters*, vol. 8, pp. 199–201, April 1983.

[34] W. E. Smith, R. G. Paxman, and H. H. Barrett, "Image reconstruction from coded data: I. Reconstruction algorithms and experimental results," *Journal of the Optical Society of America*, vol. 2, pp. 491–500, April 1985.

[35] H. Haneishi, T. Masuda, N. Ohyama, T. Honda, and J. Tsujiuchi, "Analysis of the cost function used in simulated annealing for CT image reconstruction," *Applied Optics*, vol. 29, pp. 259–265, January 1990.

[36] L. Garnero, A. Franchois, J.-P. Hugonin, C. Pichot, and N. Joachimowicz, "Microwave imaging-complex permittivity reconstruction-by simulated annealing," *IEEE Transactions on Microwave Theory and Techniques*, vol. 39, pp. 1801 –1807, November 1991.

[37] T. Cornwell, "A novel principle for optimization of the instantaneous Fourier plane coverage of correction arrays," *IEEE Transactions on Antennas and Propagation*, vol. 36, pp. 1165 –1167, aug 1988.

[38] K. Preis and A. Ziegler, "Optimal Design of Electromagnetic Devices with Evolution Strategies," in *3-D electromagnetic field analysis: proceedings of the International Symposium and TEAM Workshop*, pp. 91–92, September 1989.

[39] A. Gottvald, "Optimal magnet design for NMR," *IEEE Transactions on Magnetics*, vol. 26, pp. 399 –402, mar 1990.

[40] K. Preis, C. Magele, and O. Biro, "FEM and evolution strategies in the optimal design of electromagnetic devices," *IEEE Transactions on Magnetics*, vol. 26, pp. 2181 –2183, sep 1990.

[41] E. Michielssen, J.-M. Sajer, S. Ranjithan, and R. Mittra, "Design of lightweight, broadband microwave absorbers using genetic algorithms," *IEEE Transactions on Microwave Theory and Techniques*, vol. 41, pp. 1024 –1031, June/July 1993.

[42] J. Johnson and V. Rahmat-Samii, "Genetic algorithms in engineering electromagnetics," *IEEE Antennas and Propagation Magazine*, vol. 39, pp. 7 –21, August 1997.

[43] D. Weile and E. Michielssen, "Genetic algorithm optimization applied to electromagnetics: a review," *IEEE Transactions on Antennas and Propagation*, vol. 45, pp. 343 –353, March 1997.

[44] R. Haupt, "An introduction to genetic algorithms for electromagnetics," *IEEE Antennas and Propagation Magazine*, vol. 37, pp. 7 –15, April 1995.

[45] A. Tennant, M. Dawoud, and A. Anderson, "Array pattern nulling by element position perturbations using a genetic algorithm," *Electronics Letters*, vol. 30, pp. 174 –176, February 1994.

[46] R. Haupt, “Phase-only adaptive nulling with a genetic algorithm,” *IEEE Transactions on Antennas and Propagation*, vol. 45, pp. 1009 –1015, June 1997.

[47] R. Haupt, “Thinned arrays using genetic algorithms,” *IEEE Transactions on Antennas and Propagation*, vol. 42, pp. 993 –999, July 1994.

[48] N. Jin and Y. Rahmat-Samii, “Hybrid Real-Binary Particle Swarm Optimization (HPSO) in Engineering Electromagnetics,” *IEEE Transactions on Antennas and Propagation*, vol. 58, pp. 3786 –3794, December 2010.

[49] Z. Altman, R. Mittra, and A. Boag, “New designs of ultra wide-band communication antennas using a genetic algorithm,” *IEEE Transactions on Antennas and Propagation*, vol. 45, pp. 1494 –1501, October 1997.

[50] L. Lizzi, F. Viani, R. Azaro, and A. Massa, “Optimization of a Spline-Shaped UWB Antenna by PSO,” *IEEE Antennas and Wireless Propagation Letters*, vol. 6, pp. 182 –185, 2007.

[51] D. Gies, “Particle swarm optimization: applications in electromagnetic design,” Master’s thesis, University of California Los Angeles, 2004.

[52] S. Xu, Y. Rahmat-Samii, and D. Gies, “Shaped-reflector antenna designs using particle swarm optimization: An example of a direct-broadcast satellite antenna,” *Microwave and Optical Technology Letters*, vol. 48, no. 7, pp. 1341–1347, 2006.

[53] E. Jones and W. Joines, “Design of Yagi-Uda antennas using genetic algorithms,” *IEEE Transactions on Antennas and Propagation*, vol. 45, pp. 1386 –1392, September 1997.

[54] Z. Bayraktar, P. Werner, and D. Werner, “The Design of Miniature Three-Element Stochastic Yagi-Uda Arrays Using Particle Swarm Optimization,” *IEEE Antennas and Wireless Propagation Letters*, vol. 5, pp. 22 –26, December 2006.

[55] H. Mosallaei and Y. Rahmat-Samii, “RCS reduction of canonical targets using genetic algorithm synthesized RAM,” *IEEE Transactions on Antennas and Propagation*, vol. 48, pp. 1594 –1606, October 2000.

[56] J. Kovitz and Y. Rahmat-Samii, “Micro-actuated pixel patch antenna design using particle swarm optimization,” in *2011 IEEE International Symposium on Antennas and Propagation (APSURSI)*, pp. 2415 –2418, July 2011.

[57] S. Chakravarty and R. Mittra, “Design of a frequency selective surface (FSS) with very low cross-polarization discrimination via the parallel micro-genetic algorithm (PMGA),” *IEEE Transactions on Antennas and Propagation*, vol. 51, pp. 1664 – 1668, July 2003.

[58] A. Fallahi, M. Mishrikey, C. Hafner, and R. Vahldieck, “Efficient Procedures for the Optimization of Frequency Selective Surfaces,” *IEEE Transactions on Antennas and Propagation*, vol. 56, pp. 1340 –1349, May 2008.

[59] N. Jin and Y. Rahmat-Samii, "Parallel particle swarm optimization and finite- difference time-domain (PSO/FDTD) algorithm for multiband and wide-band patch antenna designs," *IEEE Transactions on Antennas and Propagation*, vol. 53, pp. 3459 – 3468, November 2005.

[60] W.-C. Liu, "Design of a multiband CPW-fed monopole antenna using a particle swarm optimization approach," *IEEE Transactions on Antennas and Propagation*, vol. 53, pp. 3273 – 3279, October 2005.

[61] R. McDonough and A. Whalen, *Detection of Signals in Noise.* Academic Press, 2 ed., 1995.

[62] L. C. Shen and J. A. Kong, *Applied Electromagnetism.* PWS Publishers, 2 ed., 1987.

[63] D. Pozar, *Microwave Engineering.* John Wiley, 3 ed., 2005.

[64] Ansoft, "Ansoft HFSS," 2007.

[65] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge University Press, March 2004.

[66] Y.-W. Shang and Y.-H. Qiu, "A Note on the Extended Rosenbrock Function," *Evolutionary Computation*, vol. 14, pp. 119–126, March 2006.

[67] M. Locatelli, "A Note on the Griewank Test Function," *Journal of Global Optimization*, vol. 25, pp. 169–174, February 2003.

[68] N. Hansen, "Tutorial: Covariance Matrix Adaptation (CMA) Evolution Strategy." `http://www.lri.fr/~hansen/handout2006.pdf`, September 2006.

[69] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4, pp. 1942 –1948 vol.4, November/December 1995.

[70] J. Robinson and Y. Rahmat-Samii, "Particle swarm optimization in electromagnetics," *IEEE Transactions on Antennas and Propagation*, vol. 52, pp. 397 – 407, February 2004.

[71] N. Jin, "Particle swarm optimization in engineering electromagnetics," Master's thesis, University of California Los Angeles, 2007.

[72] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *The 1998 IEEE International Conference on Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence.*, pp. 69 –73, May 1998.

[73] Y. Shi and R. Eberhart, "Empirical study of particle swarm optimization," in *Proceedings of the 1999 Congress on Evolutionary Computation*, vol. 3, p. 3, 1999.

[74] Y. Shi and R. Eberhart, "Parameter selection in particle swarm optimization," in *Evolutionary Programming VII* (V. Porto, N. Saravanan, D. Waagen, and A. Eiben, eds.), vol. 1447 of *Lecture Notes in Computer Science*, pp. 591–600, Springer Berlin / Heidelberg, 1998.

[75] N. Jin and Y. Rahmat-Samii, "Advances in Particle Swarm Optimization for Antenna Designs: Real-Number, Binary, Single-Objective and Multiobjective Implementations," *IEEE Transactions on Antennas and Propagation*, vol. 55, pp. 556 –567, March 2007.

[76] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pp. 39 –43, October 1995.

[77] P. Angeline, "Using selection to improve particle swarm optimization," in *The 1998 IEEE International Conference on Evolutionary Computation Proceedings*, IEEE World Congress on Computational Intelligence, pp. 84 –89, May 1998.

[78] S. Xu and Y. Rahmat-Samii, "Boundary Conditions in Particle Swarm Optimization Revisited," *IEEE Transactions on Antennas and Propagation*, vol. 55, pp. 760 –765, March 2007.

[79] T. Bck, F. Hoffmeister, and H.-P. Schwefel, "A Survey of Evolution Strategies," in *Proceedings of the Fourth International Conference on Genetic Algorithms*, pp. 2–9, Morgan Kaufmann, 1991.

[80] H.-G. Beyer and H.-P. Schwefel, "Evolution Strategies - A Comprehensive Introduction," vol. 1, pp. 3–52, May 2002.

[81] M. Gregory, Z. Bayraktar, and D. Werner, "Fast Optimization of Electromagnetic Design Problems Using the Covariance Matrix Adaptation Evolutionary Strategy," *IEEE Transactions on Antennas and Propagation*, vol. 59, pp. 1275 –1285, April 2011.

[82] H.-P. Schwefel and G. Rudolph, "Contemporary Evolution Strategies," pp. 893–907, Springer-Verlag, 1995.

[83] N. Hansen, "The CMA Evolution Strategy: A Tutorial." `http://www.lri.fr/~hansen/cmatutorial.pdf`, June 2011.

[84] N. Hansen and A. Ostermeier, "Completely Derandomized Self-Adaptation in Evolution Strategies," *Evolutionary Computation*, vol. 9, pp. 159–195, 2001.

[85] Z. Michalewicz and M. Schoenauer, "Evolutionary Algorithms for Constrained Parameter Optimization Problems," *Evolutionary Computation*, vol. 4, pp. 1–32, 1996.

[86] K. E. Parsopoulos and M. N. Vrahatis, "Particle Swarm Optimization Method for Constrained Optimization Problems," in *In Proceedings of the Euro-International Symposium on Computational Intelligence 2002*, pp. 214–220, IOS Press, 2002.

[87] Z. Michalewicz, “A survey of constraint handling techniques in evolutionary computation methods,” in *Proceedings of the 4th Annual Conference on Evolutionary Programming*, pp. 135–155, MIT Press, 1995.

[88] S. Koziel and Z. Michalewicz, “Evolutionary Algorithms, Homomorphous Mappings, and Constrained Parameter Optimization,” *Evolutionary Computation*, vol. 7, p. pp., 1999.

[89] C. A. C. Coello, “Theoretical and Numerical Constraint-Handling Techniques used with Evolutionary Algorithms: A Survey of the State of the Art,” 2002.

[90] Z. Michalewicz and M. Schoenauer, “Evolutionary Algorithms for Constrained Parameter Optimization Problems,” *Evolutionary Computation*, vol. 4, pp. 1–32, 1996.

[91] X. Hu, R. Eberhart, and Y. Shi, “Engineering optimization with particle swarm,” in *Proceedings of the 2003 IEEE Swarm Intelligence Symposium*, pp. 53 – 57, April 2003.

[92] N. Hansen and S. Kern, “Evaluating the CMA Evolution Strategy on Multimodal Test Functions,” in *Parallel Problem Solving from Nature - PPSN VIII* (X. Yao, E. Burke, J. Lozano, J. Smith, J. Merelo-Guervs, J. Bullinaria, J. Rowe, P. Tino, A. Kabn, and H.-P. Schwefel, eds.), vol. 3242 of *Lecture Notes in Computer Science*, pp. 282–291, Springer Berlin / Heidelberg, 2004.

[93] A. Hoorfar. Private Communication. 2011.

[94] N. Hansen, S. Finck, R. Ros, and A. Auger, “Real-Parameter Black-Box Optimization Benchmarking 2009: Noiseless Functions Definitions,” Research Report RR-6829, INRIA, 2009.

[95] M. Khodier and C. Christodoulou, “Linear Array Geometry Synthesis With Minimum Sidelobe Level and Null Control Using Particle Swarm Optimization,” *IEEE Transactions on Antennas and Propagation*, vol. 53, pp. 2674 – 2679, August 2005.

[96] NOAA, “Polarimetric Doppler Radar.” `http://www.nssl.noaa.gov/research/radar/dualpol.php`.

[97] “Dual Polarisation Radar.” `http://www.radartutorial.eu/15.weather/wx23.en.html`.

[98] D. Rice, “U.S. weather radar network gets upgrade.” `http://www.usatoday.com/weather/news/story/2012-01-17/national-weather-service-new-radar-system/52746738/1`, January 2012.

[99] Y. Wang and V. Chandrasekar, “Polarization Isolation Requirements for Linear Dual-Polarization Weather Radar in Simultaneous Transmission Mode of Operation,” *IEEE Transactions on Geoscience Remote Sensing*, no. 44, pp. 2019–2028, 2006.

[100] K. V. Beard and C. Chuang, “A New Model for the Equilibrium Shape of Raindrops,” *Journal of the Atmospheric Sciences*, vol. 44, no. 11, pp. 1509–1524, 1987.

[101] “Differential Reflectivity from Raindrops and WSR-88D Rainfall Estimation.” `http://www.chill.colostate.edu/w/Differential_Reflectivity_from_Raindrops_and_WSR-88D_Rainfall_Estimation`.

[102] L. Liu, V. N. Bringi, V. Chandrasekar, E. A. Mueller, and A. Mudukutore, “Analysis of the Copolar Correlation Coefficient between Horizontal and Vertical Polarizations,” *Journal of Atmospheric and Oceanic Technology*, vol. 11, pp. 950–963, August 1994.

[103] G. Zhang, R. J. Doviak, D. S. Zrnic, R. Palmer, L. Lei, and Y. Al-Rashid, “Polarimetric phased-array radar for weather measurement: a planar or cylindrical configuration?,” *Journal of Atmospheric and Oceanic Technology, pages = 63-73,*, vol. 28, 2011.

[104] D. Zrnic, R. Doviak, G. Zhang, and A. Ryzhkov, “Bias in Differential Reflectivity due to Cross Coupling through the Radiation Patterns of Polarimetric Weather Radars,” *Journal of Atmospheric and Oceanic Technology*, no. 27, pp. 1624–1637, 2009.

[105] A. Ludwig, “The definition of cross polarization,” *IEEE Transactions on Antennas and Propagation*, vol. 21, pp. 116 – 119, January 1973.

[106] G. Zhang. Private Communication. 2011.

[107] S. Bhardwaj, “Design and Development of Dual Polarized, Stacked Patch Antenna Element for S-Band Dual-Pol Weather Radar Array,” Master’s thesis, University of California Los Angeles, 2012.

[108] S. Bhardwaj. Private Communication. 2011.

[109] M. Sanad and N. Hassan, “Compact Internal Multi-Band Microstrip Antennas for Portable GPS, PCS, Cellular and Satellite Phones,” *Microwave Journal*, vol. 42, pp. 90–98, August 1999.

[110] E. Kaivanto, M. Berg, E. Salonen, and P. de Maagt, “Wearable Circularly Polarized Antenna for Personal Satellite Communication and Navigation,” *IEEE Transactions on Antennas and Propagation*, vol. 59, pp. 4490–4496, December 2011.

[111] Z. Li and Y. Rahmat-Samii, “Optimization of PIFA-IFA combination in handset antenna designs,” *IEEE Transactions on Antennas and Propagation*, vol. 53, pp. 1770 – 1778, May 2005.

[112] G. Coutts, “Wideband Diagonal Quadruple-Ridge Orthomode Transducer for Circular Polarization Detection,” *IEEE Transactions on Antennas and Propagation*, vol. 59, pp. 1902 –1909, June 2011.

[113] S. Skinner and G. James, “Wide-band orthomode transducers,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 39, pp. 294 –300, February 1991.

[114] P. Lecia and M. Kasal, “X band septum polarizer as feed for parabolic antenna,” in *2010 15th International Conference on Microwave Techniques (COMITE)*, pp. 35–38, April 2010.

[115] D. Davis, O. Digiondomenico, and J. Kempic, “A new type of circularly polarized antenna element,” in *Antennas and Propagation Society International Symposium, 1967*, vol. 5, pp. 26 – 33, October 1967.

[116] M. Chen and G. Tsandoulas, “A wide-band square-waveguide array polarizer,” *IEEE Transactions on Antennas and Propagation*, vol. 21, pp. 389 – 391, may 1973.

[117] M. Franco, “A High-Performance Dual-Mode Feed Horn for Parabolic Reflectors with a Stepped-Septum Polarizer in a Circular Waveguide [Antenna Designer’s Notebook],” *IEEE Antennas and Propagation Magazine*, vol. 53, pp. 142 –146, June 2011.

[118] I. Kim and Y. Rahmat-Samii, “Revisiting stepped septum circular polarizer using full-wave simulations,” in *2011 IEEE International Symposium on Antennas and Propagation (APSURSI)*, pp. 919 –921, July 2011.

[119] N. Albertsen and P. Skov-Madsen, “A Compact Septum Polarizer,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 31, pp. 654 –660, August 1983.

[120] J. Bornemann and V. Labay, “Ridge waveguide polarizer with finite and stepped-thickness septum,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, pp. 1782 – 1787, August 1995.

[121] T. Ege and P. Mcandrew, “Analysis of stepped septum polarisers,” *Electronics Letters*, vol. 21, pp. 1166 –1168, 21 1985.

[122] S.-W. Wang, C.-H. Chien, C.-L. Wang, and R.-B. Wu, “A circular polarizer designed with a dielectric septum loading,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, pp. 1719 – 1723, July 2004.

[123] G. Schennum and T. Skiver, “Antenna feed element for low circular cross-polarization,” in *Proceedings of the IEEE Aerospace Conference*, vol. 3, pp. 135 –150 vol.3, February 1997.

[124] S. Hasan, “Design amp; measurement techniques for innovative, high performance, circularly polarized, ultra wideband corrugated horn antenna with septum polarizer for space applications,” in *2011 IEEE International Conference on Microwave Technology Computational Electromagnetics (ICMTCE)*, pp. 323 –326, May 2011.

[125] B. Streetman, *Solid state electronic devices.* Prentice-Hall series in solid state physical electronics, Prentice-Hall, 1972.

[126] “File:Circular.Polarization.Circularly.Polarized.Light Without.Components Right.Handed.svg.” `http://en.wikipedia.org/wiki/File:Circular.Polarization.Circularly.Polarized.Light_Without.Components_Right.Handed.svg`, March 2010. Wikimedia Commons File.

[127] “File:Circular.Polarization.Circularly.Polarized.Light Without.Components Left.Handed.svg.” `http://en.wikipedia.org/wiki/File:Circular.Polarization.Circularly.Polarized.Light_Without.Components_Left.Handed.svg`, March 2010. Wikimedia Commons File.

[128] J. T. Bernhard, *Reconfigurable Antennas.* Antennas and Propagation Series, Morgan & Claypool Publishers, 2007.

[129] T. S. Rappaport, *Wireless Communications: Principles and Practice.* Prentice Hall communications engineering and emerging technologies series, Pearson Education, 2009.

[130] F. Yang, X.-X. Zhang, X. Ye, and Y. Rahmat-Samii, "Wide-band E-shaped patch antennas for wireless communications," *IEEE Transactions on Antennas and Propagation*, vol. 49, pp. 1094 –1100, July 2001.

[131] D. R. Jackson, *Antenna Engineering Handbook*, ch. Microstrip Antennas. McGraw Hill, 2007.

[132] S. Long and M. Walton, "A dual-frequency stacked circular-disc antenna," *IEEE Transactions on Antennas and Propagation*, vol. 27, pp. 270 – 273, March 1979.

[133] K. Luk, C. Mak, Y. Chow, and K. Lee, "Broadband microstrip patch antenna," *Electronics Letters*, vol. 34, pp. 1442 –1443, July 1998.

[134] T. Huynh and K.-F. Lee, "Single-layer single-patch wideband microstrip antenna," *Electronics Letters*, vol. 31, pp. 1310–1312, August 1995.

[135] S. Weigand, G. Huff, K. Pan, and J. Bernhard, "Analysis and design of broad-band single-layer rectangular U-slot microstrip patch antennas," *IEEE Transactions on Antennas and Propagation*, vol. 51, pp. 457 – 468, March 2003.

[136] N. Jin and Y. Rahmat-Samii, "Parallel particle swarm optimization and finite- difference time-domain (PSO/FDTD) algorithm for multiband and wide-band patch antenna designs," *IEEE Transactions on Antennas and Propagation*, vol. 53, pp. 3459 – 3468, November 2005.

[137] S. Bhardwaj and Y. Rahmat-Samii, "A comparative study of c-shaped, e-shaped, and u-slotted patch antennas," *Microwave and Optical Technology Letters*, vol. 54, no. 7, pp. 1746–1757, 2012.

[138] G. M. Rebeiz, *RF MEMS: Theory, Design and Technology.* New York: Wiley, 2003.

[139] RadantMEMS, "RMSW100HP Datasheet." `http://www.radantmems.com/radantmems.data/Library/Radant-Datasheet100HP_1.4.pdf`, 2012.

[140] H. Rajagopalan, Y. Rahmat-Samii, and W. Imbriale, "RF MEMS Actuated Reconfigurable Reflectarray Patch-Slot Element," *IEEE Transactions on Antennas and Propagation*, vol. 56, pp. 3689 –3699, December 2008.

[141] K. Entesari, K. Obeidat, A. Brown, and G. Rebeiz, "A 25-75 MHz RF MEMS Tunable Filter," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, pp. 2399 –2405, November 2007.

[142] A. Kajiwara, "Line-of-sight indoor radio communication using circular polarized waves," *IEEE Transactions on Vehicular Technology*, vol. 44, pp. 487 –493, August 1995.

[143] T. Rappaport and D. Hawbaker, “Wide-band microwave propagation parameters using circular and linear polarized antennas for indoor wireless channels,” *IEEE Transactions on Communications*, vol. 40, pp. 240 –245, February 1992.

[144] K.-F. Tong and T.-P. Wong, “Circularly Polarized U-Slot Antenna,” *IEEE Transactions on Antennas and Propagation*, vol. 55, pp. 2382 –2385, August 2007.

[145] A. Khidre, K. F. Lee, F. Yang, and A. Eisherbeni, “Wideband Circularly Polarized E-Shaped Patch Antenna for Wireless Applications ,” *IEEE Antennas and Propagation Magazine*, vol. 52, pp. 219 –229, October 2010.

[146] J. Kovitz, H. Rajagopalan, and Y. Rahmat-Samii, “A Novel Optimized Broadband Reconfigurable RHCP/LHCP E-shaped Patch Antenna,” in *2012 IEEE International Symposium on Antennas and Propagation*, July 2012. Accepted for presentation.

[147] J. Kovitz, H. Rajagopalan, and Y. Rahmat-Samii, “Next Generation MEMS Reconfigurable E-Shaped Patch Antenna Design using Particle Swarm Optimization,” in *2012 North American Radio Science Meeting (URSI)*, January 2012.

[148] Y.-C. Liang, K.-C. Chen, G. Li, and P. Mahonen, “Cognitive Radio Networking and Communications: An Overview,” *IEEE Transactions on Vehicular Technology*, vol. 60, pp. 3386 –3407, September 2011.

[149] B. Wang and K. Liu, “Advances in cognitive radio networks: A survey,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, pp. 5 –23, February 2011.

[150] E. Ebrahimi and P. Hall, “A dual port wide-narrowband antenna for cognitive radio,” in *3rd European Conference on Antennas and Propagation*, pp. 809 –812, March 2009.

[151] P. Hall, P. Gardner, J. Kelly, E. Ebrahimi, M. Hamid, F. Ghanem, F. Herraiz-Martinez, and D. Segovia-Vargas, “Reconfigurable antenna challenges for future radio systems,” in *3rd European Conference on Antennas and Propagation*, pp. 949 –955, March 2009.

[152] Y. Tawk, M. Bkassiny, G. El-Howayek, S. Jayaweera, K. Avery, and C. Christodoulou, “Reconfigurable front-end antennas for cognitive radio applications,” *IET Microwaves, Antennas & Propagation*, vol. 5, no. 8, pp. 985–992, 2011.

[153] J. Freeman, B. Lamberty, and G. Andrews, “Optoelectronically reconfigurable monopole antenna,” *Electronics Letters*, vol. 28, pp. 1502 –1503, July 1992.

[154] C. Panagamuwa, A. Chauraya, and J. Vardaxoglou, “Frequency and beam reconfigurable antenna using photoconducting switches,” *IEEE Transactions on Antennas and Propagation*, vol. 54, pp. 449 – 454, February 2006.

[155] J. Bernhard, E. Kiely, and G. Washington, “A smart mechanically actuated two-layer electromagnetically coupled microstrip antenna with variable frequency, bandwidth, and antenna gain,” *IEEE Transactions on Antennas and Propagation*, vol. 49, pp. 597 –601, April 2001.

[156] J.-C. Langer, J. Zou, C. Liu, and J. Bernhard, “Micromachined reconfigurable out-of-plane microstrip patch antenna using plastic deformation magnetic actuation,” *IEEE Microwave and Wireless Components Letters*, vol. 13, pp. 120 –122, March 2003.

[157] F. Yang and Y. Rahmat-Samii, “Patch antennas with switchable slots (PASS) in wireless communications: concepts, designs, and applications,” *IEEE Antennas and Propagation Magazine*, vol. 47, pp. 13–29, April 2005.

[158] I. Kim and Y. Rahmat-Samii, “RF MEMS Switchable Slot Patch Antenna Integrated With Bias Network,” *IEEE Transactions on Antennas and Propagation*, vol. 59, pp. 4811 –4815, December 2011.

[159] B.-Z. Wang, S. Xiao, and J. Wang, “Reconfigurable patch-antenna design for wideband wireless communication systems,” *IET Microwaves, Antennas, & Propagation*, vol. 1, pp. 414–419, April 2007.

[160] Vishay Intertechnology, Inc., *Frequency Response of Thin Film Chip Resistors*, February 2009. `http://www.vishay.com/docs/60107/freqresp.pdf`.

[161] B. Cetiner and N. Biyikli, “Penta-Band Planar Inverted F-Antenna (PIFA) Integrated by RF-NEMS Switches,” in *17th Biennial University/Government/Industry Micro/Nano Symposium*, pp. 116 –119, July 2008.

[162] B. Cetiner, G. Crusats, L. Jofre, and N. Biyikli, “RF MEMS Integrated Frequency Reconfigurable Annular Slot Antenna,” *IEEE Transactions on Antennas and Propagation*, vol. 58, pp. 626 –632, March 2010.

[163] “Technical Notes, PELCO Conductive Graphite Isopropanol Based.” `http://www.tedpella.com/technote_html/16053%20TN.pdf`.