THE UNIVERSITY
of LIVERPOOL

GENETIC ALGORITHMS IN ANTENNA DESIGN WITH APPLICATIONS TO NVIS SYSTEMS

Thesis submitted in accordance with the
requirements of the University of Liverpool
for the degree of Doctor in Philosophy
by

Wen-Chung Liu

Department of
Electrical Engineering and Electronics

August 1999

To

My Mum and Dad.

My wife Mei-Hua

and

My son Cheng-Han

ABSTRACT

Optimising antennas to obtain specified performance requires both practical experience and analytical skill because of the complexity of the problem. It generally involves a large number of parameters and thus becomes a challenge to manipulate them by using traditional optimization techniques. This study investigates an alternative approach to antenna optimization using the Genetic Algorithm (GA) in conjunction with the Method of Moments (MoM). This allows the antenna configuration that most closely produces the desired electromagnetic characteristics to be synthesised without any prior knowledge on its physical details. In the work described the use of this novel optimization procedure is focused on the design of vehicular antennas for the applications of *Near-Vertical-Incidence Skywave* (NVIS) propagation, which is a mode of propagation relying solely on the ionosphere. The successful application of this technique predicts that optimal NVIS performance can be obtained by reactively loading a defined vehicle-mounted loop antenna with an appropriate capacitive element. The matching network designed using the GA and a transmission-line model for this optimised vehicle-antenna system are also developed.

Acknowledgements

I would like to express my gratitude to my supervisor, Dr B.A. Austin, who has been supervising my work throughout this research programme. His suggestions, comments and advice have always been invaluable. The project support from the Defence Evaluation and Research Agency (DERA) is also gratefully acknowledged.

My special thanks go also to Mr. D.C. Kemp whose assistance in correcting the writing of this thesis could not be forgotten.

Additionally, I must express my gratitude to all members of my family for their full support to take my duty of looking after my parents during these years.

Finally, I would like to dedicate this thesis to the memory of my young brother.

Table of Contents

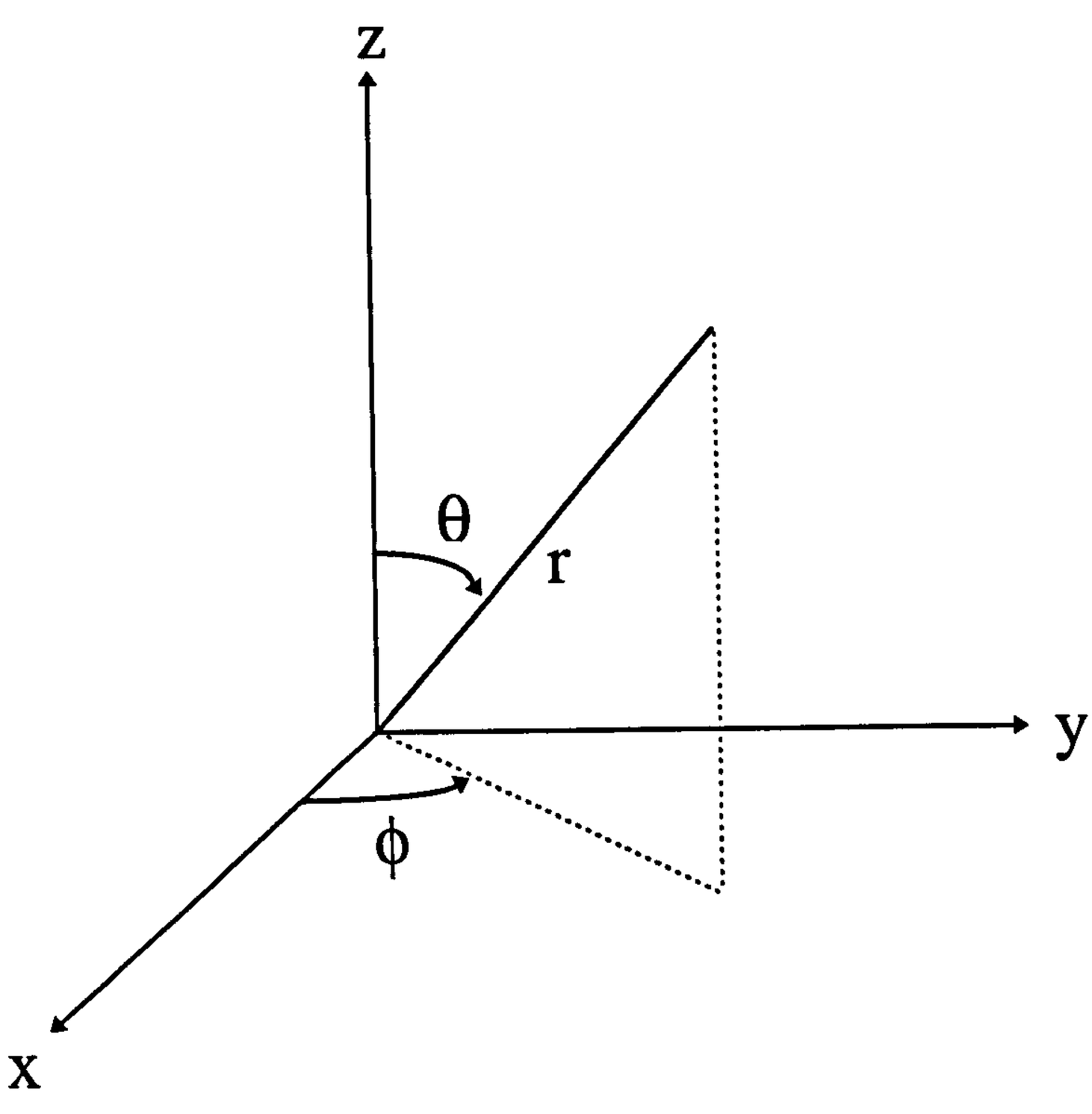
List of Figures

List of Tables

Glossary of Terms

m	meters
mm	millimetre
c	speed of light in free space (2.998×10^{8} m/s)
f	frequency (Hz)
λ	wavelength (m)
ω	angular frequency ($2\pi f$)
k_0, β	wave number ($2\pi/\lambda$)
σ	electrical conductivity (S/m)
ε	electric permittivity (F/m)
ε_0	electric permittivity of free space (8.85×10^{-12} F/m)
μ	magnetic permeability (H/m)
μ_0	magnetic permeability of free space ($4\pi\times10^{-7}$ H/m)
E	electric field intensity (V/m)
H	magnetic field intensity (A/m)
GA	Genetic Alogorithm
F/B	Front-to-Back Ratio
MoM	Method of Moments
NVIS	Near-Vertical-Incidence Skywave

The Coordinate System



CHAPTER 1

INTRODUCTION

Optimising antennas to obtain some desired performance requires both practical experience and analytical skills because of the general complexity of the problem. The shape, size, loads and the matching networks used to transfer the power can significantly affect the antenna's ability to meet a set of specified performance requirements. These include gain, radiation pattern, front-to-back ratio (F/B), relative sidelobe level (rsll) and input impedance (or VSWR). Traditional optimisation techniques such as the gradient descent method search for the desired solution by surveying a region of the domain-space around an initial guess, while the random walk method does the same by a complete randomisation strategy. Both are well suited to solving problems that have a small number of variables because they are able to search a significant portion of the entire solution-space and thus frequently achieve a local optimum solution. However, electromagnetic optimisation problems generally involve a large number of parameters which may be either continuous, discontinuous or both. These often include limitations in the allowable values and sometimes exist within the non-differentiable region of the solution domain. Such problems are frequently the most challenging to traditional optimisation techniques used for antenna investigations.

The ultimate goal of any optimisation problem is to find a solution that represents a *global* maximum or minimum. An optimiser with the capability of implementing some optimisation procedure reliably and efficiently will be of immense benefit in the development of electromagnetic analysis and synthesis techniques. This objective motivates the search for any potential method. Researchers have found such a method based on the process of evolution known as the Genetic Algorithm (GA) which can be applied efficiently to solve electromagnetic problems. A key reason for their importance is their ability to handle a large number of parameters simultaneously. The GA is a completely new approach to antenna design and, in various forms, has been widely applied to optimise performance.

The aims of the research work described in this thesis are three-fold: firstly, to develop general-purpose software to verify the functions of a GA on solution searching; secondly, to combine the GA with the Numerical Electromagnetic code version 2 (NEC2), which is based on the Method of Moments (MoM), to form a powerful antenna-design tool; finally, to explore this novel combination of GA/NEC2 when applied to a variety of antenna optimisation schemes from simple configurations to those mounted on complex structures. The complex structures in this work are land-based vehicles unless otherwise specified. A special requirement, well-suited to the GA/NEC2 approach are the antennas used for *near-vertical-incidence skywave* (NVIS) applications. The desired antenna performance implies maximum gain towards the zenith at HF (3-30 MHz). In addition, the GA has also been used to design the matching network for optimum power transfer to the designed NVIS antenna.

The outline of this thesis will now be described. Chapter 2 discusses the potential of the GA and also includes a review of its application to antenna performance

optimisation that is especially relevant to this research. A number of different antenna configurations optimised using the GA for obtaining specified characteristics will be introduced. These configurations include the monopole, the loop antenna, the phased array and the Yagi-Uda array with or without electrical loads; while the desired characteristics include high gain, maximum rsll and broadband, etc. The ultimate objective of this work is to investigate an NVIS antenna with high performance, therefore, another particular issue that will be also reviewed in this chapter are the special features of NVIS propagation and the antennas already used for this application.

In chapter 3, both the theoretical development and numerical implementation of a common GA are introduced. In particular, issues that affect the efficiency of the GA's implementation are addressed. These include the application of various schemes for the different steps which construct a GA procedure. It is vital that the proper schemes are employed in the GA and this necessitates that some trials are made before determining them. Finally, the chapter concludes with an example to demonstrate the capability of the GA in finding a global maximum from among an ensemble of local maxima. The comparison of results from different schemes are also included.

In this research work, the numerical code NEC2 is used to predict the performance of every GA-optimised antenna. Chapter 4 will therefore provide the theoretical background for this method. It also includes a discussion on the use of a wire-grid model to discretely represent the complex structure, as most of this work is focused on the investigation of an NVIS antenna mounted on this structure. The accuracy of the use of the NEC2 by the author is tested by comparing the predicted results of a selected system which is similar to this work with those simulated using

the Finite-Difference Time-Domain (FDTD) method, a method which solves Maxwell's equations as a function of time in discrete time steps at discrete points in space.

After the accuracy of the use of GA and NEC2 have been confirmed, chapter 5 will therefore concentrate on the practical application of the GA/NEC2 combination for the optimisation of desired antenna characteristics from appropriate configurations. To illustrate their capabilities GA/NEC2 will be used to obtain the maximum gain from a shaped parasitic antenna array and the optimisation of a loop loaded monopole antenna for omnidirectional radiation pattern.

Previously, a substantial amount of NVIS antenna design has been conducted using the Characteristic-Mode (CM) techniques. This approach is complicated, and therefore the use of the novel GA/NEC2 technique to investigate an antenna with optimum NVIS performance is detailed in chapter 6. These results are then compared with those obtained by the CM techniques that are described in literature. Various vehicular antenna configurations such as a whip, a loop and an electrically loaded loop antenna are examined. In addition, the matching network for the resulting antenna is also designed using the GA.

Chapter 7 demonstrates the development of a transmission-line model for the antenna-vehicle system designed in previous chapter. The basic transmission-line theory is used to derive the current distribution along the vehicle-loop structures used here. Also described in this chapter is the prediction of the appropriate load using this model. The results compared to those achieved in previous chapter are particularly good.

Finally, some conclusions on this work and the recommendations for future work are addressed in chapter 8.

There are two appendices in this thesis. Appendix A shows some reported results used for comparison with those obtained from this work. Also shown in the appendix B are four publications that resulted from parts of this research programme.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Ever since antennas were first used, investigation of these electromagnetic radiation devices have been by theoretical development or experimental validation. However, for some special applications there are still many challenges waiting to be overcome. This is the main motivation for antenna designers to direct much of their time to find the methods which are most accurate and efficient to satisfy the design criteria. The vast majority of research effort is related to the optimisation of the antenna's electromagnetic properties such as gain, front-to-back ratio(F/B), side lobe, bandwidth; or the antenna's configuration including the physical size, loads and matching networks. The result of the research is that the design of antennas for special applications has become feasible.

Until now, in antenna design, traditional techniques use the method of moments (MoM) to determine the current distributions on each conductor of a selected antenna configuration and, from these, the radiation characteristics are calculated and compared with a target specification and then by adjusting the factors to make a gradual approaching to the objective. This procedure is sometimes impractical and both experience and some appropriate candidate configurations are usually required to

ensure success. In addition, the electromagnetic optimisation problems are also always very nonlinear, nondifferentiable and lead to highly interactive and expensive design procedures using the conventional techniques such as the down-hill simplex method, the deterministic (gradient-based) optimisation method etc. In contrast to that, the adaptive search technique, the Genetic Algorithm (GA), can be used in conjunction with Numerical Electromagnetic Code (NEC), for a revolutionary new approach to antenna design. It is based on the processes of natural selection and the "survival of the fittest" in an evolutionary system which leads iteratively to that desired specification with minimal foresight or pre-conditioning on the part of the designer The immense power of the technique is its ability to satisfy a performance criterion without any *priori* knowledge of candidate configurations, and the facility for finding the global optimum result.

With the powerful ability to search in a multi-dimensional space, the applications of the GA technique are increasing in electromagnetics as design tools and problem solvers. In the past few years, many successful results have been reported using the GA in designing various antenna configurations satisfying the desired properties. These results can be taken as the evidence for its powerful capability to optimise antennas both faster and in terms of a global set of objectives.

The ultimate objective of this research work is the use of the GA to optimise a vehicular antenna for Near-Vertical-Incidence Skywave (NVIS) propagation. It is therefore vital and necessary to give a brief survey of the current development of antenna design using the GA and the antenna systems used for NVIS propagation before moving on to the subject of this work.

Hence this chapter commences with section 2.2 to introduce some typical examples of antenna optimisation using the GA. Various antennas optimised for different characteristic criteria will be considered in this section. Following is section 2.3 that examines the current NVIS antenna systems developed by either theoretical or experimental methods. In this section the characteristics of NVIS propagation will be introduced first and then the widely used vehicular antennas such as the whip and loop antennas are described.

2.2 Review of Antenna Design Using the Genetic Algorithm

Many applications require radiation characteristics that may not be achievable by a single element. It may, however, be possible that an aggregate of radiating elements in an electrical and geometrical arrangement (an array) will result in the desired radiation characteristics. The arrangement of the array may be such that the radiation from the elements adds up to give a radiation maximum in a particular direction or directions, minimum in others, or otherwise as desired. However, the radiation characteristics of large arrays are difficult to control because there exist many decisive parameters such as the phase of the exciting sources, the exciting elements and the position of the elements. Therefore, the early use of GA for array problem solving was by Tennant et al. (1994). For steering array nulls to the required interference directions, and to achieve any described null depth, he selected the element position along the array axis, or normal to it, as the design parameters and then these were manipulated by the GA. Two cases of a 20 element uniform array of isotropic point sources and a 20 elements Chebyshev array with sidelobe level of -30dB were optimised for producing nulls at the desired locations. The results compared to the analytic results showed that the GA provided superior

accuracy in null location to the analytic method and has maintained the required null depth. The technique also removed the restrictions of relatively small perturbations imposed by the analytic solution.

Haupt has also done much work on the investigation of array antennas using the GA (1994, 1995a,b). In 1994, he used it to determine which elements in both linear and planar arrays, with a total of 200 elements, should be turned off for obtaining the lowest maximum relative sidelobe level (rsll). This strategy is called "thinning the array". The state of each element in the array is either "on" or "off". Here, "on" and "off" represent the element is connected to the feed network and a matched or dummy load, respectively. It is easy to represent this by a binary bit and this very easily satisfies the encoding requirements of a GA. Haupt's implementation experience also found that the elements around the centre were more likely to be excited than those towards the edge in order to achieve the desired lowest rsll. Therefore a special weighting operation was set in the creation of the initial random population to speed the convergence of the algorithm. Finally, the optimal results showed that if the elements were fed by sources with a $|\sin\phi|$ element pattern, for the linear array case a maximum rsll of 23 dB can be reached when 77% of the elements were filled, while for planar array, if 78% of the elements were filled a maximum rsll of 22 dB was achieved. These GA results were particularly interesting because statistical aperiodic array theory predicts that over 1000 elements are needed to obtain a maximum rsll of about 22 dB (Lo et al. 1966, Willey 1962), whereas Haupt used only 200.

In addition, using the GA approach Haupt (1995a,b) also optimised a partially tapered array with the eight centre elements uniformly weighted and spaced and the three edge elements on either side nonuniformly spaced; and the spacings of a 24

element array for maximum rsll. In the former case he found a maximum rsll of 21.5 dB could be achieved by properly positioning the three edge elements while in the latter case the optimised rsll is 22 dB if each element was arranged in an appropriate position predicted by the GA. In addition, the rsll of a circular array has also been optimised by Yan et al. (1997) using the same technique.

Although the resonant wire antenna (such as a monopole) is usually considered to be narrowband, wire antennas were found to have potential broadband characteristics. It has been found that the operating bandwidth of a wire antenna can be enhanced by loading it with lumped circuit element (Rao et al. 1969). Boag et al. (1996) applied the GA to the systematic computer-aided design of wire antennas with lumped loads for broadband purpose. Three different wire antenna configurations including a single monopole, a twin-whip antenna and a folded monopole were examined. In addition to both the component values and positions of the shunt RLC loads, the LC matching network was also designed simultaneously using the GA. This was possible due to the powerful ability of the GA to manipulate a large number of parameters at one time. The operating bands for the three antennas are 30-450 MHz, 2-30 MHz and 2-40 MHz, respectively. Based on their numerical experience, different numbers of loads were added to the these different antennas. The numerical experiment showed that the bandwidths of these wire antennas can be markedly improved by appropriate loads and matching networks. Compared to the conventional methods, they concluded that the GA technique has the capability to lead to an overall optimal solution for the combination of antenna and matching network simultaneously. This is very unlike the conventional method which firstly approaches the optimal antennas and only then turns to the synthesis of the matching network.

Similar to the work of Boag et al., Altman et al. (1997) selected the loaded kite antenna as the candidate configuration to be optimised using the GA for ultra broadband characteristics. Two cases of four and eight arms were both loaded with RLC resonant circuits to enhance the antenna characteristics by yielding a low VSWR by modifying the current distribution. Their GA procedure changed the load components, their locations and the parameters of the matching circuit simultaneously and iteratively until the optimal bandwidth was achieved. Finally, VSWR bandwidth of 7.5:1 and 15:1 were obtained for the loaded kite antennas with four and eight arms, respectively.

Compared to the above wire antennas, a more complex wire antenna system was examined by Weile et al. (1997). In his work the GA was used to devise both a monopole loaded with the RLC shunt circuits and its matching circuit for broadband purposes. The important aspect is the monopole situated on a complex environment - a tank. This considerably increased the difficulties of the numerical experiment. However, the GA worked very well and created an optimal antenna system which produced an acceptable standing wave ratio (SWR) over a band from 9 to 60 MHz.

In addition to the wire antennas the development of broad bandwidth patch antennas has also been reported by Johnson et al. (1997). Using the natural evolution procedure they optimised the bandwidth by recomposing the patch-antenna by presenting or discarding each subsection of metal in the patch and finally obtained a marked improvement in bandwidth. The result showed the bandwidth of the patch antenna was increased from 6% to over a 20% frequency bandwidth.

Another most unusual example of antenna design using the GA is the wire antenna proposed by Linden et al. (1996) and Altshuler et al. (1997). They used this novel GA method to devise an antenna with rather remarkable characteristics. They

required an antenna that would produce hemispherical coverage with circular polarisation - a most demanding specification indeed. Conventional antennas, based on suitably-fed patches or an array of helicals could possibly achieve it (Gotthard et al. 1979) but the GA yielded most acceptable performance using a wire antenna structure of strange geometrical shape shown in figure 2.1. The experimental measurements of a full-scale antenna showed good agreements with the GA-predicted results. In addition, Altshuler et al. also examined some other antenna configurations using the same method. Based on their results using the experimental method to design a monopole loaded with a modified folded dipole for producing a hemisphere coverage, the GA was used to search for the optimal configuration of this type of antenna with the same objective. The optimal configuration of the loaded monopole as shown in figure 2.2 is asymmetrical. This could probably not have been produced analytically. The predicted results have also been proved by experiments.

Recently, a directional antenna, the Yagi-Uda array, has been investigated using the GA to obtain improvements in the gain, front-to-back ratio (F/B) and bandwidth. Altshuler and Linden (1997) used the GA to optimise two Yagi-Uda arrays for different objectives. In the first example both the element lengths and spacings between elements of a 14-element Yagi-Uda array were optimised to obtain broadband characteristics and a low sidelobe level. As excepted, the GA finally produced a configuration that was quite different from one that would have been obtained using the conventional methods. This genetic Yagi-Uda array with a boom length only one-third of a conventional Yagi of the same number of elements can produce a bandwidth of 13.6% in frequency with sidelobes more than 20 dB down.

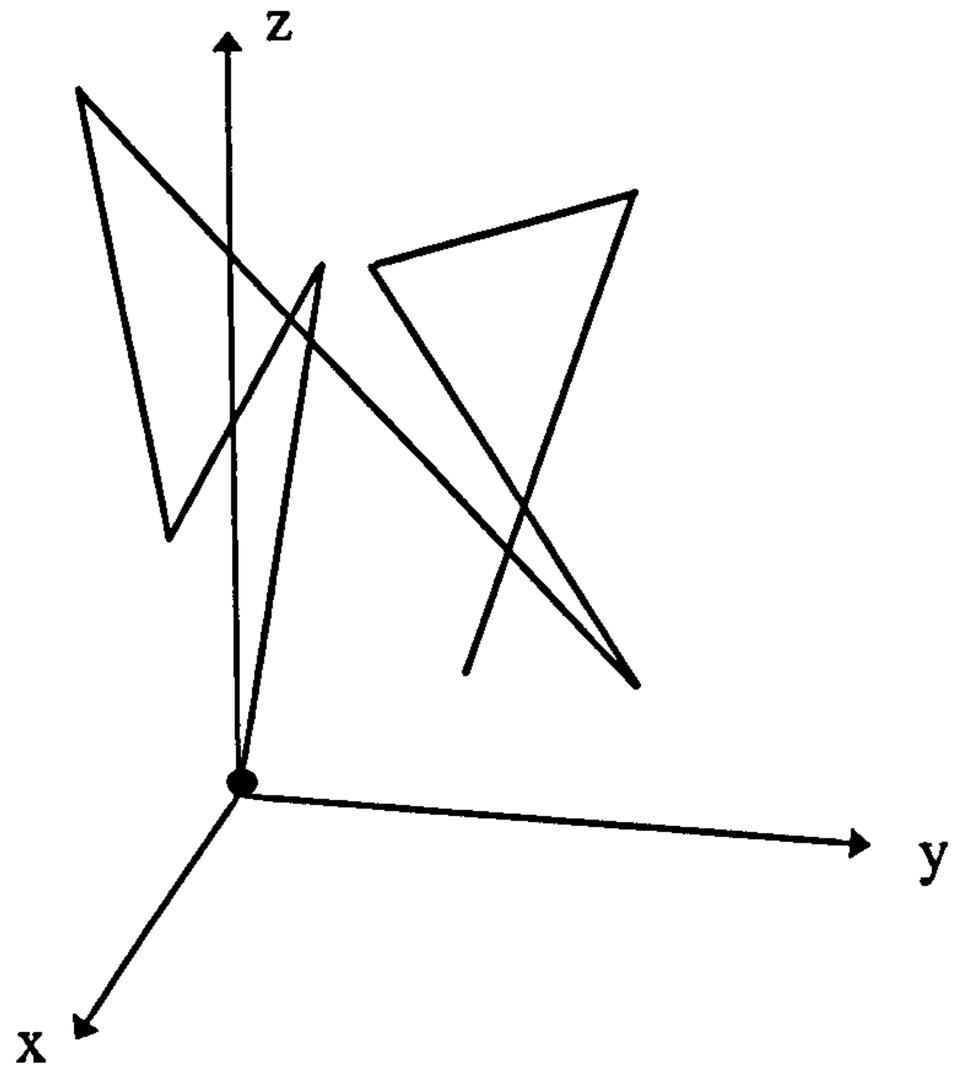


Figure 2.1 Optimised configuration of a seven-wire antenna over a ground plane by Altshuler et al. (1997).

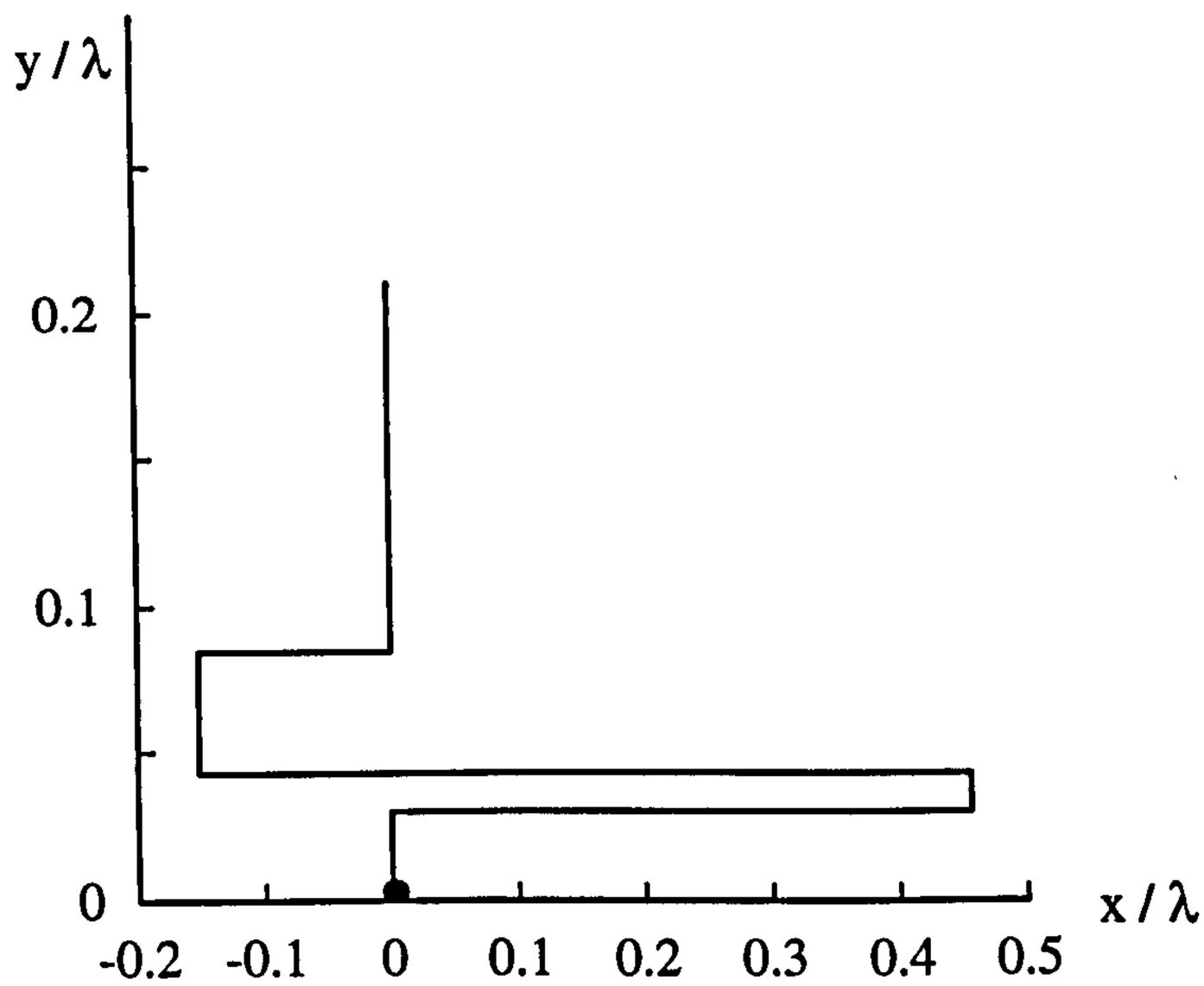


Figure 2.2 Optimised configuration of a monopole antenna loaded with a modified folded dipole by Altshuler et al. (1997).

A similar GA optimisation procedure was used in the second design example. However, in this case the Yagi-Uda array with 18 elements was required to produce high gain. In the GA approach only the length of each director element and the spacings between director elements were optimised (Thiele 1969, Chen et al. 1975, Cheng et al. 1973, 1991). Finally, the GA also created an array which had a shorter boom length but produced higher gain than a typical conventional Yagi. The resulting configuration was also quite different from the typical conventional Yagi with the same number of elements.

Jones et al. (1997) also provided some improved results from the design of different element Yagi-Uda arrays using the GA optimiser. Unlike that work, most optimisation of this array was focused on producing high gain, low sidelobe or broad bandwidth, and in some of their design cases the input impedance (except for the terms of gain and sidelobe) was also optimised simultaneously for the purpose of easy matching to a feeding system. Their numerical experiments showed that the GA is powerful enough to design a Yagi-Uda array which can produce higher gain than conventional equally spaced arrays with the same number of elements; while if the input impedance of antenna is also considered in the optimisation process, the GA still can create an optimal array with the desired impedance, as well as an even higher gain than the conventional array.

However, the growing applications of the Genetic Algorithm technique to electromagnetic problems were not only limited to the above examples. There were many other examples of the use of GA. These include Ares et al. (1996) who used it to find optimum excitations for moderate and large linear arrays with null-filling patterns; Mitchell et al. (1996) optimised the array pattern nulling by operating on the

roots of the array polynomial in the complex plane using the GA; Michielssen et al. (1993) and Weile et al. (1996) both applied it to design an optimal boradband microwave absorbers and so on.

2.3 Review of Vehicular NVIS Antennas

2.3.1 NVIS Propagation

In radio communication the use of very-high-frequency (VHF), ultra-high-frequency (UHF) and satellite systems which provide a proven reliability, seem almost to dominate the field. However, the above systems are suitable only when the path is essentially line-of-sight (LOS) and is unimpeded by obstacles. Unfortunately the terrain between the two communication terminals is not always smooth. They are often severely obstructed by large land masses such as buildings, mountains, or may lie within forested areas, jungles, deep canyons or fjords. In addition, in many cases, especially in the world of military activity, the survivability of a satellite or relay facility could never be guaranteed. Therefore, there is a need for a mode of propagation that can allow the tactical communicator to launch signals within the HF skip zone which fall between the first skywave return and the boundary of the ground-wave, up-and-over an obstacle without the need for a satellite or a man-made relay station. The region is usually always beyond the line of sight (BLOS) and is frequently obscured by terrain features. This mode of propagation is called *Near-Vertical-Incidence Skywave* (NVIS) and it requires the primary radiation and reception of transmitted energy to take place at elevation angles approaching the zenith (i.e. straight up) to ensure effective propagation (Goodman 1992 pp.217-224, Fiedler et al. 1996). Figure 2.3 depicts a typical application of this mode of propagation.

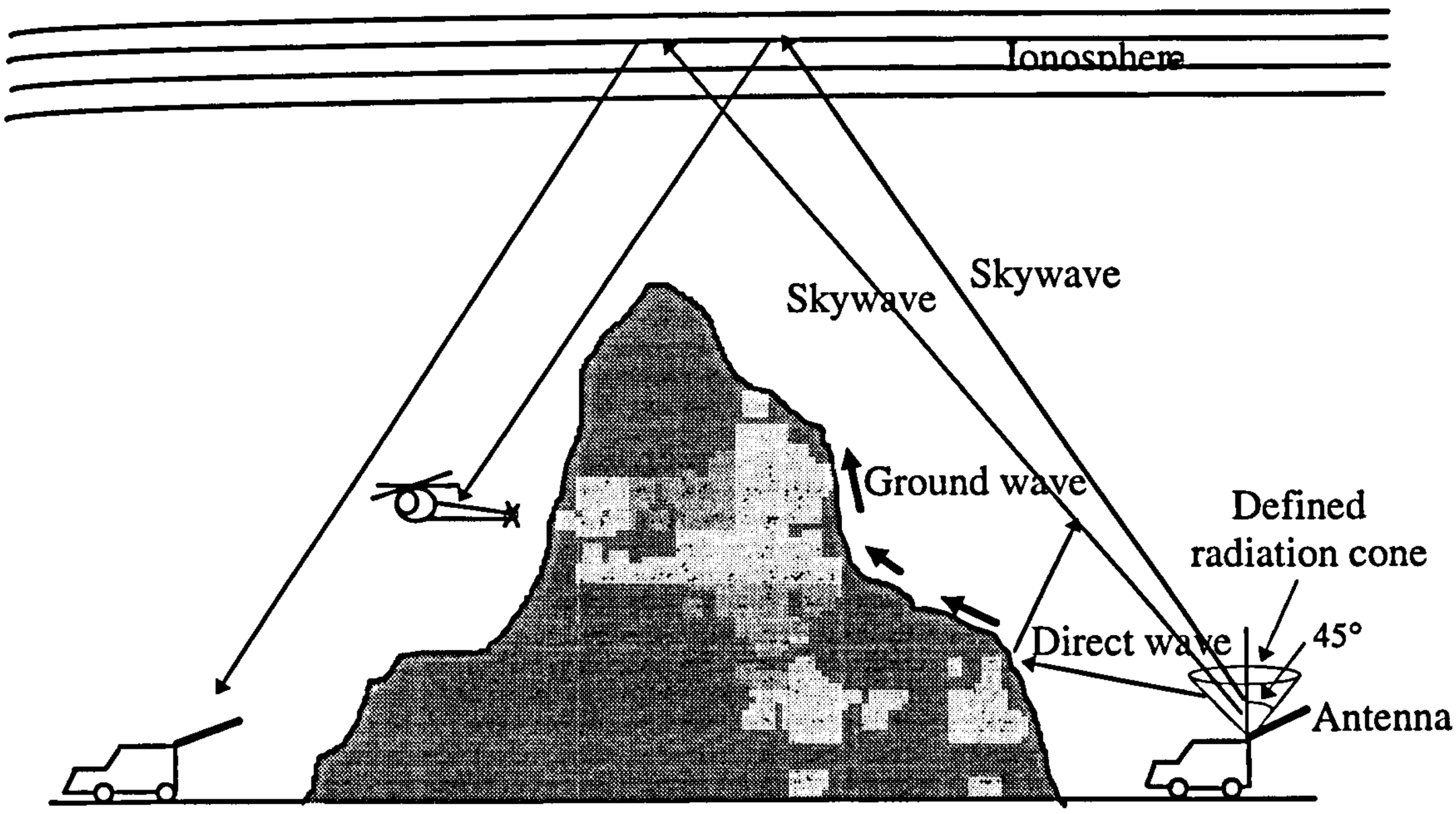


Figure 2.3 The near-vertical-incidence skywave (NVIS) mode of propagation.

A significant constraint on NVIS antenna systems is the selection of operating frequency. It depends on both the critical frequency of the particular ionospheric layer supporting this mode of propagation and the angle of incidence at that layer. The necessity for vertical incidence of the radiated signal upon the ionosphere makes the frequency of operation close to the critical frequency of the layer which supports this NVIS propagation. However, in most NVIS applications the F_2 layer is more suitable since its electron density is generally greater that in the E and F_1 layers (Davies 1990). Therefore, the typical frequency for NVIS applications is close to the critical frequency of the F_2 layer and usually lies anywhere within a fairly narrow pass band between about 2 and 10 MHz, depending upon the time of day, season, the level of sunspot activity and also the geographical location (Maslin 1987 p81).

2.3.2 NVIS Antennas

In the past, considerable work on the investigation of NVIS antennas mounted on a vehicle was done by computer modelling and experimental measurements by Hagn et al. (1970, 1973, 1974) and Eley et al. (1991). The candidate antenna in their investigations was the whip antenna which has a length very small compared to the wavelength at the typical NVIS frequency. Their results all show that though the vertical whip antenna is the most wildly used HF vehicular antenna it provides limited near-vertical radiation characteristics which reduce the effectiveness of this short-range communication mode because this vertically positioned whip produces an inherent null at the vertical direction (i.e. the direction of $\theta=0°$). This is directly opposite to the necessary criterion for NVIS propagation. However, they also proposed an approach to improve the inefficiency of such antennas for this type of propagation by tilting the whip to obtain some increase in radiated power toward the zenith. After a series of computer simulations and extensive measurements it was concluded that the best improvement on the efficiency of this mode of propagation (NVIS) was clearly obtained when the combination of the whip and the conductive vehicle itself form a nearly horizontal dipole by tilting the whip towards the rear and away from the body of vehicle. In terms of antenna performance this may be ideal while the major disadvantage are the practical difficulties associated with the whip protruding, at a nearly horizontal angle, from the rear of the vehicle, especially when the vehicle is moving.

An alternative antenna for overcoming the NVIS problem is the loop, which, if electrically small and vertically oriented, will naturally produce maximum radiation in the plane of the loop, and hence towards the zenith. Such a loop configuration

mounted on various air, land and sea mobile systems was examined by Burberry (1982). In 1991 both Baker and Cox et al. also examined the suitability of a loop antenna sited on a helicopter for NVIS by measuring the performance of a scale model of this system. The results confirmed that the loop is more suitable than the whip for NVIS in both physical and practical aspects.

Further investigation of vehicular NVIS antennas was done by Austin and Murray in 1998. The performance of a whip antenna for NVIS applications like those reported in the literature was confirmed by them again. However most of their efforts were on the examination of loop antennas by using the analytical method of Characteristic Modes (CM) which synthesise the distribution of the required mode current on the combined antenna-vehicle structure for approaching a defined NVIS radiation pattern. They found four identifiable points of mode excitation that dominate the contributions to construct an ideal NVIS radiation pattern. Two of these positions are associated with the loop antenna and two with the pillars on either side of the windshield. However there is inherent physical complexity in feeding the four points simultaneously. Therefore a suitable result was achieved by using just a dual-feed system, involving only two voltage sources, positioned at the bases of the vertical members of the loop. In addition, they also showed that the best NVIS performance of this system, which is much better than that of the conventional loop only excited by a single source, was achieved if the two sources are fed in-phase. Some measurements for the cases of whip and single-feed antennas have also been done by them and excellent agreement has been obtained with their computed results.

In the work of Austin and Murray three performance terms, D_{45}, η and G_{45} relative to the NVIS applications were defined for describing the antenna's NVIS

performance. Because these terms will be adapted in later chapters of this thesis it is necessary to give a brief introduction to them.

The term D_{45} is defined as follows,

$$D_{45} = \frac{P_{rad}(\theta < 45^{\circ})}{P_{rad}} \times 100\% \tag{6.1}$$

where P_{rad} is the total radiated power; $P_{rad}(\theta<45°)$ is the radiated power falling within a target region concentrated on the zenith, within a cone of 45° half-angle (see figure 2.3); and thus D_{45} is therefore given as the directivity within that region.

The term of η is simply defined as the radiation efficiency of the antenna-vehicle system and is given by

$$\eta = \frac{P_{rad}}{P_{in}} \times 100\% \tag{6.2}$$

Finally, the term of G_{45} is a measurement of the performance that includes all losses associated with it, and so leads naturally to the expression for the system gain (or called total efficiency) in the specified target direction. This is given by

$$G_{45} = \frac{P_{rad}(\theta < 45^{\circ})}{P_{in}} \times 100\% \tag{6.3}$$

As the general definition in the field of antenna analysis, the system gain (G_{45}) defined here is equal to the product of the directivity (D_{45}) and the radiation efficiency (η).

2.4 Summary

This chapter has reviewed the novel application of the Genetic Algorithm to the optimisation of antenna performance. The GA clearly shows great potential in antenna

design. They have been applied successfully to the synthesis of novel antennas or those which have never been obtained from the conventional analytical method.

Section 2.2 firstly demonstrated the reported optimal results of a number of linear, planar or circular arrays with various numbers of elements. The optimised radiation characteristics include the null locations and the maximum relative sidelobe level. The capability of the GA to manipulate many design parameters at one time enable these problems which cannot be solved by the conventional techniques to be overcome easily and thus to obtain satisfactory results. Secondly, various wire antennas loaded with lumped circuits have been optimised to broaden the bandwidth. In some of these examples, the design of the matching network was also included into the GA procedure. All the predicted results have provided marked improvement in operation bandwidth compared to the antennas without loads. Thirdly, with a limitation on the antenna dimensions as a defined objective, the GA has the ability to create very unconventional antennas which can produce such features as hemispherical coverage. These antennas include such oddities as the seven-wire antenna and the monopole antenna loaded with an asymmetrically modified folded dipole. Finally, the Yagi-Uda array with different numbers of elements was optimised for broad bandwidth, low sidelobe level or high gain. These results show that by properly adjusting the element lengths and spacings between the elements of a Yagi-Uda array using the GA technique can produce results which are better than a conventional equally spaced Yagi-Uda array with the same number of elements. Also introduced to end this section are some other electromagnetic problems solved by using the GA. All these results have proved that this revolutionary technique is a powerful tool in designing appropriate antenna configurations to satisfy any specified objective.

The second part of this chapter - section 2.3 has addressed the basic properties of the NVIS mode of propagation. This communication application requires the radiated energy to be propagated towards the zenith and its available operating frequency range is typically from 2 to 10 MHz. This type of propagation can only be implemented by some rather unusual antennas on vehicles. Hence, the currently used NVIS antennas, including the whip and loop, were reviewed. Three terms, D_{45}, η, G_{45} relevant to NVIS performance and defined by Austin et al. were introduced for use in this thesis.

CHAPTER 3

THE CONCEPT AND STRATEGY OF THE GENETIC ALGORITHM

3.1 Introduction

With an exploratory procedure, the iterative adaptive search technique - the Genetic Algorithm (GA) is used to design and develop high performance structures in complex task domains and this capability has been examined from the design examples shown in the previous chapter. The algorithm is like a simulation of natural evolution phenomena and provides a technique of efficiency and easy implementation for problem solving. Over the past few decades, this application to search and machine learning problems in science, engineering and other fields has been made possible by a number of theoretical developments as well as numerical experience (Goldberg, 1989). This is so, because GAs place a minimum of requirements and restrictions on the user prior to engaging the search procedure. Generally, they have two inherent characteristics, which are: (i) sufficient variability is maintained to prevent convergence to local optima and (ii) both categorisation and recombination yield powerful implicit parallelism, to make a GA highly suitable for learning by discovery. This chapter therefore discusses the fundamental theorem and control techniques of a GA and also examines its accuracy when applied to practical problem solving.

Section 3.2 introduces the concept of the GA's mimicing of natural evolution, including the description of the optimisation space in terms of strings of bits, or "chromosomes". In addition, a comparison made by Johnson et al. (1997) between the GA and some widely used techniques for optimisation or problem solving is also discussed. A summary of the GAs basic components and their characteristics is provided at the end of this section. The basic steps of implementing a GA are introduced in section 3.2; subsequently section 3.3 examines the appropriate control techniques in each of these steps. These techniques were investigated by various researchers using theoretical or numerical experiments, and different subjects to markedly increase the efficiency of the GA. It is especially useful to observe the step-by-step process as GA, and test its applicability, as it implements it step-by-step on a typical case. Therefore, in section 3.4 the peak of a multi-peak, two-dimensional function typical of an optimisation sample with both global and local maxima, is searched by the GA developed in this work to explore how powerful it is in searching for a global optimum and this provides the verification of this program for use in other applications.

3.2 Natural Concept

According to Darwinian doctrine animals and plants generate their new off-spring in a manner better to adapt themselves to their environment by the processes of natural selection, reproduction and mutation. Only the fittest survive to reproduce in the next stage of the selection and evolution processes. This procedure illustrates exactly Nature's evolution and selection process in a form not dissimilar to the GA and indicates how it could be used to design better products or solutions in different fields of applied science (Riolo 1992, Whitley 1993, Mitchell 1997).

Holland (1973) is the pioneer who introduced the concept of a GA to an optimisation scheme. In his research, two main themes were investigated and elaborated. They were: (i) the ability of simple and discrete representations (a bit string) to encode complicated structures and (ii) the power of simple transformations to improve such structures. These themes reveal the key features of a GA used as a function optimiser. Following Holland, many researchers explored the theory and control techniques of this method in associated studies to improve the computational efficiency and this resulted in remarkable progress (Quagliarella 1998). This work enabled the GA to have greater ability to exploit accumulating information from an initially unknown search space, and then to bias subsequent searches into useful subspaces.

To show the distinguishing characteristics of the GA, a comparison of some widely used optimisation techniques was made by Johnson et al. (1997). These optimisers includes conjugate gradient, random searcher and the GA itself. Table 3.1 lists the results of their comparison. It clearly shows that compared to the other two techniques, which usually only explore solutions around an arbitrary starting point and have difficulty working with quantised parameters, the GA is particularly well suited for finding a global solution in a discontinuous or non-differential solution space. In electromagnetic-design problems particularly, the rate of convergence is often not nearly as important as getting a solution. And then having found a solution, the ultimate goal is to find the best solution or global maximum.

Figure 3.1 is the flowchart of a GA. It begins by representing the design parameters in a string of binary bits called a *chromosome*. This process is very different from other optimisers that operate directly on the specific design parameters. Thus each parameter is encoded by a small string of bits, generally called a *gene,* that form the basic building

Table 3.1 Comparison of various optimisation techniques.

(Source: Johnson et al. 1997)

	Conjugated Gradient	Random Searcher	Genetic Algorithm
Global Optimisation	Poor	Average	Good
Discontinuous Object Functions	Poor	Good	Good
Non-differentiable Object Functions	Poor	Good	Good
Convergence Rate	Good	Poor	Average

block of a GA and have associated real physical properties. After encoding the parameters, a group of initial chromosomes (so-called "seed" chromosomes) are generated by random selection to form the first generation of candidates. The randomisation in this step enables the initial design parameters to be broadened to any point of the entire search space by probability. This start condition will significantly assist the GA to easily approach the solution. Then, an associated value (called "fitness" or "cost") of each chromosome will be calculated in accordance with a function named the *objective function* or *cost function*, defined and configured from a consideration of the search objective. The cost function decides the fitness of each chromosome in terms of the desired characteristics, for surviving into the next generation or being discarded. In the mean-time, a condition is also set during the following step which controls the algorithm when it stops by either setting it to obtain an acceptable solution or to run for a set number of iterations. The algorithm will then rank all chromosomes from the best-fit to the least-fit as it continues. By the process of natural selection, the chromosomes in the superior-species subset become parents to

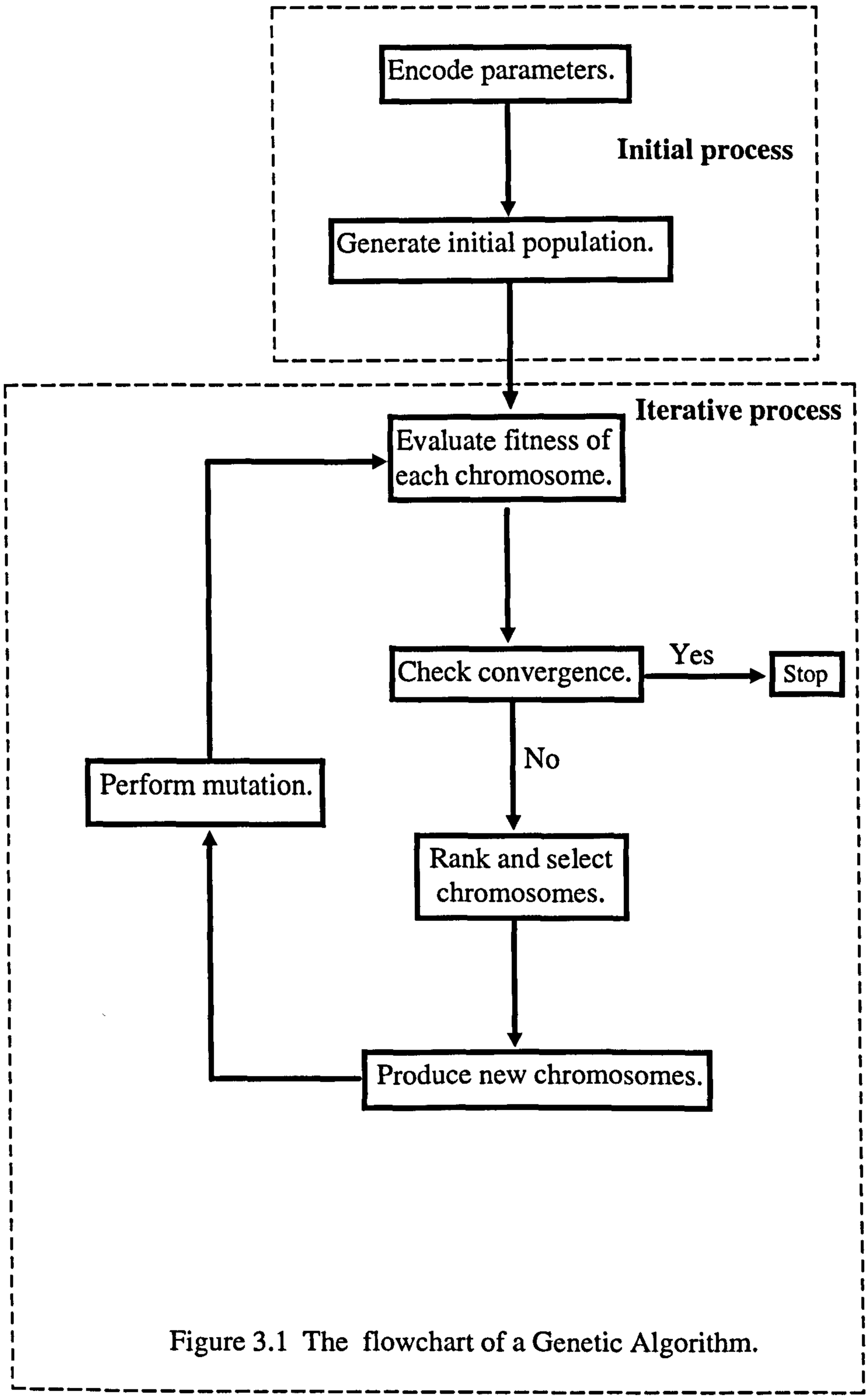


Figure 3.1 The flowchart of a Genetic Algorithm.

reproduce new chromosomes and all unacceptable ones are discarded. New chromosomes are therefore produced by the mating of surviving parents. There are thus two main steps in the mating process: first, select any two of the survivors preparing to mate, for which there are many different schemes for this process. The most important is the proportionate-selection process as in a roulette wheel. The selection of mated chromosomes is followed by the crossover procedure which causes each mated chromosome to be spilt into two or more subsections, depending on the scheme used, and then exchange these subsections between the two mated chromosomes to produce the new chromosomes. Each pair of parents will therefore produce two new chromosomes. After mating, the next generation, including the survivors from previous generation, will have the same number of chromosomes as that before 'natural selection'. This algorithm is known as "a steady-state genetic algorithm" because all the survivors are kept as the members of the new chromosome family (Johnson et al. 1997).

Then, a pre-defined rate of chromosomes of the new generation as well as a pre-defined number of bits of each of them are randomly chosen to make mutation. This operator will change the value of any selected bit inversely i.e. from "0" to "1" or vice-versa. Its function is to prevent the algorithm converging too fast and therefore converging to some local optimum.

When the mutation step is completed the procedure returns to the step of evaluating the objective function. Thus the GA runs iteratively through the applications of ranking, selection, mating and mutation until it reaches the set goal and then it stops. These processes are markedly similar to evolution in Nature.

Based on the above illustrations of the processes of a GA its basic components can now be used for problem solving and the procedure can be summarised approximately as follows:

1. The solution need not to be processed directly; a binary string called a chromosome composed of genes takes the place of the real solution whether it is continuous or discrete in nature.
2. A set of trial solutions is processed simultaneously in each iteration rather than sequentially in a one-solution search.
3. The objective function plays an important role in constructing a virtual natural environment that decides which solution should survive.
4. This reproductive process, including operator selection followed by crossover, is ultimately responsible for the algorithm's convergence.
5. The possibility of becoming stuck in a local optimum is avoided by the mutation strategy to enable a global optimum to be approached.

When using a GA as a problem solver the control techniques within each step significantly affect its operational efficiency. These techniques will be discussed in detail in the following section.

3.3 Control Techniques

3.3.1 Parameter Coding

In the GA implementation an array of real-parameter values is defined as a chromosome at the first step. For example, if N parameters, given as P_1, P_2,.... P_N, are composed to form an optimal solution, and each parameter (i.e. gene) is encoded by a

small bit string with an associated bit number of $L_{n,}$, then the chromosome could be written symbolically as follows:

$$\begin{aligned} &\{ \quad P_1 \qquad P_2 \qquad P_3 \qquad \ldots\ldots\ldots \qquad P_N \quad \} \\ &= \{ b_{11}b_{12}\ldots.b_{1L_1}b_{21}b_{22}\ldots.b_{2L_2}b_{31}b_{32}\ldots.b_{3L_3}\ldots\ldots\ldots\ldots\ldots\ldots b_{N1}b_{N2}\ldots.b_{NL_N} \} \end{aligned} \tag{3.1}$$

Where

b_{nm}: the m^{th} bit of the n^{th} gene $\quad n=1,2,\ldots\ldots\ldots,N; \quad m=1,2,\ldots\ldots\ldots L_n$ (3.2)

The relationship between the bit strings and real-parameter values can be mapped by a simple linear transformation denoted as

$$P_n = P_{n\,(min)} + \sum_{m=1}^{L_n} b_{nm} \cdot 2^{m-1} \cdot \Delta P_n, \quad n=1,2,\ldots\ldots\ldots N \tag{3.3}$$

where

$$\Delta P_n = \frac{P_{n\,(max)} - P_{n\,(min)}}{2^{L_n} - 1} \quad \text{is the bit resolution of parameter } P_n \tag{3.4}$$

and parameter P_n takes values ranging from $P_{n\,(min)}$ to $P_{n\,(max)}$.

The small bit strings are then concatenated to form a single string of length $\sum_{j=1}^{N} L_j$ to yield a chromosome.

Choosing the number of bits for a chromosome is always a puzzle when encoding the parameters to be searched. Experience of Haupt (1995c) and Li et al. (1996) indicates that a higher bit number can produce a more accurate approach to the problem but will take longer implementation time and occupy larger computer memory; by contrast, a faster approach but accompanied with the more deviation could be obtained when using the lower bit number. In this way, the choice of the number of bits depends on the problem at hand, and Altshuler (1997) suggested that a few trial runs are usually necessary to make the decision.

3.3.2 Initial Population

Another most useful trait of a GA is its ability to find multiple solutions, should these exist. Therefore, a group of randomly-selected chromosomes should always be generated first to cater for such flexibility of the GA. Generally, more chromosomes provide better sampling of the solution space but increase computing time. But to date no theoretical procedure has been found to aid in the choice of population size. For most practical applications a population size proportional to the chromosome length should suffice. A typical size suggested by Johnson et al. (1997) is in the range from 30 to 100 chromosomes. However, it should be noted that a proper weighting of the parameters in the generation of the initial population can achieve better results even with a smaller size of initial population. In addition, the type of probability distribution also has a significant impact on the convergence time. Sometimes, invoking a *priori* knowledge in the generation of the initial population can help the algorithm converge faster.

In addition, it should be noted that in many steps of the implementation of a GA, including the generation of initial chromosomes, the methodology of randomisation is applied. However, the random number generator of the used computer language usually produces a wider range of numbers than that of the desired numbers used in the GA. Hence, it is always necessary to add a selected weighting factor to the process of randomisation to produce a more uniform probability distribution of numbers in the desired range.

3.3.3 Objective (or Cost) Function

The required, optimised objectives are defined as the "objective function" (or "cost function") and determine the fitness (or cost) of each chromosome. This cost function takes into account all of the desired features that are to be optimised and is the only connection between the physical problem being optimised and the genetic algorithm used to optimise it. So, before evaluating the "fitness", the real parameters are first extracted from their associated chromosomes and then the physical properties of each chromosome are determined. The proximity of each to the objective is an indicator of their fitness. Depending upon the actual requirement, there are many schemes for selecting this function; for example, it could be set as either linear or non-linear combinations of different criteria. Besides, during the initial trial process the variation between the criteria is always very different, and so it is necessary to select an appropriate scaling factor for each to prevent the process of improvement from concentrating on only some individuals. A general objective function could be represented as in equation 3.5 :

$$\text{Objective function} = \sum_i C_i (X_i - O_i)^{n_i} \tag{3.5}$$

where the parameters are described as

C_i : scaling factor

X_i : the real characteristic of the chromosome

O_i : the objective characteristic

n_i : order of each term

As mentioned before, the ultimate goal of any optimisation procedure is to find a solution that represents a global maximum or minimum, thus the value of the function,

could be either the minimum or the maximum, according to the definition, and could be represented by the sign of n_i.

3.3.4 Implementation of the Selection Process

Selection introduces the influence of the cost function into the genetic-algorithm optimisation process. Basically, it is responsible for the algorithm's convergence and has to utilise the cost of a given individual, since cost is the measure of the profit, utility or goodness of that individual. However, intuitively, selection cannot be based solely on choosing the best individual from each population because the best individual may not be very close to the optimal solution. Instead, some chance that relatively unfit individuals might also be selected must be preserved to ensure that genes carried by these unfit individuals are not "lost" prematurely from the population. In general, selection involves a mechanism relating an individual's fitness to the average fitness of the population.

Once the individuals satisfy a particular strategic pass criterion corresponding to the objective cost they become the survivors (or parents) that will produce the new chromosomes. In general, for a steady-state GA, 10 - 50% of the population, ranked in order according to the cost of each chromosome, are kept and become the parents in the new generation (Altshuler et al. 1997). The appropriate percentage selected depends upon the different mating schemes. This is because the numbers of new individuals are not absolutely identical between different schemes, but the total number of individuals, including the parents and reproduced sub-population in the new generation, must equal that of the previous generation to satisfy the criterion of a steady-state GA.

The process of reproduction includes two main steps which are the selection of spouse and exchange of genes.

3.3.4.1 Spouse Selection

Some popular strategies are often used for this step. Among them, the "roulette wheel selection" is the most famous and is illustrated in figure 3.2. The probability of selecting an individual from the surviving population for reproduction is based purely on a function of the relative fitness of the individual. The function is

$$P(survivor\ i)=\frac{W_i \cdot fitness(survivor\ i)}{\sum_j W_j \cdot fitness(survivor\ j)} \tag{3.6}$$

where W_i, normally equal 1, is a scaling factor used to avoid stochastic errors when population sizes are small and the denominator represents the total scaled fitness of the surviving individuals.

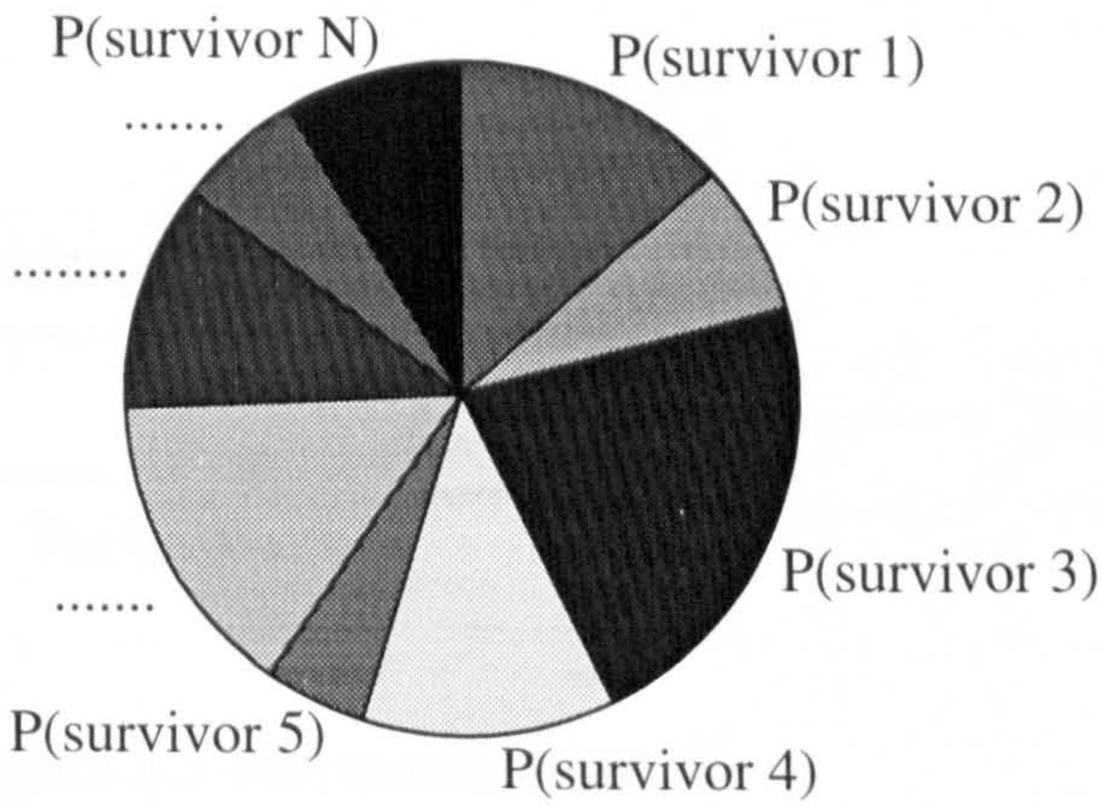


Figure 3.2 Roulette-wheel selection.

Clearly, although the fitter a chromosome and hence the larger its shares of the virtual "roulette wheel", the less-fit individual still has same probability of being selected. The wheel is "spun" to generate the first parent and then again for the second parent. If the two parents are unexpectedly the same, a third spin should be done because clones are not allowed. A possibility exists for an individual to be multiply selected in different pairings. In this way each pair is generated.

Tournament selection is another popular strategy. It is depicted in figure 3.3. In this selection a sub-population of a defined number of individuals, of which the most commonly defined number is two, is chosen at random from the surviving population. To decide which individual will be selected, the individuals of this sub-population are compared to each other based on their fitness or scaled fitness. Finally, the individual in the sub-population with the highest fitness is selected and becomes a parent. Then all the sub-population individuals, including the selected one, are placed back into the surviving population. The same process is repeated until all parents are selected. As in the "roulette wheel" selection, this selection procedure makes it possible for individuals to participate in multiple pairings, and some indeed do.

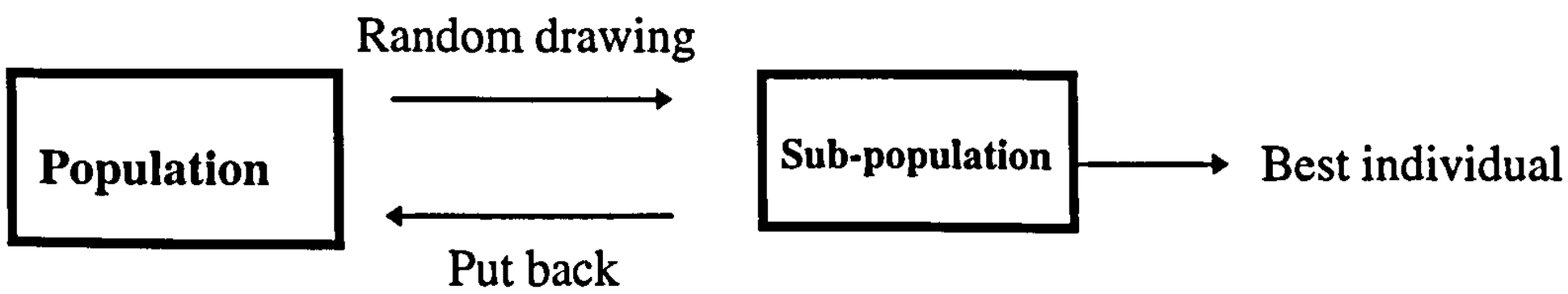
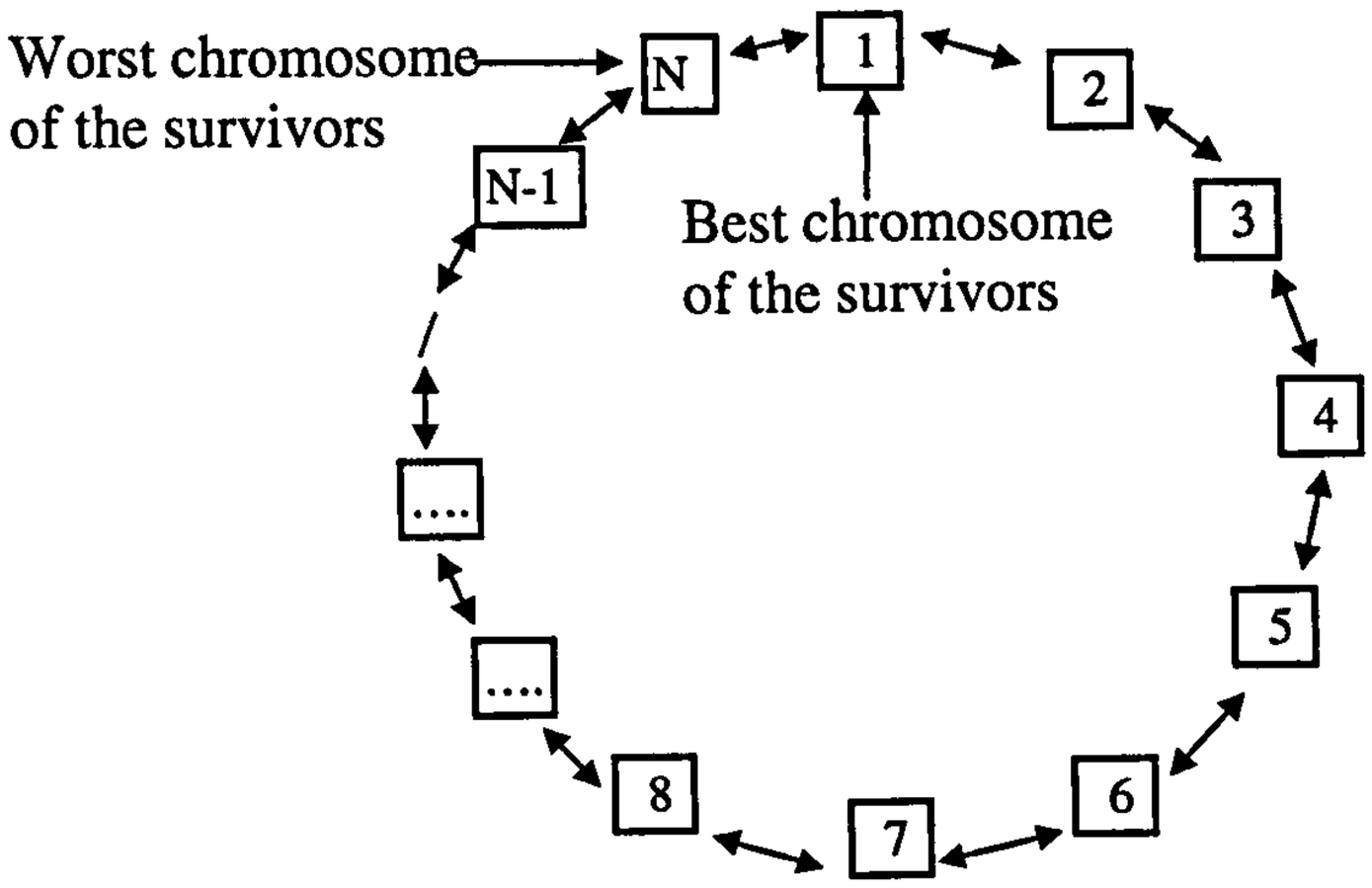


Figure 3.3 Tournament selection.

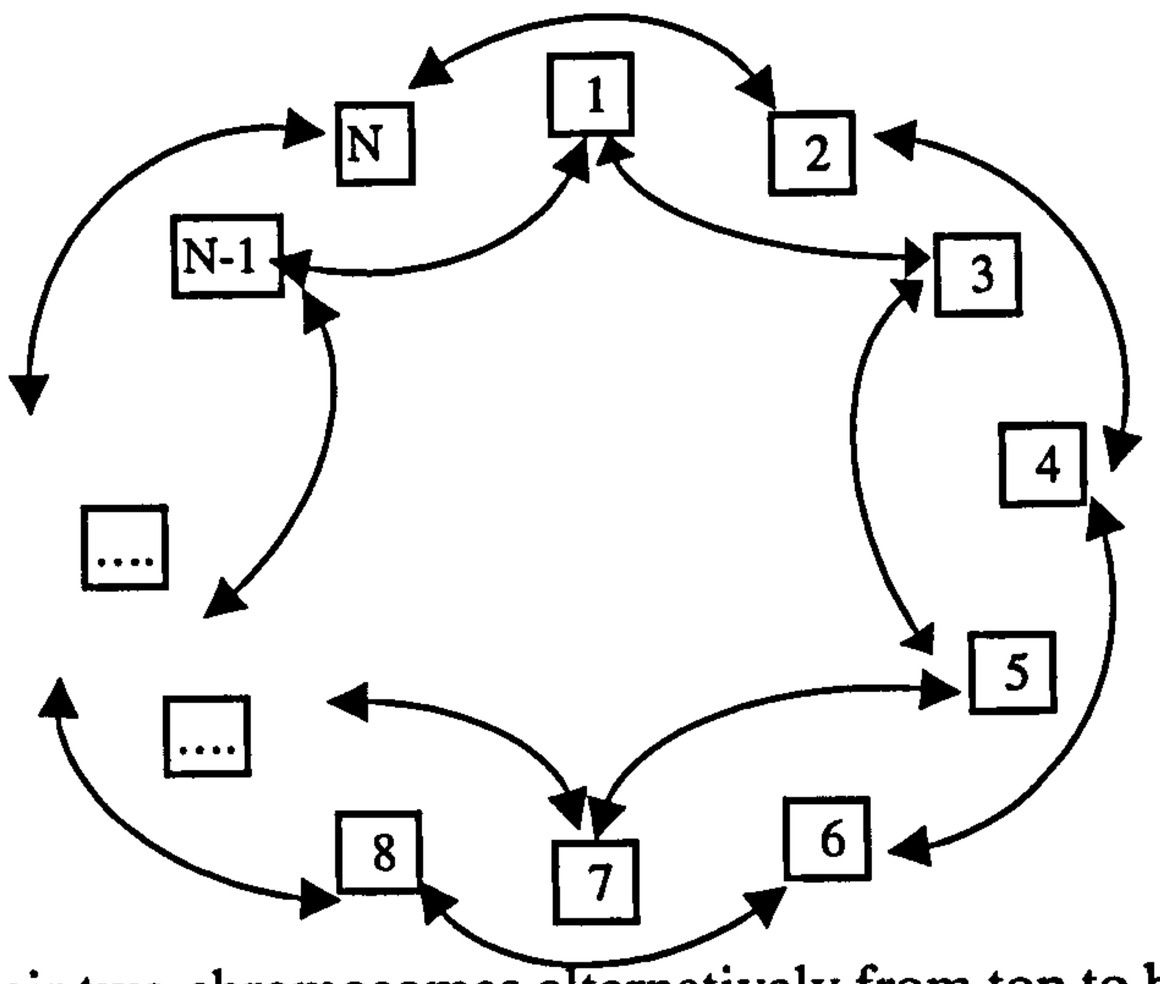
A further approach is random selection. It is unlike the above two strategies to weight the individuals with their appropriate fitness. Without considering its cost this selection scheme bestows each individual an equal probability to be selected. This strategy approaches the solution depending on more probability. Sometimes it converges very quickly but sometimes does not.

The above three selection strategies could be classified as random selection. Some other non-random selection strategies are also used frequently. In these approaches parents are paired in different but regular ways and without weighting their appropriate fitness. Three such commonly used approaches are discussed here and are shown in figure 3.4. These selections for preparing to pair are all based on the ranking list of the surviving individuals. Firstly, every two individuals are paired from top to bottom (see figure 3.4(a)). If the least-fit individual is linked to the fittest individual the list forms a circle. Each individual is allowed to participate two pairings with its neighbours. This is somewhat like consanguineous marriage, except for the least-fit individual has another chance to be paired with the fittest individual. It is due to this that the pairs could probably produce very different off-spring to widen the distribution of solutions. Secondly, to increase the difference between the parents they can be paired alternatively as shown in figure 3.4(b). Finally, the largest difference between the parents is produced when pairing the best with the worst, the second best with second worst etc. as depicted in figure 3.4(c). Equal opportunity for each individual and diverse distribution of the off-spring are two characteristics of these non-random approaches.

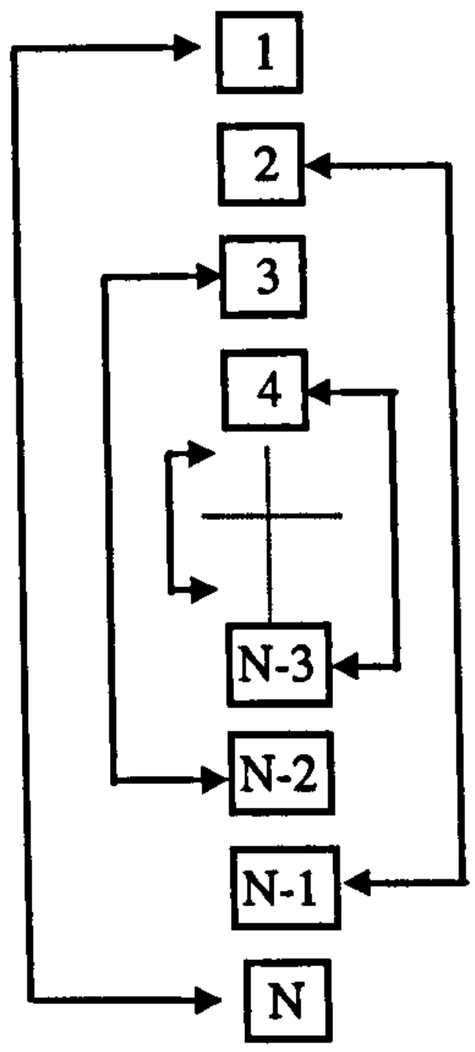
In this research, after some trials on the different schemes of spouse selection the



(a) Pair two chromosomes from top to bottom.



(b) Pair two chromosomes alternatively from top to bottom.



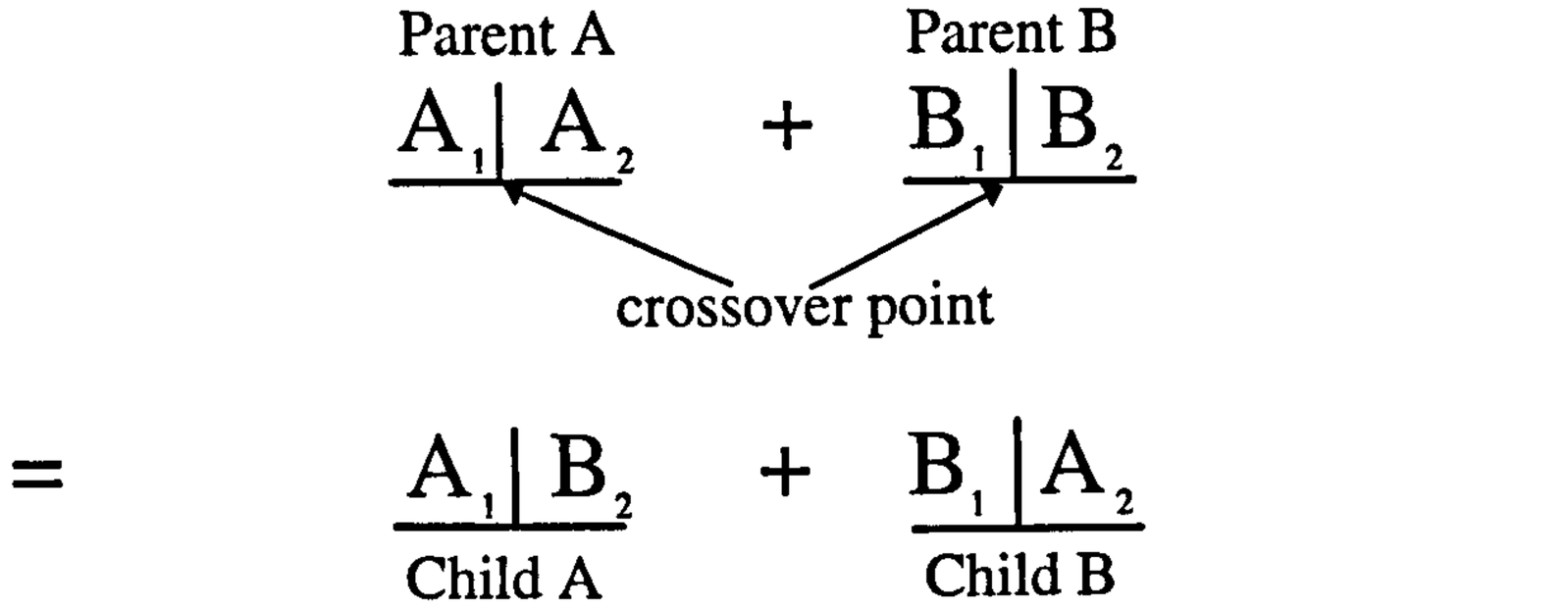
(c) Pair the best with the worst.

Figure 3.4 Three non-random selection schemes.

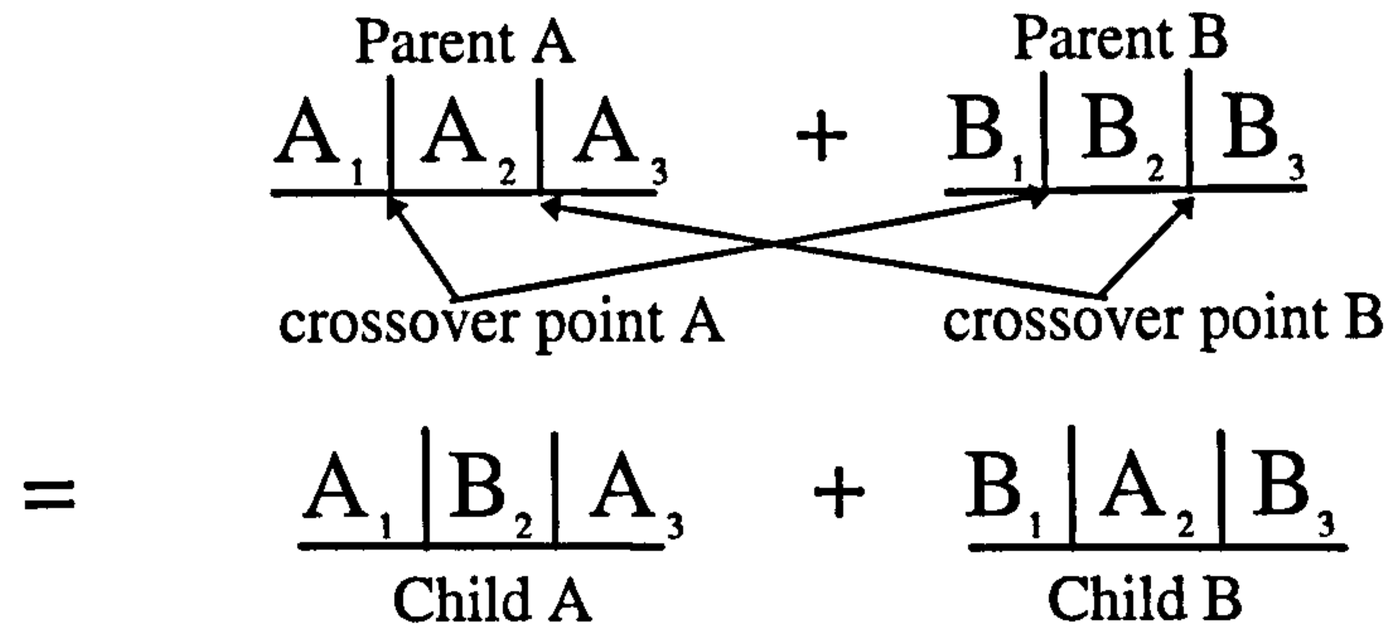
scheme of alternative pairing can provide sufficient difference between the parents to diversify the chromosomes in the new generation. This scheme will thus be used in the GA procedure for all investigations in this research work.

3.3.4.2 Gene Exchange

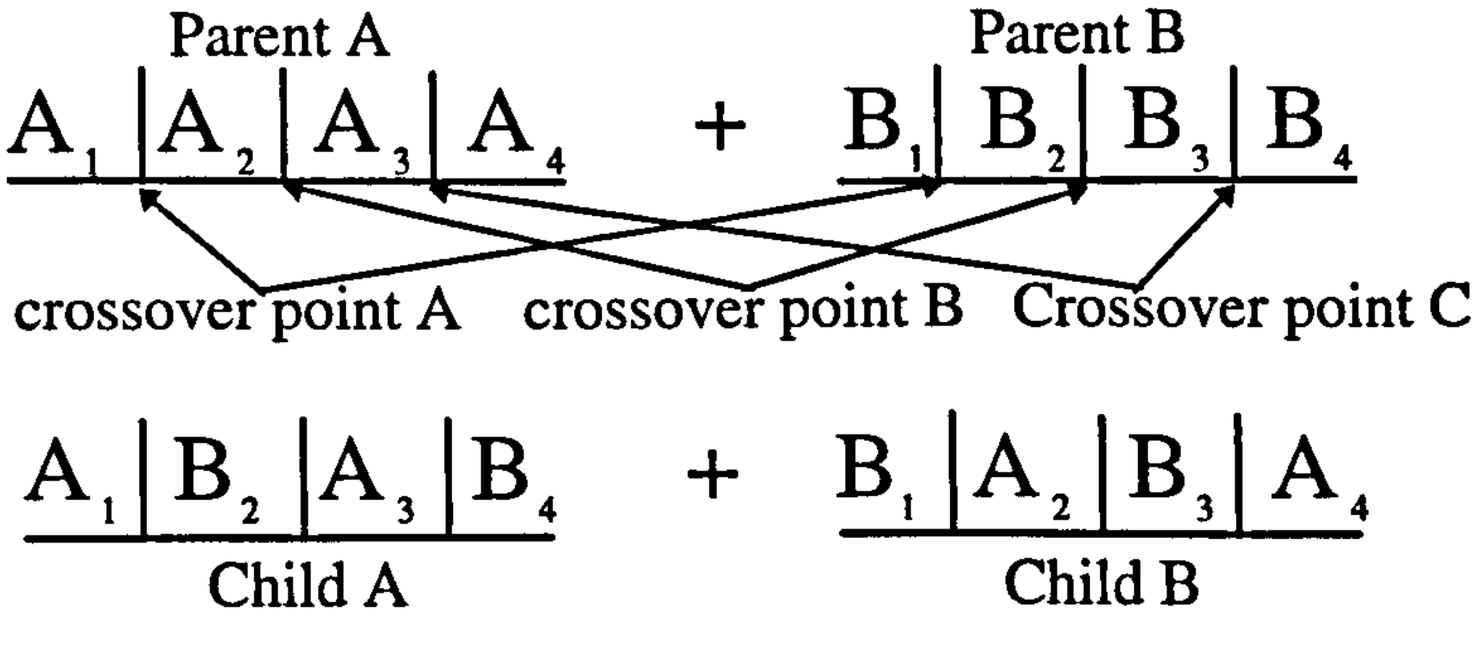
The second step for reproduction is swapping the portions of the bits between the two selected parents and this is called a crossover scheme. This operator is the main search tool of the GA (Goldberg 1989) since it rearranges the genes from the combined chromosomes. The reason is that these chromosomes all contain genetic information which is known to be useful for producing better combinations of genes, thereby resulting in more suitable individuals. The most typical scheme is the single-point crossover, shown in figure.3.5(a). In this operation, a location between the first and the last bits of a chromosome is selected by randomisation. Therefore the two paired chromosomes are both divided into two bit strings (e.g. A_1A_2 and B_1B_2) from the specific locus and then exchange the bit string with each other to produce two new chromosomes. It is clear that after the process of gene exchange even if the chromosomes A_1A_2 or B_1B_2 simultaneously carry two better end-elements each new chromosome can only inherit one from each parent. Generally, a chromosome with fewer bits will more easily inherit the end elements after several crossover processes. For this, a single-point crossover is more suitable for the condition in which the chromosome is composed of fewer bits. For the chromosome with a longer bit string, some alternative mechanisms such as two-point crossover, three-point crossover, etc have been proposed. These rules are



(a) Single-point crossover scheme.



(b) Dual-point crossover scheme.



(c) Three-point crossover scheme.

Figure 3.5 Three popular crossover schemes used for paring the parent chromosomes to produce offspring for the new generation.

shown in figure 3.5(b) and 3.5(c). Like the single-point crossover the crossover points in the latter two mechanisms are also selected randomly. Under such conditions the genes of the new chromosomes come from the rearrangement of different parts of genes provided by parent chromosomes. This enriches the possible combinations of a solution in the new generation to enable the GA to reach an accepted result faster. The rule of dual-point crossover includes the effects of the single-point crossover scheme because it is essentially similar to it. Similarly, the three-point crossover scheme implies the effects of both single- and dual-point crossover schemes.

3.3.5 Mutation Operator

In a GA after a number of runs the off-spring are always very similar to their parents. This is undoubtedly true when the number of iterations is large enough, and therefore leads to the solution becoming stuck in a local pool. For exploring considerable portions of the solution surface a certain rate of random perturbations to genes are suggested in each generation to alter the gene positions. In this process a pre-selected rate of chromosomes and a defined number of bits in each are selected by randomisation and these bits are then changed inversely i.e. from 1s to 0s or vice versa (see Figure 3.6). It is a mutation process much like that of living things in the nature. Although the search space is broadened if the number of mutations is increased, a mutation likely to occur between highly fit chromosomes, and so possibly destroy them, should be prevented. From experience, a fraction less than 1% of the total bits per iteration was suggested by Haupt (1995b) and Sarabandi et al. (1997) while the probability of mutation p_{mut} most often quoted in the literature is $0.001 \leq p_{mut} \leq 0.01$ (Weile et al. 1997).

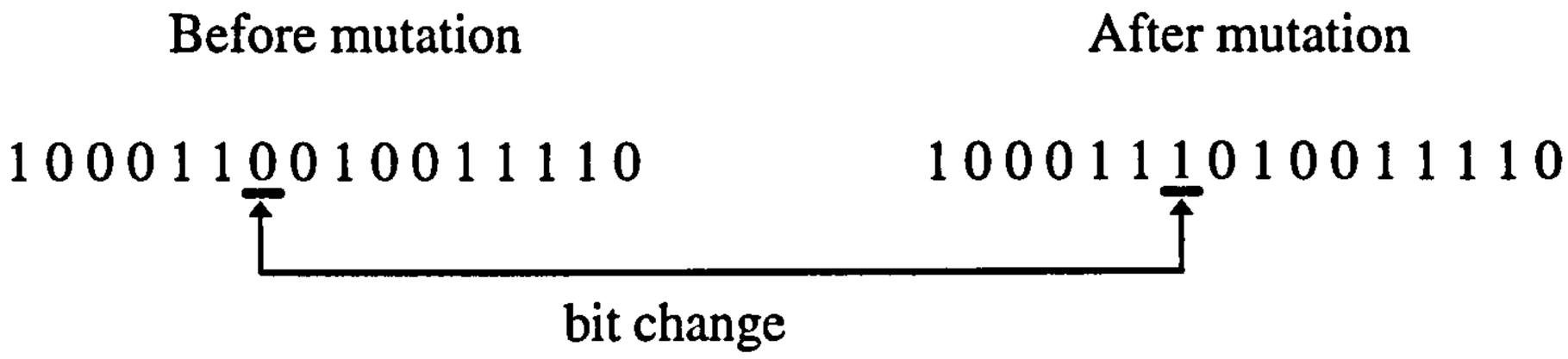


Figure 3.6 Operation of a single bit mutation.

3.3.6 Uniformity of Probability Distribution

As discussed in previous sections, during the implementation process of a GA, many steps such as the generation of initial population and the selections of crossover, swapping loci and the mutation of bits are all accomplished by randomisation following the principle of "natural evolution". The role of the randomisation process during the GA approach cannot be neglected as it can affect the GA efficiency. In this research reported here, all the GA programs are written using the C++ language which uses a multiplicative congruential generator to produce any random number. It is necessary to make the randomisation distribution uniform within the desired range as far as possible to save computer run time when generating these numbers randomly. To illustrate this we will assume a defined number of bits to be selected randomly from an example chromosome with the bit length of 100. If three random bits of this chromosome are selected to do mutation, without any processing to the first 200 numbers generated from the random generator the distribution of these numbers (as shown in figure 3.7) is usually widely spread and only two numbers (53 and 100) of them fall within the desired range from 1 to 100. It therefore will take some more iterations to produce the third number within the desired range. However, scaling the

results by an appropriate factor, 0.003 for this case, can effectively change the distribution. Figure 3.8 shows the distribution of the first 200 numbers resulting from the scaling randomisation. It only took four iterations to produce the desired three numbers which are 1, 32 and 3 (the other number is 0). Clearly it saves much time in this step with the scaling randomisation. Generally, the scaling factor is dependent upon the desired range and the program language being used. For some computer languages this step is unnecessary as they have already an installed subroutine or macro which can be defined to generate all random numbers within the desired range. However this problem should be noted to avoid too much time is consumed in this step.

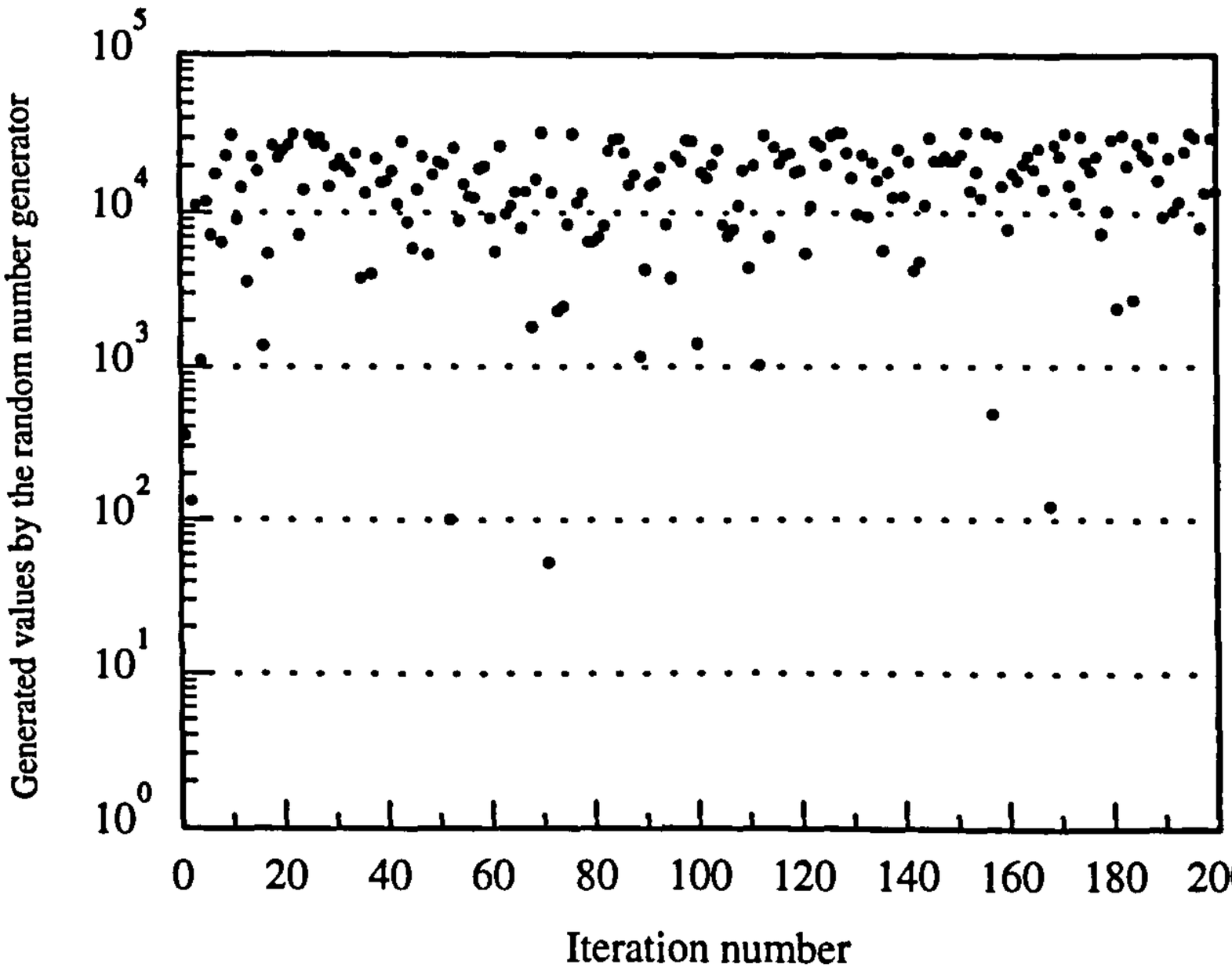


Figure 3.7 Distribution of the randomisation without scaling.

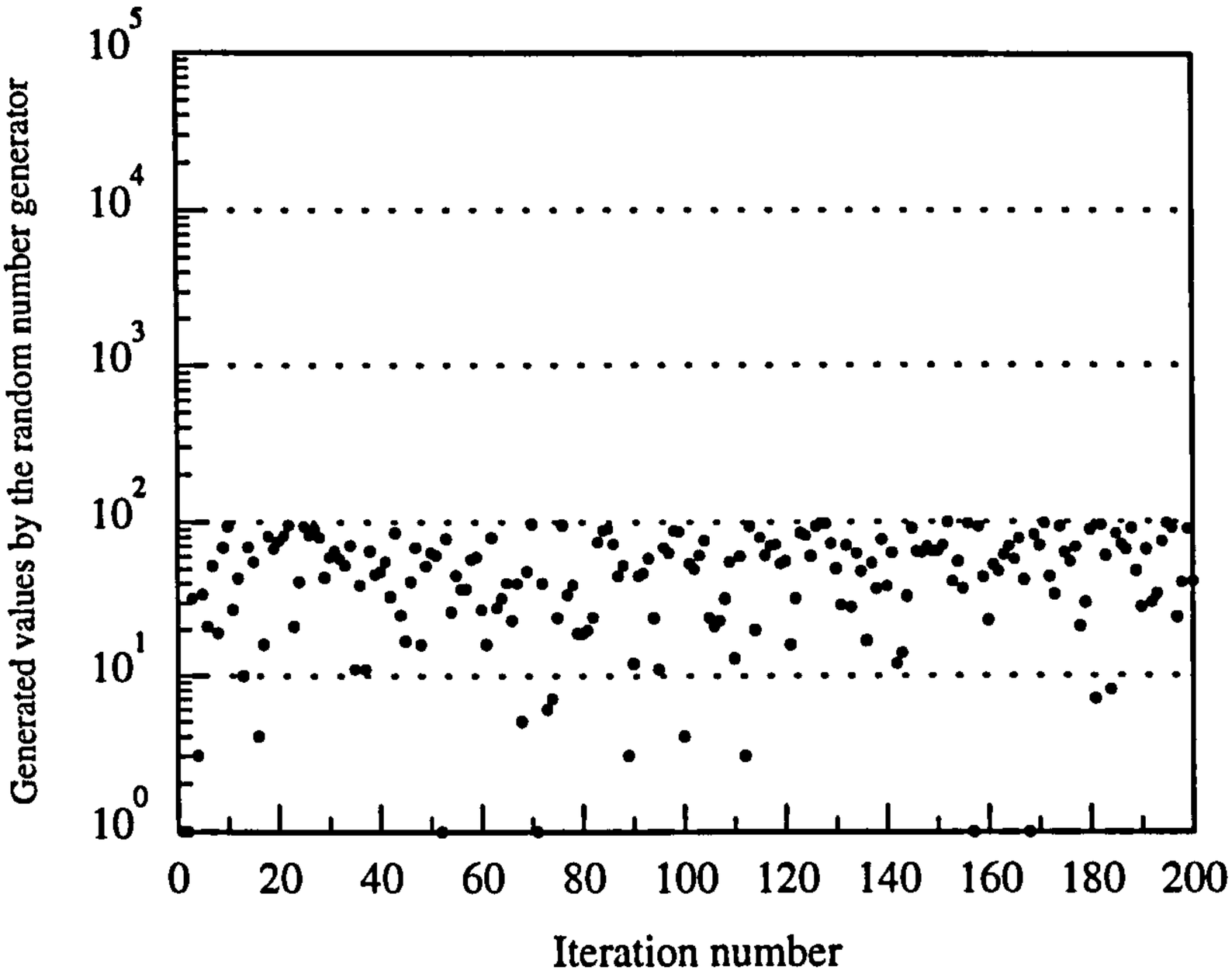


Figure 3.8 Distribution of the randomisation with scaling.

3.3.7 Global or Local Optimum ?

For most analysable problems the results from the GA approach can easily be shown to be a global optimum. By contrast, for those which are not analysable, it is difficult to determine whether the result is a global or a local optimum. Because this algorithm is based on the concept of natural evolution there is no analytical method to prove that the an exactly global optimum is always found. However, much numerical experience shows that the GA can produce the global optimum solution or at least a very close approximation to it (Johnson et al. 1997). In addition, experience also showed that different initial populations usually produce different results (Altshuler et al. 1997). This is because the algorithm always converges to a value very close to the solution and then waits for the effect of mutation before moving from any the

temporary convergence. These properties could be utilised to confirm that the global optimum has been found.

In summary, a number of trials using different initial populations is vital to help confirm the accuracy of optimisation. On comparing these, and if most are clustered closely together, then the best one of all could be regarded as the optimum.

3.4 The GA Implementation of a Case Study

The GA optimiser described above is somewhat ambiguous. In order to clearly illustrate its operation this section will introduce the GA implementation of a typical optimisation procedure in a step-by-step manner. Consider, for example, a two-dimensional function with parameters x and y that is composed of three terms of the *sinc* functions in the form written in equation 3.7. If the range within which x and y are considered is constrained to $0 \le x, y \le 10$, it can be shown that the solution surface of this equation has three main peaks accompanied by a number of small ones. The values and coordinates (x, y) of the three highest peaks from largest to smallest are: 2.0 at (4,6); 1.0 at (1,1); and 0.5 at (6,1). This equation has a well-defined global maximum within the constrained parameter range. A plot of the solution surface for this multi-peak function is shown in figure 3.9.

$$f(x, y)=2.0\left|\frac{\sin[\pi (x-4)]}{\pi (x-4)}\cdot\frac{\sin[\pi (y-6)]}{\pi (y-6)}\right|+1.0\left|\frac{\sin[\pi (x-1)]}{\pi (x-1)}\cdot\frac{\sin[\pi (y-1)]}{\pi (y-1)}\right| \qquad (3.7)$$
$$+0.5\left|\frac{\sin[\pi (x-6)]}{\pi (x-6)}\cdot\frac{\sin[\pi (y-1)]}{\pi (y-1)}\right|$$

As a candidate for an optimisation-based determination of its maximum value this equation is taxing because the three main peaks are well separated in space and it is

therefore possible to become trapped at a local optimum. However, the GA optimiser easily overcomes this problem and climbs to the highest point by its unique implementation processes of evaluation, selection, recombination and mutation after encoding the parameters x and y into a string of binary bits.

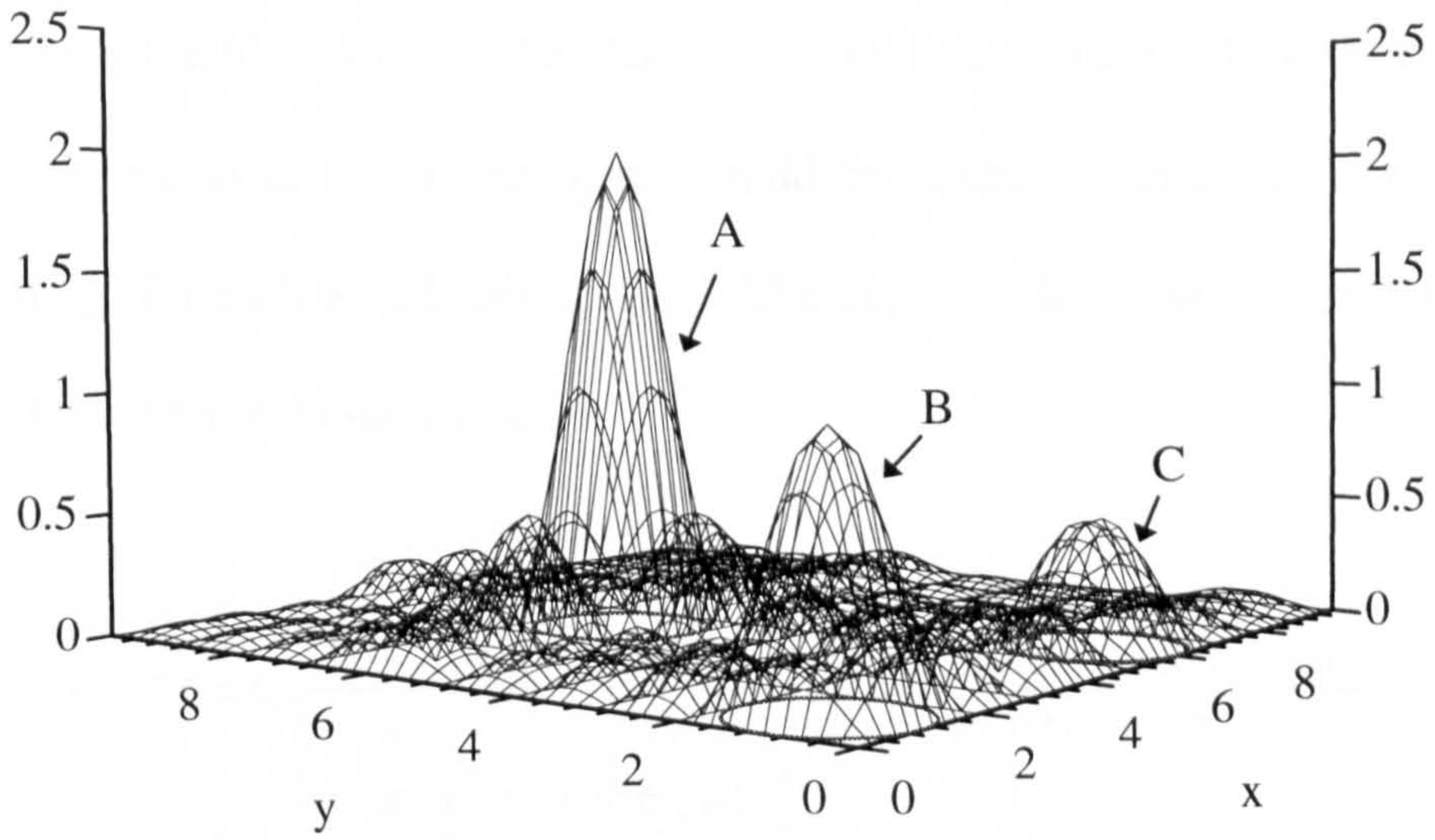


Figure 3.9 A plot of the solution surface for equation 3.7.

3.4.1 Genetic Algorithm Approach

As discussed in previous sections a GA starts by encoding each parameter into a binary string which maps to a real number in the design space. Here, for this two-dimensional equation f(x,y) two unknowns, x and y, the genes in a GA that produce a maximum value of the function are to be found. Each gene is constrained within the range from 0 to 10 and 15 bits are selected to represent its real value. This means that a variation of 10/ $(2^{15}-1)$ occurs when the binary value of each gene is changed by 1.

Therefore each chromosome composed by two genes has 30 bits and this shows that there exist 2^{30} possible solutions in the design space to be searched for the optimal value of f(x,y). After devising the scheme to encode and decode the parameters the first generation with a population of 30 chromosomes is created by randomisation. Generally, the number of chromosomes is proportional to that of the bit string thereby here the number is chosen approximately (Goldberg et al. 1991, Weile et al. 1997). If the result is not achieved it will be changed. In this case, since it is very simple to maximise the value of f(x,y), the fitness could be simply defined as the evaluated value of f(x,y) for each tested pair of (x,y). The objective function is thus defined as equation 3.8 with a maximum value of 2.

$$\begin{aligned} \textit{Objective function} = 2.0\left|\frac{\sin[\pi\,(x\text{-}4)]}{\pi\,(x\text{-}4)}\cdot\frac{\sin[\pi\,(y\text{-}6)]}{\pi\,(y\text{-}6)}\right| + 1.0\left|\frac{\sin[\pi\,(x\text{-}1)]}{\pi\,(x\text{-}1)}\cdot\frac{\sin[\pi\,(y\text{-}1)]}{\pi\,(y\text{-}1)}\right| \\ + 0.5\left|\frac{\sin[\pi\,(x\text{-}6)]}{\pi\,(x\text{-}6)}\cdot\frac{\sin[\pi\,(y\text{-}1)]}{\pi\,(y\text{-}1)}\right| \end{aligned} \tag{3.8}$$

Table 3.2 is a list of the randomly generated 30 chromosomes as the initial population and their associated values of x, y and fitness. Ranking and selection actions are carried out at the next step. To satisfy the criterion of a steady-state genetic algorithm, except for producing new chromosomes, the survivors are also left to participate in the next generation. In addition, the crossover scheme of pairing the parents from top to bottom alternatively, as shown in figure 3.4b, is used here. Synthesising these, the top one-third of the chromosomes (i.e. 10 chromosomes) have to be selected to maintain the same number of chromosomes in the new generation after reproduction as that of the previous generation. Table 3.3 shows the ranking list

Table 3.2 Schematic of the initial population.

No.	Chromosome	x , y	Fitness
1	011100101000101 100011000011101	6.35212 7.20267	0.05203
2	010010000010101 011111100101010	6.56819 3.31980	0.06245
3	000111101010000 100100100001011	0.42726 8.14753	0.02283
4	111101111011111 101000100111100	9.83886 2.36488	0.01220
5	101110110000111 011110101100110	8.81771 4.01318	0.00071
6	110111000000010 010011111110110	2.51808 4.37086	0.09912
7	110101001011000 101101010000001	1.02878 5.05295	0.01410
8	110001000110011 101010100101010	7.97967 3.30729	0.00190
9	111001101011001 101001100011001	6.04724 5.96850	0.04912
10	100100001001101 110111110100101	6.95608 6.48305	0.02124
11	110011111011111 110111111100000	9.84008 0.31098	0.01671
12	000100110010110 100010001100110	4.12366 3.98968	0.01025
13	110011011110001 001001110011100	5.60167 2.25715	0.10464
14	110100000011001 101001110111111	5.94104 9.91394	0.00617
15	110010011110111 111001010001001	9.34202 5.67614	0.09199
16	010000111100100 111101010101101	1.54363 7.08487	0.02833
17	110101000011010 011110001000111	3.45073 8.83755	0.06652
18	100011000011100 011100101110000	2.20252 0.57070	0.14380
19	000100011111101 111001000111111	7.46361 9.85595	0.00906
20	110011000110110 011111010100100	4.23444 1.46428	0.21349
21	111100000100000 111010010110001	0.16083 5.51500	0.06648
22	000110100001001 111000010110111	5.65203 9.26023	0.03750
23	100101110100100 011000010110010	1.47740 3.00974	0.00421
24	010100101110001 100001101001110	5.56963 4.48286	0.12210
25	110000110001010 101001000101110	3.18461 4.54268	0.10248
26	100011111101001 000001111011000	5.93310 1.08402	0.50505
27	001000000101010 010110110010100	3.28257 1.62908	0.11747
28	011101111001100 000001010011000	2.02582 0.98636	0.02845
29	011001101000110 101000000110001	3.85937 5.47044	1.16406
30	011101100001010 011111011111000	3.15867 1.22990	0.10499

Table 3.3 Schematic of the initial population after ranking and selection.

No.	Chromosome	Fitness	Priority	Selection
29	011001101000110 101000000110001	1.16406	1	parent 1
26	100011111101001 000001111011000	0.50505	2	parent 2
20	110011000110110 011111010100100	0.21349	3	parent 3
18	100011000011100 011100101110000	0.14380	4	parent 4
24	010100101110001 100001101001110	0.12210	5	parent 5
27	001000000101010 010110110010100	0.11747	6	parent 6
30	011101100001010 011111011111000	0.10499	7	parent 7
13	110011011110001 001001110011100	0.10464	8	parent 8
25	110000110001010 101001000101110	0.10248	9	parent 9
6	110111000000010 010011111110110	0.09912	10	parent 10
15	110010011110111 111001010001001	0.09199	11	discard
17	110101000011010 011110001000111	0.06652	12	discard
21	111100000100000 111010010110001	0.06648	13	discard
2	010010000010101 011111100101010	0.06245	14	discard
1	011100101000101 100011000011101	0.05203	15	discard
9	111001101011001 101001100011001	0.04912	16	discard
22	000110100001001 111000010110111	0.03750	17	discard
28	011101111001100 000001010011000	0.02845	18	discard
16	010000111100100 111101010101101	0.02833	19	discard
3	000111101010000 100100100001011	0.02283	20	discard
10	100100001001101 110111110100101	0.02124	21	discard
11	110011111011111 110111111100000	0.01671	22	discard
7	110101001011000 101101010000001	0.01410	23	discard
4	111101111011111 101000100111100	0.01220	24	discard
12	000100110010110 100010001100110	0.01025	25	discard
19	000100011111101 111001000111111	0.00906	26	discard
14	110100000011001 101001110111111	0.00617	27	discard
23	100101110100100 011000010110010	0.00421	28	discard
8	110001000110011 101010100101010	0.00190	29	discard
5	101110110000111 011110101100110	0.00071	30	discard

and the selected items. The relationship between the parents and the off-spring is shown in figure 3.10. The off-spring are expected to inherit both ends of the genes from their parents through the gene rearrangement. The dual-point crossover scheme which also covers the effect of single-point crossover is therefore used.

For the purpose of assisting the algorithm to escape from a local pool (i.e. preventing premature convergence), the proportion of 0.2% of the total bits was selected to make mutation. In other words, two bits in each generation were selected and their values were changed inversely. Normally, two chromosomes and one bit of each were selected by randomisation to make a mutation except that if the two selected chromosomes were the same then two bits in this chromosome will do mutation. The GA runs iteratively through the above steps and stops at a set number, say 200, of iterations.

3.4.2 Analysis of the Results

The GA starts to search for the optimal solution from a random population. Figure 3.11 shows the distribution of these individuals within the domain of the search space. Here, three circles denoted by A, B and C and centred at the three main peaks of the solution surface whose coordinates are (4, 6), (1, 1) and (6, 1), respectively, and are plotted as the reference regions to reveal how close the individuals are to the optimal solution. It is apparent, too, that the initial individuals are distributed diversely with only relatively few within the circles. As listed in table 3.2 the best individual in this generation has a fitness of only 1.16406, with the parameters x=3.85937, y=5.47044. It is still far from the optimum value of 2, with the parameters x=4, y=6. The algorithm started from this population and approach the solution iteratively.

parent 1 + parent 3 = child 1 + child 2
parent 2 + parent 4 = child 3 + child 4
parent 3 + parent 5 = child 5 + child 6
....
....
parent 8 + parent 10 = child 15 + child 16
parent 9 + parent 1 = child 17 + child 18
parent 10 + parent 2 = child 19 + child 20

Figure 3.10 The relationship between the parents and the off-spring.

After five generations most individuals apparently move towards the peaks as shown in figure 3.12. Due to the gene exchange from the similar parents, it is inevitable that some individuals were very close or even were reproduced. This results can be seen from the figure in which the solution distribution of some individuals have actually overlapped. Finally, the algorithm took 58 iterations to reach convergence and to achieve the optimal solution of f(x, y)=2.0 with x=4.000061 and y=5.999939 with the errors in the two parameters, x and y being due to bit resolution. In this situation because of repeated gene recomposition, most individuals of the population were within a very small difference of the bits in their binary strings except for the few individuals, due to mutation, that were still apart from the main peaks. These special individuals, as described before, tended to resist being grouped together with others and this action is especially important when the algorithm is approaching convergence. A plot of the solution distribution after the GA converged to the target is shown in figure 3.13.

The variation of the fitness of both the best and the average individuals of the population in each generation during the iteration process of this optimisation are illustrated in figure 3.14. The curve for average fitness fluctuated slightly as the

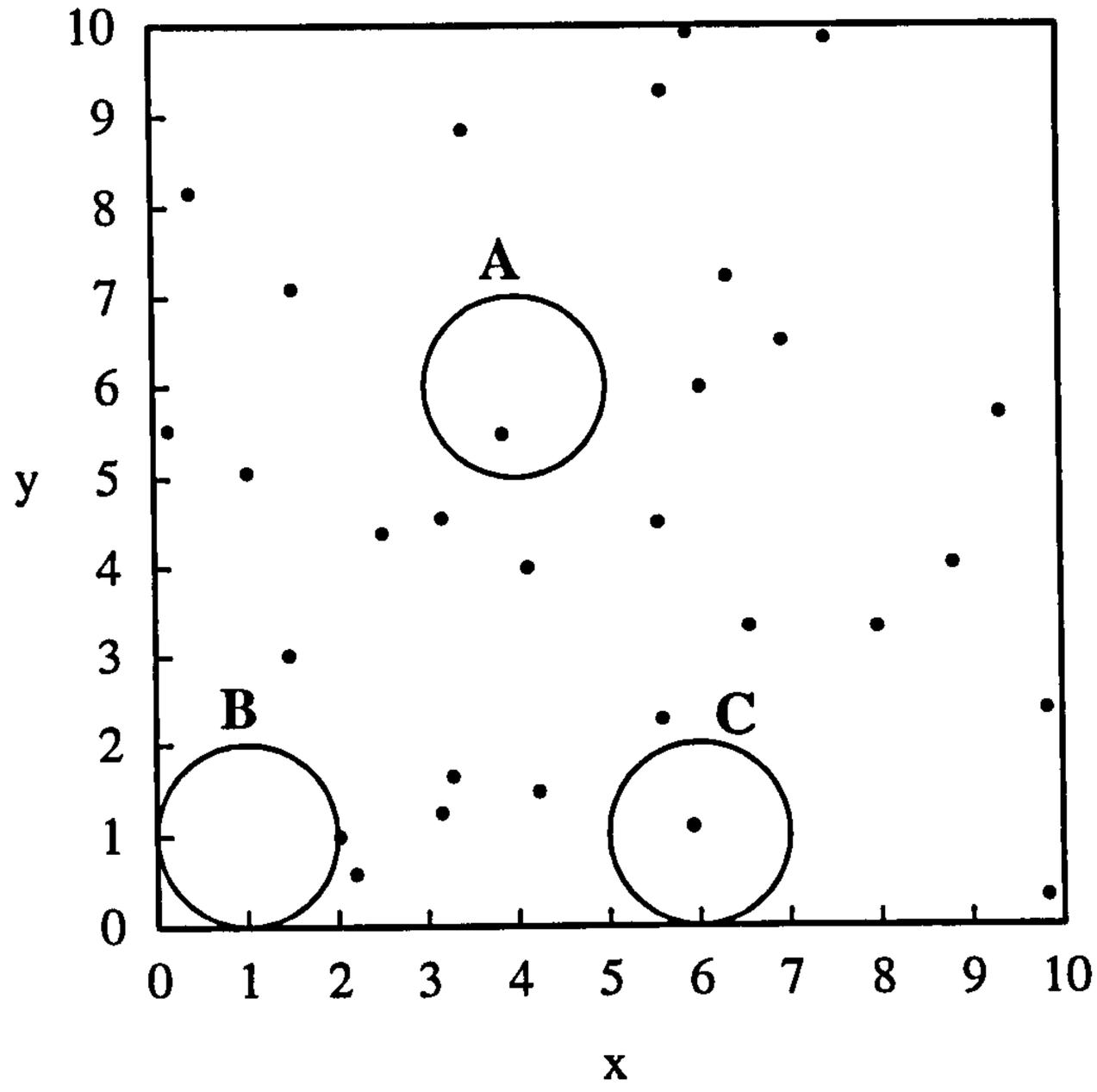


Figure 3.11 Distribution of the individuals in the initial population generated by randomisation.

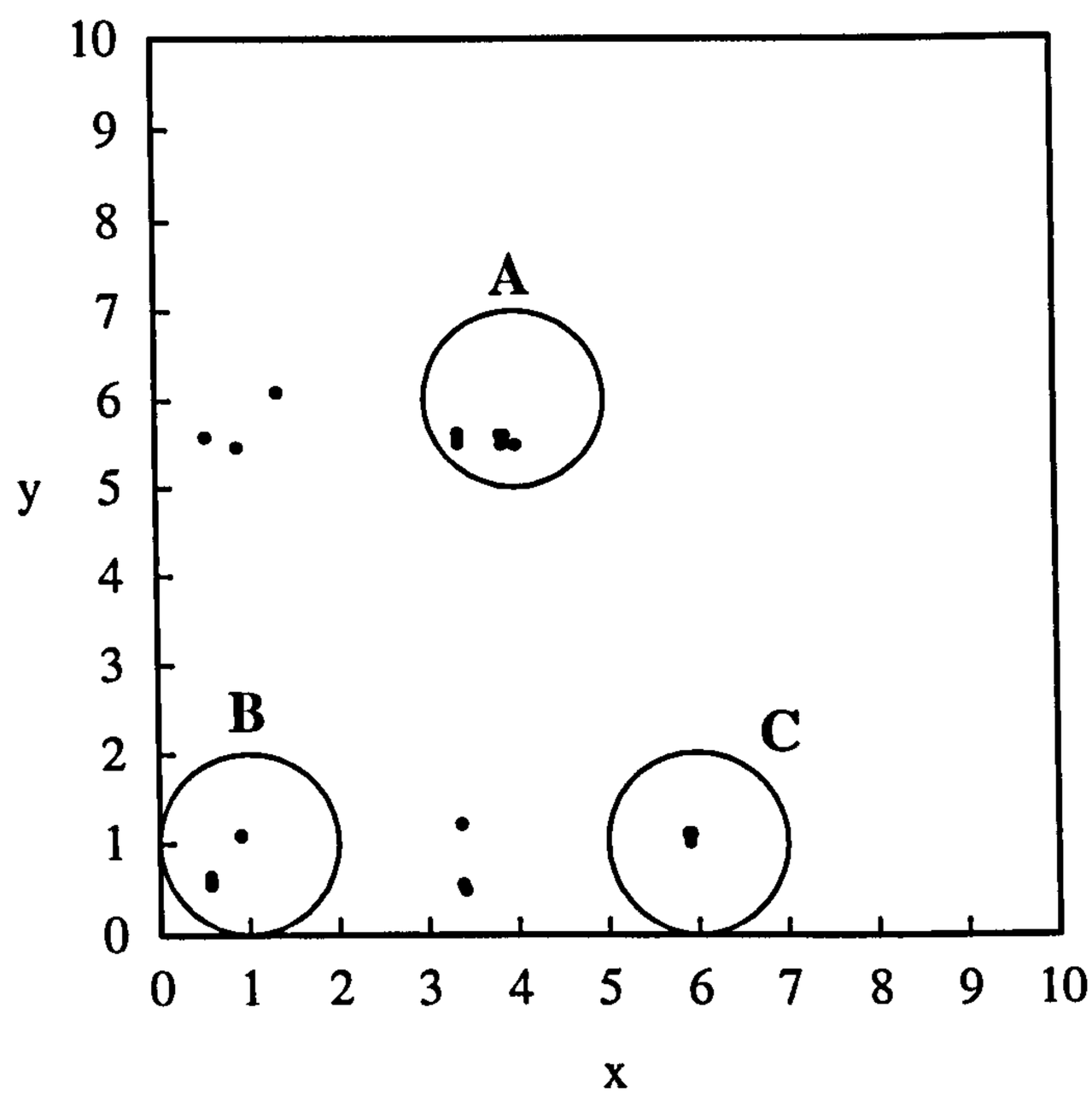


Figure 3.12 Distribution of the individuals after five generations.

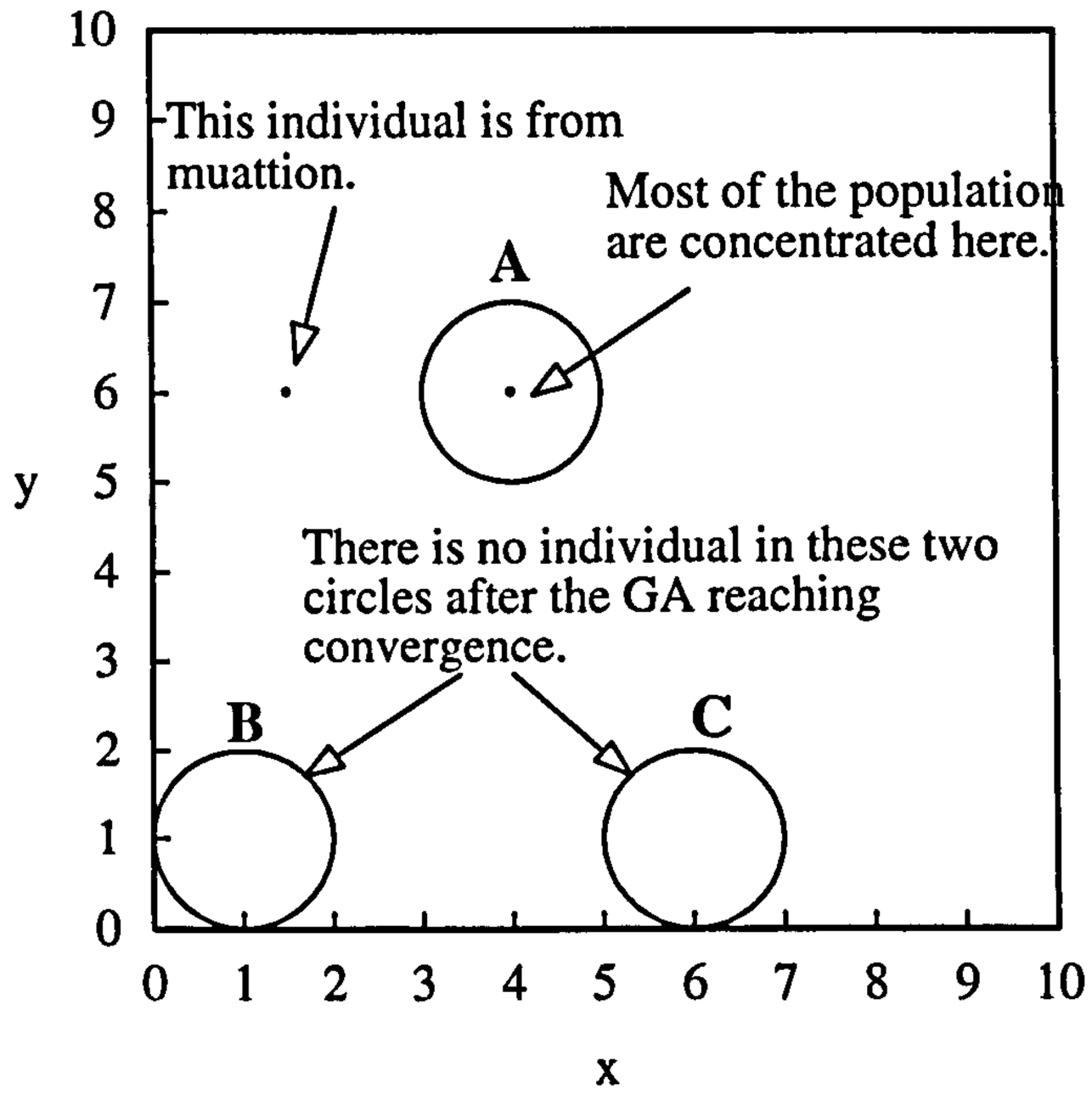


Figure 3.13 Distribution of the individuals after the GA reaching convergence.

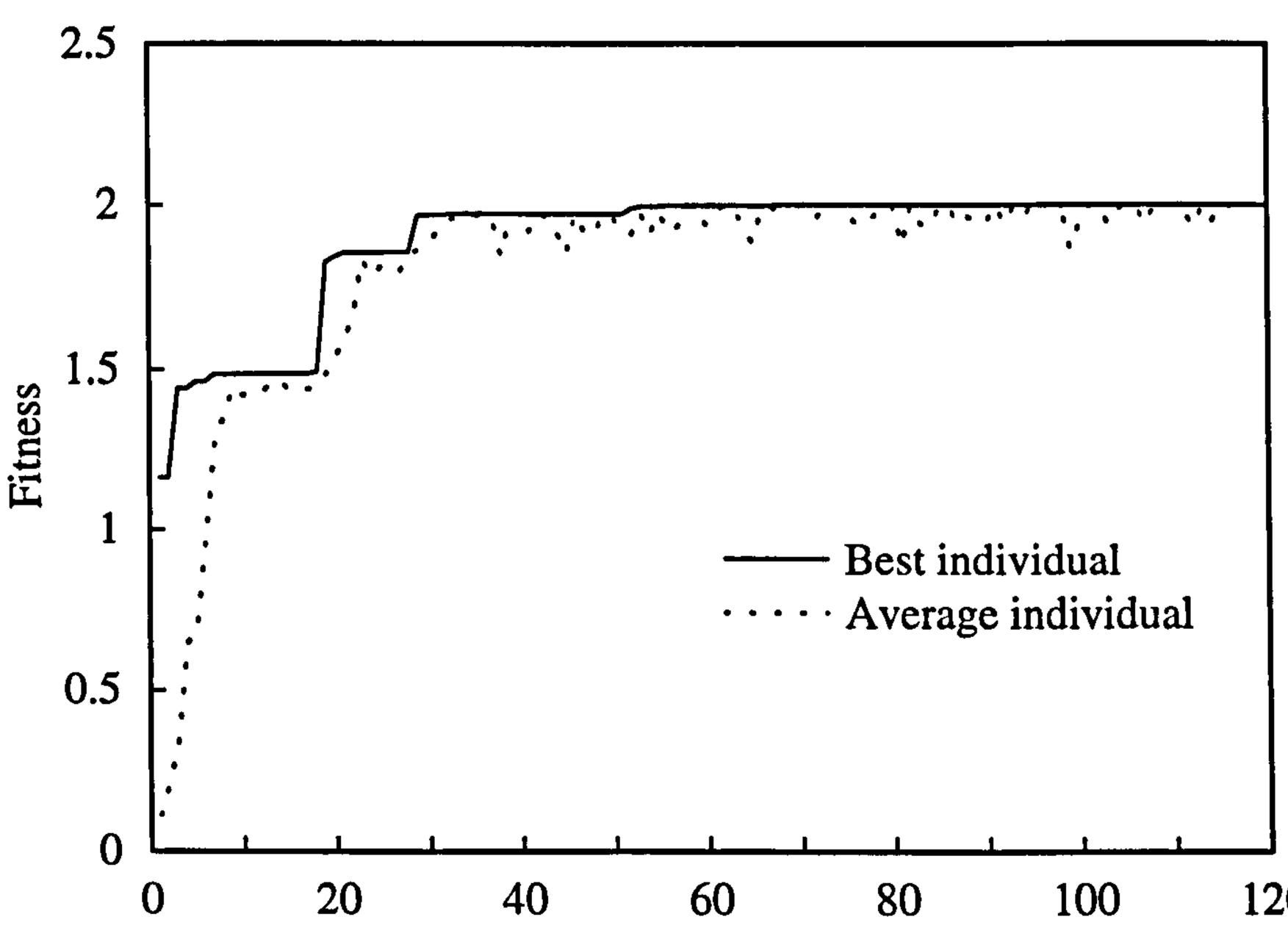


Figure 3,14 Variation of the fitness for both the best and average individual against the generation.

algorithm approached convergence and this phenomena is due to the effects of mutation.

In addition, for this case, the convergence characteristics of the GA using different numbers of both population and crossover points were compared. Four cases, structured as follows, were evaluated: (i) 30 chromosomes and single-point crossover; (ii) 30 chromosomes and dual-point crossover; (iii) 60 chromosomes and single-point crossover; and (iv) 60 chromosomes and dual-point crossover. Figure 3.15 plots the progress of the best individual against the number of generations for these cases. The differences in the observed convergence results can be attributed to several factors. First, the smaller the population, the slower the GA converges. For instance, when the population is 30 chromosomes about 60 generations are required before the single- or dual-point crossover case converges. By contrast, both cases (iii) and (iv) take less than 30 generations to reach convergence. Secondly, progress in GA optimisation occurs primarily as a result of the recombination process embodied in the crossover operator. For the same population cases, (i) and (ii); or (iii) and (iv), using the dual-point crossover scheme for gene recombination can produce higher efficiency in improving the best individual than is possible by using the single-point crossover scheme. Finally, the effect of mutation for the single-point crossover case appears more important than in the dual-point case. In cases (i) and (iii), both exhibit a relatively long period of virtual convergence (i.e. premature) before reaching the real convergence and appear to jump to the goal suddenly from this stable state. Usually, this phenomenon is due to the effect of mutation.

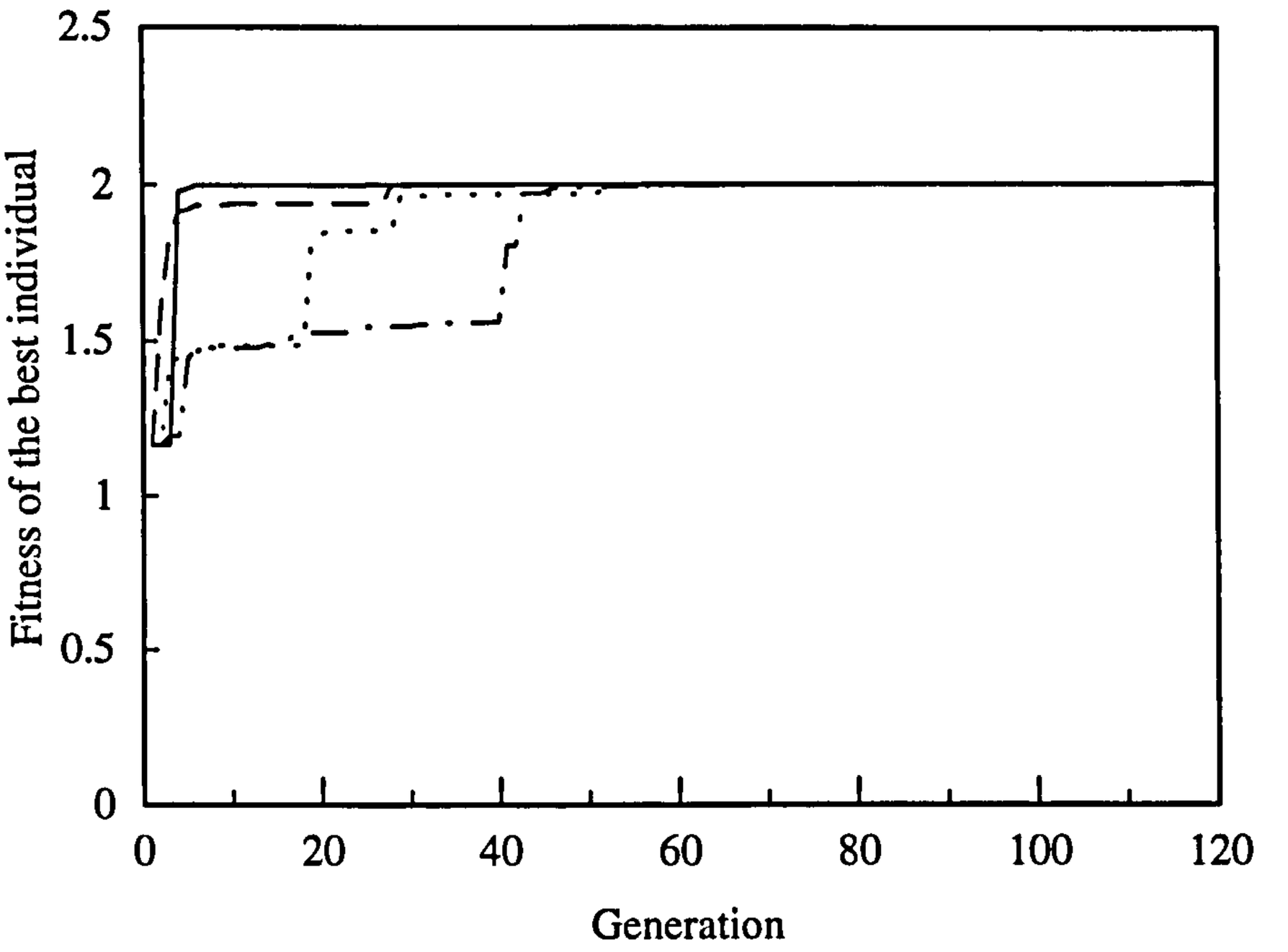


Figure 3.15 Comparisons of the convergence of a genetic algorithm using different schemes.

3.5 Summary

Genetic algorithms are well known for their ability to find excellent optimised solutions where traditional techniques fail, but they are generally slow. In this chapter the structure of a general GA and the operation of the principal GA operators have been introduced. These operators include parameter encoding, generation of the initial population, selection, crossover and mutation. Some widely used control techniques for these operators have also been discussed to provide some general guidelines for GA implementation. Finally, a step-by-step case study was then presented in an

attempt to put the discussion of the genetic algorithm into a specific context. It was shown that, for a multi-dimensional problem, the steady-state GA performed well in finding the global solution from a search space with many local optima. This can be applied to problems which are highly constrained, requiring only modifications to the objective function. Finally, the validation of the GA procedure developed for this work has also been confirmed.

CHAPTER 4

NUMERICAL ELECTROMAGNETIC COMPUTATION - THE METHOD OF MOMENTS

4.1 Introduction

In antenna design a computational method is necessary to simulate the antenna's characteristics such as current distribution, input impedance and radiation pattern etc. In the past, the applications of these kind of methods were used with many restrictions because of the limitations of computer speed and storage. These made it suitable for the analysis of antennas of simple configuration and a small size compared to the wavelength. However with great advances in speed and processing power, computers can now efficiently and effectively simulate very complex structures and electrically large antennas. One of the most important methods is the Method of Moments (MoM). Since it was developed by Harrington in 1968 the MoM has become the most widely used tool for antenna analysis, and indeed may properly be claimed as forming the basis of all numerical tools in computational electromagnetics.

The MoM is a numerical procedure used to solve integro-differential equations by transforming them into a system of simultaneous linear algebraic equations. Generally, the solution of an electromagnetic problem is achieved in two steps (Newman et al., 1991). The first is to develop an integral equation (IE) with the selection of an appropriate Green's function for the problem. The objective of this equation is to find

the unknown current density which is induced on the surface of the structure. By using the equivalence theorem and satisfying the boundary conditions applicable to the structure, the induced current density of the structure can be represented by an equivalent surface or volume current and forms a surface or volume integral equation. These equations are derived using Green's theorem. This generalises the IE so that it is suitable for solving a variety of different problems. In the second step the integral equation is then solved for the unknown induced current density using the MoM which will be introduced in this chapter.

In this work, because the MoM used for simulating the antenna characteristics is used in conjunction with the GA to design an antenna satisfying a pre-defined objective, it is vital to be familiar with its basic principles to ensure the accuracy of calculation for each antenna design. Therefore, this chapter commences, in section 4.2, by introducing the basic theory of the MoM and also by examining how an unknown continuous current distribution on an antenna can be approximated by a set of discrete functions. Also included is a discussion of the applications of different approximations to the real current distribution to determine which is most appropriate for different electromagnetic structures.

For an antenna of complex structure it is usually necessary to simplify it by means of the equivalence theorem for processing by the MoM. With this simplification a surface conductor can be replaced, depending on different approximation methods discussed in section 4.2, by small surface patches or by a mesh which is constructed of simple wires. The process of simplification must follow certain rules to assure that the results from the simplified structure closely approximate those of the real complex one. These rules are described in section 4.3.

Section 4.4 addresses the use of the Numerical Electromagnetic Code version 2-NEC2 (Burke et al. 1981a) which was developed based on the MoM and is the domain tool for this work. In this part the validity of NEC2 is verified by simulating a wire folded loop antenna mounted on a complex structure to ensure its correct use in this work. All the computed results are compared with those either from theory, experiments or other numerical methods.

4.2 The Method of Moments

In antenna analysis the determination of the current distribution on the antenna is a key step because it depends on many factors including the operating frequency, the method of excitation, the proximity to surrounding objects and the antenna's geometry. A complex geometry will increase the difficulty of computation and similarly decrease the accuracy of the computed results. Thus, the accurate approximation of the current distribution on an antenna is crucial for it is only when that is achieved that the impedance can be defined by the ratio of the electric to the magnetic field at that point. Once the accurate current distribution of an antenna is known the other characteristics such as gain, radiation efficiency etc. can all be determined easily. So far, there are many developed methods available to electromagnetic computation. However, with the following advantages, the MoM has become one of the most popular tools to simulate antenna's performance (Perez 1998).

- It is effective in modelling wire antennas and wire antennas attached to complex structures with dimensions less than several wavelengths (Burke et al. 1981b).
- It detailedly models the current distribution induced on the antenna and the structures where it is mounted.

- It can estimate the E and H fields anywhere outside the radiating structure.
- It predicts acceptable antenna parameters including the input impedance, gain and so on.

4.2.1 Mathematical Theory

The analysis of most engineering problems involves a known excitation and a derivable transfer function with the objective of determining the output. The relationships between these elements can be represented as figure 4.1. For antenna synthesis the transfer function, undoubtedly, involves both Maxwell's equations and the geometrical structure itself and the output is usually the radiation characteristics of the antenna. Correspondingly, the MoM is a method to define a numerical electromagnetic transfer function which predicts the current distribution on the structure with variable excitation.

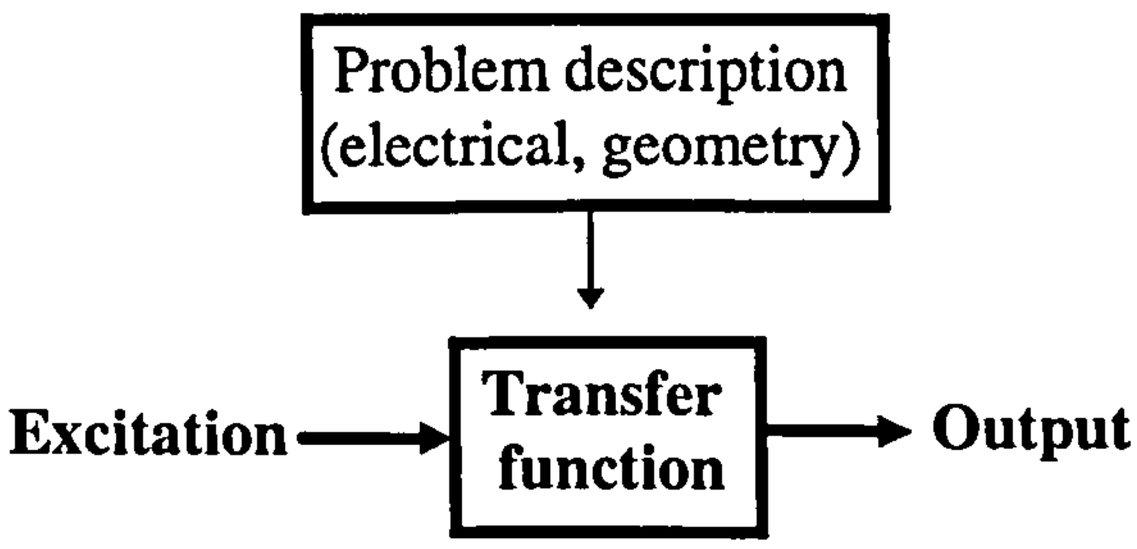


Figure 4.1 A diagram describing the relationships of an engineering or physical problem.

In the realm of functional analysis the solution of functional equations is addressed by interpreting such solutions in terms of projections onto subspaces of functional spaces, as shown in figure 4.2.

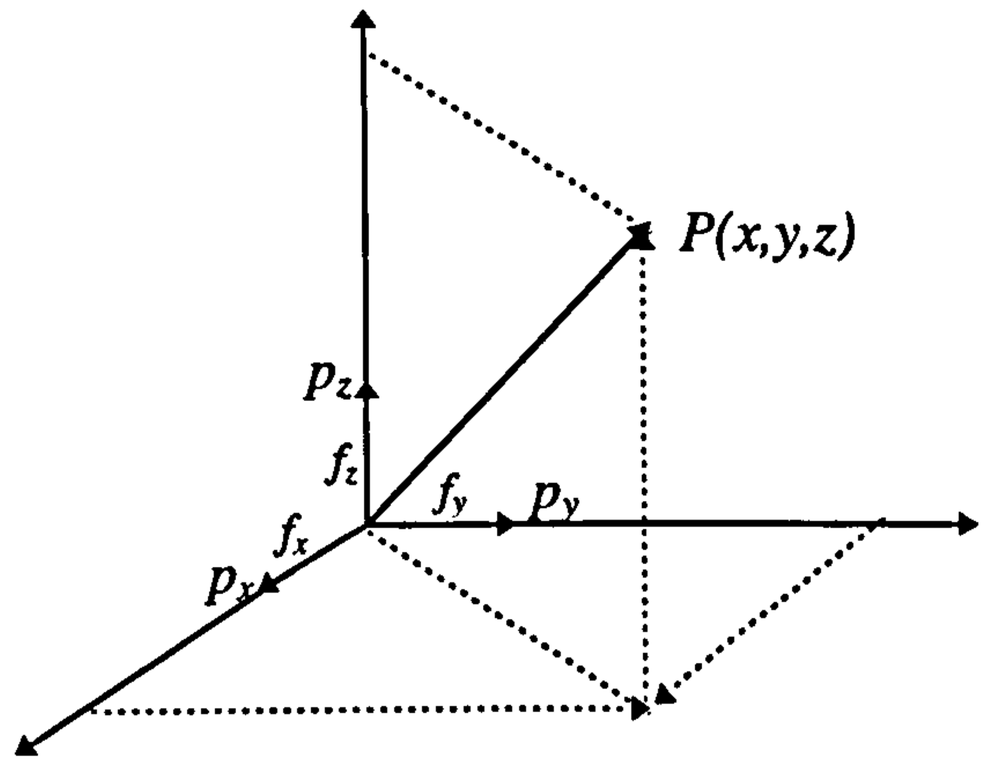


Figure 4.2 Approximate solutions by projection methods using functions.

The functional equation *p(x, y, z)* can be replaced by

$$p(x,y,z) = p_x f_x + p_y f_y + p_z f_z \tag{4.1}$$

where, f_x, f_y and f_z are the subspaces of the function space. It can be summarised as the equation of

$$L(p) = f \tag{4.2}$$

Where L is a known linear operator i.e a transfer function, f is a known excitation function and p is a response function left to be determined. For an electromagnetic radiation problem it can always be expressed as an integral, but inhomogeneous, equation with a general form as in equation 4.2. In this case, once L and f are

specified, then p can be determined easily. The linearity of the operator L makes it possible to solve this equation numerically. Let us assume that the unknown response function p could be expanded approximately as a linear combination of N terms and written as

$$p \cong c_1 p_1 + c_2 p_2 + c_3 p_3 + \cdots\cdots + c_N p_N = \sum_{i=1}^{N} c_i p_i \tag{4.3}$$

where the c_i are constant, but unknown, coefficients and p_i, known functions usually referred as a basis or expansion functions, on the assumption that the solution of p will be an approximation if N is not infinite. We substituting 4.3 into 4.2 and obtain

$$\sum_{i=1}^{N} c_i L(p_i) \cong f \tag{4.4}$$

which reduces the continuous equation 4.2 to a discrete one with N unknowns. Based on the basic knowledge of the linear algebra the solution of N unknowns requires at least N relative equations. Therefore the next step to solve this problem is to introduce the method of an inner product. A general inner product of functions w and f can be defined as

$$\langle w, f \rangle = \iint_s w^* \cdot f \, ds \tag{4.5}$$

where S is the surface of the structure being discussed. Now define a set of so-called test or weighting functions w_m according to the appropriate boundary conditions and then take the inner product of equation 4.4 with w_m to get

$$\langle w_m, f \rangle \cong \langle w_m, \sum_{i=1}^{N} c_i L(p_i) \rangle, \qquad m = 1,2,3,\cdots\cdots,N \tag{4.6}$$

In this equation, because c_i are constant and L is a linear operator, the operation of summation and constant coefficient c_i can be removed from the inner product. Thus equation 4.6 can be rewritten as

$$\langle w_m, f \rangle \cong \sum_{i=1}^{N} c_i \langle w_m, L(p_i) \rangle \tag{4.7}$$

or it can be expanded into a form of matrix as follows

$$\begin{bmatrix} \langle w_1, L(p_1)\rangle & \langle w_1, L(p_2)\rangle & \cdots & \cdots & \langle w_1, L(p_{1N})\rangle \\ \langle w_2, L(p_1)\rangle & \langle w_2, L(p_2)\rangle & \cdots & \cdots & \langle w_2, L(p_{1N})\rangle \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ \langle w_N, L(p_1)\rangle & \langle w_N, L(p_2)\rangle & \cdots & \cdots & \langle w_N, L(p_{1N})\rangle \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ \vdots \\ c_N \end{bmatrix} \cong \begin{bmatrix} \langle w_1, f \rangle \\ \langle w_2, f \rangle \\ \vdots \\ \vdots \\ \langle w_N, f \rangle \end{bmatrix} \tag{4.8}$$

From basic circuit theory, an N node circuit can be solved by Kirchhoff's theory through defining N equations with N unknown currents. The equations become

$$\sum_{i=1}^{N} Z_{ji} I_i = V_j, \quad j=1,2,3,\cdots, N \tag{4.9}$$

Clearly, the form of equation 4.8 is essentially analogous to that of equation 4.9.

Thus we can approximate the integral but inhomogeneous problem to an N mesh circuit problem and just represent Z_{ji}, I_i, V_j by

$$[Z_{ji}] = \begin{bmatrix} \langle w_1, L(p_1)\rangle & \langle w_1, L(p_2)\rangle & \cdots & \cdots & \langle w_1, L(p_{1N})\rangle \\ \langle w_2, L(p_1)\rangle & \langle w_2, L(p_2)\rangle & \cdots & \cdots & \langle w_2, L(p_{1N})\rangle \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ \langle w_N, L(p_1)\rangle & \langle w_N, L(p_2)\rangle & \cdots & \cdots & \langle w_N, L(p_{1N})\rangle \end{bmatrix} \tag{4.10}$$

$$[I_i]=\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ \vdots \\ c_N \end{bmatrix} \tag{4.11}$$

Similarly to solve equation 4.9 the unknown coefficients c_j can now be determined by solving equation 4.10 using matrix inversion techniques, as

$$[I_i]=[Z_{ji}]^{-1}[V_j] \tag{4.12}$$

$$[V_j]=\begin{bmatrix} \langle w_1, f \rangle \\ \langle w_2, f \rangle \\ \vdots \\ \vdots \\ \langle w_N, f \rangle \end{bmatrix} \tag{4.13}$$

These derivations described above are the theoretical basis of the MoM. Clearly, the number of samples N, the basis function p_i and the weighting functions w_m will govern how accurately the method approximates the continuous form by a discrete form. The following section introduces the application of this method to electromagnetic problems and discusses the selection of the different factors, which determine the accuracy of the method.

4.2.2 Applications to Electromagnetic Problems

The electromagnetic problem of an antenna consists of solving for the electric field E and magnetic field H that are created by an impressed current distribution J. In the simplest approach this current distribution is obtained during the solution process. The manipulation of Maxwell's equations is the practical method for solving any electromagnetic problem. In order to simplify the solution for E and H with a given J

the electric scalar potential function ϕ and magnetic vector potential function A are introduced into the process. Considering any source volume v' (see figure 4.3) composed of a homogeneous material the expressions relating these two quantities to charge and current density, ρ and J respectively, may be defined as (Stutzman and Thiele 1981 p11):

$$\nabla^2\phi + \omega^2\mu\varepsilon\phi = -\frac{\rho}{\varepsilon} \tag{4.14}$$

$$\nabla^2 A + \omega^2\mu\varepsilon A = -J \tag{4.15}$$

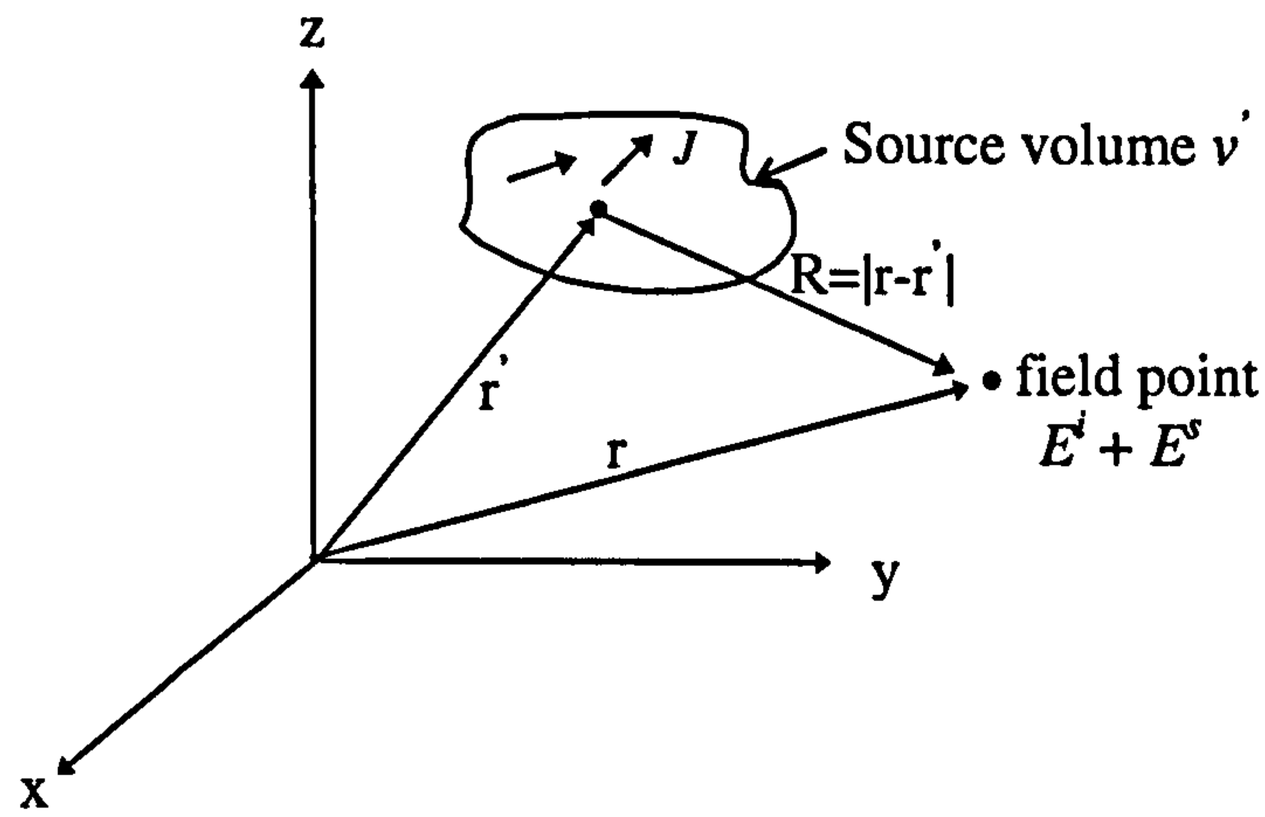


Figure 4.3 Vectors used to solve radiation problems.

where ω is the radian frequency, μ the permeability and ε the permittivity. For that the forms of equation 4.14 and 4.15 are similar, the solutions of them must be identical in form. The solutions can be summarised and written as (Stutzman and Thiele 1981 p13)

$$\phi(r') = \frac{1}{\varepsilon}\iiint_{v'} \rho(r')\frac{e^{-j\beta R}}{4\pi R}dv' \tag{4.16}$$

$$A(r') = \iiint_{v'} J(r')\frac{e^{-j\beta R}}{4\pi R}dv' \tag{4.17}$$

where r and r' are two points in the volume v', termed the observation point and the source point respectively, β is the wave number and is equal to $\omega\sqrt{\mu\varepsilon}$.

To implement the principles presented above and to suit the type of antenna used in this thesis, the following is an analysis of a cylindrical, straight wire antenna situated in free space. The antenna geometry is shown in figure 4.4.

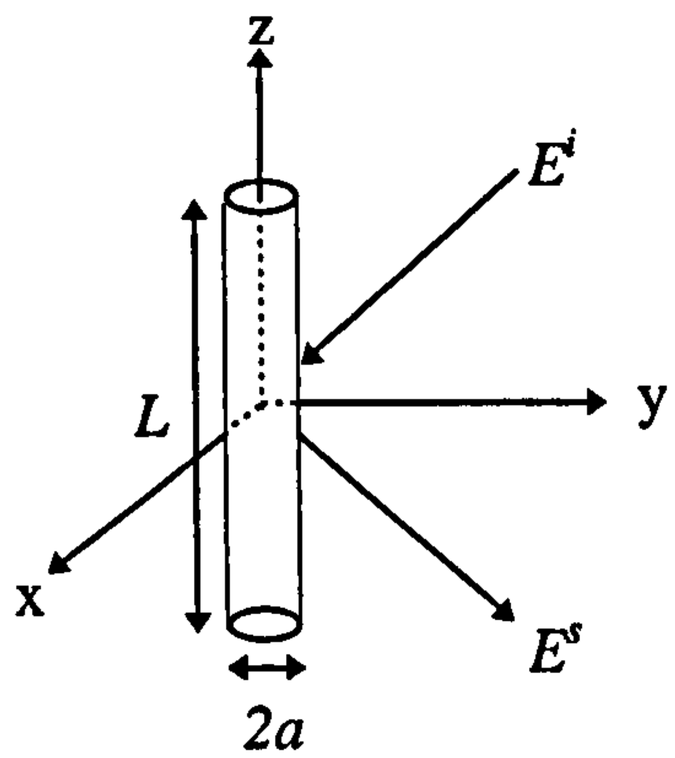


Figure 4.4 Uniform plane wave obliquely incident on a conducting wire.

Here a and L are the radius and length of the wire respectively. In order to reduce the difficulty of the required manipulations some assumptions are made in advance. Firstly, assume the wire radius to be much less than the operating wavelength i.e. the wire is thin. Secondly, the transverse current may be assumed to be negligible compared to the current flowing axially along the wire. This is especially true when the wire radius is much less than the wavelength. Thirdly, any circumferential variation of the current may be ignored and the current may therefore be assumed to flow along a filament on the axis of the wire.

Hence for the geometry shown in figure 4.4 the current at a point z' may be defined as:

$$I(z') = J(z') \cdot 2\pi a \tag{4.18}$$

Substituting 4.18 into 4.17, the equation is reduced to

$$\begin{aligned} \mathrm{A}(z) &= \int_{-\frac{L}{2}}^{\frac{L}{2}} \int_0^{2\pi} \frac{I(z')}{2\pi a} \cdot \frac{e^{-j\beta R}}{4\pi R} a \, d\phi \, dz' \\ &= \int_{-\frac{L}{2}}^{\frac{L}{2}} I(Z') \cdot G(z,z') dz' \end{aligned} \tag{4.19}$$

where G(z, z') is the free space Green's function given as

$$G(z,z') = \frac{e^{-j\beta R}}{4\pi R} \tag{4.20}$$

According to the continuity equation, the relationship between current distribution and charge density is

$$\frac{dI}{dz'} = -j\omega\rho(z') \tag{4.21}$$

so equation 4.16 could be rewritten as

$$\begin{aligned} \phi(z) &= \frac{1}{\varepsilon_0} \int_{-\frac{L}{2}}^{\frac{L}{2}} \int_0^{2\pi} \frac{1}{-j\omega} \cdot \frac{dI(z')}{dz'} G(z,z') a \, d\phi \, dz' \\ &= \frac{1}{-j\omega\varepsilon_0} \int_{-\frac{L}{2}}^{\frac{L}{2}} \frac{dI(z')}{dz'} \cdot G(z,z') \, dz' \end{aligned} \tag{4.22}$$

From Maxwell's equations, the electric field of this z-directed wire, due to A can be defined as (Stutzman and Thiele 1981, p10)

$$\begin{aligned} E &= -j\omega\mu_0 A + \frac{\nabla(\nabla \cdot A)}{j\omega\varepsilon_0} \\ \Rightarrow E_z &= \frac{1}{j\omega\varepsilon_0} \left(\frac{\partial^2 A_z}{\partial z^2} + \beta^2 A_z\right), \qquad a << \lambda \end{aligned} \tag{4.23}$$

Combining equations 4.19 and 4.23 gives the electric field E_z as

$$E_z = \frac{1}{j\omega\varepsilon_0}\int_{-\frac{l}{2}}^{\frac{l}{2}}[\frac{\partial^2 G(z,z')}{\partial z^2} + \beta^2 G(Z,Z')]I(z')dz' \tag{4.24}$$

This equation is known as Pocklington's Electric Field Integral Equation (EFIE), first derived by Pocklington (1897) to show that the current distribution on the thin wires is approximately sinusoidal and propagation speed of the current is near the velocity of light. It is the general form suitable for use with the MoM approach. Although, in this case the wire is z-directed and placed in free space, this theory is easily extended to a much more complex structure such as a vehicle, as is discussed in this thesis, with or without a ground plane. In addition, in accordance with the surface equivalence principle (Harrington 1961), the quantity E_z radiated in free space by the equivalent current $I(z')$ can be denoted as the scattered field E_z^s. When the wire is perfectly conducting the boundary condition forces the total tangential electric field on the surface and also interior to the wire to zero. This results in

$$E_z^s = -E_z^i \tag{4.25}$$

where $$\frac{1}{j\omega\varepsilon_0}\int_{-\frac{l}{2}}^{\frac{l}{2}}[\frac{\partial^2 G(z,z')}{\partial z^2} + \beta^2 G(Z,Z')]I(z')dz' = -E_z^i \tag{4.26}$$

Here E_z^i is referred to as the incident electric field, which is impressed on the surface of the conducting wire by the incident wave. Hence, if the current $I(z')$ is approximated by a series of basis functions $p_n(z')$ and the wire is divided into N segments such that

$$I(z') \cong \sum_{n=1}^{N} I_n p_n(z') \tag{4.27}$$

then

$$\sum_{n=1}^{N} I_n \frac{1}{j\omega\varepsilon_0}\int_{-\frac{L}{2}}^{\frac{L}{2}}[\frac{\partial^2 G(z_m,z')}{\partial z^2}+\beta^2 G(z_m,z')]p_n(z')dz' \cong -E_z^i(z_m) \qquad (4.28)$$

$$m=1,\ 2,\ \cdots N$$

In equation 4.28, at the centre of the m^{th} segment, the sum of the scattered fields from all N segments is set equal to the negative of the incident field at the point z_m. This is a special case of the more general MoM, referred to as point-matching or collocation, and simplifies the calculation. By comparing this equation with equation 4.9, they are seen to be similar and may be solved for p_n by using the MoM.

For deriving a more general MoM , an approach known as *the method of weighted residuals* is used. It defines the total tangential electric field consisting of both scattered and incident fields to be a residual and denoted by R, as follows:

$$R(z)=\sum_{n=1}^{N} I_n \frac{1}{j\omega\varepsilon_0}\int_{-\frac{L}{2}}^{\frac{L}{2}}[\frac{\partial^2 G(z,z')}{\partial z^2}+\beta^2 G(z,z')]p_n(z')dz' + E_z^i(z) \qquad (4.29)$$

This residual can be forced to zero in an average sense by weighting it with a series of appropriate functions, w_m, called weighting or testing functions. This is different from the point-matching method for which the residual is zero at each segment. Hence the equation of a general MoM can be illustrated as

$$\int w_m(z)R(z)dz = 0$$

$$\Rightarrow \int_{-\frac{L}{2}}^{\frac{L}{2}} w_m(z)\sum_{n=1}^{N} I_n \frac{1}{j\omega\varepsilon_0}\int_{-\frac{L}{2}}^{\frac{L}{2}}[\frac{\partial^2 G(z,z')}{\partial z^2}+\beta^2 G(z,z')]p_n(z')dz'dz + \int_{-\frac{L}{2}}^{\frac{L}{2}} w_m(z)E_z^i(z)dz = 0$$

$$m = 1,2,3,\cdots\cdots,N \qquad (4.30)$$

The current obtained from solving equation 4.30 will not necessarily be such that the residual is zero everywhere along the surface of the wire, but the average over the wire will tend to zero. This gives a more accurate current distribution for a given N

segments than the point-matching method, in which the direct delta functions are used as the weighting functions.

Meanwhile, for linear elements, in addition to the Pocklington's EFIE, there are some other methods such as Pocklington's Magnetic Field Integral Equation (MFIE) and Hallen's Integral Equation (Hallen 1938, Balanis 1997 p.392), which can be used conveniently to describe the relationship between the fields and the current distribution on conducting wires. All these methods have a common distinguishing feature IN that each is an integral equation and has a form similar to equation 4.25 and thus their solutions can also be found using the MoM.

4.2.3 Basis and Weighting Functions

When using the MoM for a numerical solution of the current distribution along the wire, one important step is the choice of basis functions. A correct selection of this function can achieve an efficient and accurate solution. Generally, it is desirable to choose those functions that accurately resemble the anticipated form of the current on the wire. In addition, these functions should be linear independent, i.e. they lie within the domain of L and satisfy Kirchoff's current law at a wire junction.

Theoretically, there are many possible basis sets. However, only a limited number are used in practice. These can be classified into two groups: subsectional functions and entire domain functions.

a) **Subsectional functions**

These functions are only suitable over a small portion of the whole structure (e.g. segments for wires, patches for surfaces and cells for volumes) and are the ones

commonly used in MoM codes. Four commonly used sets of subsectional current basis functions are:

1) Pulse functions: This is the most common and simplest one which results in a staircase approximation to the current distribution on the wire. Each pulse is nonzero only over a single segment length of one wire and half-segments are used and assigned to zero to fulfil the boundary condition at all open ends.

2) Triangular functions: When these functions are used, the result consists of a piecewise linear approximation of the current in the wires and each function covers two connected segments and each segment contains two piecewise triangular functions except for those that have an unconnected end point. An example of a piecewise triangle approximation to a current distribution is given in figure 4.5. P_n are the triangular basis functions and I_n are the appropriate scaling factors for P_n. The resulting current representation is smoother than that for "pulse functions", but at the cost of increased computational complexity.

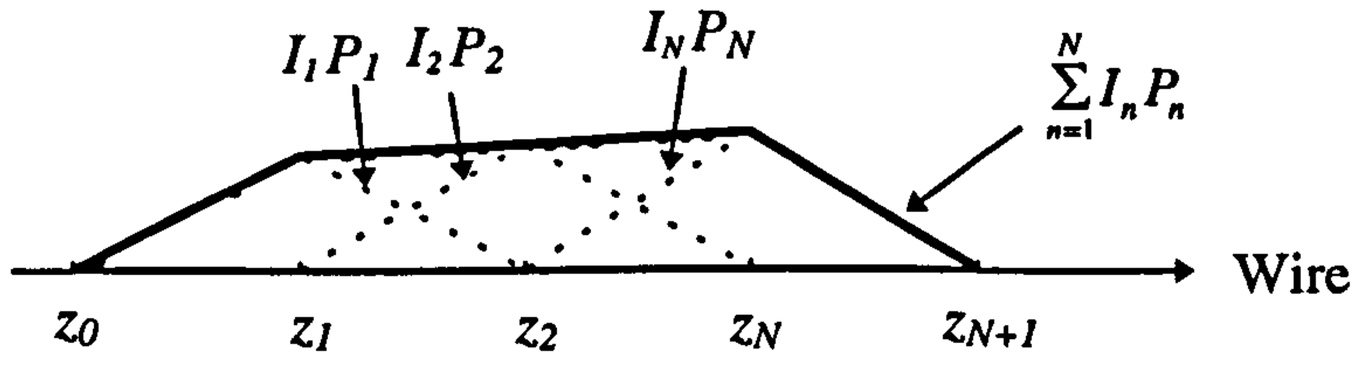


Figure 4.5 Piecewise triangle basis functions.

3) Sinusoidal functions: The piecewise sinusoidal functions were first introduced by Richmond (1974). Similar to the triangular functions, in this approximation, a set of sinusoidal functions are chosen as the elements of basis functions so that, in

some cases, the integral operators may be evaluated without numerical integration when their integrands are multiplied by a sine or cosine function. Hence, considerable advantages in both computation time and resistance to errors can be gained by using basis functions like the piecewise sinusoid.

4) Three-term functions: In this approximation, the current on each segment is represented by three terms - a constant, a sine and a cosine. Such expansions were first used by Yeh et al. (1967) and has been shown to provide rapid solution convergence and good accuracy.

b) Entire domain functions

As the name implies, this is a single, usually multi-term, function that is defined to be nonzero over the entire length of the structure being considered. Thus no segmentation is involved but an unknown factor will be associated with each term in the function. Sinusoidal functions are commonly used and this is similar to expanding arbitrary functions using a Fourier series (Balanis 1997 p400). The use of entire domain basis functions is useful when the unknown function is expected to take a particular pattern. For example, the current distribution on wire antennas is known to be essentially sinusoidal so an entire domain basis function with sinusoidal terms can be used. When the structure becomes electrically large or complicated and the current distribution departs from the simply sinusoidal, the use of this type of basis function becomes difficult and impractical; which is why it is not used in general-purpose MoM computer codes.

As discussed in section 4.2.2, when using a linear combination of the basis functions to approximate the current distribution along the wire, it is not sufficient to

solve for all unknown coefficients that scale each term of the basis functions by using only one equation. In order to have enough linearly independent equations the weighting, or test functions must satisfy the boundary conditions in an average sense over the entire surface. Furthermore, the choice of a proper form of weighting function can generally provide an added advantage to simplify the computations required for evaluating the inner product. The condition of linear independence between elements and the advantage of computation simplicity are also important characteristics of basis functions. Because of this, when the domain of weighting and basis functions are the same, similar types of functions are often used for them both. This special case is referred to as Galerkin's method (Finlayson 1972) and has been shown to be especially useful when solving problems that involve symmetrical impedance matrices (Austin and Murray 1991).

4.3 Wire-grid Modelling Techniques

As described in the previous section, the Method of Moments is used to solve an electromagnetic integral equation, which is generally the EFIE or the MFIE, and is derived in accordance with the defined geometry which could be a two-dimensional or three-dimensional body. When using the MoM some method of representing the surface of the body in a wire mesh form is usually necessary. This procedure is known as wire-grid modelling and was first developed by Richmond (1966) for solving some canonical scattering problems. It involves the realistic assumption that the wire mesh will behave electromagnetically in roughly the same way as the modelled solid surface under similar conditions. However, it should also be noted that if the approach is the

MFIE the surface can only be modelled by small patches (or cells) each having a continuous surface because of the requirement that it only applies to a close surface.

Fortunately thin-wire modelling codes are really available and are also easy to develop. However, when using this technique, the successful substitution of the resultant current for the actual current flow on the continuous solid surface depends upon three factors which are: (i) the shape of the cells within the mesh; (ii) the size of the cells and (iii) the radius of the wire segments used in the cells. Some guidelines on determining these are illustrated as follows.

a) **Cell shape**: In the literature different cell shapes such as rectangular (Trueman et al. 1991), triangular (Moore et al. 1986; Taboada et al. 1999) and polygonal (Ferguson et al. 1988) have been used to build the wire-grid models for different structures. However, the rectangular shape is the most widely used configuration for it is relatively easy to generate and gives satisfactory results. In this formulation, the current flow is always limited to orthogonal directions in the mesh. This can mean that for certain special structures the current flow is not modelled correctly. Therefore an alternative formulation – triangular-shaped cells - are introduced to overcome this problem (Moore et al. 1986). Being non-orthogonal in shape has the advantage of simulating the current more realistically and it is suitable for the case when the direction of the current flow is unknown. However, compared to the quasi-orthogonal mesh, the triangular mesh is more difficult to generate. It is also more cumbersome to decide on the wire radius by using simple formulas which will be introduced below. Meanwhile, for complicated structures, some parts also need to be modelled with irregular-shaped cells in order to follow the details of the surface topology. Overall though, it is

apparent from the literature that the rectangular or quadrilateral grid cells are adequate, provided that the other wire grid modelling design criteria, grid size and wire radius are chosen with care.

b) **Cell size:** As the methodology of the wire-grid model is only an approximation to the continuous surface the cell size should be small enough compared to a wavelength to ensure that the representation is reasonable. However, this will increase computation time and requires more storage space. Hence various methods aimed at obtaining the maximum available cell size are investigated. A maximum cell area of $0.02\lambda^2$ to $0.03\lambda^2$ was suggested by Moore et al. (1986, p107). For a square cell this would correspond to a maximum grid spacing of approximately $\lambda/7$ to $\lambda/6$. Lin et al. (1975) concluded that a grid spacing of less than $\lambda/4$ can accurately predict the radar cross section (RCS) of aircraft. Ramzi et al. (1998) used wire-grid modelling to model a vehicle in the AM broadcast band and concluded that a grid spacing of less than $\lambda/100$ can provided accurate results. Nishikawa (1984) also empirically selected a wire length of $\lambda/200$ to model motor vehicles and obtained good results. However, the generally accepted value which can achieve a good compromise between the computational cost and numerical accuracy for most applications is in the range from $\lambda/20$ to $\lambda/10$ at the highest frequency of interest (Poggio 1973, Miller et al. 1992). In addition, it is also suggested that a scheme keeping the cell area constant throughout the structure, except for the edges or the areas with high curvature in which a fine mesh with wire along the edges, should be used (Perez 1998 p315).

c) **Radii of wires:** Another important factor which affects the accuracy of the wire-grid model is the wire radius. In the first uses of this technique, an empirical value of 0.005λ was used for different studies (Richmond 1966, Nishikawa 1984, Taniguchi et al. 1985). Later some other methods for determining the wire radius were proposed. Among them was the "twice surface area rule" which states that the sum of the surface area of the wires in one direction must be double the total area of the surface replaced by the wire grid (Lee et al. 1976, Burke and Poggio 1981a, Ludwig 1987). Experience has shown it to be most appropriate. This correct selection of wire radius can avoid an extra self-inductance from being added to the solid surface and this has been proved theoretically by Lee et al. (1976).

4.4 The Numerical Electromagnetic Code – NEC2

The Numerical Electromagnetic Code (NEC) is a popular MoM code developed by Burke and Poggio (1981a) for analysing complex metallic structures by means of wires.

Currently, the most widely used version is NEC version 2 (NEC2). It is the main code used in this work for the analysis of the electromagnetic response of the antennas under investigation. As with other electromagnetic codes it is built around the numerical solution of integral equations for the currents induced on the structure by sources or incident fields. The solution of the EFIE for wires and that of the MFIE for smooth closed surface are both included in this program. In the first case the wires are modelled by short straight segments with the current on each segment represented by

the three-term functions as introduced in section 4.2.3, while in the second, the surfaces are modelled by small patches with the surface current in each patch expanded in a set of pulse functions except in the region of wire connection. The schemes for simplifying the geometry provides for convenient and accurate modelling of a wide range of structures. Delta functions (point matching) are selected as the weighting functions for this MoM code. However, unlike other solution methods requiring many assumptions to simplify the approach, NEC2 approximates the problem very closely and can therefore provide a more accurate solution.

Essentially, this code has the capability to simulate a model which includes perfect or imperfect conductors, lumped-element loading and both nonradiating networks and transmission lines which connect parts of the structure together. In addition, a structure in free space, over a perfectly conducting ground plane or over a real ground plane can also be modelled accurately; even if the structure is very close to the ground surface.

Following the guidelines discussed in section 4.3, and with excitation by voltage sources on the structure or by an incident plane wave, NEC2 can accurately compute parameters such as induced current and charges, input impedance, near electric or magnetic fields, far electric or electromagnetic fields, gain, directivity and radar cross sections.

4.5 Validation Using a Handset Model

The NEC2 code has been shown to be successful in modelling a variety of antennas, however, the accuracy of the simulated results is dependent on the user following the guidelines which have been discussed in previous sections. Therefore,

before applying NEC2 with confidence to this research programme some confirmation or validation of its correct use must be obtained by comparing the predicted results with those published for a representative example. It is the purpose of this section to assess the accuracy of both the input impedance and radiation pattern results obtained using NEC2 when modelling a number of complex continuous surfaces represented as wire grid models.

4.5.1 Structure Definition

The input impedance and radiation patterns of a vertical folded loop antenna centrally mounted on a conducting box, simulating a small hand-held portable telephone with and without dielectric coating, have been experimentally and analytically investigated by Katsibas et al. (1998). In their work, the analytical method of the Finite-Difference Time-Domain (FDTD) (Taflove 1995) was used to simulate the electromagnetic properties of this antenna system. Furthermore, for the case of the box without coating, the results were also predicted by using NEC. These analytical results were all compared with measurements and very good agreement was obtained. However, it should be noted that in their illustrations of the antenna system, (except for the results) there is no comment on certain key parameters such as the mesh size and the radii of the wires used in generating a wire-grid model of the box. For a number of reasons this example was chosen to validate the author's use of NEC2. Firstly, the paper by Katsibas includes results from measurements as well as from a totally different modelling technique from NEC2. Secondly, it involves a geometry whose complexity is very similar to that in this research work.

Figure 4.6 shows the geometrical configuration of this antenna system. The vertical wire folded loop antenna with dimensions of 2 cm and 4 cm in height and in length, respectively, was mounted centrally on the top plate of the conducting box which has a three-dimensional structure of 3×6×10 cm in the x, y, z directions, respectively. The excitation was fed to the base of the left side of the folded loop and the other end of the loop is short-circuited to the box. This antenna is intended to work in free space in the cellular radio band, typically 900 MHz. In Katsibas's investigation two cases with and without dielectric coating on the box surfaces were considered. However, the use of NEC2 is not suitable for the analysis of a box with dielectric coating. So only the box with perfectly conducting surfaces was considered here.

In order to model this geometrical configuration using NEC2 the box was represented by wire grids. In Katsibas's results, the dominant (first) resonance occurs at the frequency of around 1 GHz. The wire-grid model of this box is therefore developed on the basis of this frequency. Since the largest dimension of this box is small compared to the wavelength at 1 GHz it is reasonable to divide each surface of the box into meshes with a side length less than 0.1λ which is usually suggested for acceptable accuracy and minimal computation cost. The grid spacing was therefore chosen to be 1 cm (i.e. 0.033λ at 1 GHz) and was based on the so called "twice surface area rule"; the radius of each wire being 0.125 cm. The resultant wire length is much smaller than the wavelength and the ratio of each wire's length to its diameter is equal to four. These all satisfy the criteria for a thin-wire approximation. In addition, the loop antenna was also segmented by the same length as used for grid spacing. These values were maintained for all calculations over the frequency band of interest. Figure 4.7 shows the completed model for this handset system.

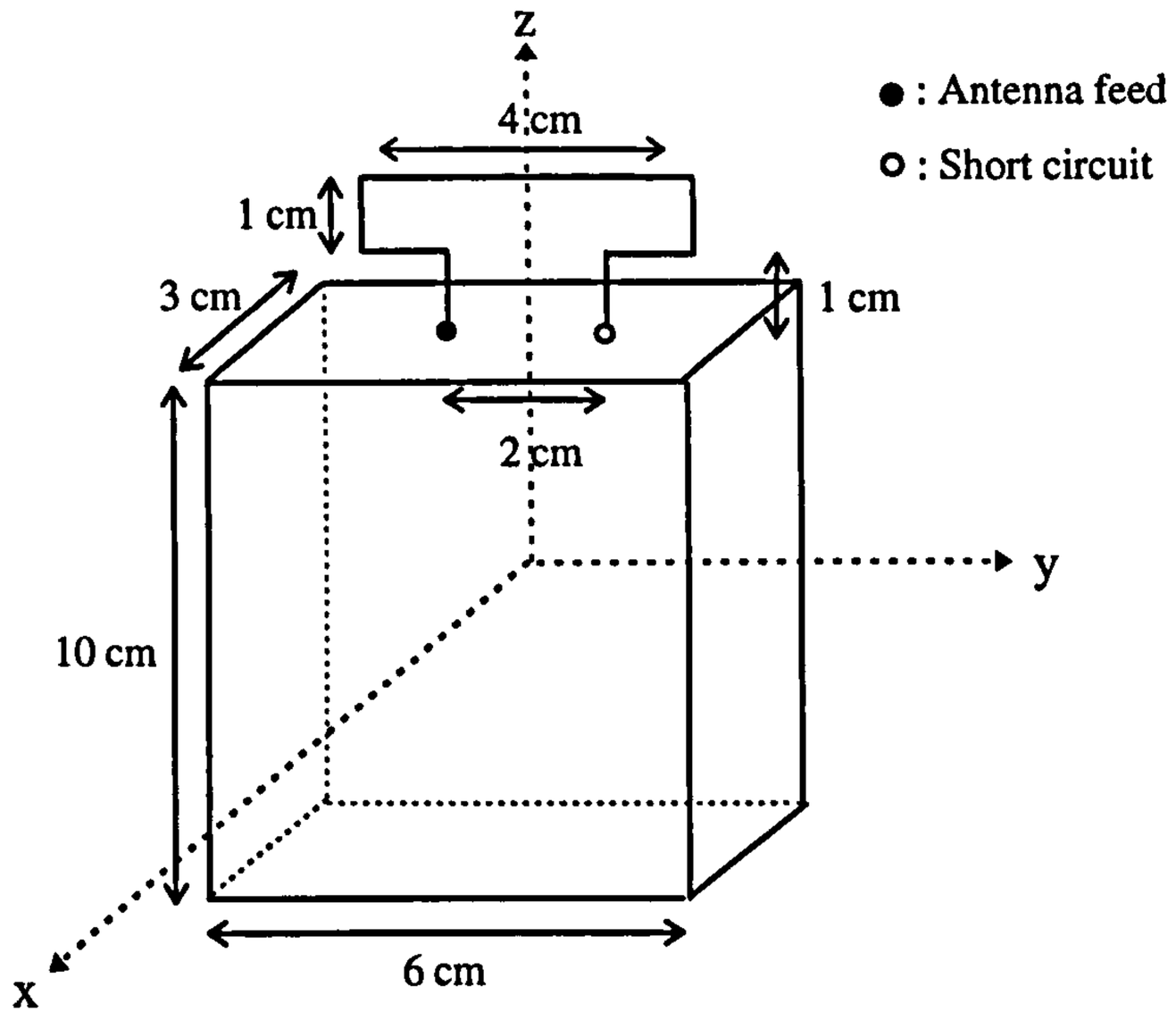


Figure 4.6 Vertical folded loop antenna centrally mounted on a perfectly conducting box.

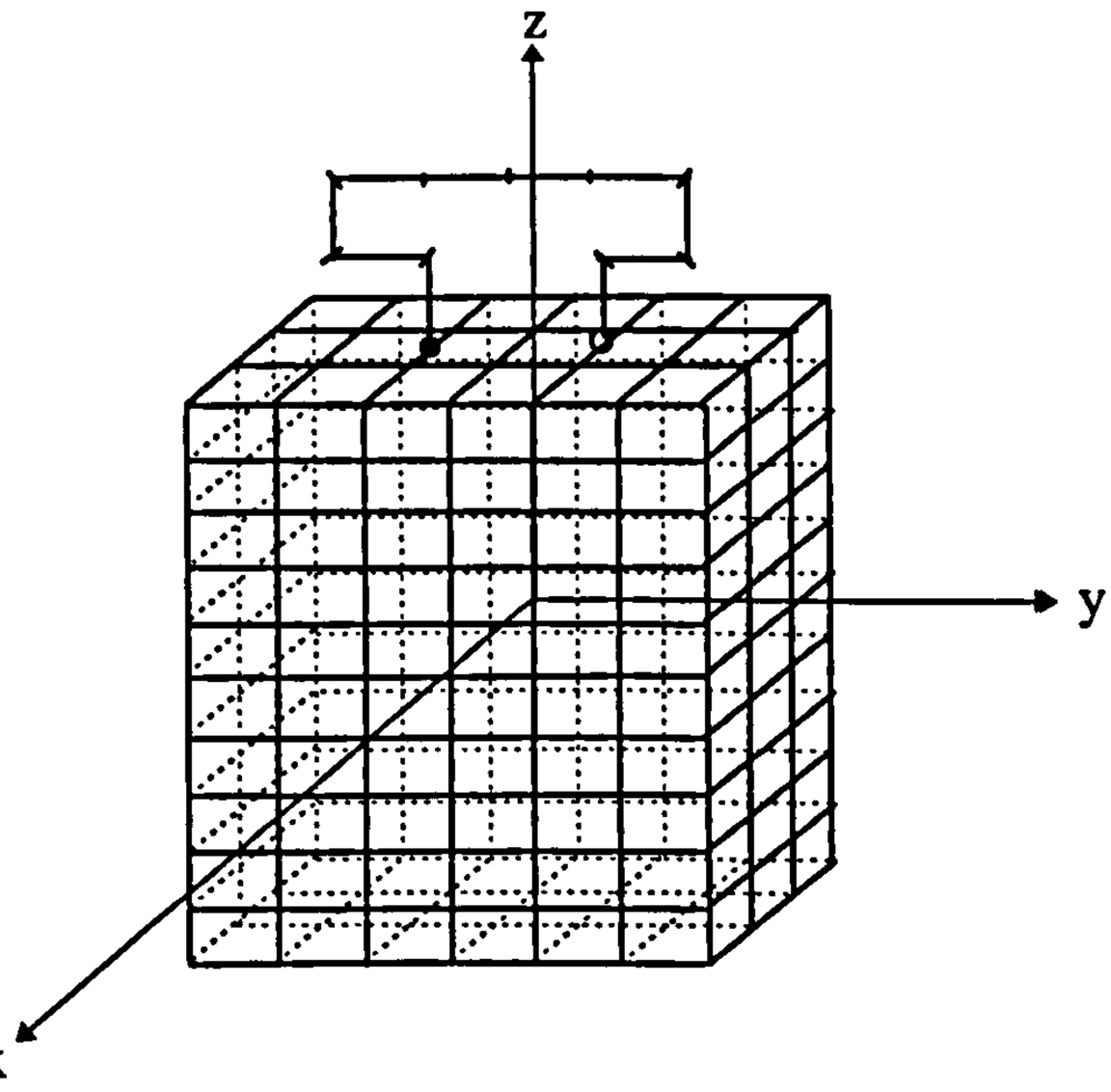


Figure 4.7 Wire-grid model of the vertical folded loop antenna centrally mounted on a perfectly conducting box.

4.5.2 Comparison of Predicted and Published Results

The results from NEC2 for the input resistance and reactance of the antenna configuration over the frequency range below 5 GHz are shown in figures 4.8 and 4.9. From basic inspection, it is clear that the results are in very good agreement to Katsibas's results (see figures A.1 and A.2 in Appendix A) obtained by both measurement and calculation using the FDTD method, over the frequency band. The first resonant frequency occurs at 945 MHz. This value is exactly same as that obtained from measurement and the FDTD method. From the plot of reactance frequencies for the second and third resonance are approximately at 2050 and 2625 MHz, respectively. They are of the series type with smaller values of resistance. These two values are, respectively, very close to the 2150 and 2600 MHz results of Katsibas.

The far-field radiation patterns for the co-polar E_θ components of this antenna in the principal planes were also analysed at the frequency of 900 MHz by NEC2 to allow more comparison with Katsibas's results (see figures A.3-5). These results are illustrated in figure 4.10-12. Figure 4.10 presents the pattern of the E_θ components in the azimuth plane. Although the gain at the feed side is little higher than that at the short-circuited side, it exhibits nearly omnidirectional radiation as is desirable for mobile communication system coverage. The gain shift is caused by the fact that the antenna feed is not positioned at the centre of the box. The pattern in the xz-plane are shown in figure 4.11. As expected it is symmetric since the structure is also symmetric in this plane. The effect of the feed position on the radiation patterns is shown in figure 4.12. The left lobe, which is located at the same side as the feed, is dominant in the yz-plane. There is also a deep null which is positioned at 30° from the z-axis on the opposite side to the antenna feed. This occurs because the antenna is fed

asymmetrically from the centre. These results once again are all in excellent agreement with the work of Katsibas.

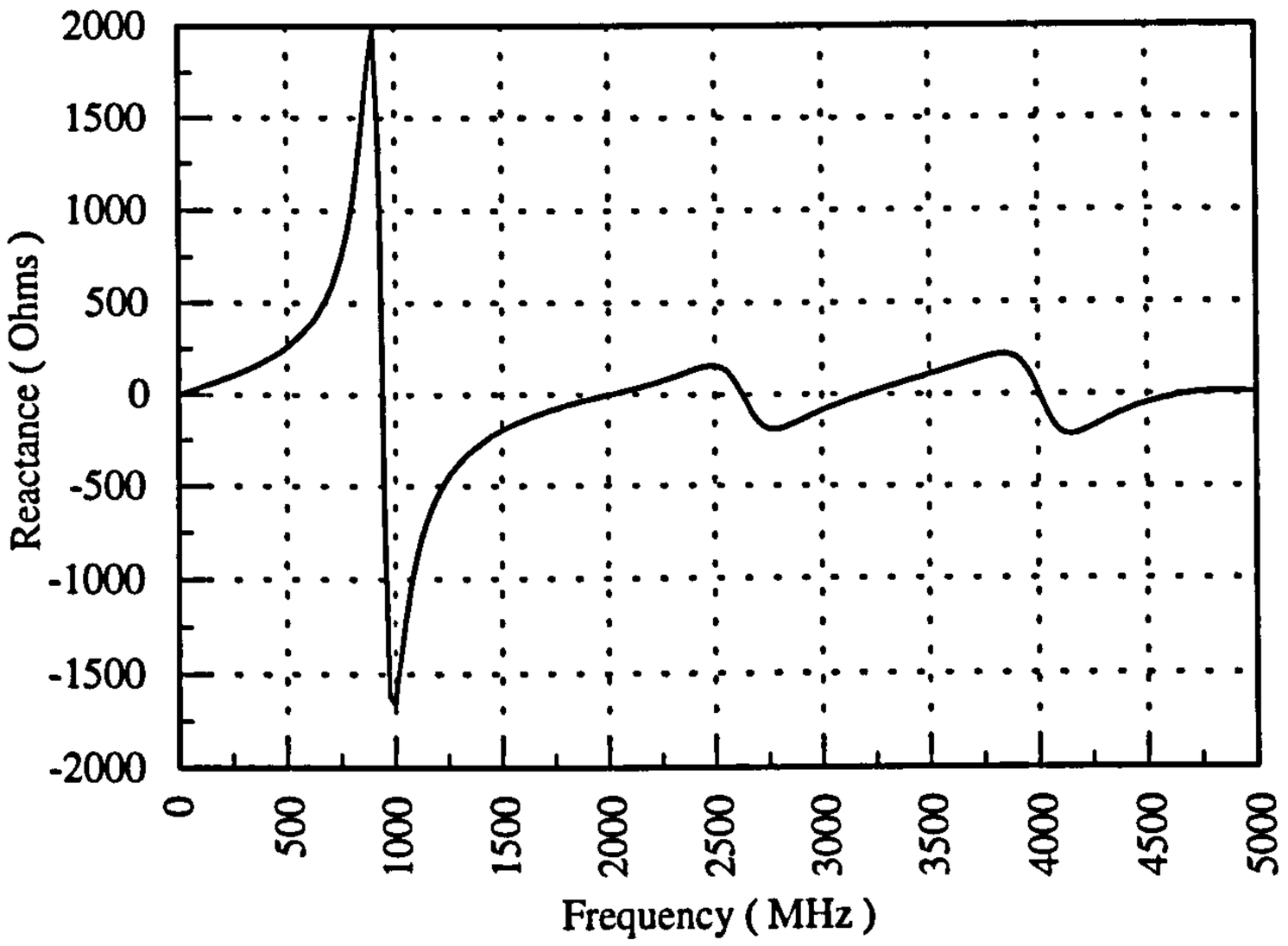


Figure 4.8 Input reactance of wire folded loop antenna centrally mounted on a perfectly conducting box.

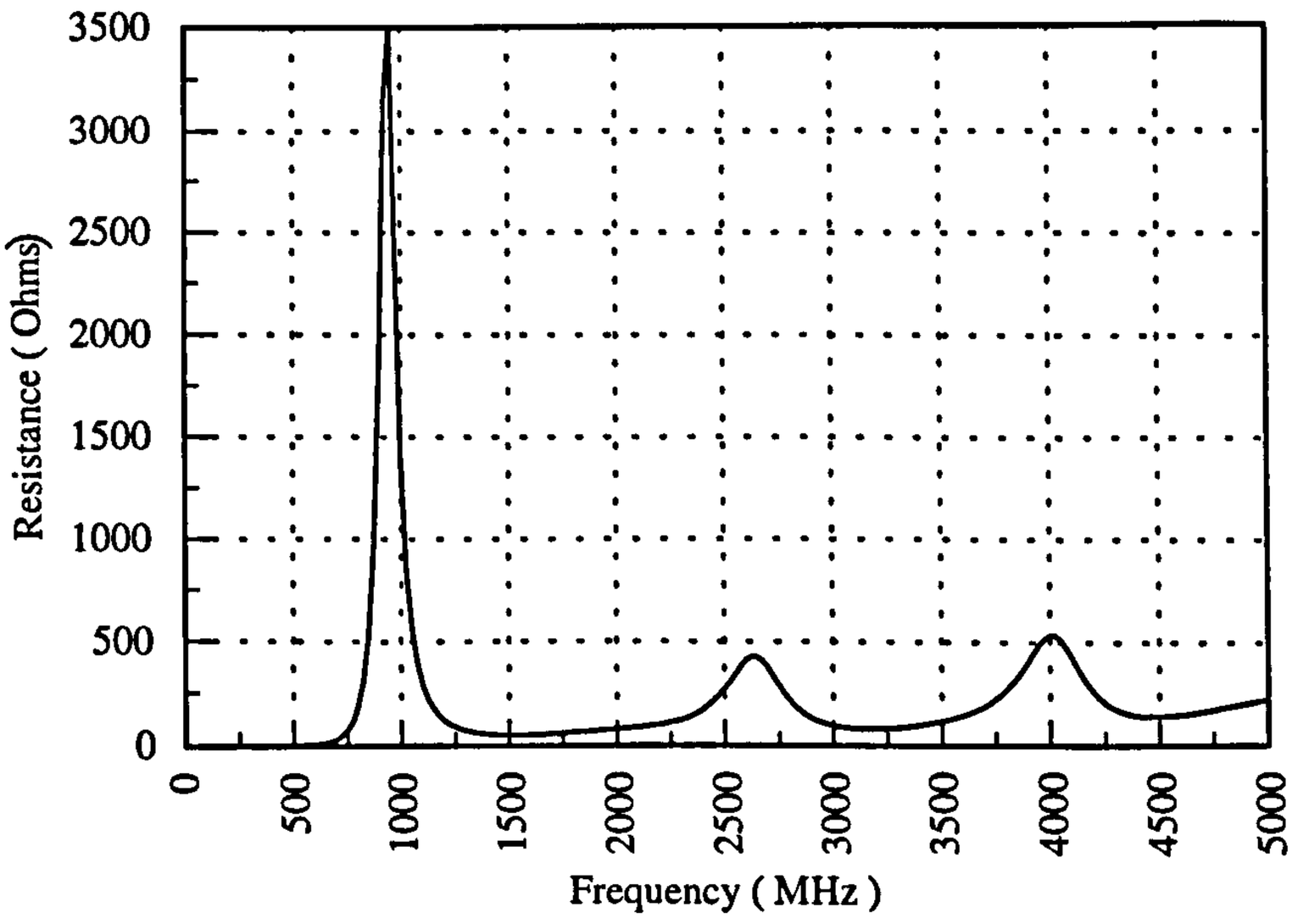


Figure 4.9 Input resistance of wire folded loop antenna centrally mounted on a perfectly conducting box.

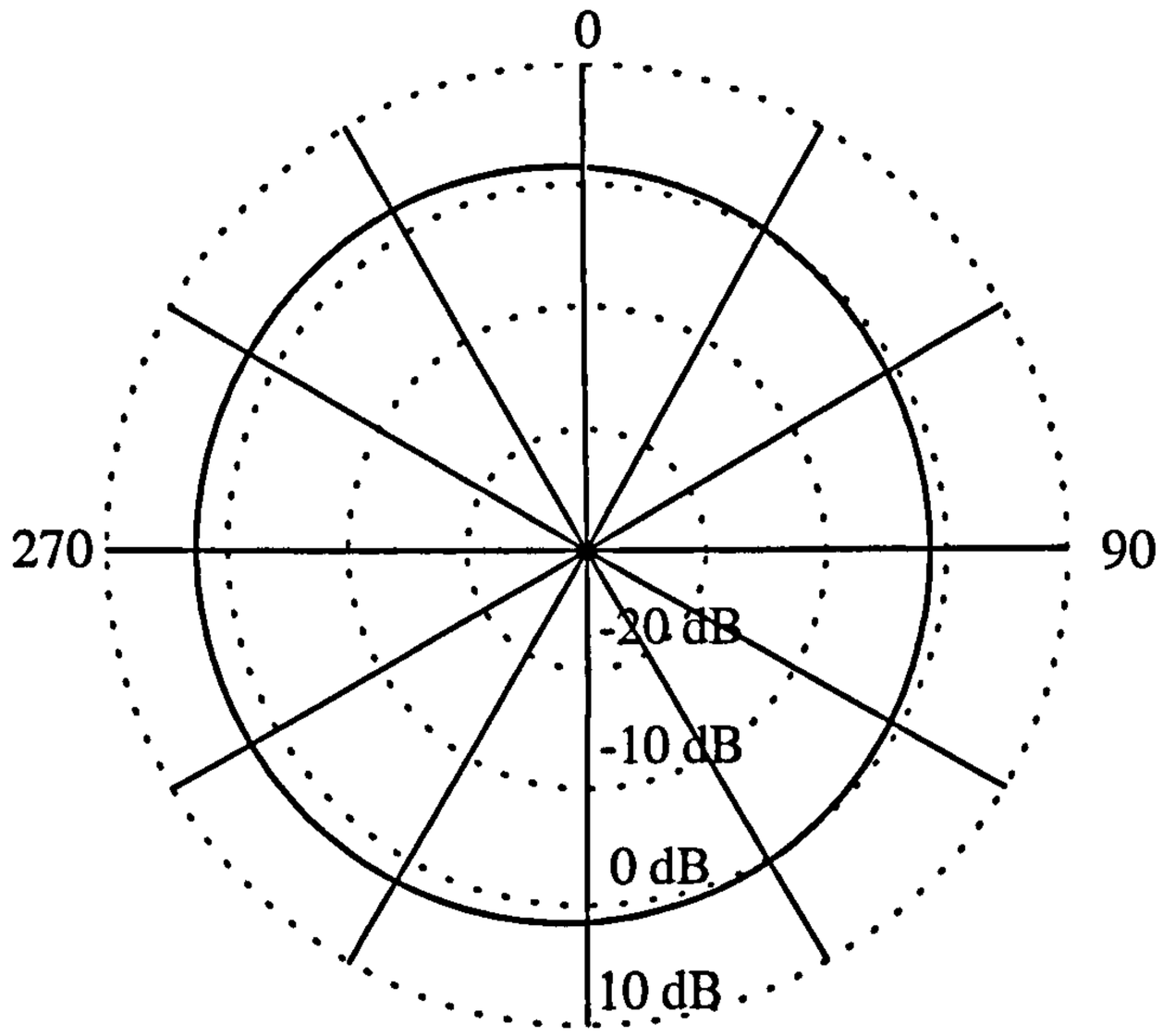


Figure 4.10 Predicted azimuth plane pattern on xy-plane of wire folded loop antenna mounted on a perfectly conducting box

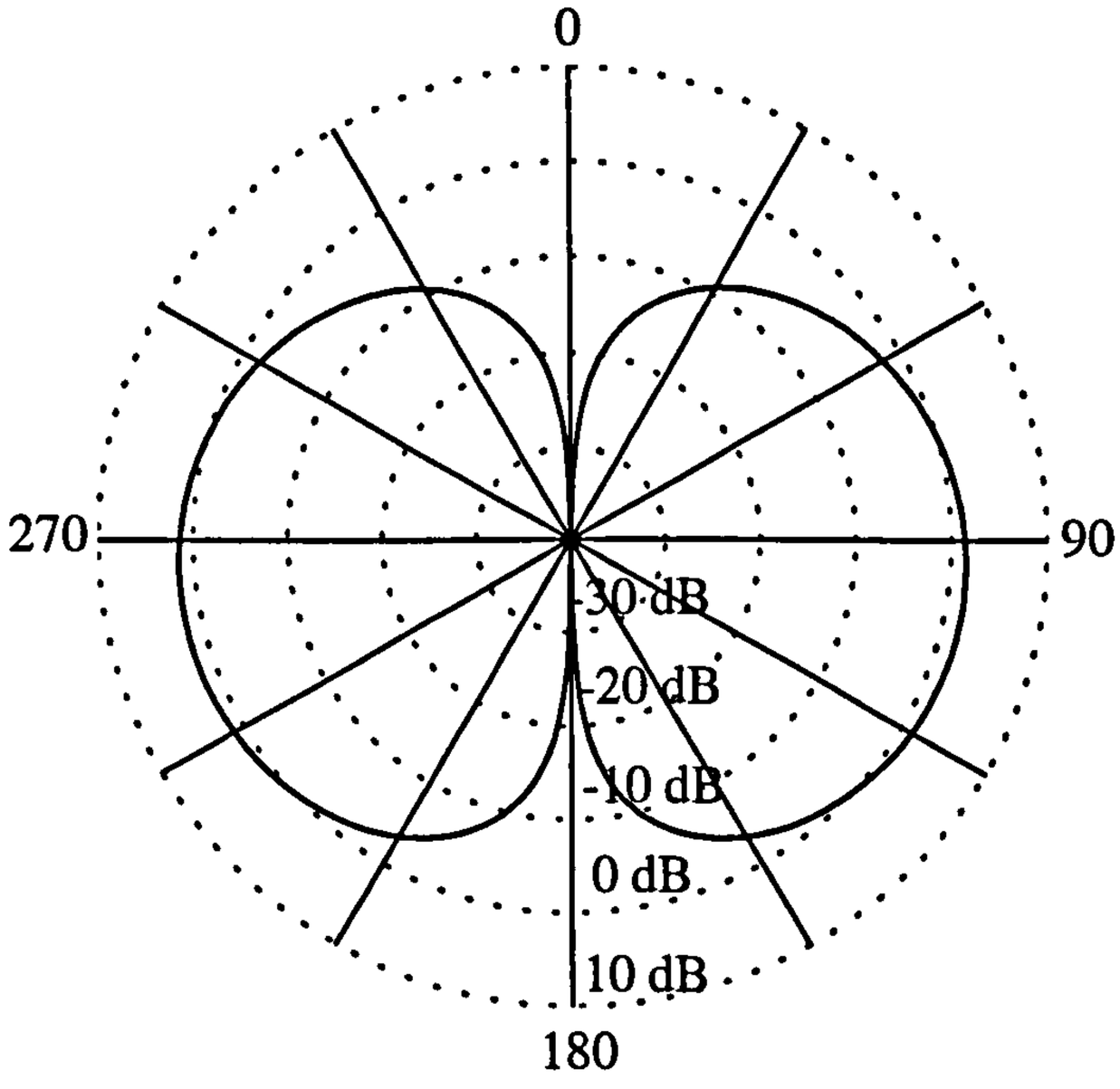


Figure 4.11 Predicted elevation plane pattern on xz-plane of wire folded loop antenna mounted on a perfectly conducting box

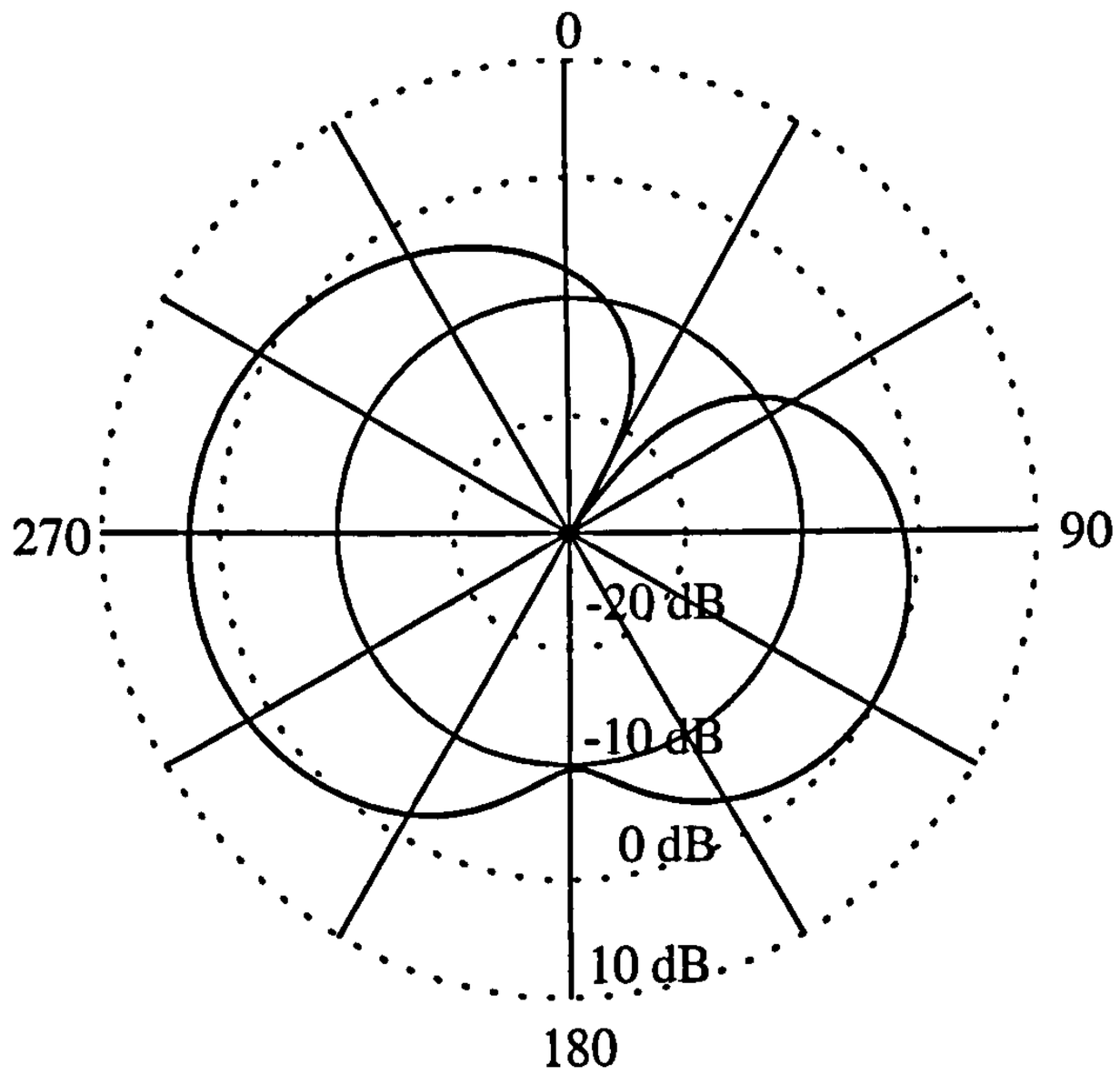


Figure 4.12 Predicted elevation plane pattern on yz-plane of wire folded loop antenna mounted on a perfectly conducting box

4.6 Summary

This chapter has introduced the underlying theory of the Method of Moments since it is the main numerical technique used for the calculation of antenna performance in this thesis.

The basis and initial formulation of the theory were reviewed in section 4.2, particularly as applied to electromagnetic problems. The importance of the integral representation and the expansion and testing functions used have also been investigated.

Section 4.3 has introduced the wire-grid model which allows the simulation of continuous surfaces. Some of the most important modelling guidelines that are necessary to achieve the most accurate results have been described. These include both

analytically obtained guidelines and the practically inferred, so called, "rules of thumb" that stem from others modelling experiences. Sequentially, in section 4.4 the main functions of the popular code - NEC2 were demonstrated. It is especially important to familiarise oneself with this code since it is the main tool used to simulate the designed antennas in this work.

Finally, in section 4.5, NEC2 has been validated by simulating a wire folded loop antenna mounted on a perfectly conducting box. The analytical results, including the input impedance and the radiation patterns, were compared with those from measurements and FDTD method both investigated by Katsibas et al. (1998) and very good agreement has been obtained. These results confirm the correct use of both the wire-grid model and the NEC2 code for analysing a complex structure.

CHAPTER 5

WIRE ANTENNA DESIGN USING THE GENETIC ALGORITHM AND MOMENT METHOD

5.1 Introduction

The previous chapter examined the use of the Method of Moments code - NEC2 to accurately predict the radiation characteristics of a folded loop antenna mounted on a complex conducting structure. Excellent overall agreement was obtained with measured data and also with FDTD-predicted data from the literature. The correct use of this code has thus been confirmed. Chapter 3 demonstrated how the novel technique of the Genetic Algorithm effectively searched for the global optimum solution of an equation with many local optima. Also examined were the effects of various control techniques on the efficiency of the GA. Finally, the implementation of a GA has also been achieved successfully.

The purpose of this chapter therefore is focused on the practical application of the combination of the GA and NEC2 to the synthesis of novel antennas. Thus, we describe, in section 5.2, the implementation of this combined method. These routines are general-purpose and are applicable to a very wide variety of antenna designs. A flow diagram is included to illustrate the features of each routine.

Wire antennas of simple configuration and low cost are the oldest antenna structures and have been designed to function as electromagnetic radiators or receivers for various applications. Any wire antenna with some change to its geometric configuration will result in a different current distribution and therefore a change of electromagnetic performance. This makes wire antennas most suitable for various applications with only some modification of shape. Therefore, for each application, the question remains how best to make the antenna achieve its optimum performance. Traditional techniques used to optimise these antennas are based on a knowledge of antenna properties and considerable design experience. They are always inefficient and make it difficult to explore a variety of wire antennas in a search for the desired characteristics. Because of this, the GA/NEC2 combination tool developed in this work is therefore used to optimise this type of antenna to achieve some appropriate electromagnetic performance.

As an example, in section 5.3, the geometrical configuration of a Yagi-Uda array with only three elements (one driver, one reflector, and one director) is optimised to obtain maximum gain. This section then goes on to present an investigation of this three-element parasitic array but with curved elements which has been reported as the optimal configuration with maximum gain. Two approaches using the GA/NEC2 technique are presented to optimise the different antenna shapes for maximum gain and front-to-back ratio F/B. In the first approach each element of the curved antenna is considered to be composed by several short wires while, in the second, the vee-shaped array is examined. The radiation characteristics of these optimised arrays will be compared with those of the antenna with curved elements.

Finally section 5.4 examines the application of GA/NEC2 for the optimisation of a monopole loaded with a folded dipole antenna. This antenna is first investigated by Altshuler (1993) using the methods of both numerical calculation and measurement to obtain circular polarisation and near-hemisphere coverage over the ground plane. Later he and Linden (1997) had also used a GA to optimise this antenna and obtained impressive results. However, in this section a different objective function from that used by them is used in the GA procedure to optimise the configuration of this antenna for the same goal. The results obtained for this antenna configuration and its radiation performance will be compared to those from Altshuler and Linden.

5.2 GA/NEC2 Combination Tool

Figure 5.1 shows a flow chart of the resulting GA/NEC2 optimisation routine. Most of the steps are similar to those introduced in chapter 3. A notable difference here is the wire-grid model and the NEC2 code are now included in the procedure of problem solving. In each iteration, each chromosome which represents the created antenna in some binary form is first decoded into its equivalent physical antenna configuration and then the entire antenna system composed by the antenna itself and the conducting structure on which it is sited is modelled by the wire-grid technique. If there is no such supporting structure associated with the antenna then the wire-grid model is not included. The resulting radiation performance of the antenna system, due to the GA optimisation can be calculated using the NEC2 code. The specified antenna performance is then compared with the objective function. If it is satisfactory the selected stop criteria will have been achieved and the routine stops; otherwise the process continues. The genetic operators including ranking, selecting, mating and

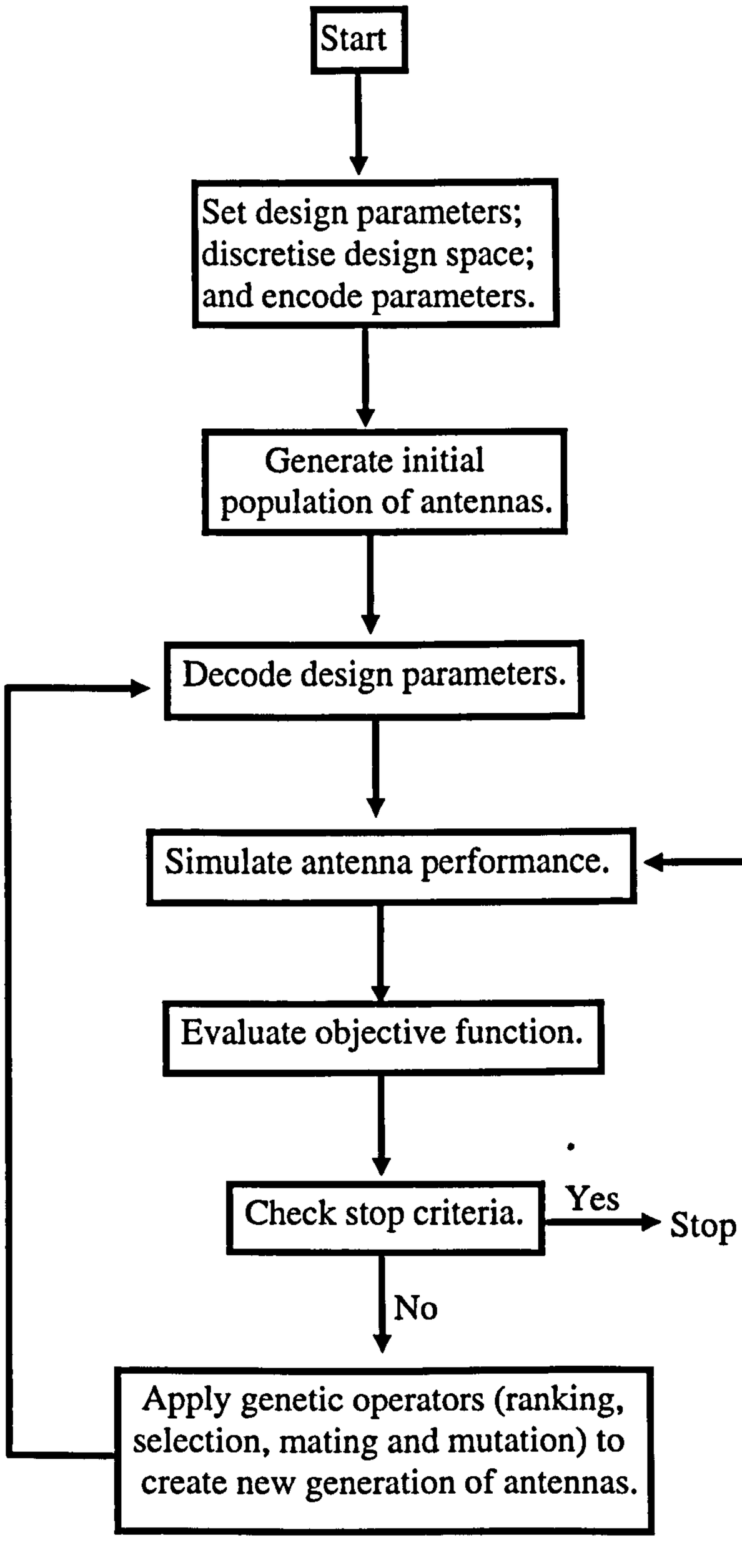


Figure 5.1 A flowchart of the GA/NEC2 antenna optimisation routine.

mutation are therefore applied to select the surviving antennas and to then create a new generation. These are all represented by so-called chromosomes again. It should be noted, unless otherwise specified, that the steady-state GA is used for all cases of "carrying over" the best antenna configurations from the previous generation to the next. The new antennas are then decoded and simulated by NEC2 and this iterative procedure is continued until the stop criteria are satisfied.

5.3 High-gain Parasitic Antenna Design

5.3.1 Curve-shaped Yagi-Uda Antennas: a Review

A Yagi-Uda array with the properties of a simpe feed structure and relative high gain is a practical configuration for many applications. It is usually used in the VHF and UHF frequency bands, but HF Yagis are also common. The basic unit of a Yagi-Uda array is composed of three straight dipoles named the reflector, driver and director as shown in figure 5.2. Normally, the number of directors can be one or more depending on the practical requirement of gain. In this antenna only the driver is fed and the others are deemed to be parasites. This is different to a general array antenna in which each element is fed from a source. So this Yagi-Uda array is also called a parasitic array.

Generally, the driver is normally tuned to a length little less than 0.5λ to reach resonance and the reflector as well as the director are respectively a little longer and a little shorter than the driver. In practical applications, increasing the number of directors usually produces a higher gain (Ehrenspeck 1959, Bojsen 1971), however this is at the expense of increased mechanical complexity. Therefore, another way to

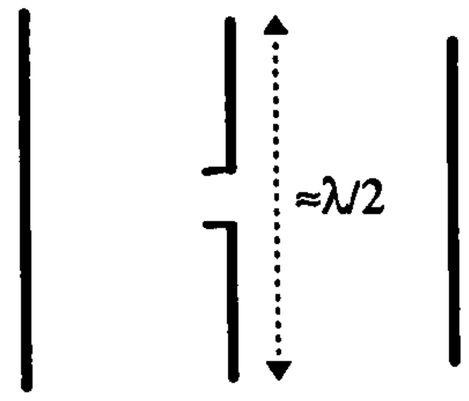


Figure 5.2 Configuration of a three-element Yagi-Uda array.

improve the gain, with just one director is to increase the length of each of the three linear elements. Unfortunately, the results showed that when the elements are lengthened much beyond the wavelength, the gain normal to the elements begins to diminish due to increasing sidelobes. However it has also been found that this usually undesired characteristic can be eliminated if the shape of each element, as well as the spacings between elements, are adjusted properly. This is because the spatial differentials of the various sections of a properly shaped antenna element could compensate for the phase reversals of the current distribution on the elements. Therefore, Landstorfer (1976, 1977, 1979) and Liang and Cheng (1983) have done a substantial amount of work using theoretical or experimental methods to optimise the element shapes and the spacings between elements for maximising the gain. In Lanstorfer's investigation, the driven element was regarded as a composite of a series of straight-line sections and the current distribution along the wire antenna was assumed to be sinusoidal irrespective of changes in antenna shape or wire radius. Then, the gradient optimisation technique was applied to find the included angle of each section and thus to generate the optimal shape for the driven element with maximum gain. Secondly, two parasites - a reflector and a director - were added to the

driven element to construct a parasitic array. The shapes of these two parasitic elements were thus optimised experimentally to achieve an optimal array with a global maximum of gain. The full length for each element was set to be a constant of 1.5 λ. Landstorfer's results also showed that the optimal shape to produce maximum gain of 11.5 dBi is a configuration with fairly complex curvature. In addition, it was also found that for best results, the coupling between elements was approximately uniform i.e. the relative spacing between the driver and reflector or director is approximately constant. However, the actual values of the spacing of elements were not mentioned in Landstorfer's work. The optimum shape for this Yagi-Uda array is shown in figure 5.3.

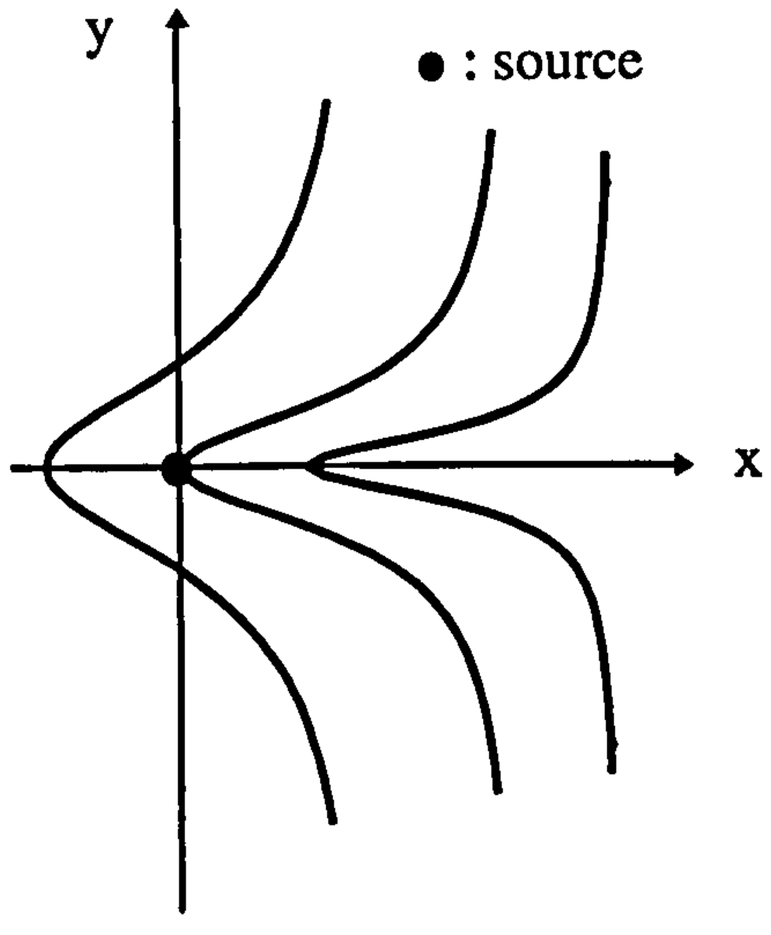


Figure 5.3 Geometry of the shaped Yagi-Uda array investigated by Lanstofer (1976,1977,1979).

Based on the work of Landstorfer, Liang and Cheng (1983) theoretically approximated the shape of the same Yagi-Uda array by a non-linear function. Different coefficients were used for the three elements. The function they used is shown in equation 5.1.

$$x = a_i\left(1 - \frac{1}{1 + b_i y^2}\right) + c_i \qquad (5.1)$$

$$i = 1,\ 2,\ 3$$

where the subscript i denotes the different elements.

In their case the current distribution along the elements was not assumed to be sinusoidal. Both the MoM and the simplex method for function optimisation were used to solve for the coefficients that yielded maximum gain. The analysed results for this antenna configuration showed that: (i) the curved array can produce a maximum gain and, (ii) the optimum length for each element is about 1.5λ. These results were also in a good agreement with those of Landstorfer. In addition, the optimised spacings between the driven element and the reflector as well as between the driven element and the director are 0.162λ and 0.151λ, respectively. However a marginally higher gain of 11.8 dBi has been obtained by Liang. The relevant radiation characteristics of the two arrays under investigation are shown in table 5.1.

For the purpose of comparison with the results which will be presented in the following sections the antenna of Liang et al. is simulated again using NEC2. The results are also included in table 5.1 and figure 5.4 illustrates the simulated radiation pattern.

Table 5.1 Radiation characteristics of the curve-shaped parasitic arrays resulted from both Landstorfer and Liang, and from Liang using NEC2.

Method	Max. Gain (dBi)	Beamwidth (deg.)	F/B (dB)	Input Impedance (Ω)
Landstorfer	11.50	--------	26.0	----------
Liang et al.	11.80	32	26.47	14.24+j32.77
Liang et al. (using NEC2)	11.90	32	28.0	20.42+j2.96

The above important pattern characteristics of such an optimised array could only be achieved with a conventional, straight element Yagi-Uda array with many more elements. Figure 5.5 illustrates the variations of gain and F/B against the number of elements for a conventional equally spaced Yagi-Uda array with the spacing and diameter of elements of 0.2λ and 0.005λ, respectively (Stutzman et al. 1981). This conventional Yagi-Uda array requires at least seven elements to produce the gain of 11.8 dBi.

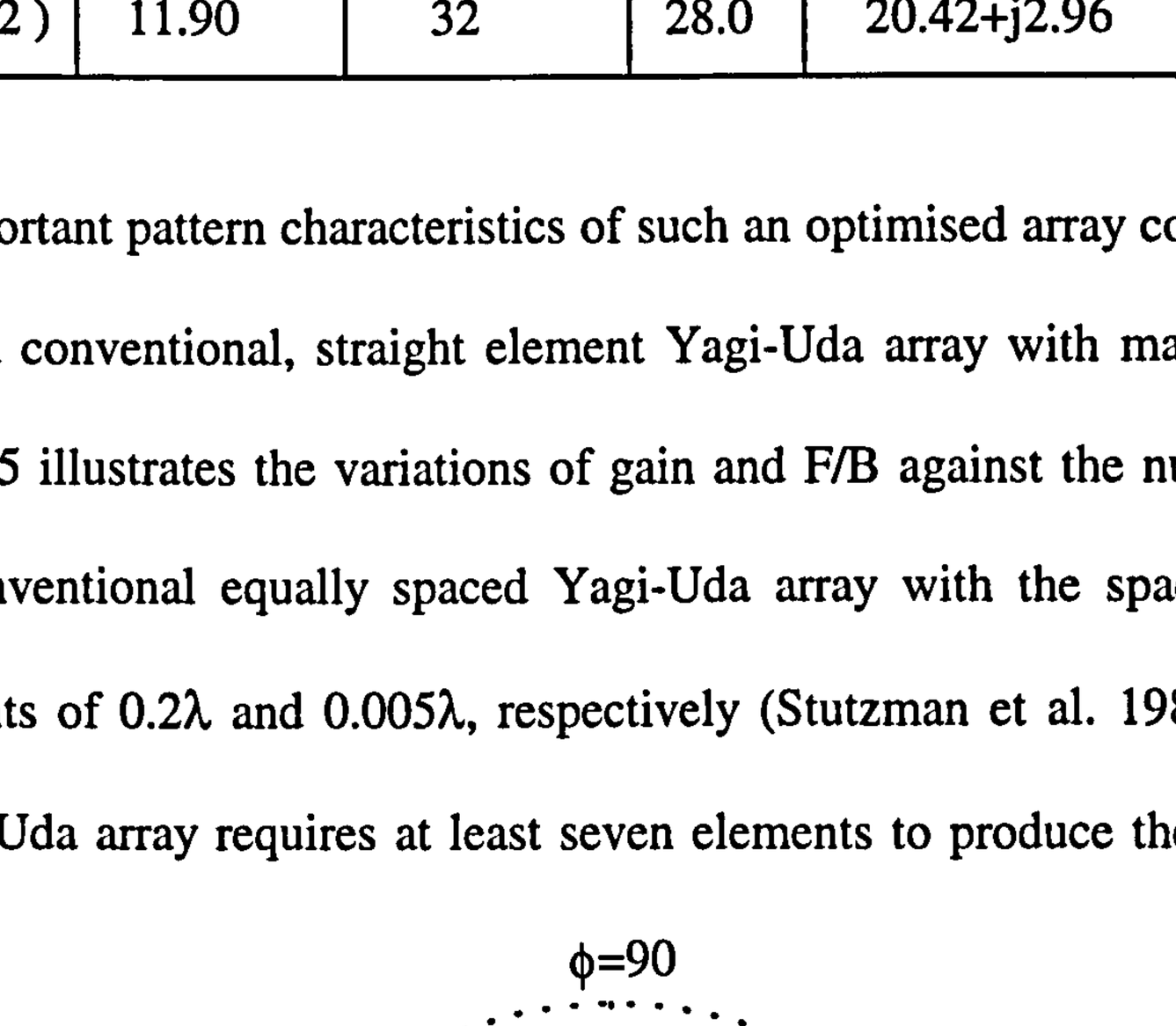


Figure 5.4 Predicted radiation pattern using NEC2 for the optimised antenna of Liang and Cheng.

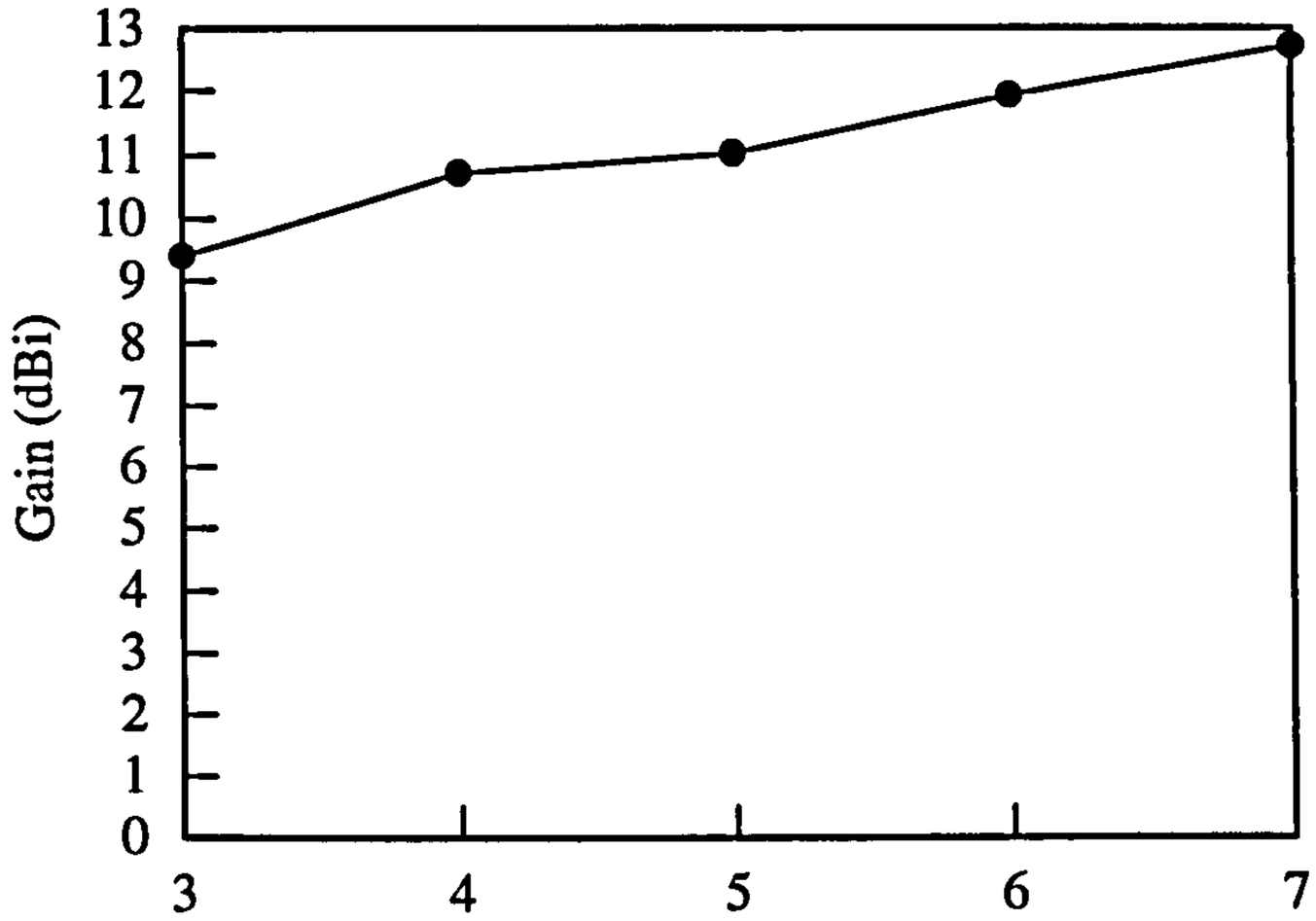


Figure 5.5 Gain of a typical equally-spaced Yagi-Uda antenna versus the total number of elements (Source :Stutzman & Thiele, "Antenna Theory and Design", Table 5-4, p.226).

5.3.2 Auto-shaped Parasitic Array

5.3.2.1 GA Approach

The optimisation of this particular type of antenna is a problem for which the GA is ideally suited. As an example of its use the shape of such a three-element parasitic array will be optimised to obtain both high gain and front-to-back ratio (F/B) using the GA/NEC2 technique. Figure 5.6 shows a priori (initial) geometry of the objective antenna to be optimised. It is symmetrical and the feed point is placed at the origin on the x-y plane. The half length of each element consists of ten straight wires, equal in length and connected end-to-end. This number of wires was selected after referring to Liang's results in which each half length of the element was divided into eleven similar segments, for the purpose of optimising the shape precisely. The radius of each

element is 0.01λ as used in Liang's work. Thus the wire length, ℓ_i, of each element, the spacing between elements, d_i, and inclination angles, ϕ_{ij}, between wires are all left as the design variables for random manipulation by the GA.

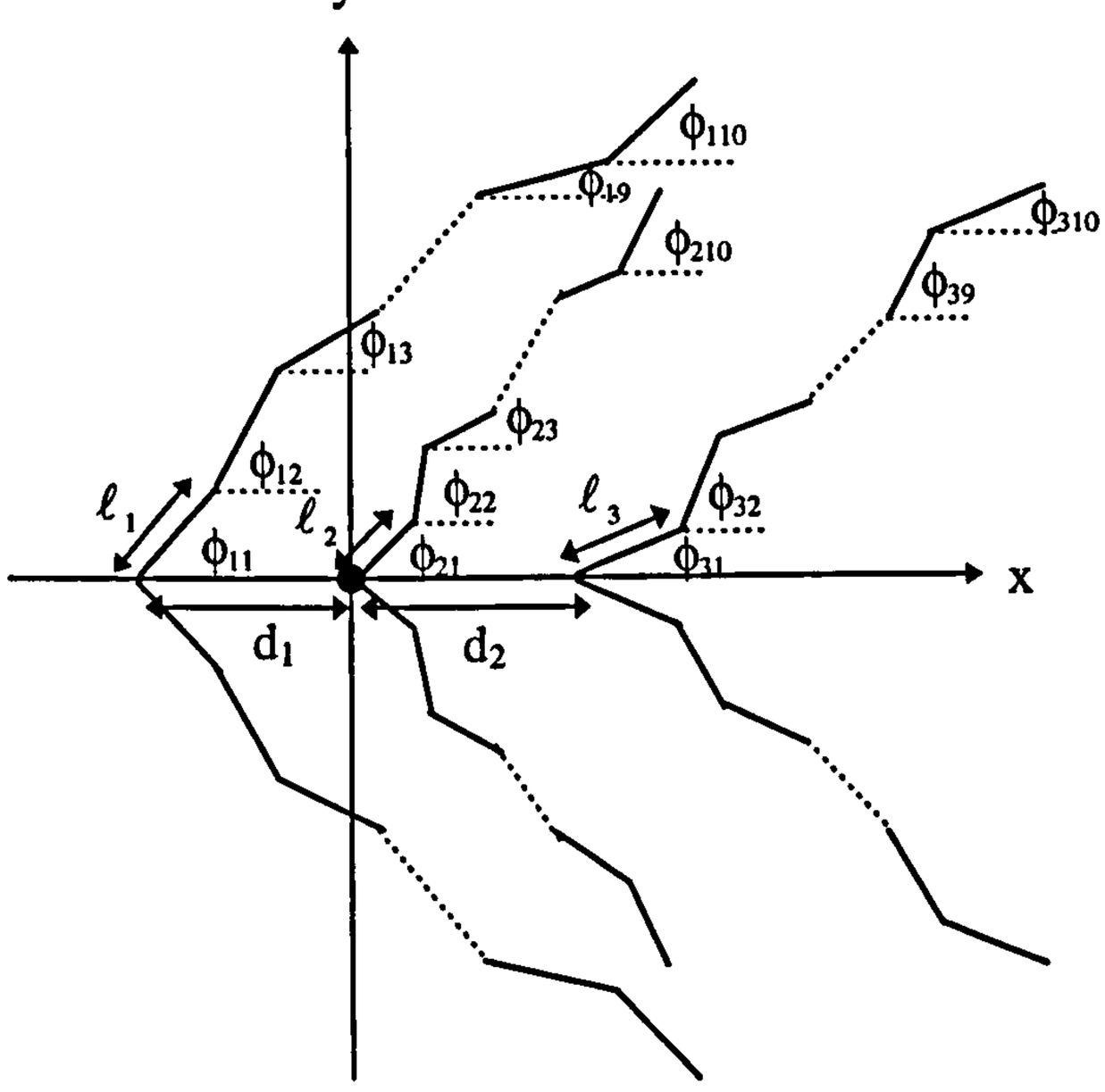


Figure 5.6 Geometry of a Yagi-Uda antenna to be optimised by using Genetic Algorithm.

Before implementing the GA some limits on the ranges of the design parameters must be defined to prevent the algorithm from getting into an unmanageable search space. Because of the symmetry of this configuration the GA only needs to determine the wire lengths and inclination angles for the half length of each element. The range of each wire length, ℓ_i, was allowed to vary between 0.065λ and 0.085λ (i.e. the total length of each element is between 1.3λ and 1.7λ); element spacing, d_i, between 0.1λ

and 0.4λ; and the inclination angle, ϕ_i, between 10° and 90°. These ranges were all based on the reported results of Liang et al. As discussed before, the GA begins by encoding the design parameters into a string of binary bits called a *chromosome*. Four, five and seven bits respectively were chosen for the variables: wire length, element spacing and included angle. Each antenna therefore requires a *chromosome* with (4×3)+(5×2)+(7×10×3)=232 bits to store the real configuration. This definition of a chromosome creates a solution domain which has 2^{232} possible solutions and the GA will search for the best one within it. A population of 60 candidate antennas was assessed in each generation. This number was selected arbitrarily but followed the suggestion discussed in chapter 2 and did produce good results for this case. To monitor the evolution process of the GA 50 generations was set for a run. If, after this number of generations, the algorithm does not converge, it will continue for a further 50, and so on. For each selected antenna the performance is simulated by NEC2. The goal is to obtain both high gain and F/B so the objective function must force all new configurations towards this target. It therefore consists of three components: gain, F/B and the maximum tolerable beamwidth set at + and -15°, with the maximum radiation in the ϕ=0° direction. In view of the wide variation of initial values obtained, each of the three parameters was normalised by an appropriate factor (Jones et al. 1997). The values of Gain=13 dBi and F/B=30 dB serve as goals since they exceeded typical results reported for similar configurations reported in the literature, while the radiated power in the angle ($-15° \le \phi \le 15°$), is normalised by the total radiated power. Numerical experiments showed that the three normalised terms made very different contributions to the value of the objective function. To avoid any of them from dominating the

iteration process, each was weighted by an associated constant. In this investigation the improvement in gain and the radiated power concentrated within the specified angle changed much more slowly than did the F/B so they must be more heavily weighted. The values of 16, 12, 1 respectively were selected from a number of preliminary runs. The objective function was then:

$$\text{Objective function} = 16\frac{Gain}{13} + 12\frac{\sum_{-15^\circ \le \phi \le 15^\circ} P(\phi)}{\sum_{-180^\circ \le \phi \le 180^\circ} P(\phi)} + \frac{F/B}{30} \qquad (5.2)$$

It is clear that in each generation any antenna with a higher value evaluated from this function than the other antennas will get better chance to survive the selection mechanism. The survivors in each generation are paired to produce new offspring and so participate to yield to the next generation. Numerical experiment showed that applying the schemes of alternative pairing and dual-point crossover can provide good results, at least for this example. An alternative pairing scheme, as introduced in section 3.3.4, allows each survivor two opportunities to pair with its neighbours and each pairing can create two new offspring. To keep the population constant in each generation therefore in the deterministic approach the top one-third of the ranked chromosomes (i.e. 20 chromosomes) are selected as the survivors. They are paired to produce new offspring and participate to the next generation. Then 0.1% of the total bits in the population was selected randomly to mutate to produce the necessary population diversity among the new generations.

Based upon the above criteria this GA started to produce the optimal antenna from an initial population created by randomisation. The iterative process continued until convergence was achieved.

5.3.2.2 Analysis of Results

Figure 5.7 reveals the progress of the GA optimisation routine as a function of the number of generations. Both gain and F/B of the best antenna in each generation are shown on the plot. Note that the best individual of the initial generation only has a gain of 9.84 dBi and F/B of 8.93 dB. It is clearly small compared to the optimised results shown in table 5.1. However, the results improved very quickly especially, as expected, the significant improvement in gain appears within the earliest generations. This is because gain term was weighted by a larger factor than that of the F/B. Subsequently, after the gain reaches about 11.5 dBi its rate of increase slows and most improvement is concentrated on the F/B. In addition, the weighting of the three specific terms also produces an apparent compensation effect between gain and F/B. The decrease in gain always accompanies a considerable increase in F/B. In the first run of 50 generations the algorithm seemed to converge after about 45. To confirm the validity of this convergence another run of 50 iterations was made. This indeed confirmed that the algorithm converged after 45 generations. Therefore, only the process of the first 60 generations is shown in the figure below.

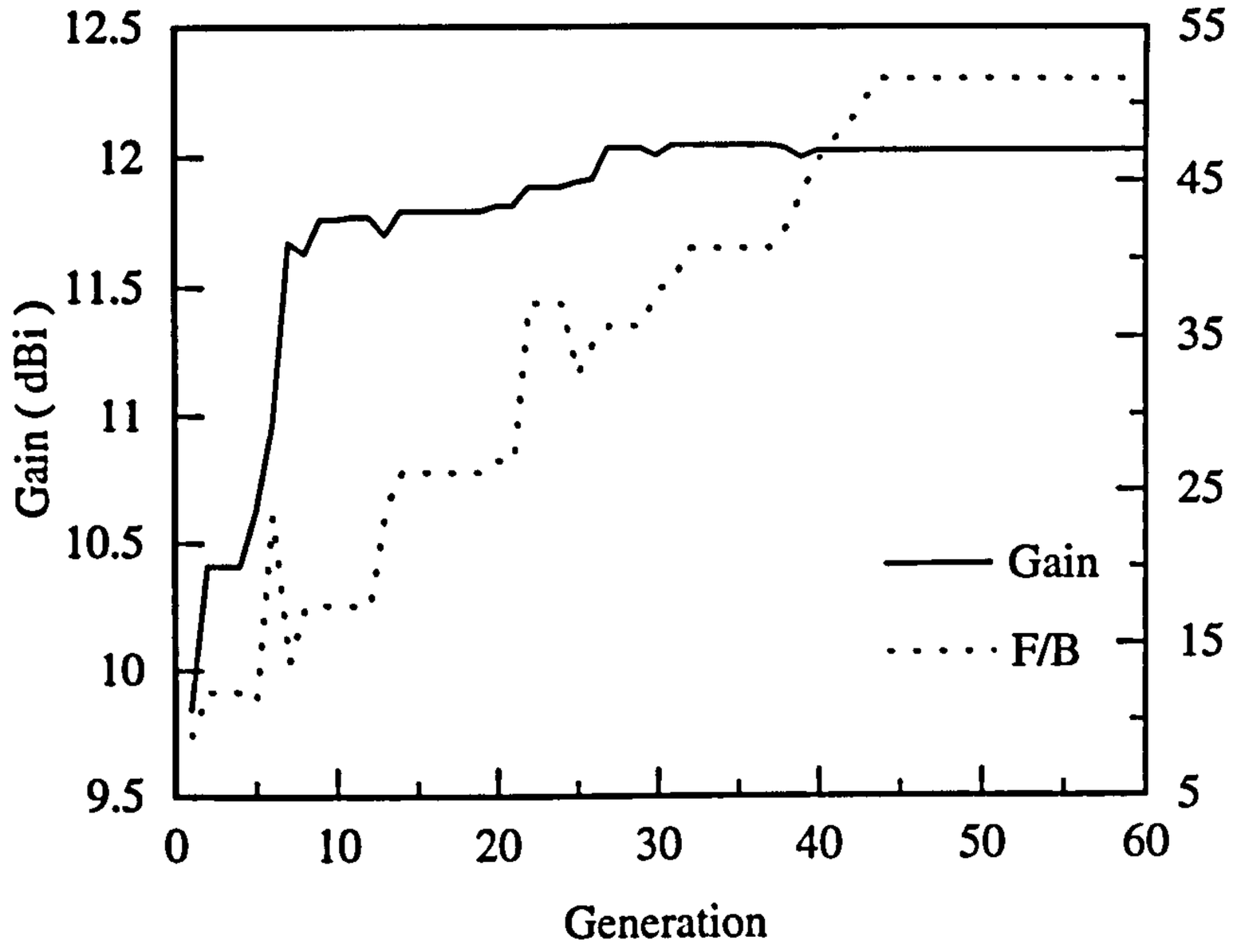


Figure 5.7 Gain and F/B of the best antenna as a function of the number of generations.

The GA certainly created an antenna which has a very unconventional geometrical shape. It is shown in figure 5.8. It is found that each element of this array has the length of about 1.6λ which is close to that of 1.5λ reported in the literature, and an elaborate shape to produce a maximum gain of 12.0 dBi and F/B of 50.8 dB. These radiation characteristics are both better than those of Landsterfor and Liang et al. In addition, the half power beamwidth meets the specified 30°. The results also show that the input impedance at the feed point of the driven element and the front-to-side ratio (F/S) are 20.7+j79 Ω and 15.76 dB, respectively. The predicted radiation pattern of this antenna is shown in figure 5.9. Because the shape of this array is determined by the GA, in an arbitrary way, it could therefore be called an "auto-shaped" array.

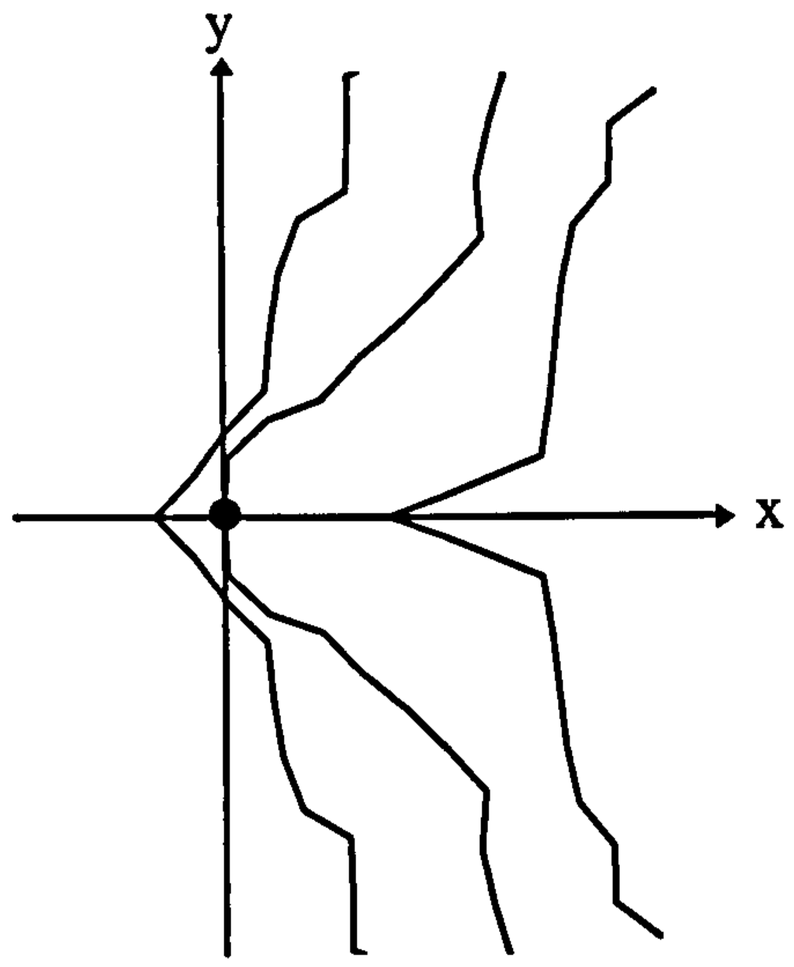


Figure 5.8 Geometry of the optimised three-element, auto-shaped Yagi-Uda antenna using a Genetic Algorithm.

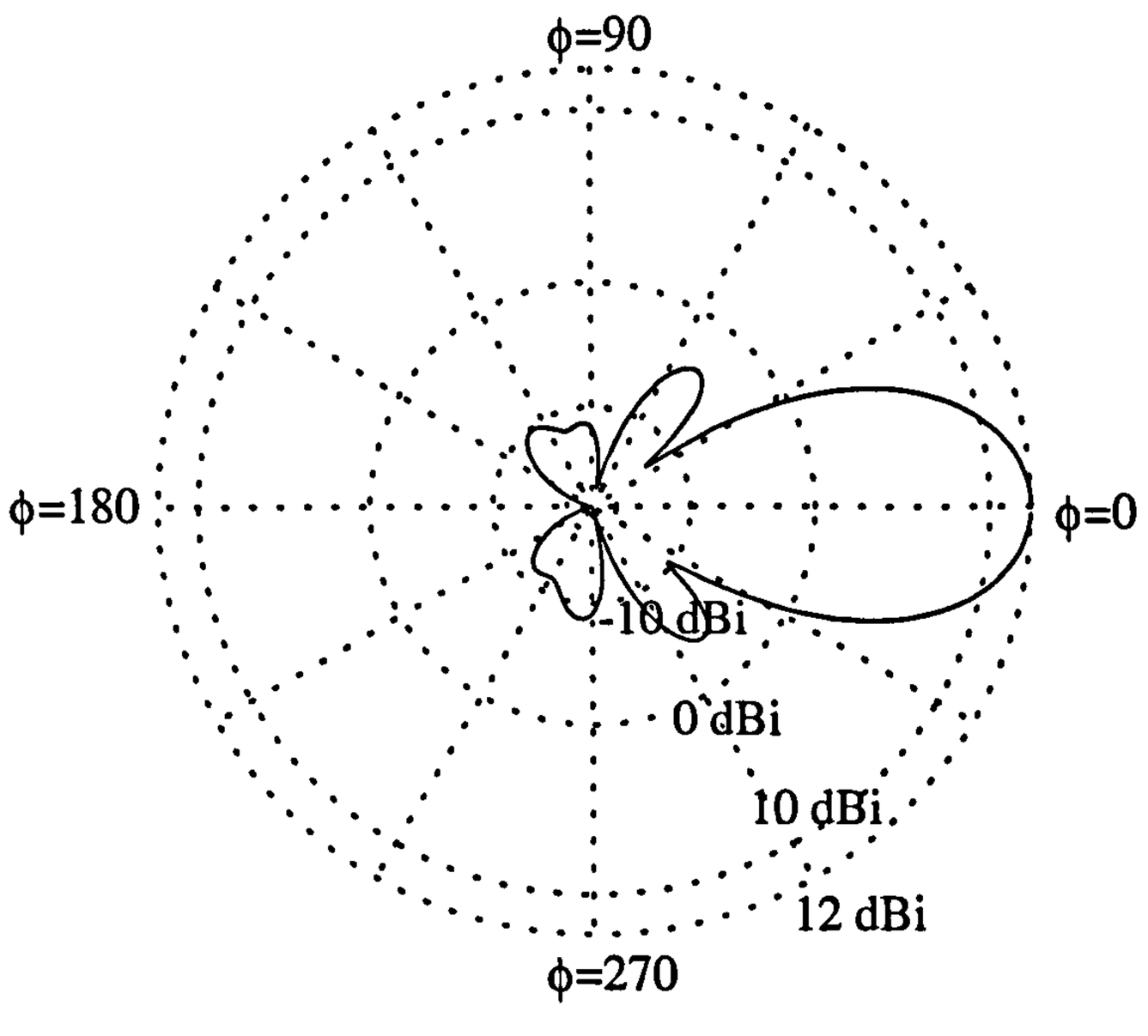


Figure 5.9 Radiation pattern for the optimised three-element, auto-shaped parasitic array.

5.3.3 Vee-shaped Parasitic Array

Although the above investigations have proved that a three-element parasitic array of complex shape can produce a particularly high gain it is difficult to construct such complex shaping and the requirement for peculiar curvature would also add to manufacturing costs. These considerations therefore became the motivation in this section to use the GA/NEC2 technique to synthesise a three-element parasitic array of relatively simple shape, and with high gain. Because each element predicted by the GA in the previous section suggests a form of V, this shape was therefore assumed to be a possible candidate for an antenna configuration of maximum gain. Another reason for the selection of this shape is that a simple vee-dipole antenna, of appropriate length and included angle, can result in significant gain (Thiele 1980). Thus, if another parasites, of similar shape, are added an array with increased gain should result.

5.3.3.1 GA Approach

Figure 5.10 shows the shape of the objective antenna. Essentially similar to the previous example, it is symmetrical and is centrally sited around the origin on the x-y plane. Each element is assumed to be a perfect conductor with the radius of 0.01λ. The input power is fed to the central point of the driven element. Thus the components to be manipulated are the lengths, spacing and included angles of the three elements. These parameters are therefore left as the design variables and included in a chromosome for recombination by the GA.

As in the procedure of a general GA the constraints of the variable ranges for the design parameters must be defined before implementing the GA. The range of ℓ_i was allowed to vary between 0.3λ and 1.1λ; d_i between 0.1λ and 0.5λ; and the included half-angle, ϕ_i, between 10° and 90°. To represent these parameters as a chromosome seven, six and eight bits have been allocated for mapping the physical values of element length ℓ_i, element spacing d_i and included half-angle ϕ_i, respectively. With this discretisation, a chromosome is thus composed of (7×3+6×2+8×3)= 57 bits and this creates a search space with 2^{57} possible solutions. Table 5.2 lists the range and bit number of all design parameters defined here for use in the GA.

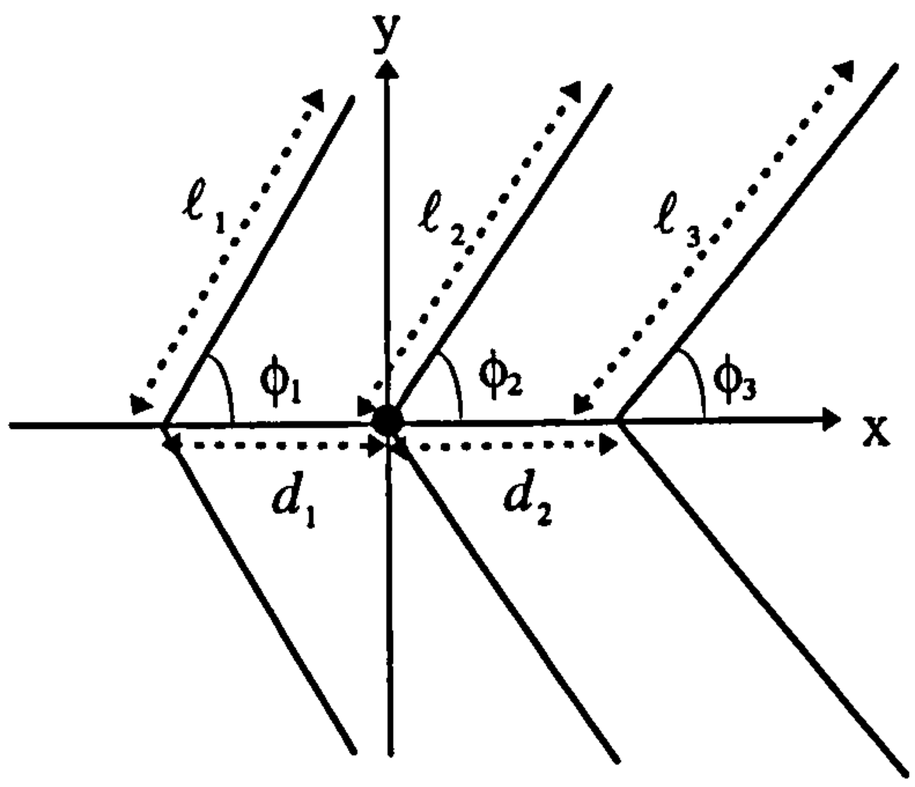


Figure 5.10 The geometry of a three-element vee-shaped parasitic array.

Table 5.2 Range and bit number of the design parameters defined for use in the GA.

Chromosome = $[\ell_1\phi_1 d_1 \ell_2\phi_2 d_2 \ell_3\phi_3]$			
	ℓ_i	d_i	ϕ_i
Range	0.3 ~ 1.1λ	0.1 ~ 0.5λ	10° ~ 90°
bits	7	6	8

Next, a population of 90 candidate configurations has been considered in each generation. A few trials have been examined to help decide this number of population. Actually, from the numerical experiments with this design the range from 60 to 90 for the number of population does not produce much difference in both the results and the convergence speed. Therefore the number of 90 was selected to diversify the candidate antennas. Since the goal of maximum gain and F/B is the same as in the previous example, equation 5.3 was therefore used again as the objective function. The other GA operators including the selection, crossover, mutation used in previous example were also applied to this design since they performed well in that case. It must be noted, of course, that if these selected schemes do not efficiently approach an acceptable solution some changes will be necessary.

5.3.3.2 Analysis of Results

Figure 5.11 depicts the performance curves for the improvement of gain and F/B from the best individual in each generation. The curves show clearly the natural evolution in antenna performance. Both the gain and F/B improved very quickly due to the recombination of the surviving antennas. Comparing the two curves, because the gain term in the objective function was weighted by a larger factor the improvement was, undoubtedly, faster than that for F/B. However, once the value of gain approximates to a threshold value of about 11.6 dBi, for this case, it improved slowly and then the F/B changed more rapidly. This GA process took about 50 generation to converge

Table 5.3 shows the geometry parameters of the optimised three-element vee-shaped parasitic array. Comparing the geometry of this GA - optimised antenna with

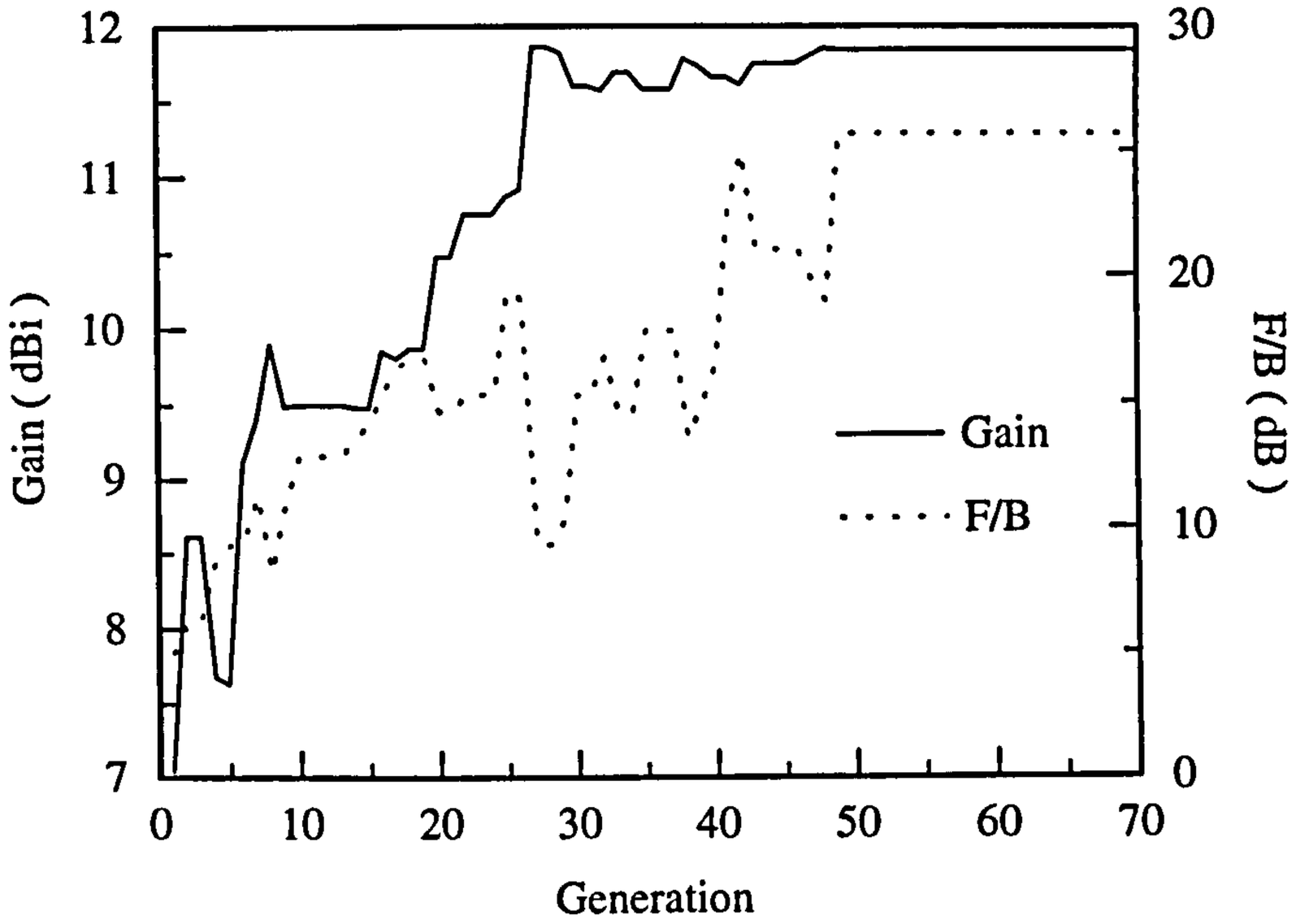


Figure 5.11 Gain and F/B of the best antenna as a function of the number of generations.

Table 5.3 Geometry parameters for the optimum three-element vee-shaped parasitic array.

ℓ_1	ℓ_2	ℓ_3	ϕ_1	ϕ_2	ϕ_3	d_1	d_2
0.747λ	0.728λ	0.697λ	49°	50°	50°	0.113λ	0.233λ

the antennas which resulted from the optimisations of Landstorfer and Liang et al. we note the following: firstly, the length of the elements were all very close to 1.5λ as in those cases; secondly, the general vee-shape of the array is another common characteristic. However, the GA technique produced an optimised array of three straight elements with included angles of the order of 100°, in marked contrast to the elaborate element curvature that was found to be necessary by Landstorfer and Liang et al. In addition, this vee-shaped antenna can produce a maximum gain of 11.84 dBi and F/B of 25.7 dB. They are both similar to those of Liang while being better than

those of Landstorfer. Figure 5.12 is a plot of the radiation pattern for this array. The half power beamwidth also meets the specified 30° defined in the objective function. In addition, the calculated results illustrate that the worst case F/S is 12.7 dB and the computed input impedance is 10.0-j4.7 Ω. It should be noted that although the optimal included half-angle of the reflector is 49° there is hardly any change in performance if it is changed to 50° for the reasons of ease of manufacture.

This investigation has proved that a simple, 1.5λ, vee-shaped, parasitic array can achieve markedly increased gain over its straight, parallel-element counterpart while not having to resort to the elaborate element curvature, as reported by Landstorfer and Liang et al. to achieve the same result.

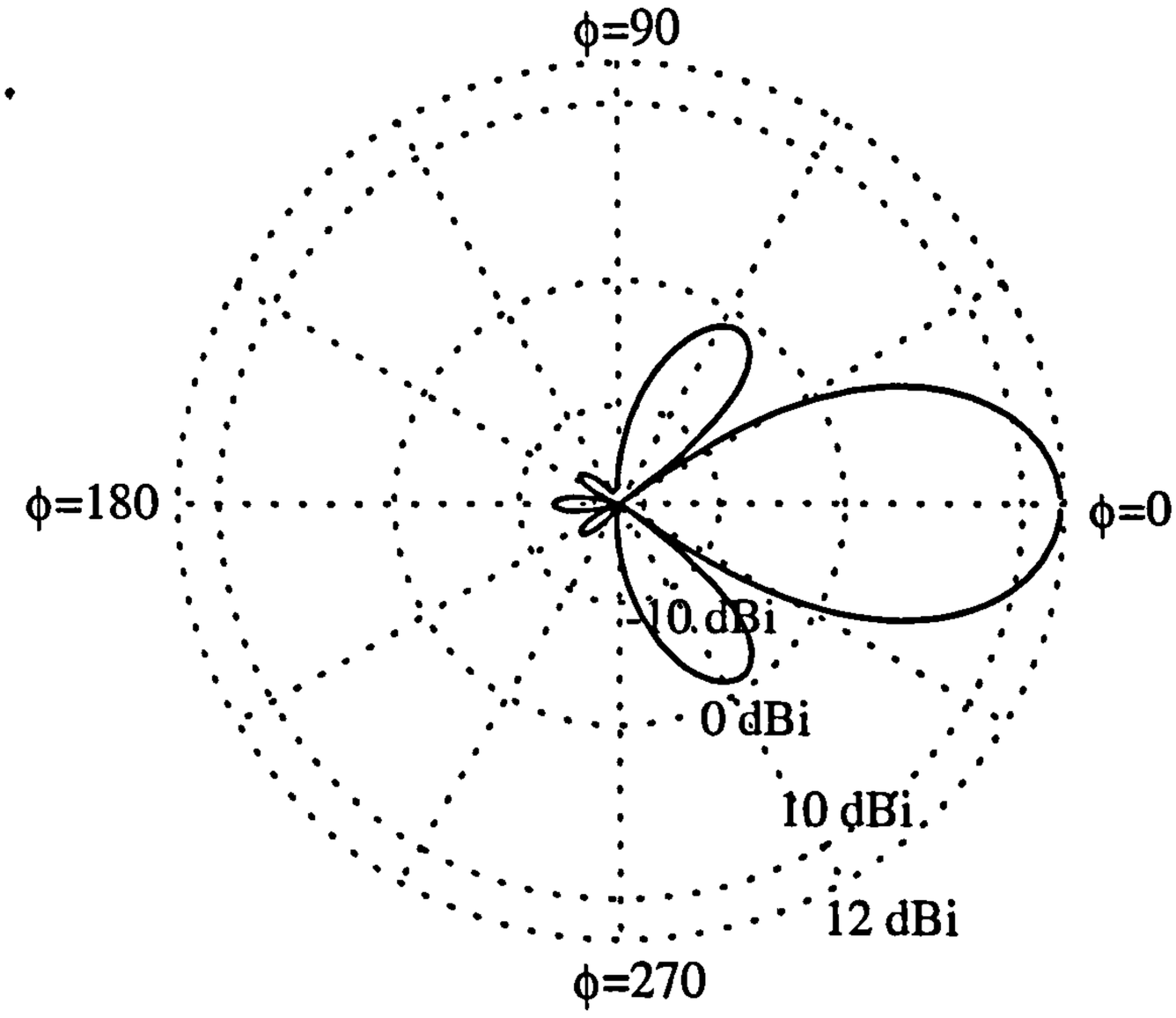


Figure 5.12 Radiation pattern for optimised three-element, vee-shaped parasitic array.

5.4 Loop-loaded Monopole Antenna

The monopole antenna is one of the most basic forms of antennas. Its performance has been intensively analysed and well documented in the literature. Generally, the monopole antenna can be deemed to be a counter-part of the dipole antenna and therefore their characteristics are very similar. For the physical trait of a linear, vertical element mounted on the ground plane, a monopole produces vertical, linear polarisation and provides all-round azimuthal coverage at the low elevation angles. By contrast, a suitably dimensioned loop antenna produces horizontal polarisation with coverage at the higher elevation angles. If properly configured the two types of antenna can combine to form one that produces both vertical and horizontal polarisation and can thus achieve near-hemispherical coverage. Such an antenna will effectively detect the depolarised signals caused by the effect of multiple reflections from a cluttered environment and thus will have application in many modern communication scenarios.

Since 1993, Altshuler has done much work using the methods of both numerical calculation and experimental measurement on this type of antenna and has obtained many valuable results (Altshuler 1993, 1996a,b). Recently, he and Linden (1997) also used the GA for optimisation of this type of antenna and obtained an optimal antenna configuration with a uniform power radiation over the hemisphere. However, in GA optimisation the objective function is defined in terms of the desired features and so using a different function usually produces different results. Therefore the GA/NEC2 technique with a different objective function to that of Altshuler (1997) is used to optimise this antenna and the results are compared with his.

5.4.1 Implementation of Genetic Algorithm

Figure 5.13 is the configuration of the so-called loop-loaded monopole antenna analysed here. It is based on that used in Altshuler's work (1997). In this configuration, the monopole is divided into two vertical wires (Z_1 and Z_4) and is connected in series with a folded loop antenna composed of six wires ($X_1 \rightarrow Z_2 \rightarrow X_1 \rightarrow X_2 \rightarrow Z_3 \rightarrow X_2$). Each wire of this antenna was assumed to be a perfect conductor with the diameter of 2 mm. From the symmetry, only six wires, that are X_1, X_2, Z_1, Z_2, Z_3 and Z_4 are left to be decided by the GA for an optimal configuration. Meanwhile, in order to implement the GA and make it converge, as well as to save computer run-time, the range of lengths for each wire of this loaded antenna has to be defined carefully. It should be large enough such that the optimal length is likely to be included, yet not too large, such that the search space becomes unmanageable, resulting in the GA taking too long to converge. From the literature (Altshuler 1997) the appropriate range for each wire was defined and is shown in table 5.4.

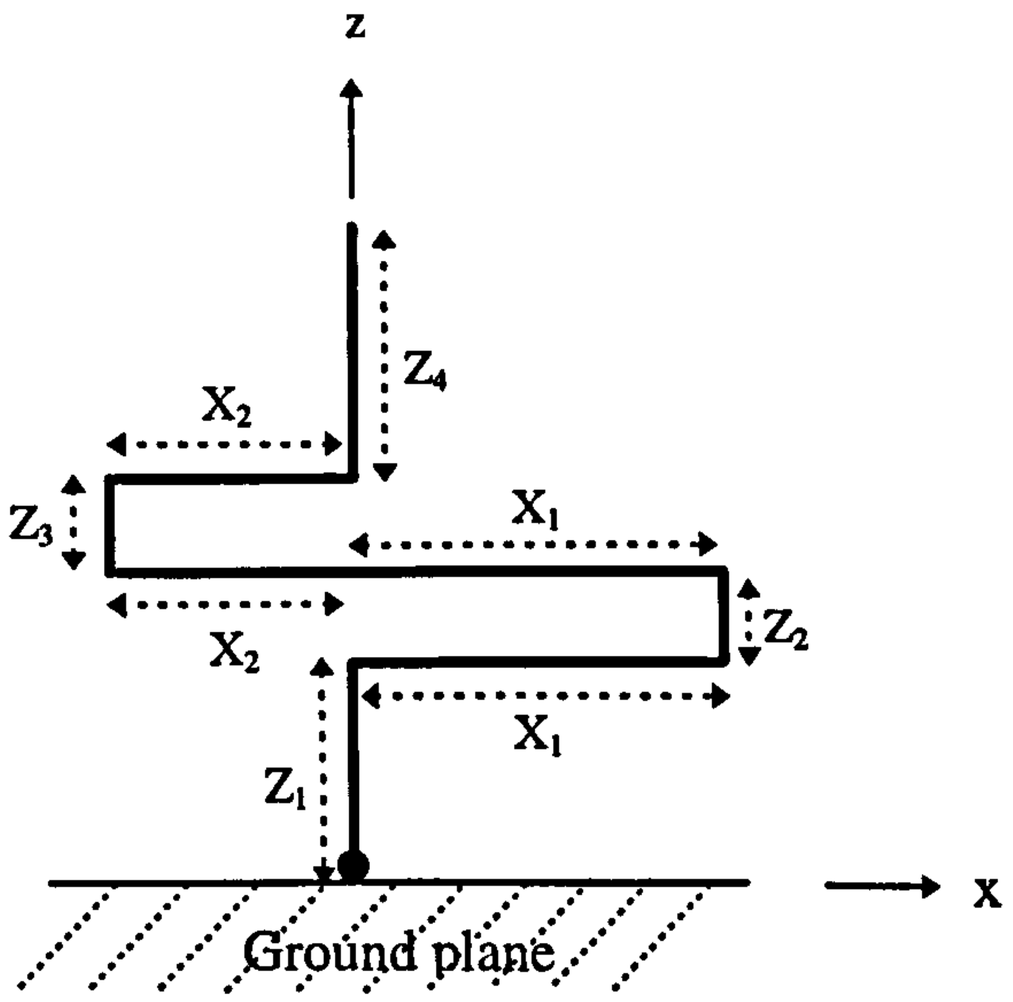


Figure 5.13 A monopole antenna loaded with a modified folded dipole.

Table 5.4 Ranges and bit numbers of the design parameters defined for the GA.

Chromosome = [Z_1 Z_2 X_1 X_2 Z_3 Z_4]	
Parameter	Range
Z_1	0.03 ~ 0.35λ
Z_2	0.01 ~ 0.10λ
X_1	0.05 ~ 0.50λ
X_2	0.05 ~ 0.50λ
Z_3	0.01 ~ 0.10λ
Z_4	0.05 ~ 0.50λ

To facilitate the practical GA implementation, each wire length of the designed monopole was discretised and represented by a string of binary bits. The number of bits selected for each wire ought to be decided appropriately to enable the GA to work efficiently. However, increasing the number of bits to improve resolution inevitably involves more computer time and can retard the convergence process. By contrast, the use of a small number of bits for a parameter always speeds up convergence but reduces the resolution. Therefore, in this design different trials on bit number for each wire length were tested to decide on a number commensurate with accuracy and computational efficiency.

The GA results show that when the bit number is five, the algorithm approaches a acceptable result. This bit number results in that each wire could have 2^5 possible lengths and the smallest change of wire length (i.e. bit resolution) is between 0.0029λ and 0.0145λ, depending the range of wire length. In this way, each antenna configuration was determined by a 30-bit string and a number of 2^{30} possible designs exist in the search space. It is indeed a challenge to find the optimal design from this widespread space. An initial population of 90 random designs was used.

In this investigation the goal is to search for the optimal configuration for this loaded monopole with near hemispherical radiation coverage at the defined operating frequency of 1600 MHz. The objective function was therefore set as the least squares deviation of the radiation pattern from the desired uniform pattern to evaluate the appropriate fitness of each design to the goal. This function, as shown in equation 5.3, was defined to consider the hemispherical surface 10° above the ground plane but was not limited to consider only the three θ-plane cuts, corresponding to ϕ angles of 0° (in the plane of the folded element), 45° and 90° as used in Altshuler's work. From the definition of the objective function, it is clear that for an ideal configuration the value of this function should be zero however this is impossible from the physical electromagnetic mechanism. Therefore, the lower the value, the better the design. In the interests of saving computer execution time as well as maintaining precision, the increment of θ and ϕ are both set to 5° in this work.

$$Objective\ function = \sum_{0^\circ \le \phi \le 360^\circ} \sum_{0^\circ \le \theta \le 80^\circ} [Gain(\theta,\phi) - Gain_{ave}]^2 \qquad (5.3)$$

After the fitness of each antenna configuration, generated by this process, had been evaluated they were then ranked. The top one-third of the population (i.e. 30 chromosomes) was retained for reproduction and the others are discarded. In the reproduction process, the schemes "from top to bottom" and "dual crossover-points" as described in chapter 3.3 were used for selecting and pairing the survivors to produce the new 60 chromosomes. The rate of 0.5% of the total bits in each generation was selected to perform mutation to maintain a more diverse population.

5.4.2 Optimal Results

After an adequate number of iterations, the GA reached convergence and produced a configuration that satisfied the radiation patterns of a near uniform coverage in the upper half-space. In this example the optimised value estimated by the objective function is 42.5. Although it is still larger than the unreachable ideal value of 0, however it is small enough to enable the distribution of the radiation fields produced by this optimised antenna to be nearly uniform over the hemisphere. Table 5.5 illustrates the resulting bit-string and its associated dimension of each wire for this optimised antenna. Its configuration as shown in figure 5.14 is apparently different to Altshuler's because of the difference between the two objective functions. The far-field properties such as E_θ, E_ϕ and the total electric field E_T were also computed at the frequency of 1600 MHz using NEC2. In order to compare the results with those of Altshuler (shown in Appendix figure A.6 and A.7), the computations were focused on the fields of three cuts in the θ plane corresponding to azimuth angles of 0°, 45°, and 90°. The results as shown in figure 5.15 indicate that even if the fields of E_θ and E_ϕ are variable the total electric fields are all nearly uniform over the three θ planes. These results are all very similar to those of Altshuler (figure A.6). However, the objective function defined here is to optimise the flatness of E_T over the almost complete hemisphere, therefore, undoubtedly, the resulting pattern should be better than that just calculated using the square deviation of the three cuts in the θ plane. The comparative curves are plotted in figure 5.16. In this plot, the horizontal axis represents the θ plane in the variation of ϕ while the vertical axis shows the square deviation for different θ planes. It is clear that the deviation is lower than Altshuler's over the almost entire

ring of ϕ, except for that when ϕ is around 90°. The least square deviation shows that the radiation pattern is more uniform. In addition, the frequency dependence of this antenna was also examined and compared with that of Altshuler's design (figure A.7). The frequency dependence of the power gain of this monopole was computed for θ-plane cuts at ϕ=45° in increments of 100 MHz, over the range from 1400 to 1800 MHz. The results are shown figure 5.17. Referring to the curve computed at 1600 MHz, the operating frequency at which the antenna was optimised, the other patterns all fluctuated but with a variation of less than 6 dB. Again, this result is similar to Altshuler's shown in figure A.7. This is a truly remarkable result which illustrates the power of the GA to produce a far from expected geometrical configuration capable of satisfying a given objective.

Table 5.5 The optimal dimensions and their associated bit strings for a monopole loaded with a loop antenna.

Wire	Z_1	Z_2	Z_3	Z_4	X_1	X_2
Bit-string	00000	00001	11010	00010	10100	00111
Length (λ)	0.0300	0.0129	0.0855	0.0790	0.3403	-0.1516
Length (m)	0.0056	0.0024	0.0160	0.0148	0.0638	-0.0284

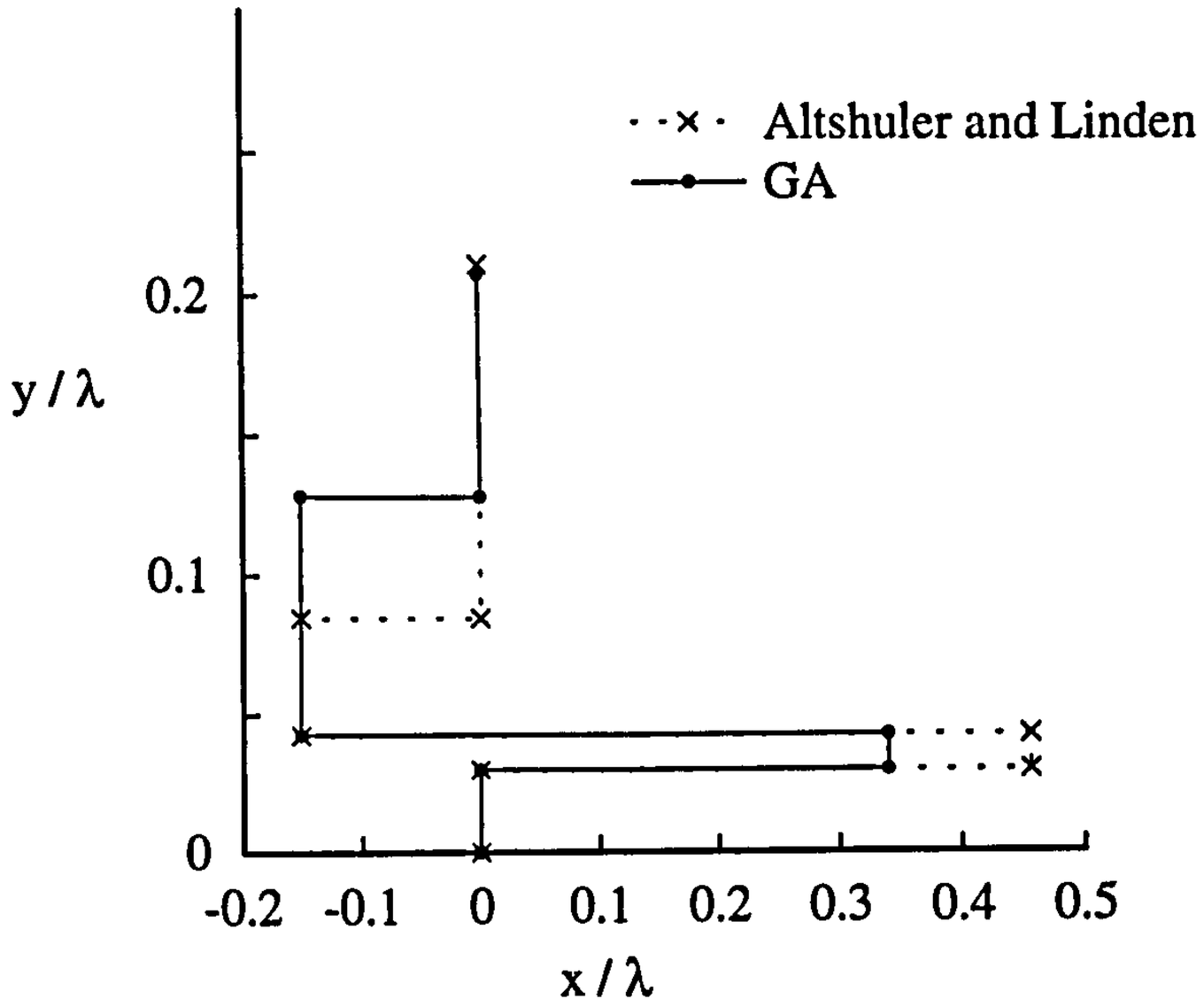


Figure 5.14 Comparison of the GA predicted optimal configuration of a monopole loaded loop antenna and that from Altshuler and Linden (1997).

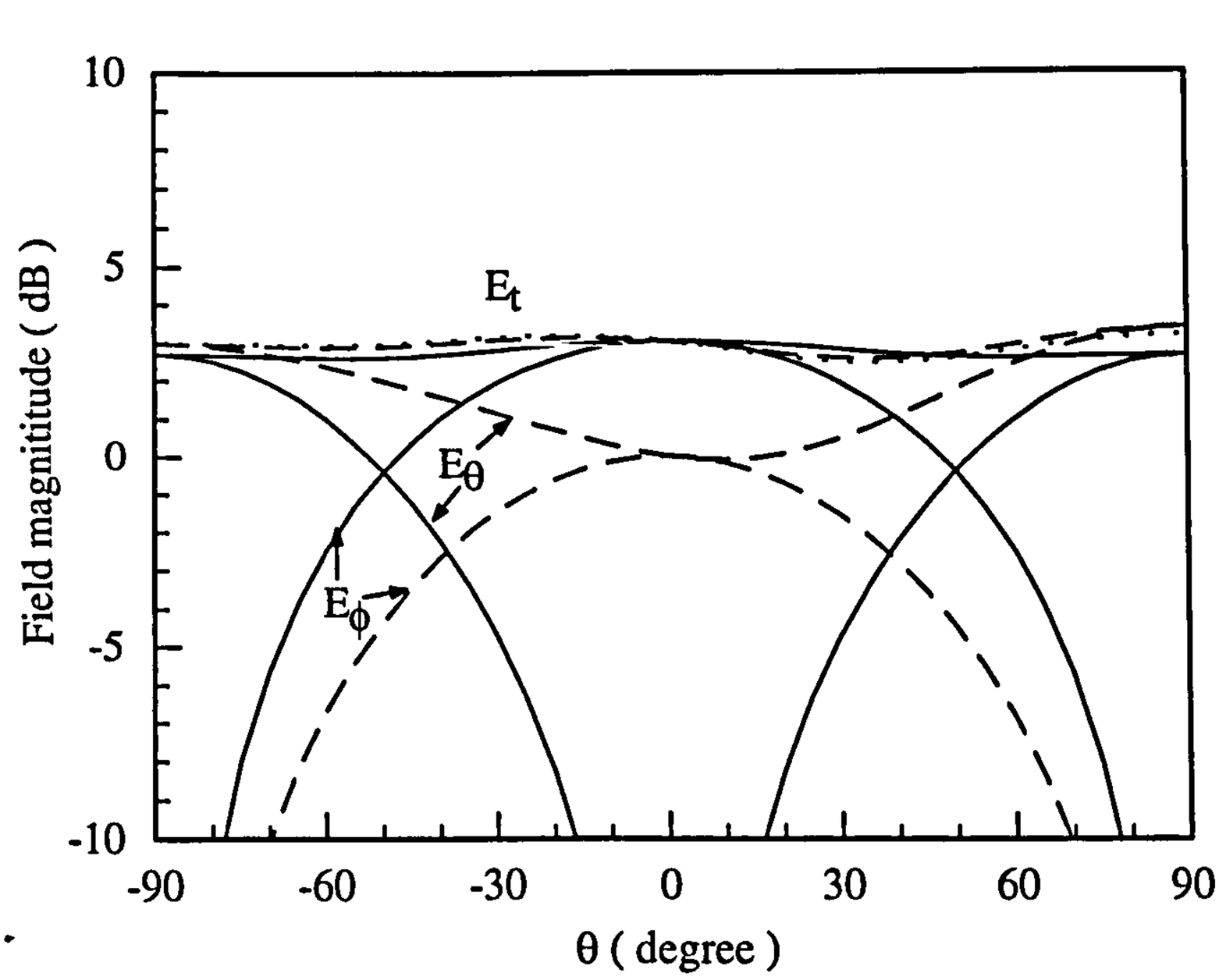


Figure 5.15 Predicated field amplitudes of E_θ, E_ϕ and E_T against the variation of θ for ϕ=0°, 45° and 90° at 1600MHz.

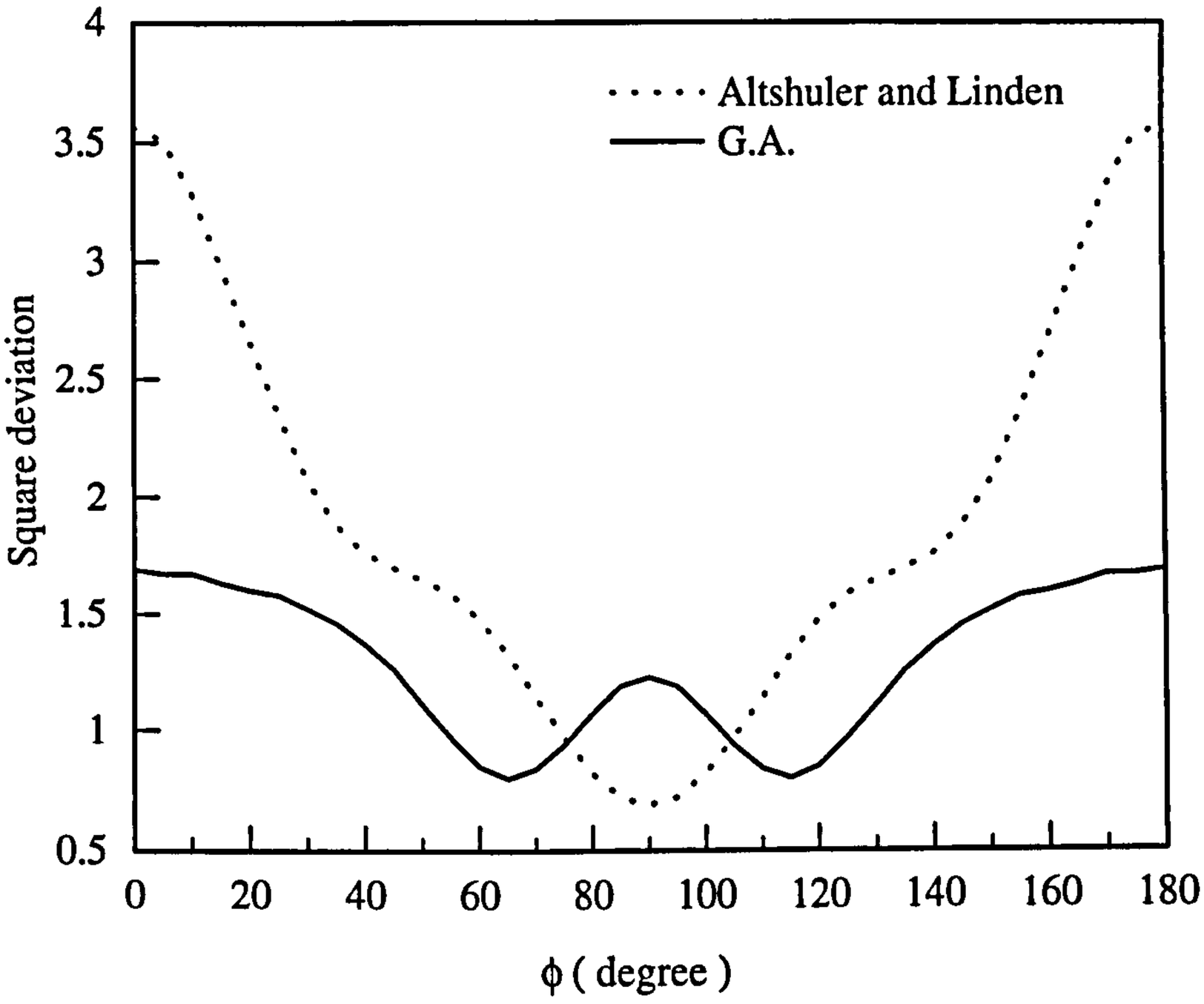


Figure 5.16 Evaluated total square deviation for the θ-plane at different ϕ.

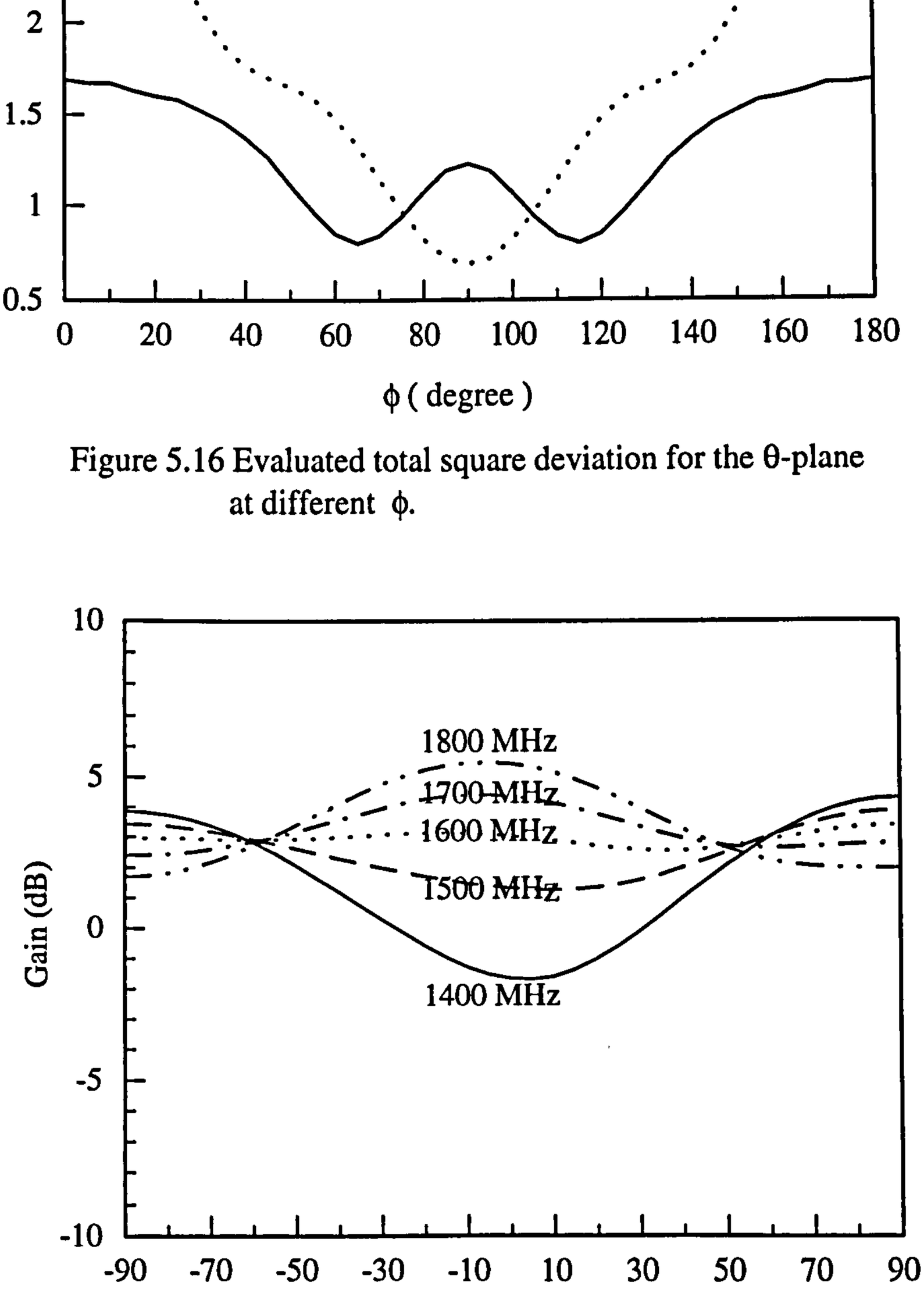


Figure 5.17 Predicated field amplitude of E_T against the variation of θ for ϕ=45° at frequencies from 1400 to 1800 MHz.

5.5 Summary

This chapter has shown the application of the combined GA/NEC2 technique - which was discussed in chapters 3 and 4 to the optimisation of a number of practical antennas.

Firstly, section 5.2 described the implementation of various necessary operating routines for this novel method. This included the general GA routines and the use of the NEC2 code and wire-grid modelling if a conducting structure is included.

In section 5.3 the element shapes of the three-element parasitic array have been optimised for obtaining both high gain and F/B. In the first subsection, section 5.3.1, the published optimal three-element Yagi-Uda array with elaborate curvature of the elements has been reviewed. Then, based on these results the GA/NEC2 method has therefore been used, in section 5.3.2, to optimise the shape of a similar array with 20 wires in each element. The configuration of this array was constructed automatically by the GA and thus the resulting antenna could be called an auto-shaped antenna. Finally a shape with better performance than that shown in the literature has been achieved.

In consideration of the difficulty in manufacturing the above antennas with the peculiar curvature section 5.3.3 therefore has tried to optimise a simple, vee-shape, parasitic array for the same goal by using the GA/NEC2 tool. The finally predicted results showed that the vee-shape array with linear elements can achieve the similar performance as the array with curvature element which had been assumed previously to be the optimal shape. This investigation provides of great advantage to manufacture of this array with a such high gain.

Finally, in section 5.4 further wire antenna design using GA/NEC2 was illustrated by optimising a monopole loaded with a modified dipole to obtain a near hemisphere coverage. A different objective function from that used for the same design shown in the literature has been applied to the GA procedure. The results have provided improvement on both the desired radiation performance and the antenna dimensions.

In summary, the GA/NEC2 design procedure, of simply defining constraints on the design variables and the objective function according to the desired features, and without any *a priori* knowledge of candidate configurations is capable of finding an unconventional antenna able to effectively solve the antenna problems. This produced antenna is usually difficult to be explored by using the method of theoretical analysis.

CHAPTER 6

APPLICATIONS OF THE GA/MOMENT METHOD TO NVIS VEHICLE–ANTENNA DESIGN

6.1 Introduction

The preceding chapter described the principles of using the combination tool - GA/NEC2 for optimisation of various wire antennas with simple structures for appropriate objectives. The results show that this tool is robust in finding the optimal configurations with desired radiation performance. Therefore, in this chapter the technique will be further applied to optimise antennas mounted on more complex vehicular structures for a special application of an HF communication scheme known as Near Vertical Incidence Skywave, or **NVIS**. As introduced in chapter 2 NVIS propagation requires predominant radiation towards the zenith. It is a particularly important mode for certain specialised applications where distance, ground-type and topography preclude the use of either ground wave or line-of-sight techniques.

Since the objective in this chapter is to demonstrate the use of the new technique to produce the optimal NVIS pattern from a vehicle-mounted antenna system, the conductive vehicle with a solid structure should be modelled following the guidelines of wire-grid modelling which has been discussed in section 4.3 to ensure that the accurate electromagnetic properties of the vehicle can be simulated by using the

Moment Method code - NEC2. The dimensions and environment of the vehicle and the details of the resulting vehicle model will be demonstrated in section 6.2.

After defining the wire-grid model of the vehicle, section 6.3 begins by using the GA/NEC2 tool to design the optimised vehicle-antenna systems. A number of different antenna systems involving either a change of antenna configuration or some variation of both the position and number of sources were tested using the GA to obtain a specified antenna configuration which yields optimal performance when radiating towards the zenith. Firstly, in section 6.3.1 the radiation performance of an electrically short vertical monopole (whip) antenna (Virani 1988, Wheeler 1975) was examined. This type of antenna is most frequently used for vehicular HF communications. However, as it is most commonly deployed (mounted vertically) it is a particularly unsatisfactory radiator for NVIS applications. Therefore, in this section the GA is applied to find the best configuration of this antenna to provide maximum radiation efficiency, and hence maximum gain, for NVIS propagation and also to compare those results with published details in the literature. In section 6.3.2 another common type of antenna configuration which will be tested for the same purpose: the vertical loop. Several cases are included. The simplest, the vehicular loop with a single base source will be considered first. In this case, the loop size as determined by the point where the antenna is connected to the vehicle structure will be examined by the GA to yield the best NVIS performance. Then this case will be followed by examining the effect of feeding the system with dual sources. The considered parameters include the positions and phases of the two sources.

Another possible configuration of loading the selected antenna with a lumped circuit will be examined in the following section - 6.4. This technique had been used

on a short dipole antenna by adding properly selected and located loads to change the current distribution along the antenna and thus to improve its radiation efficiency. For the loop considered here, if the maximum current amplitude on the candidate antenna could be concentrated along its horizontal sections by means of suitable loading the NVIS performance of the antenna should undoubtedly be improved. The GA is therefore used to determine a suitable set of the lumped load and its correct position along the loop in order to obtain a significant increment of the current distribution on the horizontal sections so as to enhance radiation upwards.

Finally, section 6.5 will go into the design of a matching network for the resulting NVIS antenna system using the GA method and then compare it with that from the theoretical analysis.

6.2 Wire-grid Model of the Land Rover Vehicle

As discussed in chapter 4 to ensure that the radiation characteristics predicted by using the Moment Method are accurate, the analysed structure has to be correctly represented by a wire-grid model. In this study the complex structure on which the antennas will be mounted is a standard long-wheel-base Land Rover, typical of the type of vehicle from which NVIS communications are required. The conductivity of the vehicle surface is approximated by that of aluminium. The distance between the bottom of the vehicle and the ground plane was assumed to be 0.5m. Figure 6.1 shows the dimensions and the produced wire-grid model of this vehicle. It was developed by Murray (1993 p201) using the accepted modelling criteria that have been discussed extensively in chapter 4. The box-like configuration, without fine detail, was found to be entirely acceptable at the frequencies of interest, so the added complexity of

surfaces with small radii of curvature was unnecessary. Except for the case of the whip antenna in which the vehicle was assumed to be over "typical" ground ($\sigma=10^{-2}$ Sm^{-1}, $\epsilon_r=10$), all the other computations to determine the NVIS performance of the antenna-vehicle systems were carried out with the vehicle positioned above an infinite, perfectly conducting ground plane.

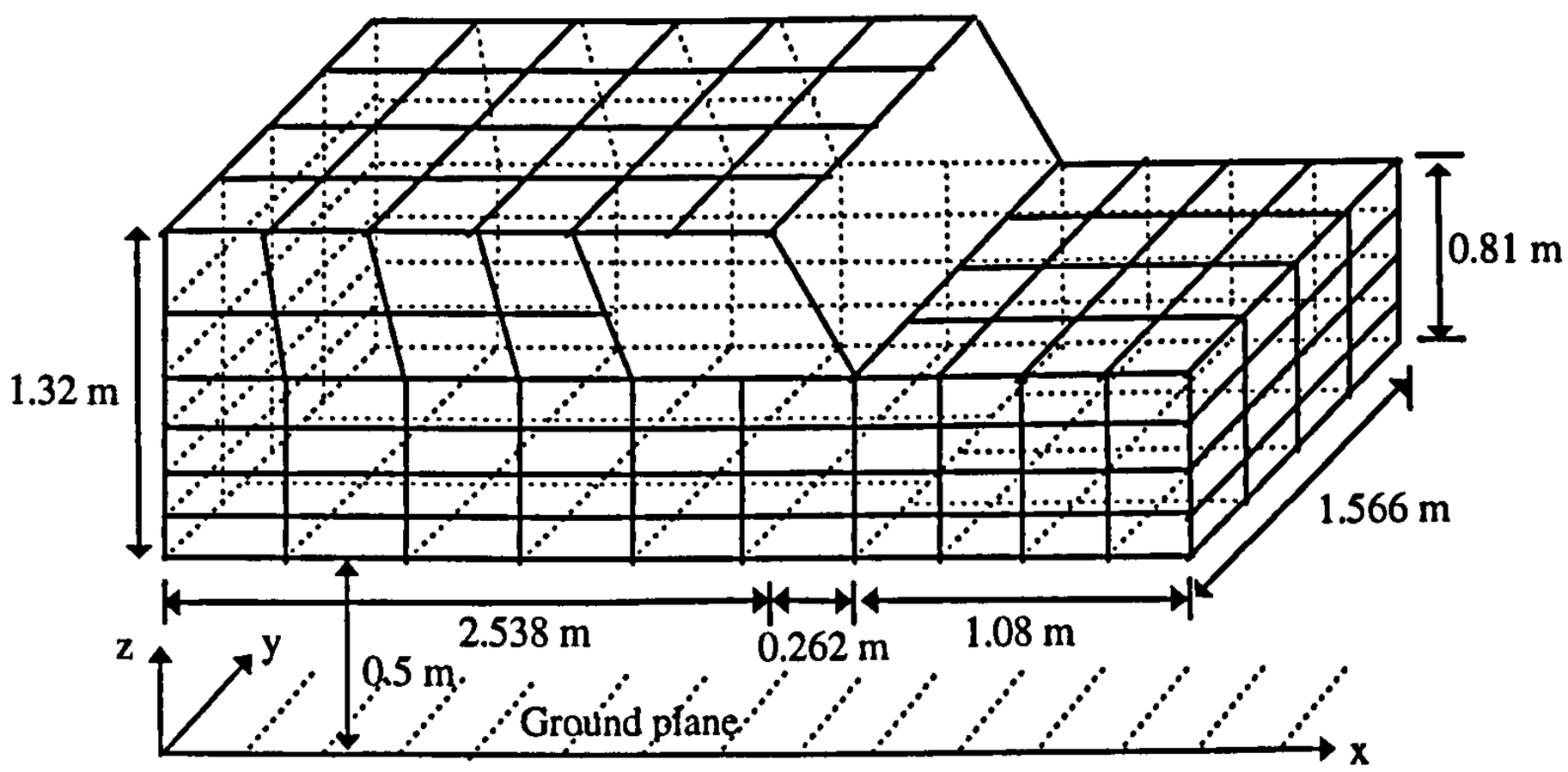


Figure 6.1 A wire-grid model of the Land Rover vehicle.

6.3 NVIS Vehicular Antenna Design Using GA/NEC2

6.3.1 Vehicular Monopole (Whip) Antenna

In this section, the NVIS performance of a thin-wire, base-fed vertical monopole antenna (whip) mounted on the vehicle above a good ground was optimised using the GA/NEC2. In the past, many entirely analytical or experimental techniques, as

reviewed in chapter 2, have been used in this investigation. However the examined antenna orientation angle in those cases was all limited to only the zenith angle θ. This motivated the use of the GA to examine a wider range of orientation angles including zenith angle θ and azimuth angle ϕ for searching the solution which may exist but has still not been found so far. The dimensions of the antenna, based on the work of Austin and Murray (1998) were therefore chosen as: antenna length of 4.8 m and radius a=0.005m. Compared to the wavelength of the typical NVIS frequency range from 2 to 10 MHz this antenna is electrically very short. In addition, in order to prevent the antenna from overhanging vehicle body when entirely horizontal and to let the fed source be independent to the tilt angle, a 0.1m vertical section at the rear right corner of the roof was used to accommodate the source. The system is shown in figure 6.2.

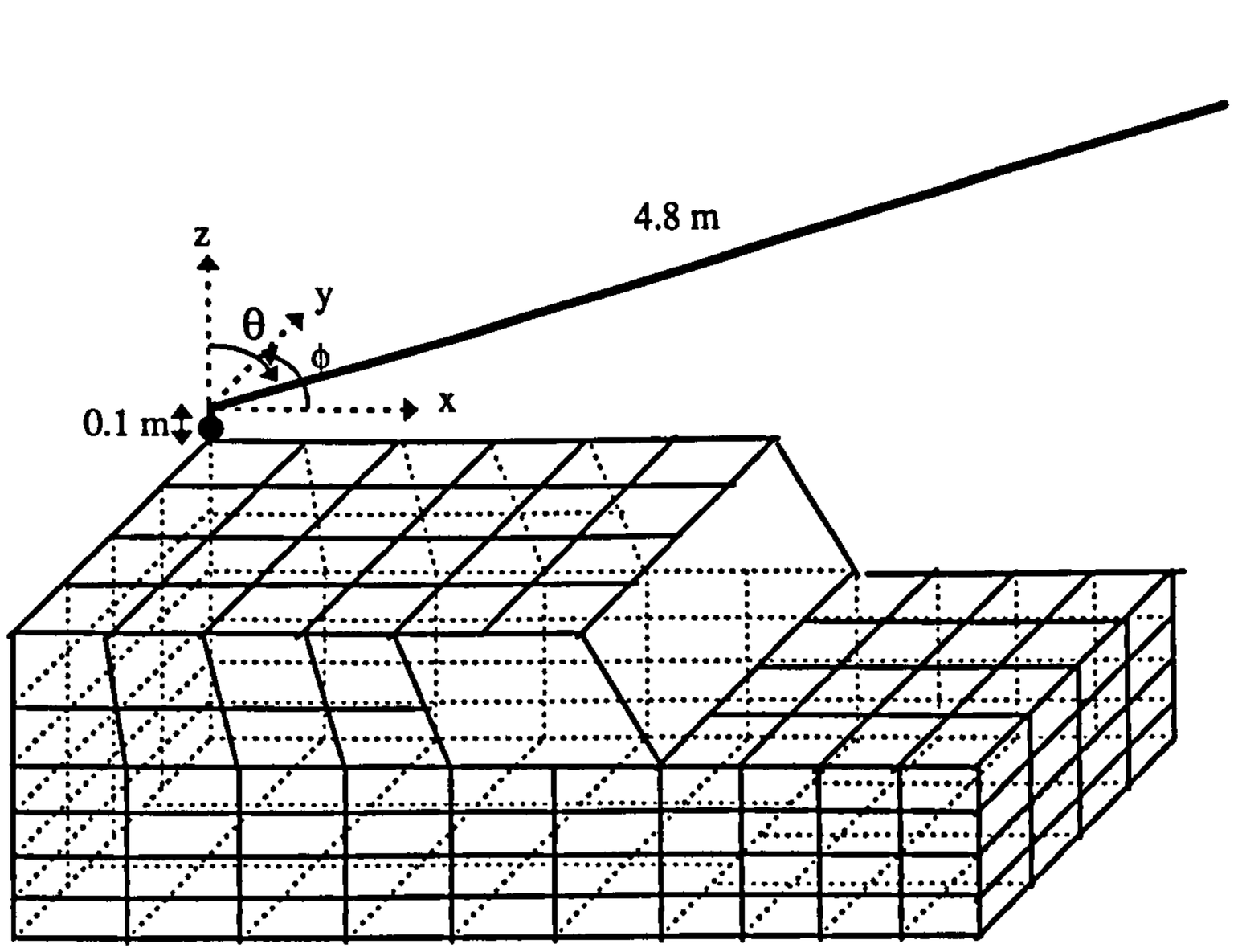


Figure 6.2 The Land Rover vehicle mounted with a short vertical monopole.

6.3.1.1 GA Implementation

For this design, concern was focused on the effect of the orientation of the antenna on the improvement of NVIS performance. Thus the zenith angle θ and azimuth angle ϕ of the antenna were set as the design variables to optimise the NVIS performance of the system. These angles were calculated referring to the position of feeding source as indicated in figure 6.2. For practical purposes, the variations of θ and ϕ were limited only within the appropriate range of $0°\leq\theta\leq90°$ and $-180°\leq\phi\leq0°$ respectively. In the encoding procedure θ and ϕ were discretely represented by eight and nine binary bits, respectively. Thus each chromosome was composed by a string of binary bits with a total number of 17. It can be represented as

$$\text{a chromosome} = \{\, b_1^{\theta} b_2^{\theta} \cdots\cdots b_8^{\theta} b_1^{\phi} b_2^{\phi} \cdots\cdots b_9^{\phi} \,\} \tag{6.1}$$

where, b_i^{θ} is the bit for θ; and b_i^{ϕ} is for ϕ.

The mapping relation between the binary string and the real parameter can also be illustrated as:

$$\theta = \Delta\theta \cdot \sum_{i=1}^{8} b_i^{\theta} \cdot 2^{i-1} \tag{6.2}$$

$$\phi = \Delta\phi \cdot \sum_{i=1}^{9} b_i^{\phi} \cdot 2^{i-1} \tag{6.3}$$

Here, $\Delta\theta$ and $\Delta\phi$ are the bit resolutions of the variables θ and ϕ, respectively. They are defined as

$$\Delta\theta = [90 / (2^8 - 1)]^{\circ} \tag{6.4}$$

$$\Delta\phi = [-180 / (2^9 - 1)]^{\circ} \tag{6.5}$$

In a GA approach, the fitness of the created antennas is assessed by the defined objective function. In this example, the goal was to obtain a maximum percentage of the input power which was radiated within a target region concentrated on the zenith, within a cone of 45° half-angle, thus it seems appropriate to designate the term G_{45} and set it as the objective function. In equation form,

$$objective\ function = G_{45}$$
$$\text{or } G_{45} = \frac{P_{rad}(\theta < 45^{\circ})}{P_{in}} \times 100\% \tag{6.6}$$

The GA's goal was to maximise the value of this function by adjusting the orientation of the whip. A population of 60 chromosomes was generated randomly as the first generation and then the better 20 chromosomes in each generation were kept and used to pair with each other to produce the 40 new chromosomes for use in the next generation. These were then paired alternatively and the two-point crossover rule, as well as 1% of mutation rate, were applied to this example.

6.3.1.2 Discussion of Results

The process of the algorithm was monitored to see when it converged. This procedure was repeated several times with different members of the initial generation at each operating frequency within the NVIS band. To illustrate the GA progress in searching for a solution for this example the case of 10 MHz was selected to be representative. Figure 6.3 shows the progress of tilt angles, θ and ϕ, in the GA optimisation as a function of the number of iterations when the antenna was operated at this frequency. After some initial fluctuation the values of θ and ϕ stabilised after

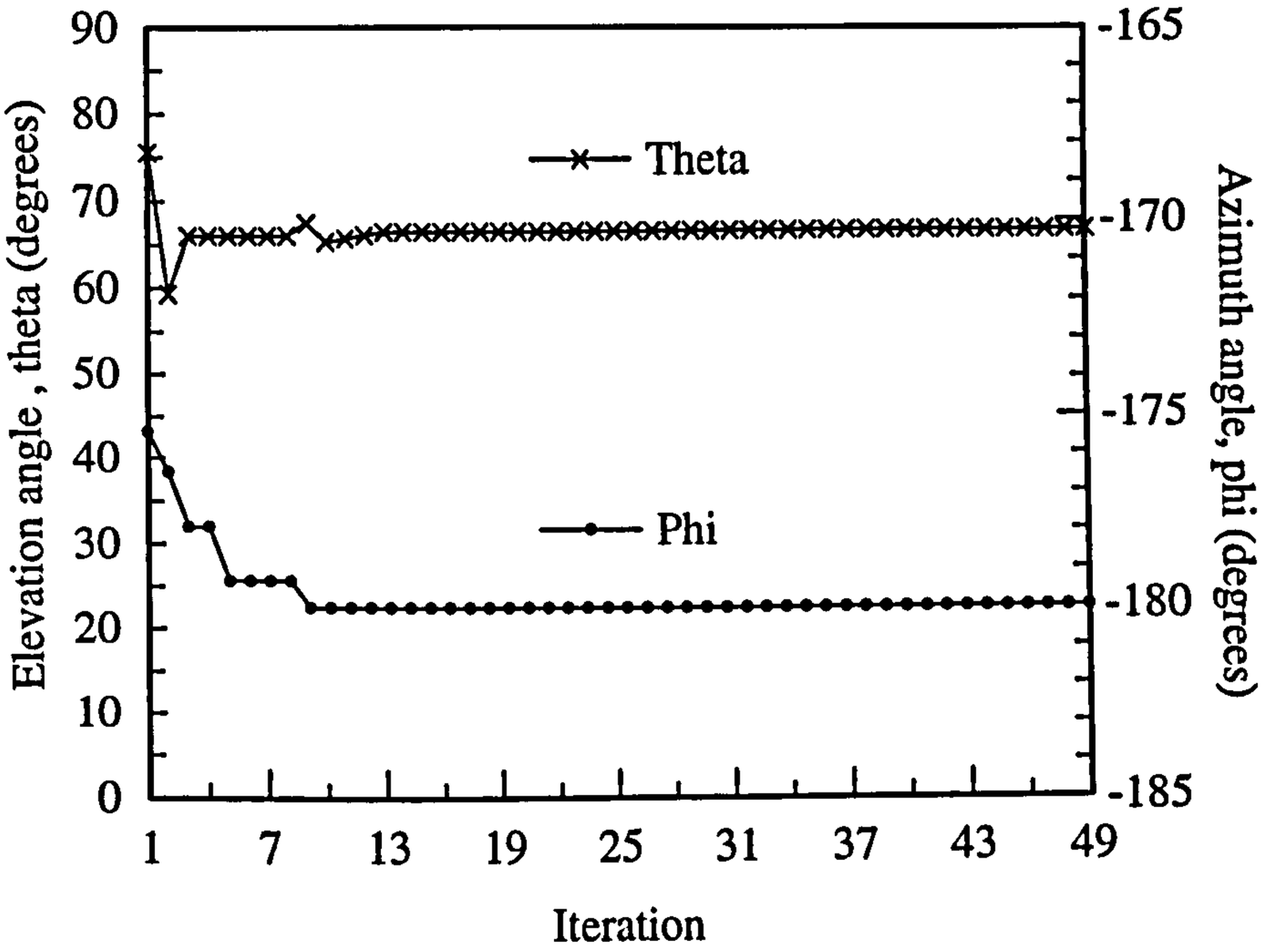


Figure 6.3 The progress of tilt angles, θ and ϕ, in the GA optimisation as a function of the number of iterations at 10 MHz.

about 13 iterations when they reached the values of 66.4° (θ) and -180° (ϕ), respectively. This result proves that when the monopole is tilted towards the horizon and away from the vehicle, an optimal percentage of the input power falls within the defined cone. It agrees broadly with those from the simulations of Eley et al. (1991) and the Characteristic Mode calculations of Austin et al. (1998). The improvement of the G_{45} term as a function of the number of the iterations is shown in figure 6.4. The individual contributions that produce this optimal $G_{45}=\eta D_{45} = 9.2\%$ are $D_{45} =42.8\%$ and $\eta = 21.6\%$.

Figure 6.5 is a plot of the predicted optimal total "gain", G_{45}, of this vehicular whip antenna with frequency and figure 6.6 is the corresponding tilt angles of θ and ϕ. Clearly, the total efficiency is proportional to the frequency. This is reasonable for as the frequency increases the whip of fixed length becomes more efficient, while the

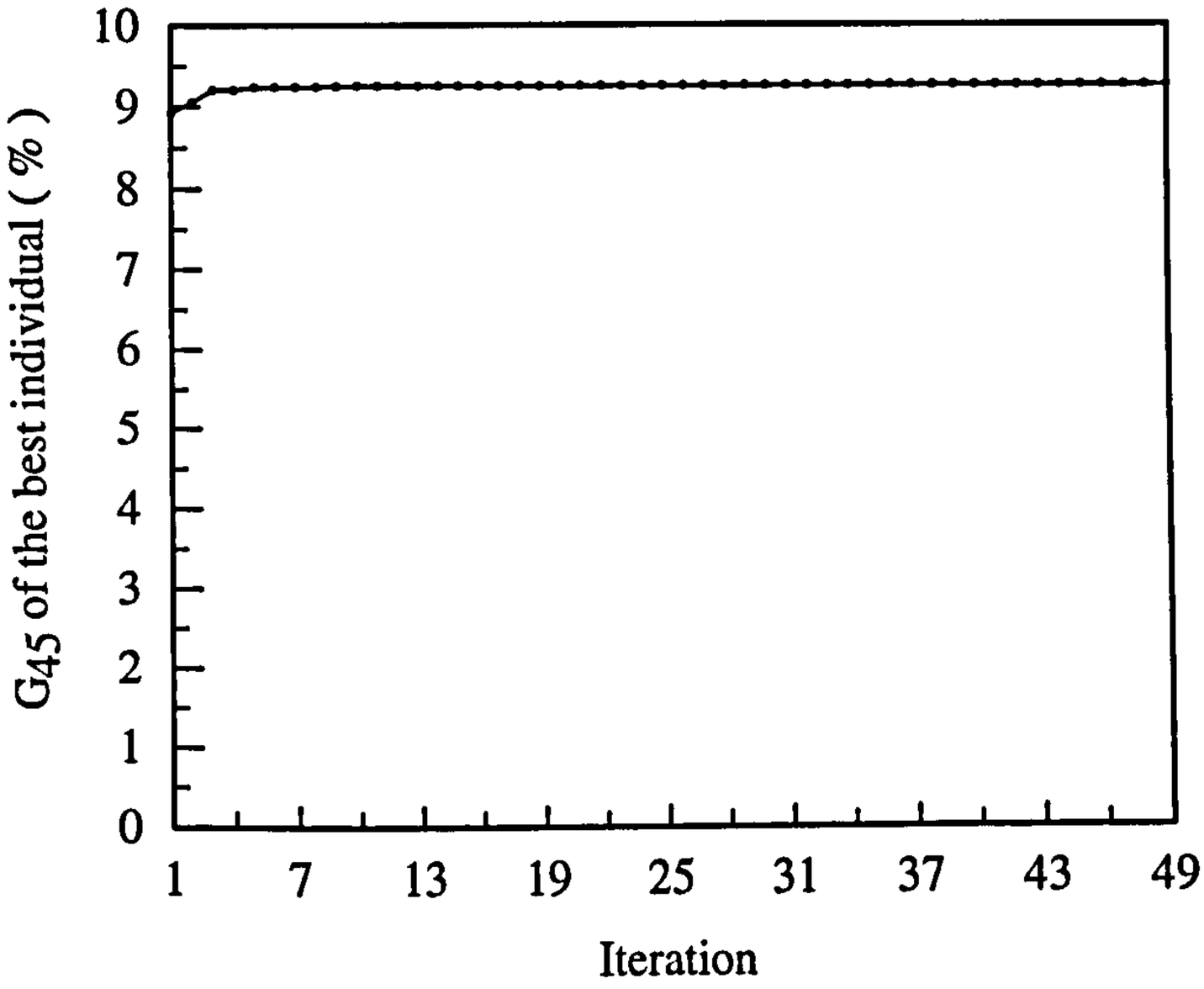


Figure 6.4 Total efficiency, G_{45}, of the best individual as a function of the number of iterations of the GA for a vehicular whip antenna at 10 MHz.

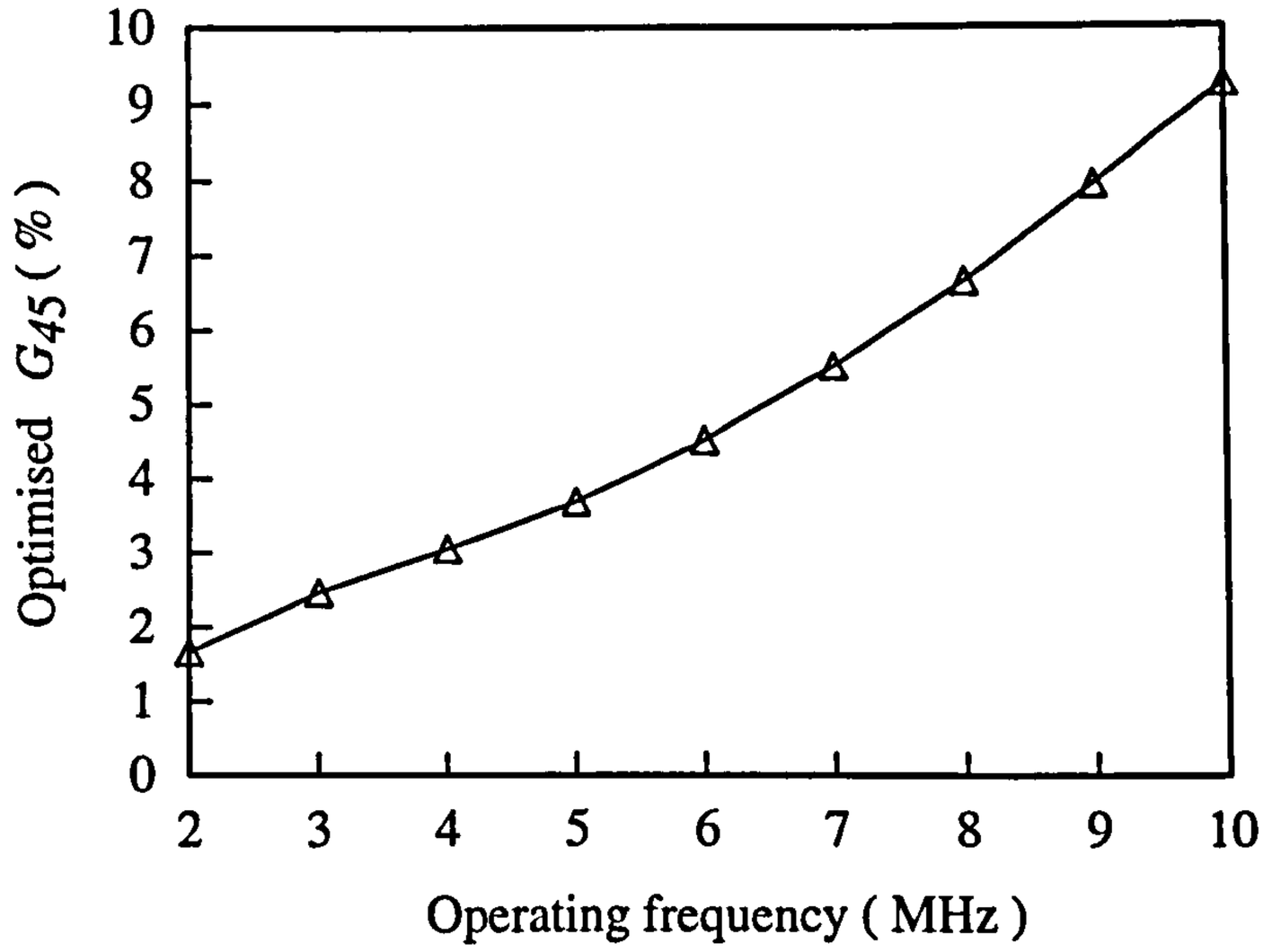


Figure 6.5 Variation of the optimised G_{45} against the operating frequency for a vehicular whip antenna.

tilting improves the skywave radiation component. At 2-3 MHz the effect of tilting the whip is to decrease its radiation resistance so severely that its efficiency falls dramatically. Thus G_{45} is minimal. But at ~8 - 10 MHz the effect of increasing θ is to optimise D_{45} while not severely decreasing η, hence G_{45} can be maximised.

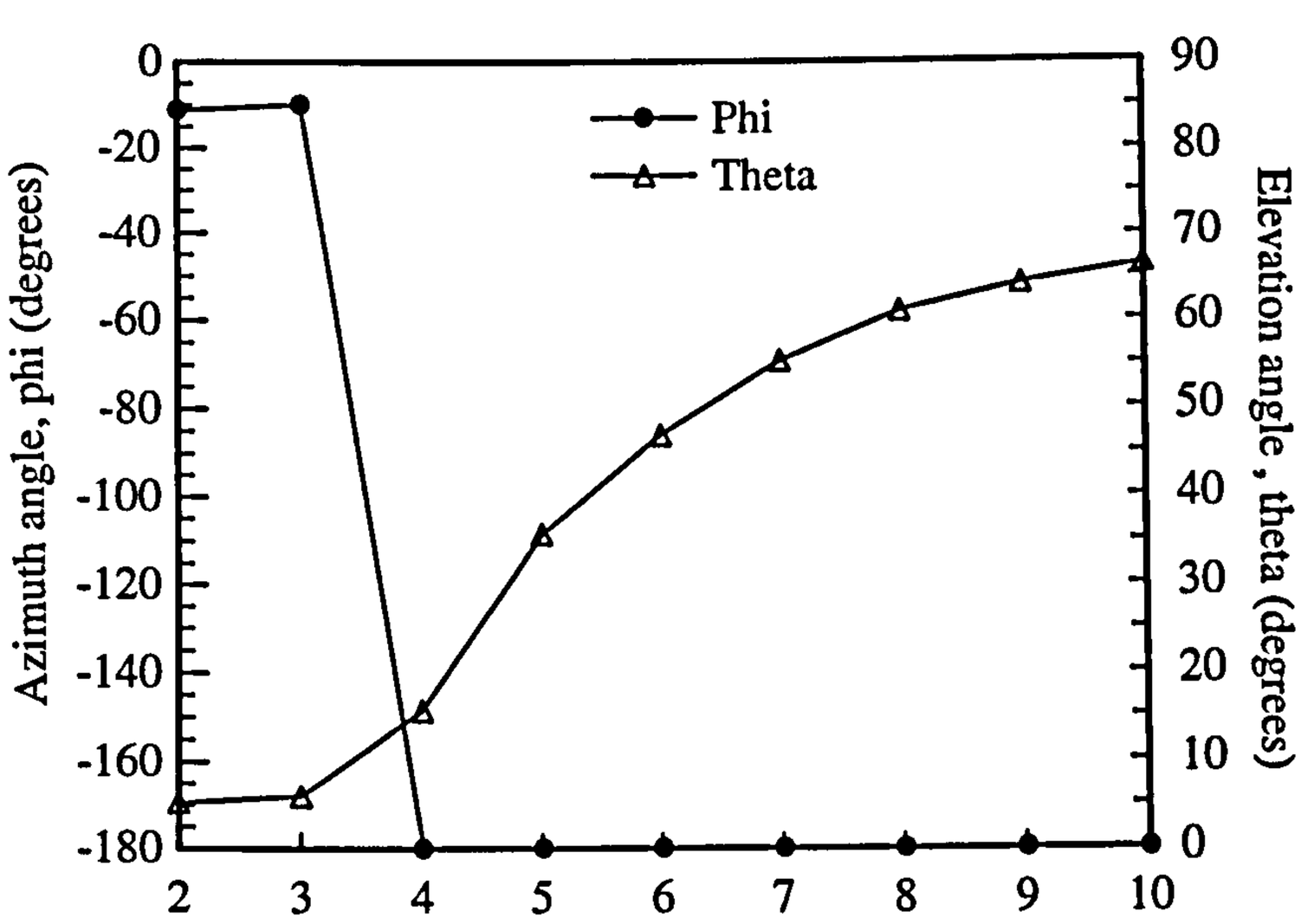


Figure 6.6 The resulting tilt angles, θ and ϕ of the vehicular whip for optimal NVIS radiation over the interested band.

6.3.2 Vehicular Loop Antennas

The results indicated in the previous section show that tilting the whip horizontally away from the vehicle can achieve optimal NVIS performance for that antenna system. However the resulting G_{45}, $1.7 \leq G_{45} \leq 9.2\%$, may not provide satisfactory performance for such applications. This is a thus strong motivation to

investigate another type of antenna - a loop antenna - for the same application. In addition, there are also some other reasons for choosing a loop as the candidate. Firstly, the operational and siting considerations give the loop antenna some mechanical advantages over other types in this application. Secondly, for electrical consideration, by shorting out portions of the loop it can be made to exhibit an inductive reactance which does not vary greatly across the operating band and this allows the use of high Q elements (capacitors) for the matching network. In most of the previous investigations, as reviewed in section 2.3.2, the configuration of the loop and the point of termination to the vehicle body were limited to some fixed criteria for problem simplification and by practical considerations. With the potential of searching for a solution from a broad solution domain the GA can be applied to optimise the configurations of vehicle-mounted loop antennas from a wider range of dimensions for optimal NVIS performance. The configurations of the loop mounted on the vehicle including a loop with single-feed source and with dual-feed sources will be examined. In these cases, the conductivity of aluminium used for the vehicle structure was also applied to the loop antenna when computing the electromagnetic properties of the system over the NVIS frequency band. In addition, a practically reasonable selection of 0.025m for antenna radius was applied and the ground was assumed to be an infinitely, perfectly conducting plane.

6.3.2.1 Single-feed Loop Antennas

The first example will be the optimisation of the loop antenna fed by a single source. Figure 6.7 shows this objective configuration. The rear end of the antenna, denoted as A, was fixed at the rear right corner of the vehicle and fed by a voltage

source. The coordinates of x, y and z for point A are (0.0, 1.566, 1.82) in units of metres. In the GA approach of this design, some constraints were put on the height (denoted as h) of the loop measured from the source end to the top of the loop and the location of the front end (denoted as B) of the loop. For practical consideration, the ranges of h and B are clearly limited. Referring to the solid structure of this vehicle, the former was confined into the range from 0.1m to 1.5m while the latter was allowed to be positioned at any place on the hood or roof of the vehicle.

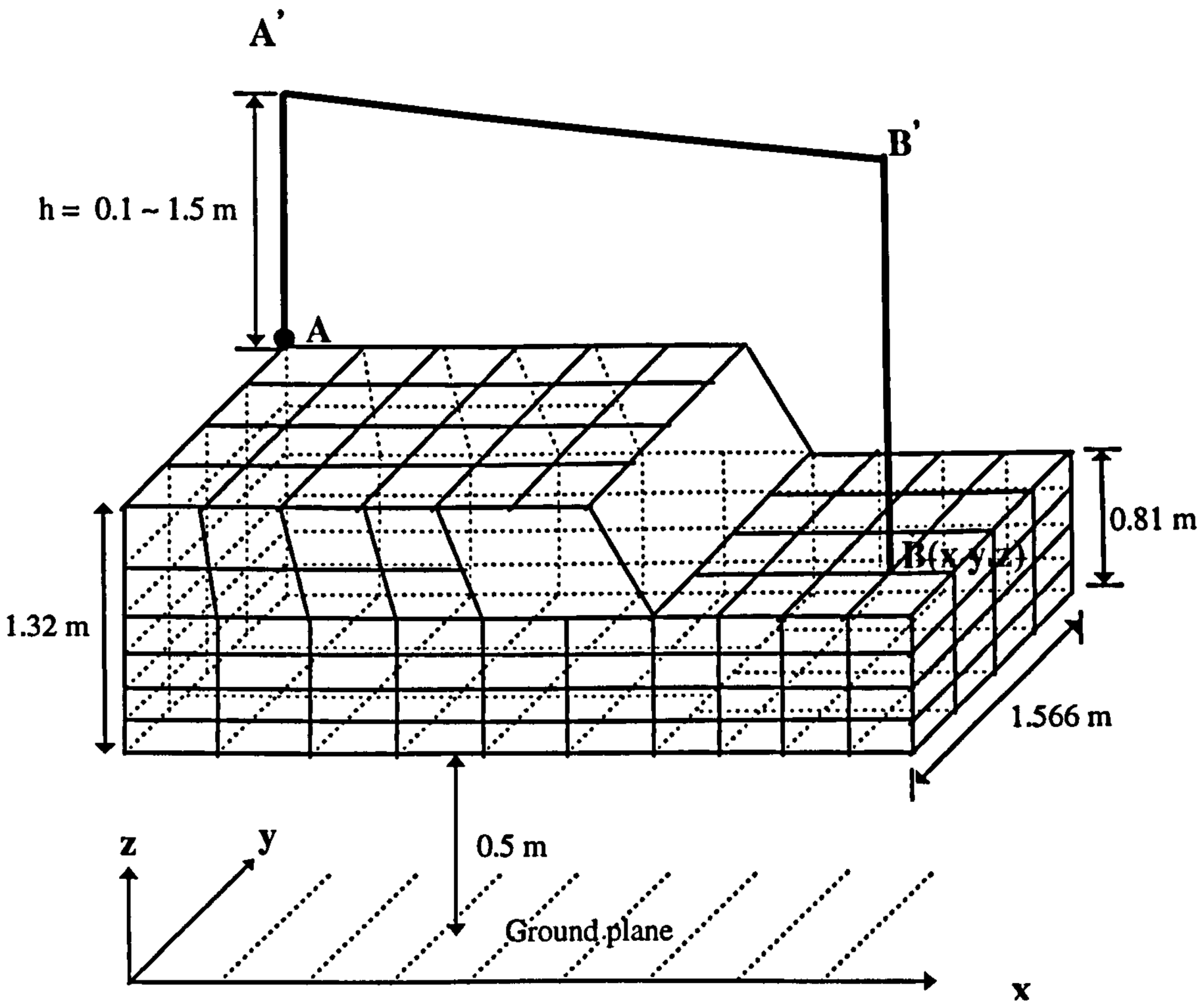


Figure 6.7 The objective configuration of a loop antenna mounted on the Land Rover vehicle.

a. GA Implementation

Since the point B must be connected to the plate of the hood or the roof of the vehicle to form a closed loop, the z-coordinate of point B could thus be only either 1.31 m for the hood plate or 1.82 m for the rear roof such that it is unnecessary to include the z-coordinate of point B into the design variables for the GA procedure. Therefore, the parameters which were manipulated by the GA were the loop height (h) and the coordinates (x, y) of the point B. In the encoding process, four, four and three bits were selected to map the values of h, x and y, respectively. Thus each chromosome was represented by a general form with eleven bits and can be expressed in a form as

$$\text{a chromosome} = \{\, b_1^h\ b_2^h\ b_3^h\ b_4^h\ b_1^x\ b_2^x\ b_3^x\ b_4^x\ b_1^y\ b_2^y\ b_3^y \,\} \tag{6.7}$$

where the superscript of b_i depicted the associated design parameters.

Thus, the relation between the genes and real parameters were expressed as

$$h = 0.1 + \left(\frac{1.5 - 0.1}{2^4 - 1}\right) \cdot \sum_{i=1}^{4} b_i^h \cdot 2^{i-1}\ \text{m} \tag{6.8}$$

$$x = \left(\frac{3.88}{2^4 - 1}\right) \cdot \sum_{i=1}^{4} b_i^x \cdot 2^{i-1}\ \text{m} \tag{6.9}$$

$$y = \left(\frac{1.566}{2^3 - 1}\right) \cdot \sum_{i=1}^{3} b_i^y \cdot 2^{i-1}\ \text{m} \tag{6.10}$$

For each generated design the predicted percentage of the power radiated within the cone of half-angle 45° was considered as its fitness for surviving or being discarded in the next generation. This means that the objective function used for this case is the same as equation 6.6. The first generation was composed of 60 chromosomes generated by randomisation. Referring to the methodology of the alternative pairing scheme the top one-third chromosomes in the ranking list will

survive and will be used to create the new offspring. Meanwhile, among the reproduction processes the dual-point crossover rule and the mutation rate of 0.5% of the total bits in each generation were applied to the basic genetic operations: crossover and mutation, respectively. The same procedure was applied to optimise the loop configurations operated within the frequency range from 2 to 10 MHz.

b. Results

For the 10 MHz case, after the various iterative calculations typical of the GA technique, the optimal configuration of the loop was produced. Indeed, it took only five iterations for the algorithm to converge. The iterative process is shown in table 6.1. The results showed that the size of the loop will mainly affect the total efficiency: the larger the loop area the greater the power radiated within the defined cone. In other words, when the loop height reached its maximum value of 1.5 m, which was set by practical considerations, and the front base was positioned at the front left corner of the vehicle, the antenna area reached its maximum and therefore the total efficiency G_{45} approached its maximum. This, of course, assumes uniform current distribution around the loop.

The variation of the maximum G_{45} against frequency over the NVIS band was summarised and plotted in figure 6.8. Although the optimised dimensions of the loop for operating at different frequencies are not shown here it should be realised that the same configuration was obtained for all frequencies while the maximum G_{45} value was very dependent on the actual operating frequency and actually decreased above (approximately) 5 MHz. The discrepancy mainly came from the change of current distribution on the loop. This can be clearly seen in figure 6.9 which plots the

Table 6.1 Process of the GA for the optimisation of a single–fed loop antenna at 10 MHz.

Iteration	Best chromosome h x y	Real values (m) h x y z	Fitness G_{45} (%)
1	0111 1011 010	1.41 3.36 0.45 1.31	13.7
2	1111 1111 100	1.50 3.88 0.22 1.31	16.5
3	1111 1111 000	1.50 3.88 0.00 1.31	17.0
4	1111 1111 000	1.50 3.88 0.00 1.31	17.0
5	1111 1111 000	1.50 3.88 0.00 1.31	17.0
...	....	....	.....

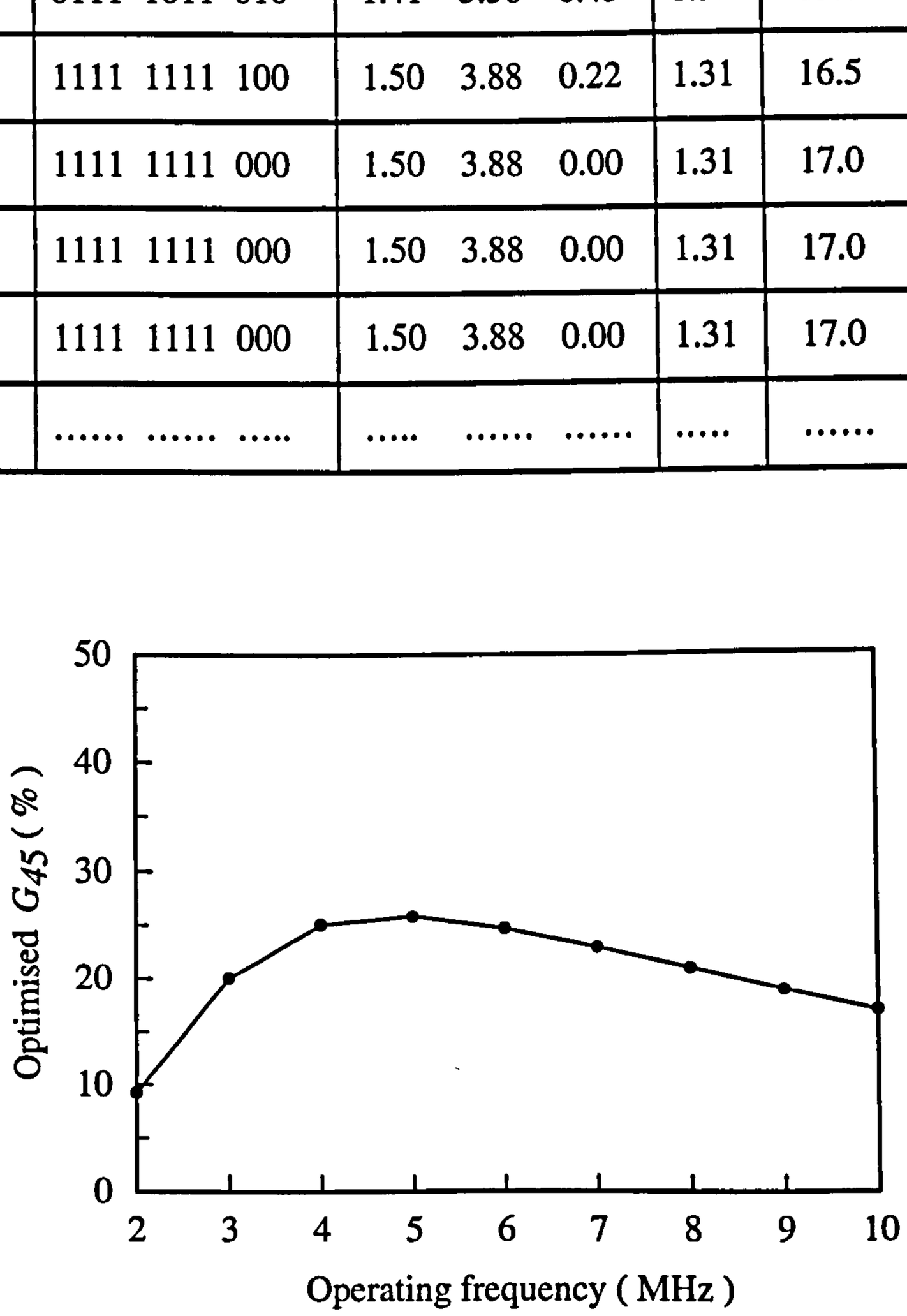


Figure 6.8 Variation of the optimised G_{45} against the operating frequency for a single-fed vehicular loop.

normalised current distribution on the loop against the operating frequency. At each frequency the current was normalised by the maximum amplitude. As those marked in figure 6.7, A, A', B' and B which appear in this figure are the critical positions on the loop. In essence, below 4 MHz, the current is uniformly distributed around the loop because it is electrically very small. Above that frequency, however, the amplitude of the current is no longer constant but varies with position around the loop as the frequency increases towards 10 MHz, at which point the loop perimeter is about 0.4λ, which no longer satisfies the small loop criterion of constant current (King 1969). The consequence of this is that instead of maximum power being radiated upwards it begins to shift away from the zenith which leads to the performance as the curve shows in figure 6.8. As to the distribution of the current phase on the loop figure 6.10 illustrates the results calculated by the NEC2. Undoubtedly, they could be predicted with hardly any change as the perimeter of the loop is very small compared to the wavelength. However a significant phase change exists between 9 and 10 MHz which will imply that the resonance frequency is around there (9.4 MHz, in fact). This can be proved from the change of reactance as will be shown later in the section 6.5.

Moreover, the effect of the position of the source on the directivity (D_{45}), the radiation efficiency (η) and the total efficiency (G_{45}) was also evaluated by using the GA approach. Here the previously determined optimal loop configuration was used and then only the position of the feeding source was included into the design space. It was allowed to vary in the range from point A to B along the antenna section. Five bits were selected to digitally represent this position and this selection of bit number was such that the loop was therefore divided into 31 segments. This implied that the position of 1.6% indicates that the source is positioned at the rear base (i.e. point A);

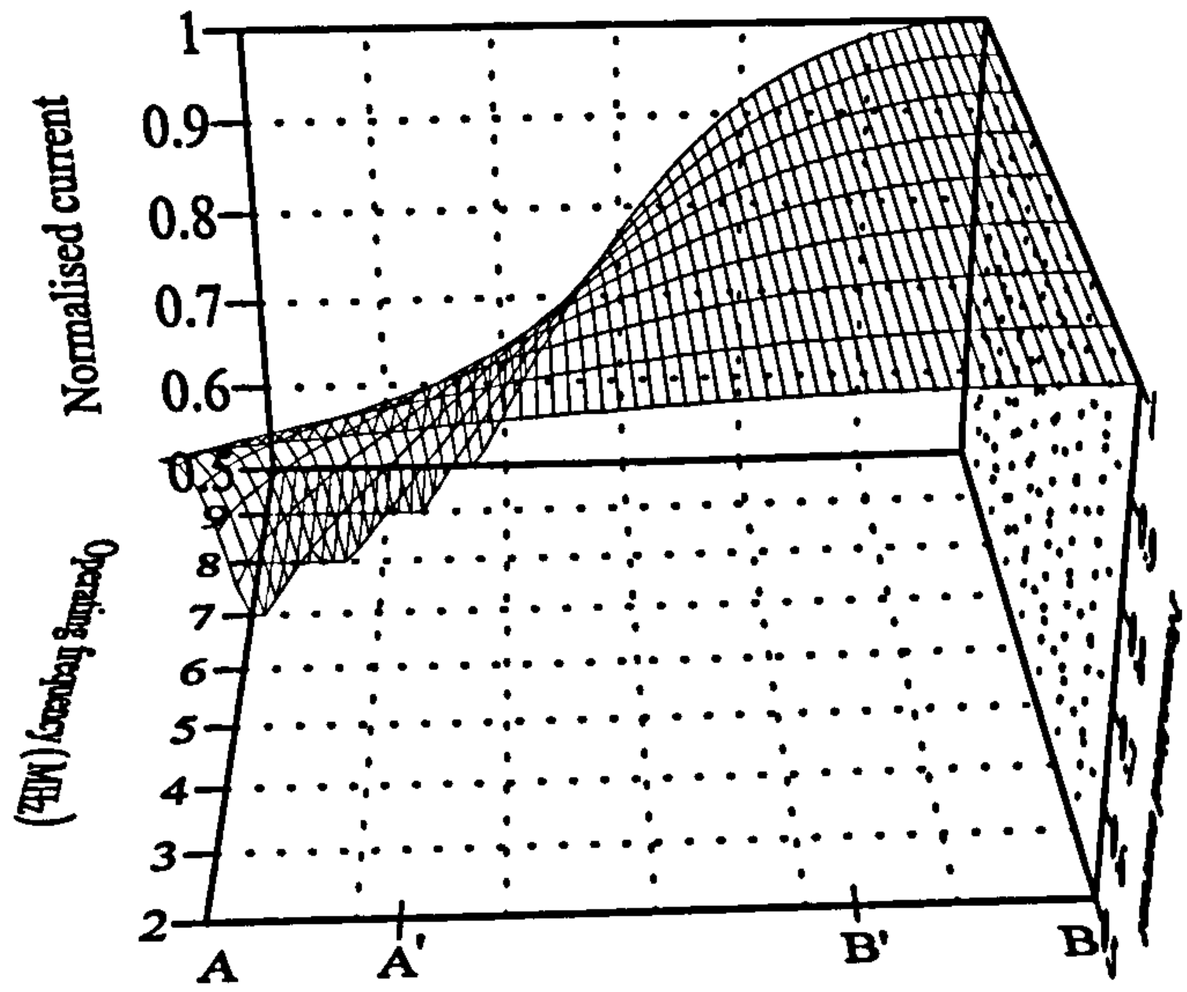


Figure 6.9 The variation of normalised current distribution on the loop driven by a single source as a function of position and frequency.

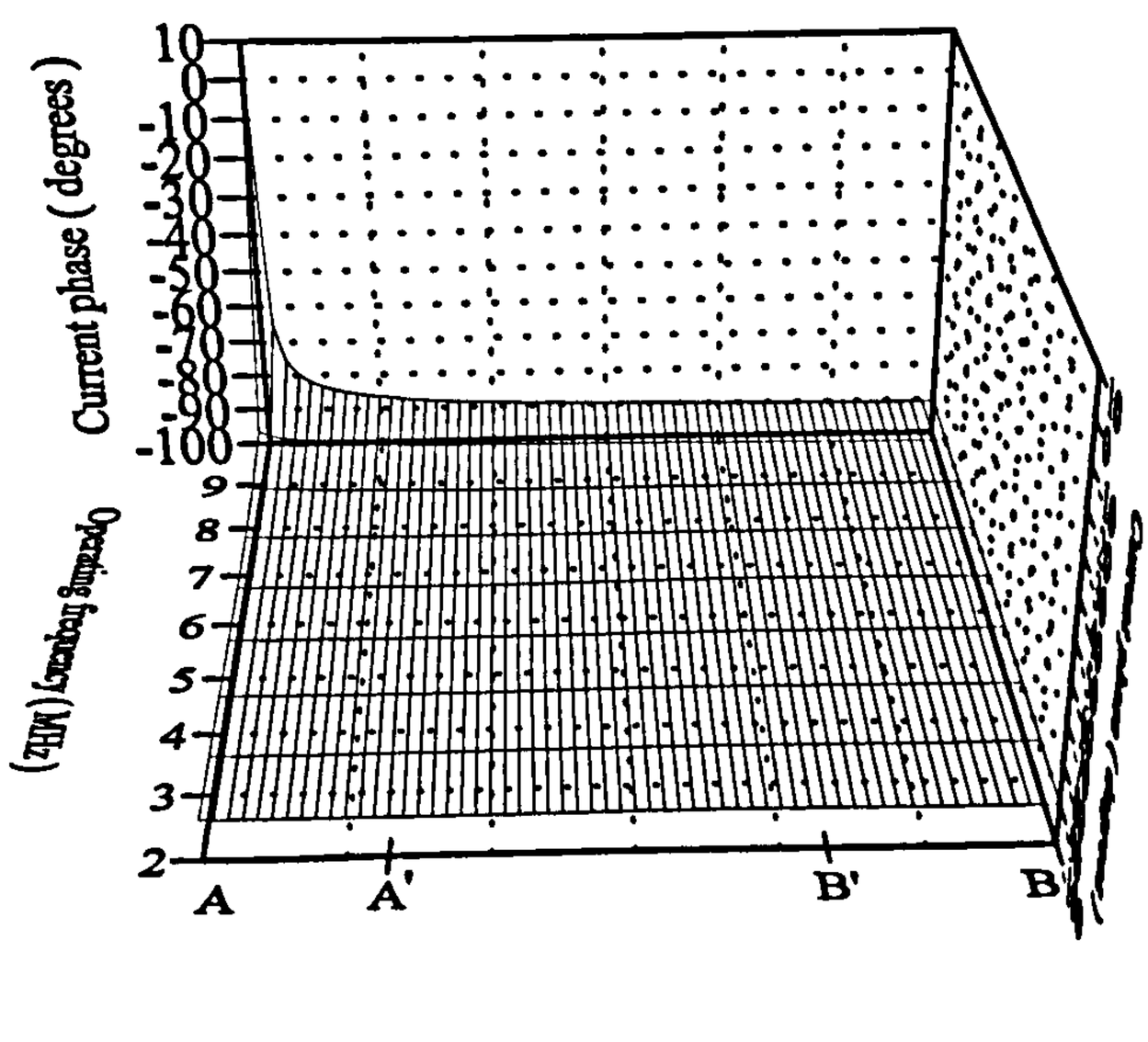


Figure 6.10 The variation of current phase on the loop driven by a single source as a function of position and frequency.

while 98.4% places it at the front base (point implementation the feeding source was assumed to be located at the middle of any B). The objective functions for the three cases were solely set as D_{45}, η and G_{45}, respectively. In addition, the other GA operations such as population, selection, crossover and mutation were all selected to be the same as those used in the previous case. The predicted results are listed in table 6.2.

Table 6.2 Source position for a maximum D_{45}, η, or G_{45}.

Freq. (MHz)	Source position (%) for Max. D_{45}	Source position (%) For Max. η	Source position (%) for Max. G_{45}
2	50	95.2	95.2
3	50	95.2	95.2
4	50	95.2	95.2
5	50	95.2	95.2
6	50	95.2	95.2
7	50	95.2	95.2
8	50	95.2	95.2
9	50	95.2	95.2
10	50	95.2	95.2

The result indicated that the maximum directivity (D_{45}) occurred when the source was fixed to the middle of the loop and this result was independent of the operating frequency. However, if the source was positioned slightly above the front base of the loop for all considered frequencies the optimal efficiency η and G_{45} occurred. These results were also almost independent of the operating frequency over the NVIS band. In terms of the electrical mechanism, these results indicate that the base ends and the

middle point of the loop are the critical positions to excite the effective radiation towards the zenith. This agrees with the predicted results of Austin and Murray (1998) using the CM method – a completely different approach.

However, for mechanical reason the source is usually put at the rear base and this results in hardly any change in G_{45} compared with that of positioning the source slightly above the front or rear base.

6.3.2.2 Dual-feed Loop Antenna

The GA results for the single-feed system show that although it has mechanical and electrical simplicity, the NVIS performance it provides is still low, even at the higher frequencies of the NVIS band. Since there are some mode-current maxima along the loop, as has been found by Murray (1993) using the analytic method of Characteristic Modes, it would seem feasible to add another source to improve the radiation performance for NVIS requirements. Therefore, in this section the GA will be used to examine the effect of the loop fed by an extra source on the NVIS performance compared with that fed at the rear base of the loop.

a. GA Approach

In the GA implementation of this investigation the configuration information including the loop height, the position of the front-base source and the phase of the two sources were all considered in the search space and coded into a string of 1s and 0s. The objective function was also set simply to the computed value of G_{45} of each designed system. This value will assess the fitness of each designed NVIS system.

Except for the phase of the two sources, the limitations to the other parameters such as the loop height, x, y and z were all the same as in the single-feed case. An

additional seven-bit string was chosen for each source phase and so the bit resolution and the relation between the binary code and the real design parameters are illustrated and listed in table 6.3.

Table 6.3 The relation between the binary code and the real design parameters.

	h (m)	x (m)	y (m)	θ_1	θ_2
range	0.1 ~ 1.5	0 ~ 3.88	0 ~ 1.566	0° ~360°	0° ~ 360°
bits	4	4	3	7	7
bit resolution	0.093	0.2587	0.2237	2.8°	2.8°

After a few trials with an initial population of 60 chromosomes, paring the survivors with the alternative selection and dual-point crossover, and then mutating the chromosomes at the rate of 0.5% of the total bits in each generation enabled the GA to reach the solution efficiently. The iterative process was monitored and was then suspended once the algorithm converged.

b. Results

The results showed that though two sources were used, the G_{45} reached maximum only when two situations are approached: i) the loop size is the largest in the defined range and ii) the two fed sources are in phase. Clearly, the former situation is the same as the result from the case in which only one source is fed. However, the most important aspect in all is that the two sources must be in-phase. Figure 6.11 shows the optimal G_{45} for the frequencies over the NVIS band. Comparing to the results of single-fed system, which was shown in figure 6.8, the total efficiency is significantly improved, especially at the frequencies above 4 MHz. As discussed before, the loop is

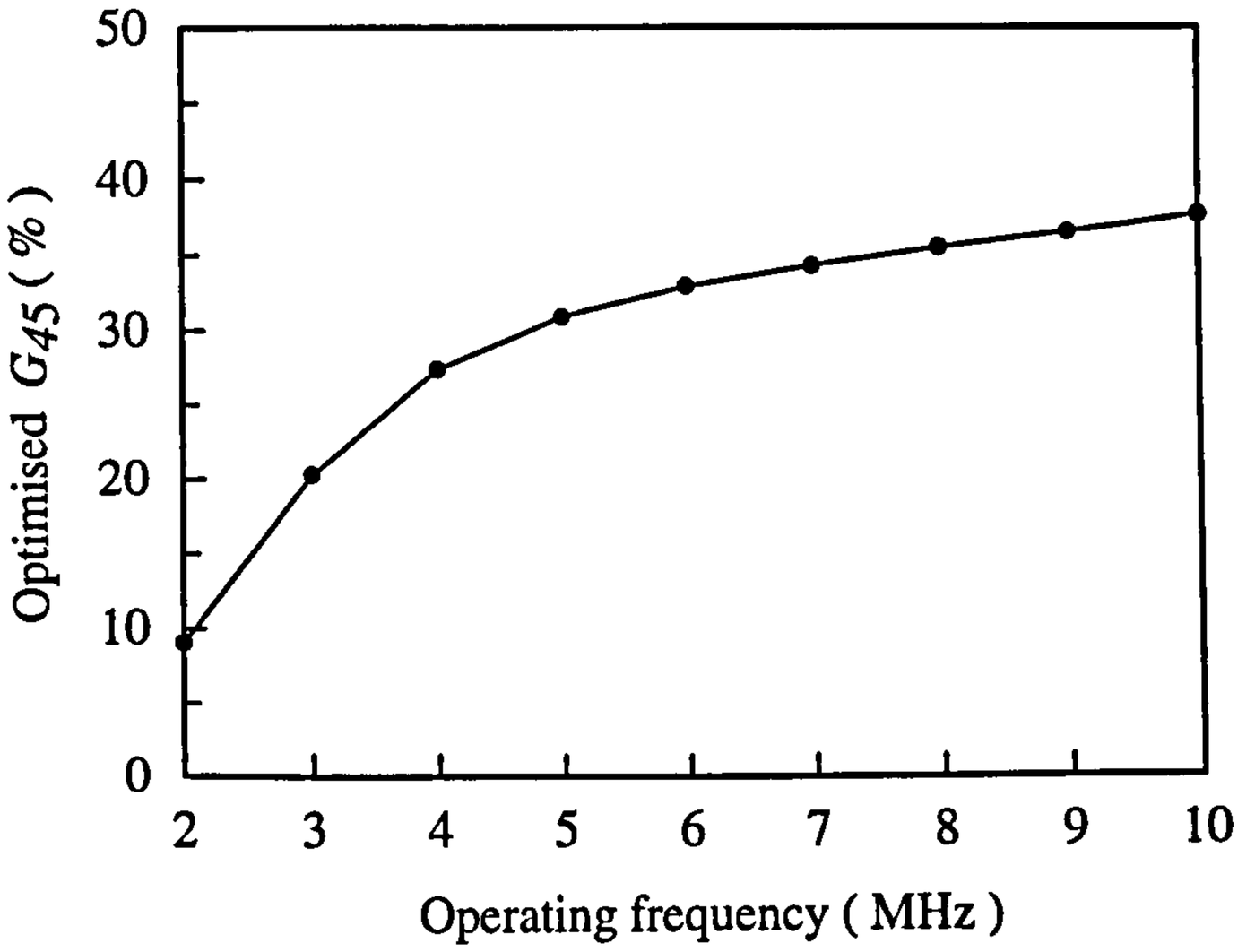


Figure 6.11 Variation of the optimised G_{45} against the operating frequency for a in-phase, dual-feed system.

no longer electrically small when the frequency exceeds this critical value and so the current distributed on the top horizontal section is no longer constant which results in a shift of radiated power away from the zenith. However, by adding another in-phase voltage source at the opposite end of the loop the radiated power can be significantly concentrated again towards the zenith and thereby obtain improved NVIS performance. Below 4 MHz, as expected, even by adding the extra in-phase source to the loop the current was still constant on such an electrically very short antenna and hence hardly any improvement was achieved from the dual-feed to single-feed system. These effects are clearly depicted in figure 6.12 which plots the normalised current distribution along the loop.

Figure 6.13 illustrates the variation of current phase on the dual-fed loop over the frequency range from 2 to 10 MHz. Whereas the phase of this system is almost

constant across the frequency range it is very unlike the single-fed system for which a natural resonance point occurs between 9 and 10 MHz and this drastically alters the current phase.

From these results it can be concluded that the added in-phase source can effectively change the current distribution to effectively concentrate the maximum value on the horizontal section of the loop. It does though not result in any change of the current phase, or the total efficiency, at frequencies below 4 MHz.

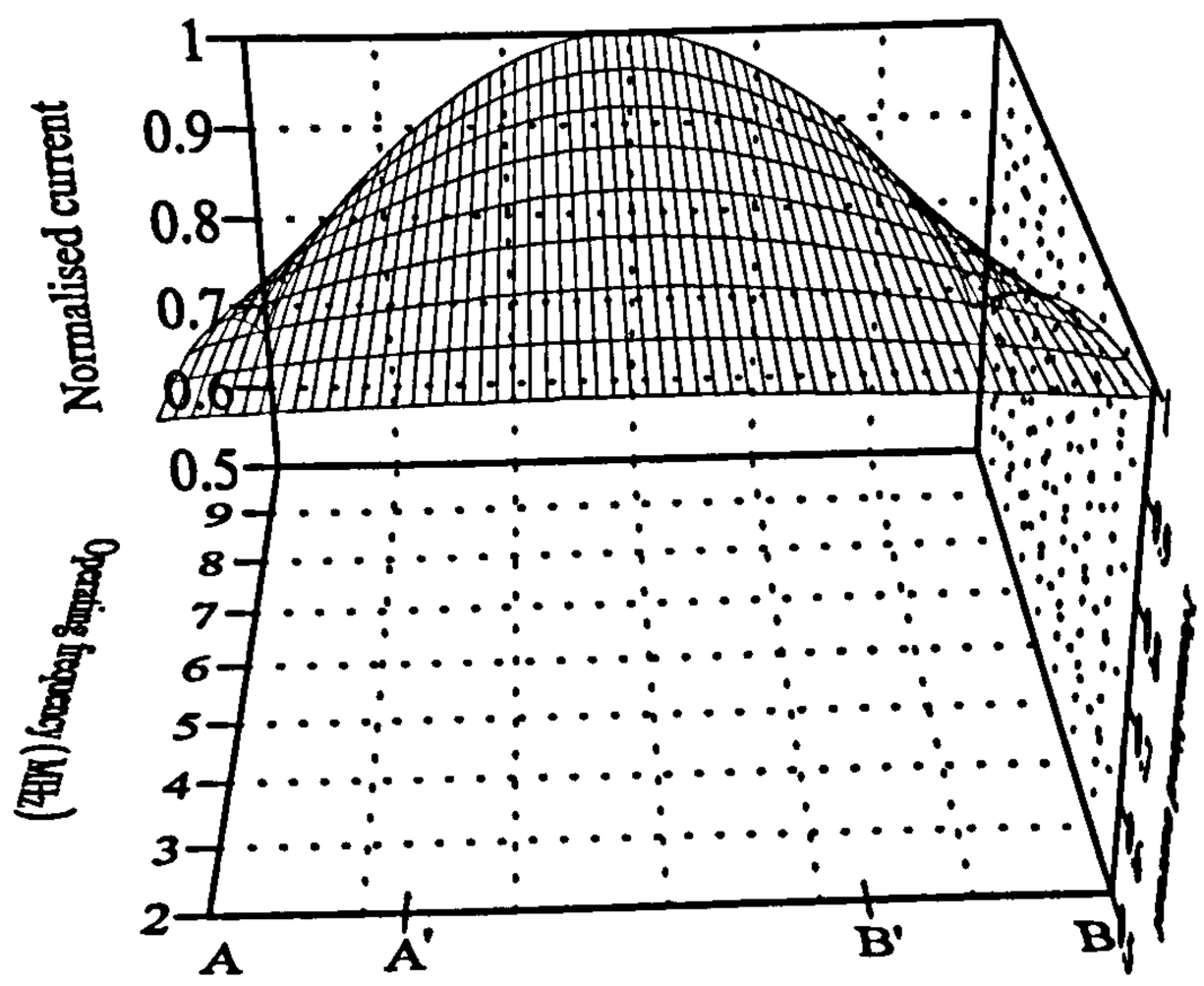


Figure 6.12 The variation of normalised current distribution on the loop driven by dual, in-phase sources as a function of position and frequency.

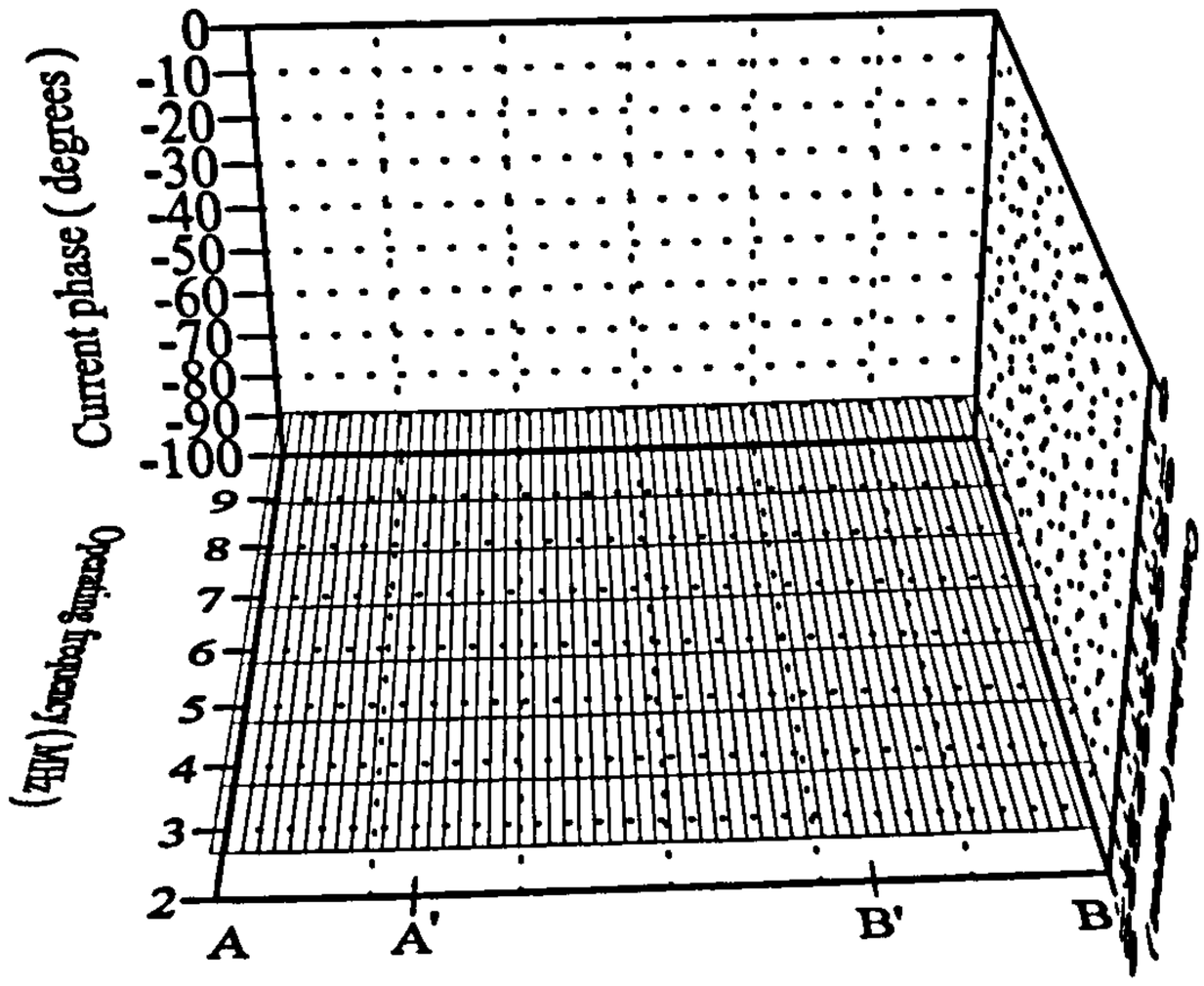


Figure 6.13 The variation of current phase on the loop driven by dual, in-phase sources as a function of position and frequency.

6.4 Electrically Loaded Loop Antenna Design for NVIS

The previous sections have optimised the NVIS performance of a loop antenna with single-feed and dual-feed sources using the combined technique of GA/NEC2. In both, and regardless of the feeding system, the GA/NEC2 technique can successfully predict the appropriate antenna configuration to produce the optimal NVIS radiation characteristics. The optimal candidate configuration is the loop antenna fed by dual, in-phase voltage sources. However, in practice it is both cumbersome and costly to implement a network for in-phase feeding the two sources which are physically well separated on the host vehicle. These factors constitute the obvious disadvantages of the dual-fed loop in practice.

An alternative method of controlling the current distribution is then required and some method using just passive techniques would clearly be ideal. A single-fed loop loaded with a passive, lumped circuit was therefore selected as a candidate to further improve the total efficiency and simplify the practical mechanism as well. In the past, considerable work was done on changing the electromagnetic characteristics of an electrically short dipole or monopole by loading it with a lumped circuit (Nyquist et al. 1968, Hansen 1975 and Austin 1988). The use of a lumped circuit is mainly to change the current distribution along the antenna such that the current on the antenna can be maximised once the value and position of the load are properly chosen. The load and its position can essentially increase the radiation efficiency by controlling current distribution. This concept is therefore applied to the single-fed loop system because if the current distribution of the loop could be concentrated along the top horizontal section by loading it with an appropriate lumped circuit the radiated power could be effectively radiated toward the zenith and hence yield better NVIS performance than could be achieved by using only a single feeding source. If this could be achieved it will significantly benefit the practical implementation of the antenna system.

6.4.1 Simplified Loaded Loop Antenna

The geometrical complexity of any vehicle or similar structure involves significant computational time when simulating the current distribution of the vehicle-antenna system using the Method of Moments. To save computing time and then to quickly survey the effect of the load on the NVIS performance it is vital to calculate the characteristics of this antenna-vehicle system from a simpler structure which can approximately represent the complex configuration of the original system. Since the

vehicle body itself could be deemed as part of this closed loop the system can therefore be approximated by a simple loop approximation of real configuration. To make the simplified model as close as possible to the original system, the position of the source and the distance of the loop from the ground are both set as the corresponding values of those in the real configuration. To be equivalent this simplification must not seriously change the loop's basic performance when compared to the real configuration. A contrast illustration is shown in figure 6.14.

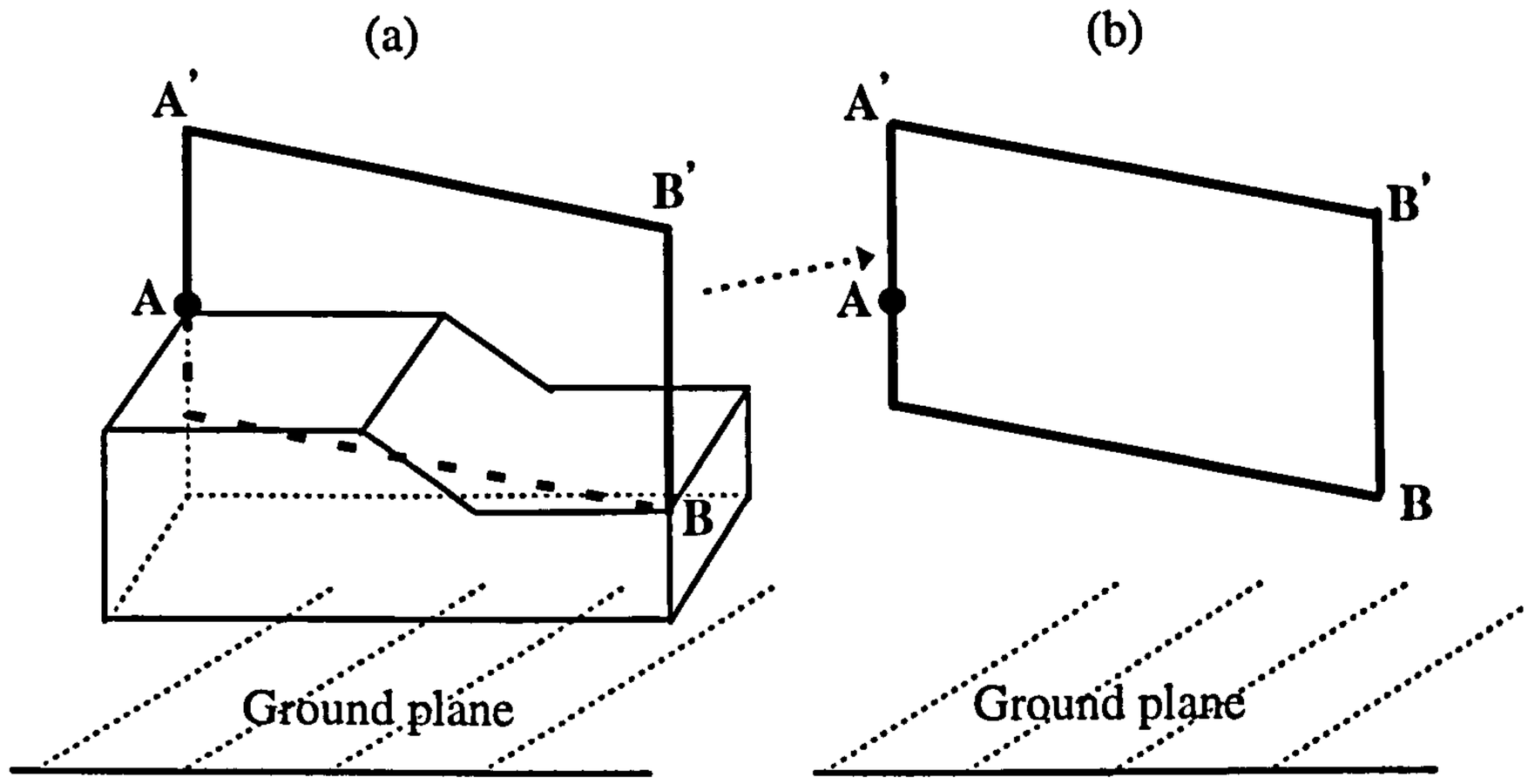


Figure 6.14 (a) The real configuration.
(b) A simple loop approximation.

The cases that were examined are shown in figure 6.15. Three different operating frequencies, 2, 6 and 10 MHz, were considered for each case. For the purpose of comparison the performance of the two typical cases – the loops fed by single and dual, in-phase sources – was evaluated and numbered as case 1 and case 2, respectively.

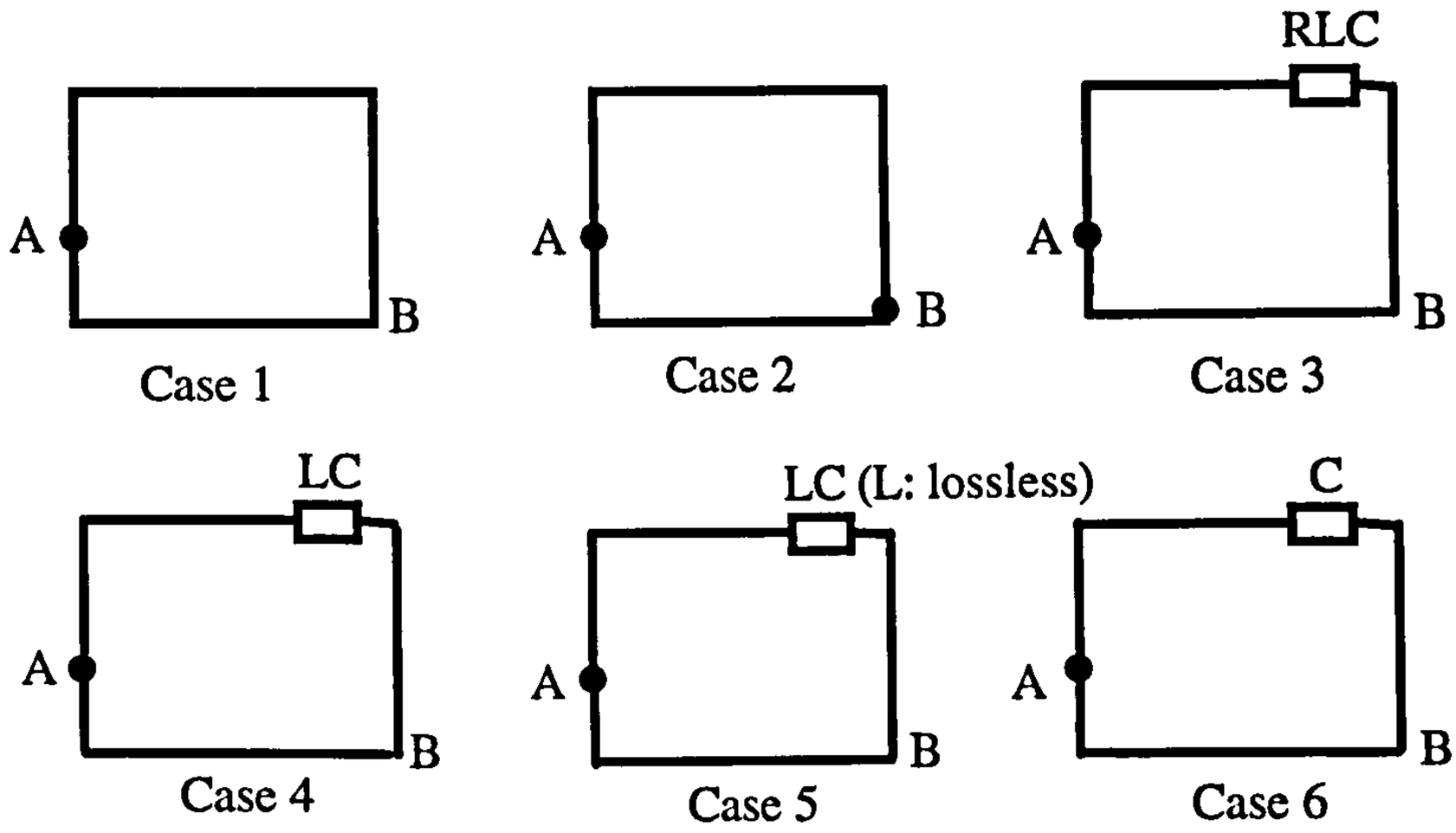


Fig.6.15 The simple loop with different sources and loads.

The lumped circuit was developed iteratively. First, consider the frequency of 10 MHz. In considering the candidate lumped circuits the quality factor, Q, of an inductor is important because, generally, it is much lower than that of a capacitor. It therefore always increases the loss resistance when it is loaded in series with the antenna. In order to avoid adding too much loss resistance to the loop and thereby reducing the radiation efficiency, this investigation commenced with case 3 by loading the loop with a parallel RLC circuit of typical Q. Since the value of Q for a capacitor is very high, it was neglected while an appropriate value of 100 was chosen for the inductor. The parallel loss resistance arising from the selected inductor was therefore considered to be equal to $Z_D=2\pi fLQ$, the so-called Dynamic Impedance of the parallel resonant network, which varies with both the operating frequency and inductor value. Therefore, the total resistance of this load is equal to the inductive loss Z_D in parallel with the loaded loss R. To find an optimum solution, the load was not restricted to a

fixed location. It could be at any position from point A to point B (see figure 6.14) along the upper path of the loop. The GA is an ideal technique for finding optimum values and position of this network. Thus the capacitor, the inductor and the resistor, as well as the position of the lumped circuit, were set as the design genes and then encoded into a chromosome. We selected 13 bits to encode each of the three component values and six bits for the position of the load. In view of practical design considerations, at each operating frequency the capacitance, inductance and resistance (only used for case 3) values, after a few trials, were allowed to vary in between 1 and 8192 pF, 1 and 8192 μH, and 5 and 40960 Ω, respectively. In addition, a population of 60 chromosomes and the schemes of alternative selection, dual-crossover and the mutation rate of 1% of the total bits in each generation were used to compose the GA procedure for this design.

Table 6.4 which should be read in conjunction with figure 6.16 shows the GA results. To interpret it consider case 3 at 10 MHz: the optimised G_{45} is 31.5%. Compared to case 1 and case 2 it is better than the former, 27.6%, while less than the latter, 39.3%. As to the load, it was found to be positioned at the front base, point B, and the inductor value was as large as 7.776 mH and R was 28.2 kΩ. The two lossy components produced a large loss resistance such that the radiation efficiency was reduced significantly. In addition, a capacitor value of 159 pF was also obtained by the GA. The resonant frequency of this RLC parallel circuit is equal to 0.14 MHz. It is below the operating frequency of 10 MHz in this case and clearly indicates that this load acted like a capacitive circuit. To increase the radiation efficiency in case 4 the load was simplified to be composed of only L and C while loss within the inductor.

Table 6.4 The NVIS performance of a simple loop with different sources or loads.

	2 MHz						6 MHz						10 MHz					
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
D_{45} (%)	36.7	31.8	36.9	37.0	37.1	37.4	32.7	36.6	33.3	38.0	38.0	38.0	27.8	39.9	33.7	41.5	41.5	41.5
η (%)	19.7	22.0	18.5	18.5	19.2	19.4	93.2	92.4	92.4	91.8	92.2	92.2	99.1	98.7	93.2	98.6	98.7	98.7
G_{45} (%)	7.2	7.0	6.8	6.8	7.1	7.2	30.5	33.8	30.8	34.9	35.0	35.0	27.6	39.3	31.5	40.9	40.9	40.9
R (kΩ)			38.4						28.4						28.2			
L (μH)			7968	4063	11				7075	4032	26				7776	2493	116	
C (pF)			8171	4583	1018	1000			2149	115	137	110			159	33	35	32
position of load	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B

The GA results clearly illustrated the total efficiency, G_{45}, increased from 31.5% to 40.9%. The resonant frequency was still below the operating frequency and the resulting dynamic impedance was as high as in an open circuit. Therefore, to further examine the effect of the LC load without loss to the loop, in case 5 the inductor was assumed to be lossless (i.e. without considering Z_D). The GA predicted almost the same NVIS performance as case 4 while the inductance was markedly reduced to only 116 μH. The load still acted like a capacitive circuit. From these results this load can be boldly assumed to be capacitive and this case was then examined explicitly in case 6. This assumption was proved conclusively from the results shown in table 6.4. Loading the loop with only a capacitor of 32 pF can achieve NVIS performance as good as the loop fed by dual, in-phase sources while much better than that fed by a single source. Another very important aspect that should be noted is the position of the load for all these cases was found to be point B of the loop.

The results predicted by the GA approach for the similar cases at 6 and 2 MHz are also listed in table 6.4. They also showed similar phenomena to the cases at 10 MHz. The three NVIS parameters were all marginally improved compared to those obtained from using dual sources alone when the loop was loaded with an appropriate capacitor at the different operating frequencies. This result has profound practical significance.

This conclusion was then applied to the examples at the intermediate frequencies between 2 and 10 MHz. The optimised NVIS performance and the appropriate capacitor value predicted by the GA are summarised and listed in table 6.5.

Table 6.5 GA predicted results of the simplified loop antenna with a capacitive load.

Freq. (MHz)	D_{45} (%)	η (%)	G_{45} (%)	Capacitor(pF)
2	37.4	19.4	7.2	1000
3	36.8	50.5	18.6	403
4	37.1	73.7	27.3	253
5	37.5	86.0	32.3	162
6	38.0	92.2	35.0	110
7	38.7	95.3	36.9	77
8	39.5	97.1	38.3	57
9	40.4	98.1	39.6	42
10	41.5	98.7	40.9	32

6.4.2 Capacitively Loaded Vehicular Loop Antenna

The previous section has examined the improvement of the NVIS performance obtained by electrically loading the simplified antenna-vehicle system. It can be concluded that an appropriate capacitive load used at the junction between the loop and the vehicle body can effectively increase the radiated power towards the zenith when compared to the unloaded, single-source loop. Thus, in this section, the NVIS characteristics of the real antenna-vehicle system with the capacitive load will be examined using the GA.

As in the previous cases the position and value of the load were left for determination by using a standard GA procedure. To avoid the GA becoming

unmanageable, the design parameters should be, firstly, represented by judiciously chosen binary-strings (genes) to satisfy the criterion for the various operations. Therefore, initially six and 16 bits were used for the position and value of the load, respectively. A change in the number of bits will be made to modify the resolution if the results are not acceptable. The limitation due to bit size of the load's position and the value of the load are illustrated in table 6.6. Although there are 2^{22} possible solutions, there are only two parameters to be determined. This enables the algorithm to work effectively without the necessity of using many chromosomes in each generation and so 60 chromosomes was arbitrarily selected as the initial population and then the same number was kept in each generation during the iterative process.

Table 6.6 The binary code and the corresponding real design parameters.

	Position (m)	capacitor (nF)
Range	0.0 ~ 7.7	0 ~ 3.28
Bits	6	16

For each frequency between 2 to 10 MHz, after a sufficient number of iterations, the algorithm converged and produced a chromosome in which the position and value of the load for an antenna configuration with optimal NVIS performance were included. As in all the cases of the simplified loop, the capacitive load is positioned at point B, the position previously was occupied by the second in-phase source. The GA also determined the value of this capacitor at all frequencies within the NVIS range to ensure maximum zenithal radiation. Figure 6.16 now shows the results obtained for all

the three cases: (i) an unloaded loop driven by a single source at A (the conventional approach); (ii) dual, in-phase sources at A and B (the optimum CM solution), and (iii) a single source at A plus a series capacitor at position B (the new approach). It will be noted that at the frequencies below approximately 4 MHz the loop perimeters are all less than $\lambda/4$, a general criterion of the electrically small antenna (Virani 1988), and become electrically smaller. This is the factor limiting the NVIS performance of the loop operated at the low-frequency end of the NVIS spectrum. However, both the dual-source loop and the capacitively-loaded loop produce markedly improved performance (better than 3.4 dB at 10 MHz, in fact) above that frequency, with there being no essential difference between them. This result is most significant because it shows that a single source at one end of such a rectangular loop, and an appropriate capacitor at the other, will produce the optimum current distribution in the loop for NVIS applications. This can be seen from the plot, figure 6.17, of the current distribution on the loop. The phase change around the loop is shown in figure 6.18 as well. Compared to the case of the loop fed by dual, in-phase sources, the current and phase distributions of the both cases are almost identical and so produce similar radiation performance. Finally, figure 6.20 gives the capacitor values calculated for the loop of the dimensions indicated in figure 6.19 between 2 and 10 MHz.

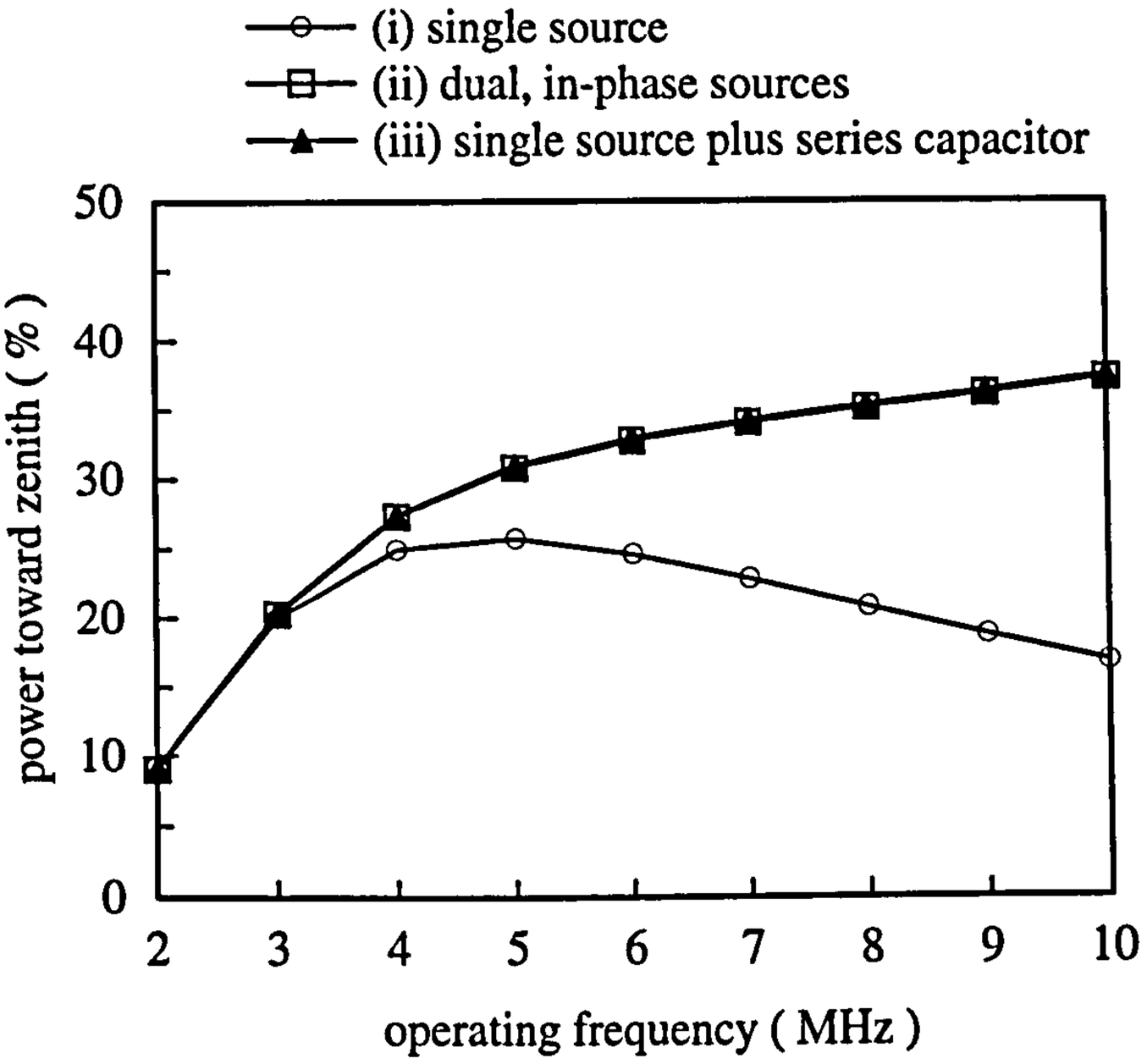


Figure 6.16 Total efficiency of the loaded and unloaded loop antennas.

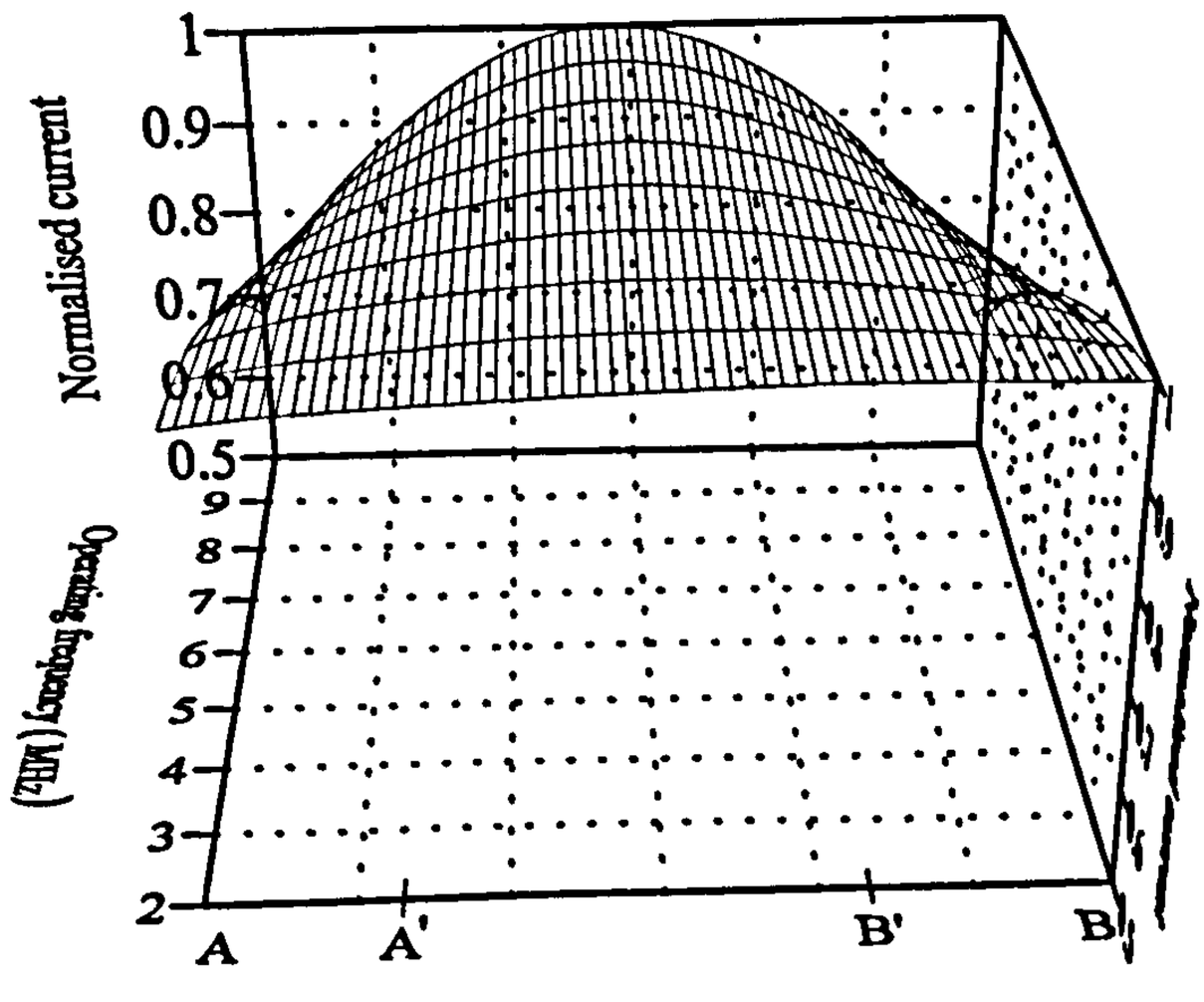


Figure 6.17 The variation of normalised current distribution on the loop driven by a single source plus a series capacitor as a function of position and frequency.

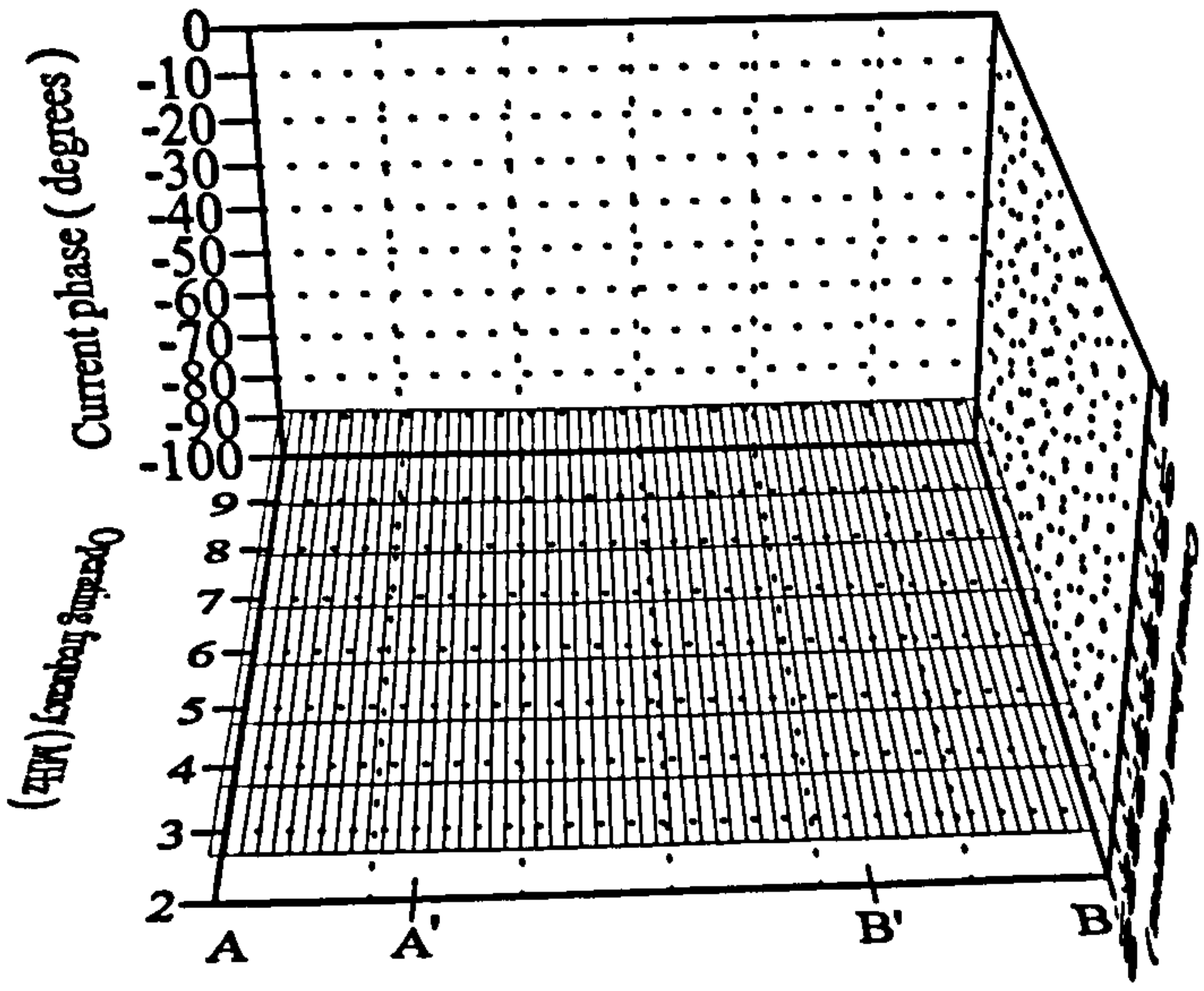


Figure 6.18 The variation of current phase on the loop driven by a single source plus a series capacitor as a function of position and frequency.

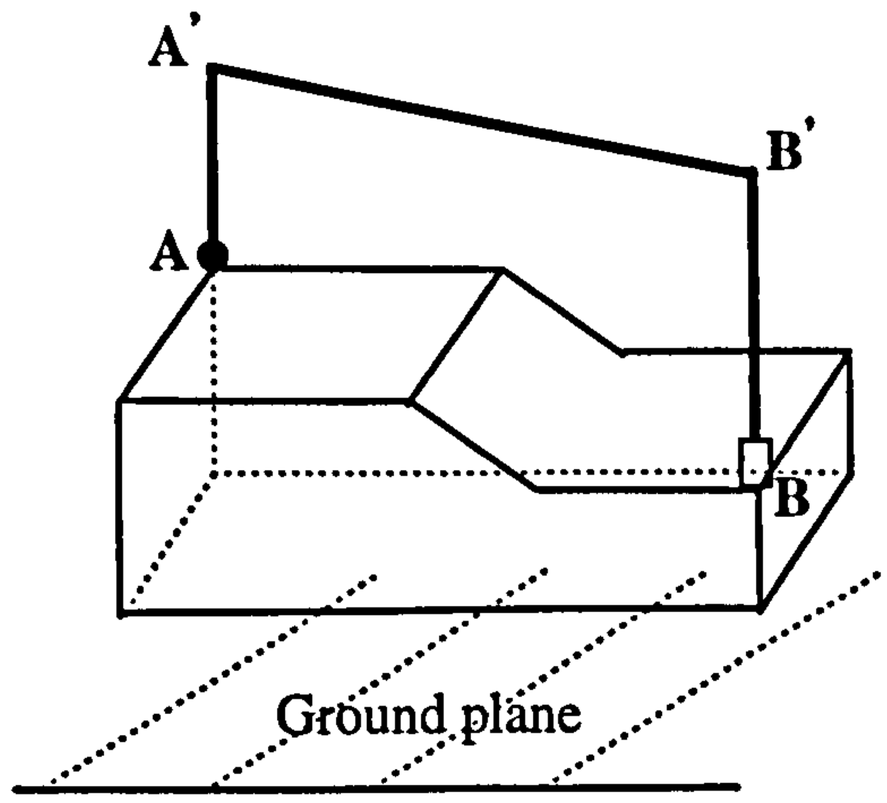


Figure 6.19 Electrically loaded loop antenna.

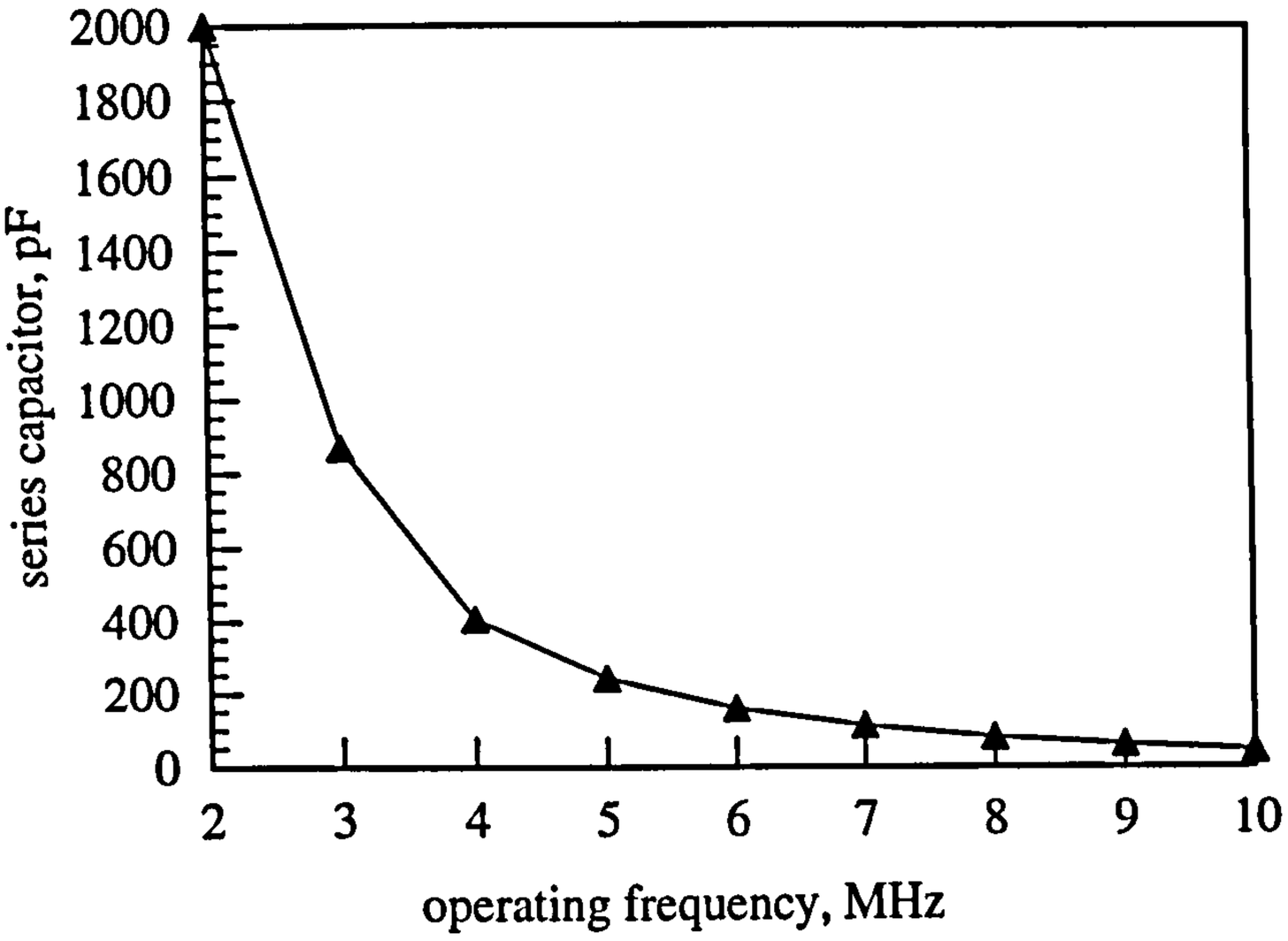


Figure 6.20 Capacitor values for optimum radiation toward the zenith.

6.5 Matching Network Design for NVIS Antennas

In antenna investigation, the very important aspect is that in the overall design of a radio communications system, the antenna is but one link in the complicated chain that accomplish the transfer of electromagnetic energy from the input point to the output point, or vice versa. It is natural to consider the antenna simply as another circuit element that must be properly matched to the rest of the network for efficient power transfer. The input impedance of the antenna is therefore a parameter related to the transfer efficiency because it describes the antenna as a circuit element and can directly be deemed as a lumped circuit. So, the operation of an antenna system over a frequency range is not completely dependent upon the frequency response of the antenna element alone but rather on the frequency characteristics of the combination

of both the transmission line and antenna element. In practice, the characteristic impedance of the transmission line is usually real and typically 50 Ω whereas that of the antenna element is complex. Generally, an antenna is equivalent to a series circuit consisting of a resistance and a reactance. It is thus of primary importance in determining a suitable matching circuit to enable a high efficiency of power transfer from the source to the antenna. Also the variation of each as a function of frequency is not the same. Thus efficient matching networks must be designed which attempt to match the characteristics of the two elements over the desired frequency range.

In this section, the matching networks for the three cases : (i) an unloaded loop driven by a single source; (ii) dual, in-phase sources at A and B, and (iii) a single source at A plus a series capacitor at position B were designed by both the theoretical analysis and the GA approach. A summary of the input impedance, Zin, of the above antennas over the NVIS band is shown in table 6.7. As discussed before, these antennas that are electrically short exhibit extra low radiation resistance R_{in} and high reactance X_{in}. It should be emphasised that the input reactance are all inductive except for that in the single-source where a change of sign on the reactance occurs at the frequency between 9 and 10 MHz. This clearly indicates that the first natural resonance of this loop occurred within that range as has been discussed in section 6.3.2.1. The predominantly inductive reactance here, fortunately, offered an extra simplicity and some advantages to the selection of configurations of the matching network. This is because the main desired function of the network is to tune out the reactive term and to present a suitable resistance to the transmission line which has usually only the real part to its characteristic impedance. Therefore a network

composed of just capacitors rather than inductors or inductors and capacitors is ideal for ensuring maximum system radiation efficiency.

Most of the antenna impedance matching problems encountered in practice can easily be handled with an L-type matching network, the so-called two-element matching network. The L-network was first described by Smith (1942). In his report eight matching networks that used only two pure reactances were examined with antennas with different input impedance. Two possible networks were selected as the candidates for the three antenna systems discussed in this work. They are both composed by two capacitors but connected to the antenna in different ways. The two transfer systems including the transmission lines, matching networks and antennas are shown in figure 6.21(a) and (b).

Table 6.7 The input impedance of the three optimised NVIS vehicular antennas.

$Z_{in}(\Omega)$	case (i)		case (ii)				case (iii)	
Freq.(MHz)	R_{in}	X_{in}	R_{in1}	X_{in1}	R_{in2}	X_{in2}	R_{in}	X_{in}
2	0.038	102.6	0.018	50.3	0.018	50.3	0.036	60.8
3	0.099	161.0	0.041	76.4	0.041	76.3	0.084	91.8
4	0.292	230.0	0.101	103.8	0.100	103.6	0.203	109.6
5	0.089	317.4	0.230	133.0	0.227	132.5	0.458	135.1
6	2.593	438.7	0.479	164.6	0.474	163.8	0.949	164.7
7	8.217	628.9	0.927	199.7	0.901	198.2	1.817	196.6
8	31.233	993.6	1.700	239.2	1.630	236.7	3.292	233.3
9	202.583	2051.9	3.066	284.8	2.833	280.7	5.735	275.6
10	13837.700	-1473.7	5.207	339.0	4.794	332.1	9.739	325.2

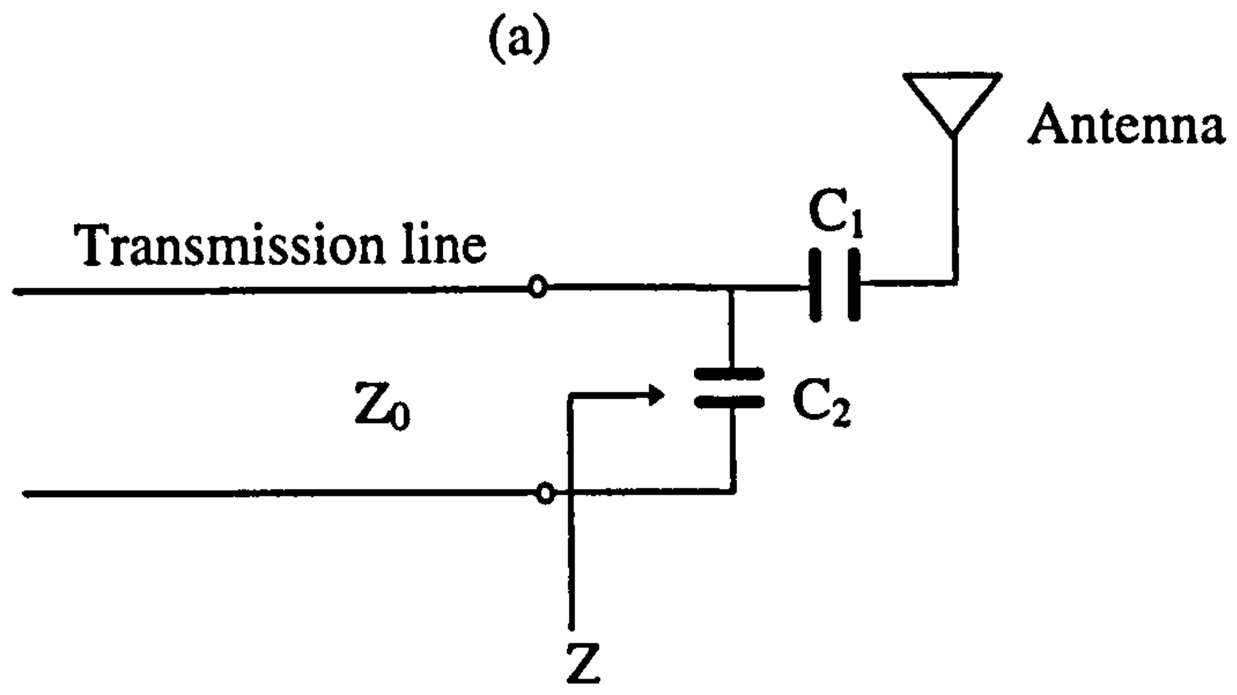


Figure 6.21 Two antenna systems including the transmission line and matching network.

6.5.1 Theoretical Design

In network (a), if the total impedance of Z, the impedance of the network, in series with the antenna is equal to Z_0, the characteristic impedance of the two-wire transmission line, the system is matched and thus the highest efficiency of power transfer from the transmission line to the antenna is achieved. The mathematical expression of the relation between these impedance can be represented as,

$$Z_0 = Z \quad \text{or} \tag{6.11}$$

$$\frac{1}{Z_0} = Y_2 + \frac{1}{Z_1 + (R_{in} + jX_{in})} \tag{6.12}$$

Where, Z_1 and Y_2 are the impedance and admittance of C_1 and C_2, respectively. They are

$$Z_1 = \frac{1}{j2\pi fC_1} = \frac{1}{Y_1} \tag{6.13}$$

$$Y_2 = j2\pi fC_2 = \frac{1}{Z_2} \tag{6.14}$$

The relations between these parameters can be derived as follows after equating real and imaginary parts,

$$Z_0 = \frac{R_{in}^2 + (X_{in} - \frac{1}{2\pi f C_1})^2}{R_{in}} \tag{6.15}$$

$$\frac{1}{2\pi f C_2} = \frac{R_{in}^2 + (X_{in} - \frac{1}{2\pi f C_1})^2}{(X_{in} - \frac{1}{2\pi f C_1})} \tag{6.16}$$

The respective C_1 and C_2 over the NVIS band can be obtained from equations 6.15 and 6.16 once the input impedance of the antenna and the characteristic impedance of the transmission line are known. In this work, a transmission line with characteristic impedance of 50 Ω was used.

Similarly, for network (b) the matching situation occurred when the following equation exists.

$$Z_0 = Z_2 + \frac{1}{Y_1 + Y_{in}} \tag{6.17}$$

Where

$$Y_{in} = \frac{1}{Z_{in}} = \frac{1}{R_{in} + jX_{in}} \tag{6.18}$$

Again, solving yields

$$Z_0 = \frac{R_{in}(\frac{1}{2\pi f C_1})^2}{R_{in}^2 + (X_{in} - \frac{1}{2\pi f C_1})^2} \tag{6.19}$$

$$\frac{1}{2\pi f C_2} = -\frac{\frac{1}{2\pi f C_1}(R_{in}^2 + X_{in}^2 - \frac{X_{in}}{2\pi f C_1})}{R_{in}^2 + (X_{in} - \frac{1}{2\pi f C_1})^2} \tag{6.20}$$

Thus C_1 and C_2, of the network (b) can also be obtained from equations 6.19 and 20. The values of C_1 and C_2 for the two matching networks over the NVIS band are presented in figure 6.22-6.24.

6.5.2 GA Design

In chapter 2 we discussed the work of Boag et al. (1996), Weile et al. (1996, 1997) and Altman et al. (1997) in successfully using the GA to design matching networks for different electrically loaded antennas. Owing to the powerful features of a GA, this approach was used again in this work for designing the matching networks suitable for the developed NVIS vehicular antenna systems. Although the GA has the capability to deal with the problem without any *a priori* knowledge of candidate configurations a pre-selected possible configuration can always enable it to approach the solution successfully and efficiently. Therefore the networks shown in figure 6.21 were selected as the objective circuits. To compose a chromosome each capacitor

value was encoded by a string of 20 bits with a bit resolution of 0.1 pF. In addition, a population of 60 chromosomes and the schemes of alternative selection, dual-crossover and 0.2% mutation rate were used to compose the GA procedure for this design. These GA criteria were all determined from a few trials. Because the objective of the designed networks was to perfectly link the antenna and the transmission line, the objective function was therefore defined to evaluate the value of voltage standing wave ratio (VSWR) of each designed circuit. A target value of unity was selected.

This GA optimisation procedure was firstly used for case (iii). Both networks (a) and (b) were designed by GA and also by the theoretical derivation. Figures 6.22 and 6.23 show the predicted capacitor values against the frequency for the optimal matching networks (a) and (b), respectively. The VSWRs obtained all tended to unity. Clearly, the GA can correctly produce the circuit without the elaborate computation from the theoretical design. In this particular applications network (b) is better than (a) in the aspect of implementation because the variation of C_2 in it is very small over the considered frequency range, except for operating at 2 MHz. For network (b) if 26 pF is selected for C_2 the value of C_1 can easily be determined by a GA for an optimal VSWR. Figure 6.24 shows the resulting values and VSWR from 2 to 10 MHz. The VSWR all less than 2 over the NVIS band which indicates that the appropriate network provides good matching between the antenna and the transmission line.

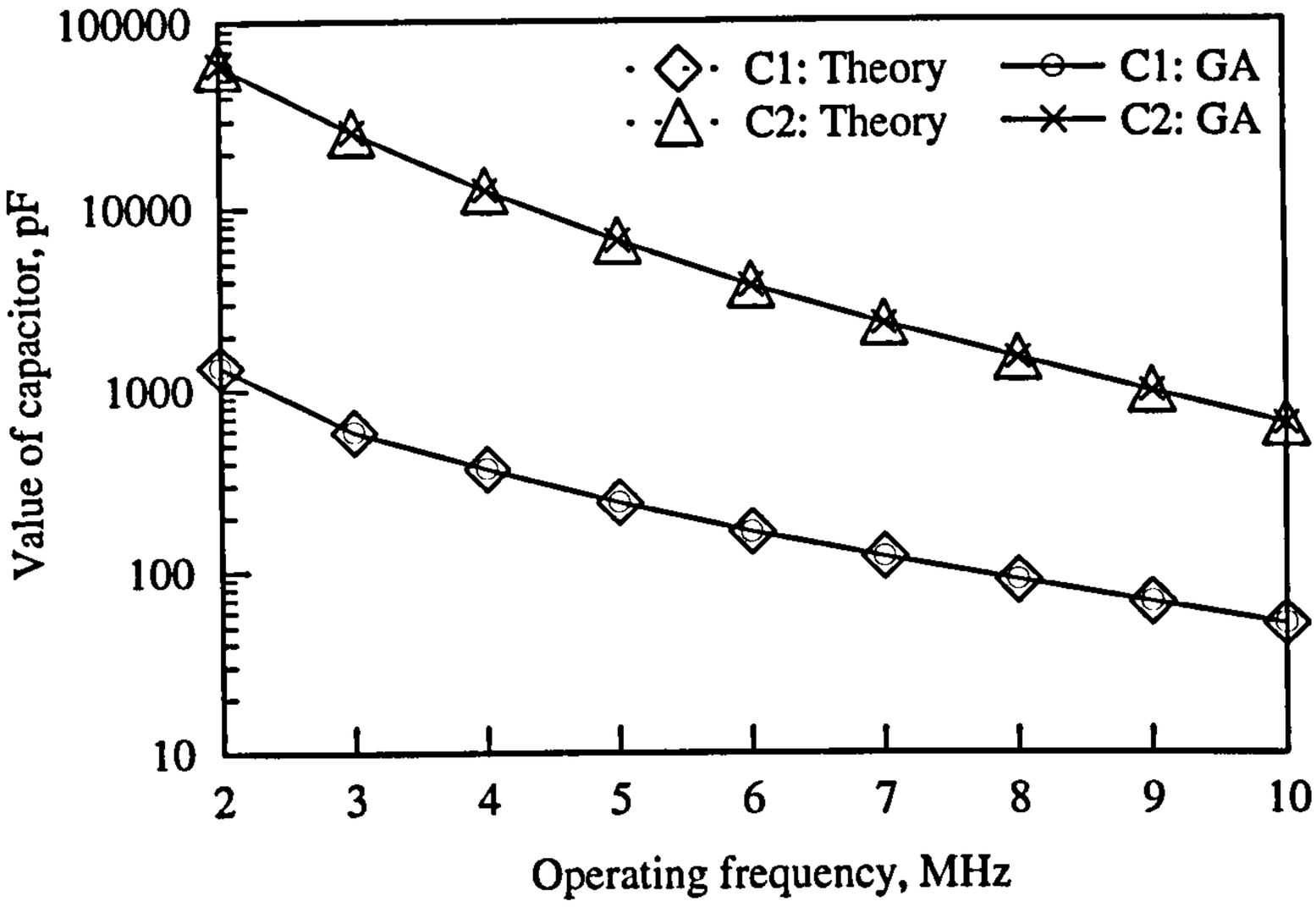


Figure 6.22 Values of capacitors for the matching network (a) of the loop driven by a single source series with a load.

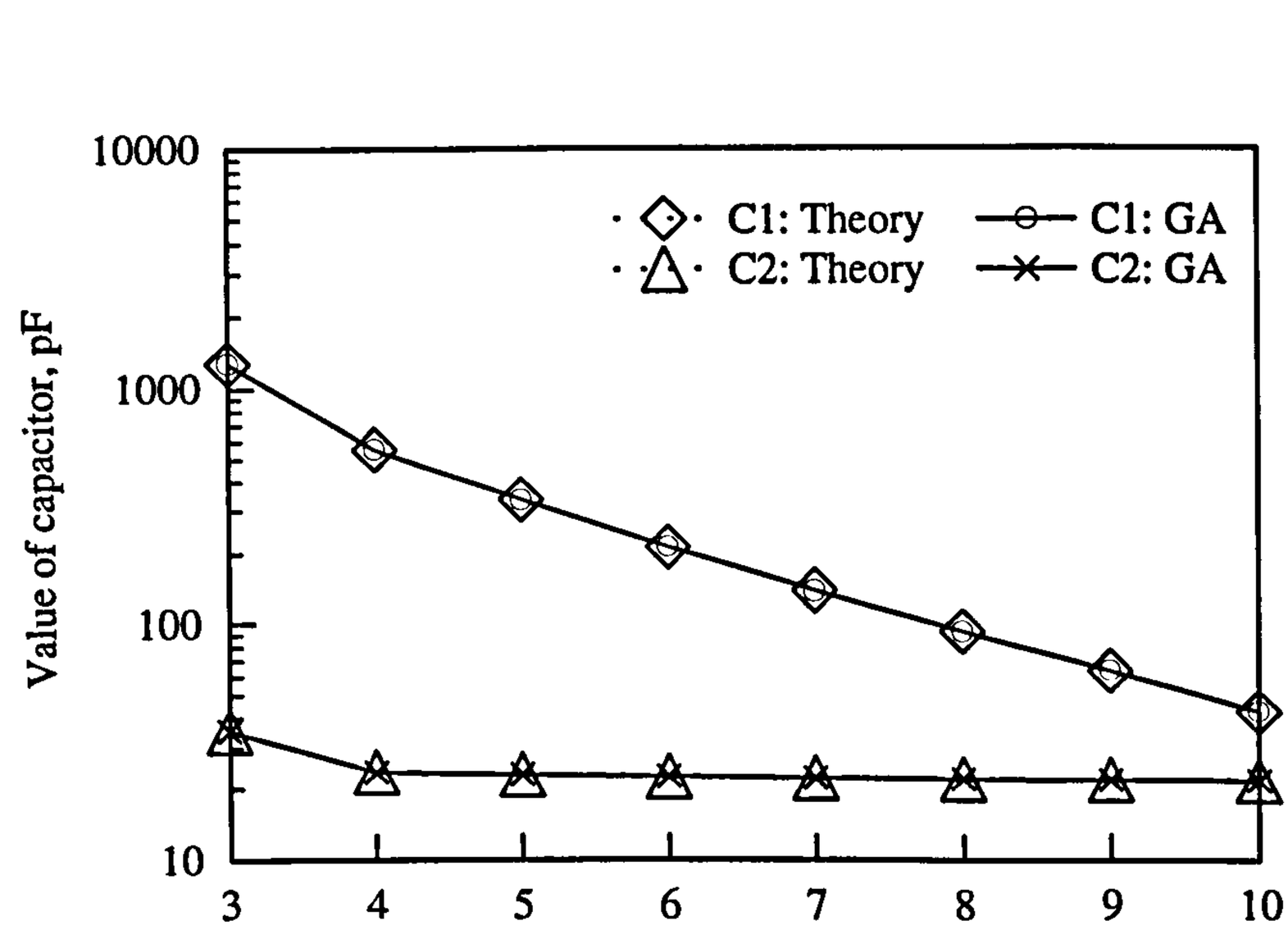


Figure 6.23 Values of capacitors for the matching network (b) of the loop driven by a single source series with a load.

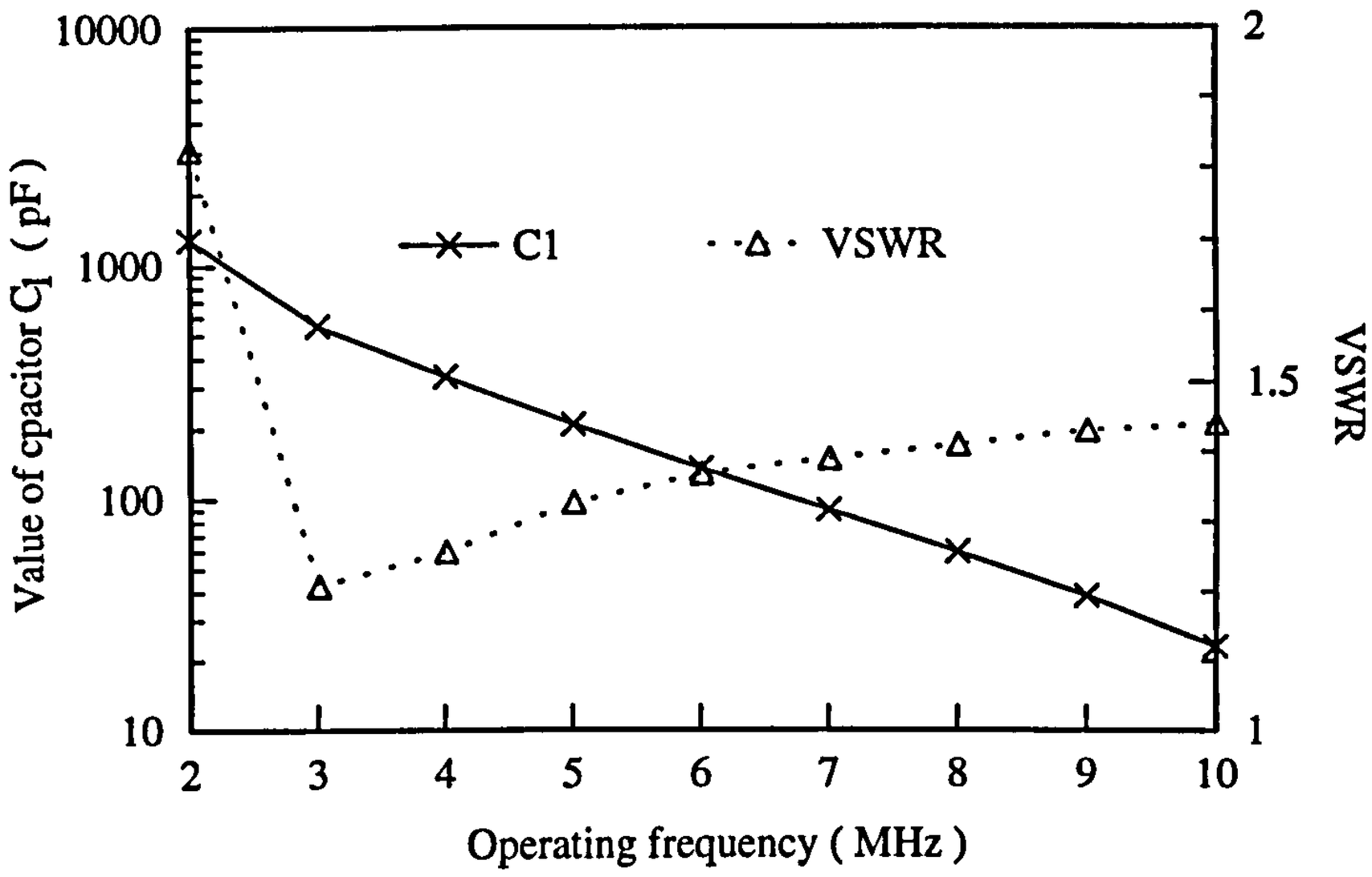


Figure 6.24 GA predicted optimal VSWR and the corresponding value of capacitor C_1 for the matching network (b) used for case (iii) when C_2 was fixed and equal to 26 pF.

6.5.3 Bandwidth

Another very important aspect in communication is the operating bandwidth. A typical acceptable bandwidth is defined to be the frequency difference over a range in which the VSWR is less than 2 (i.e. bandwidth for VSWR of 2:1). Following this definition the associated bandwidth for this transfer system was evaluated and plotted in figure 6.25. Undoubtedly, the overall bandwidth is narrow and this is because the low radiation resistance of such an electrically small antenna always produces an extra high system quality factor (Q) and thus resulted in a particularly narrow bandwidth. As shown in the plot, this is particularly true at the lower frequencies of this band.

However the bandwidth increases with the increase in operating frequency. The maximum bandwidth of this system is about 220 kHz at 10 MHz, while the minimum is 0.3 kHz at 2 MHz.

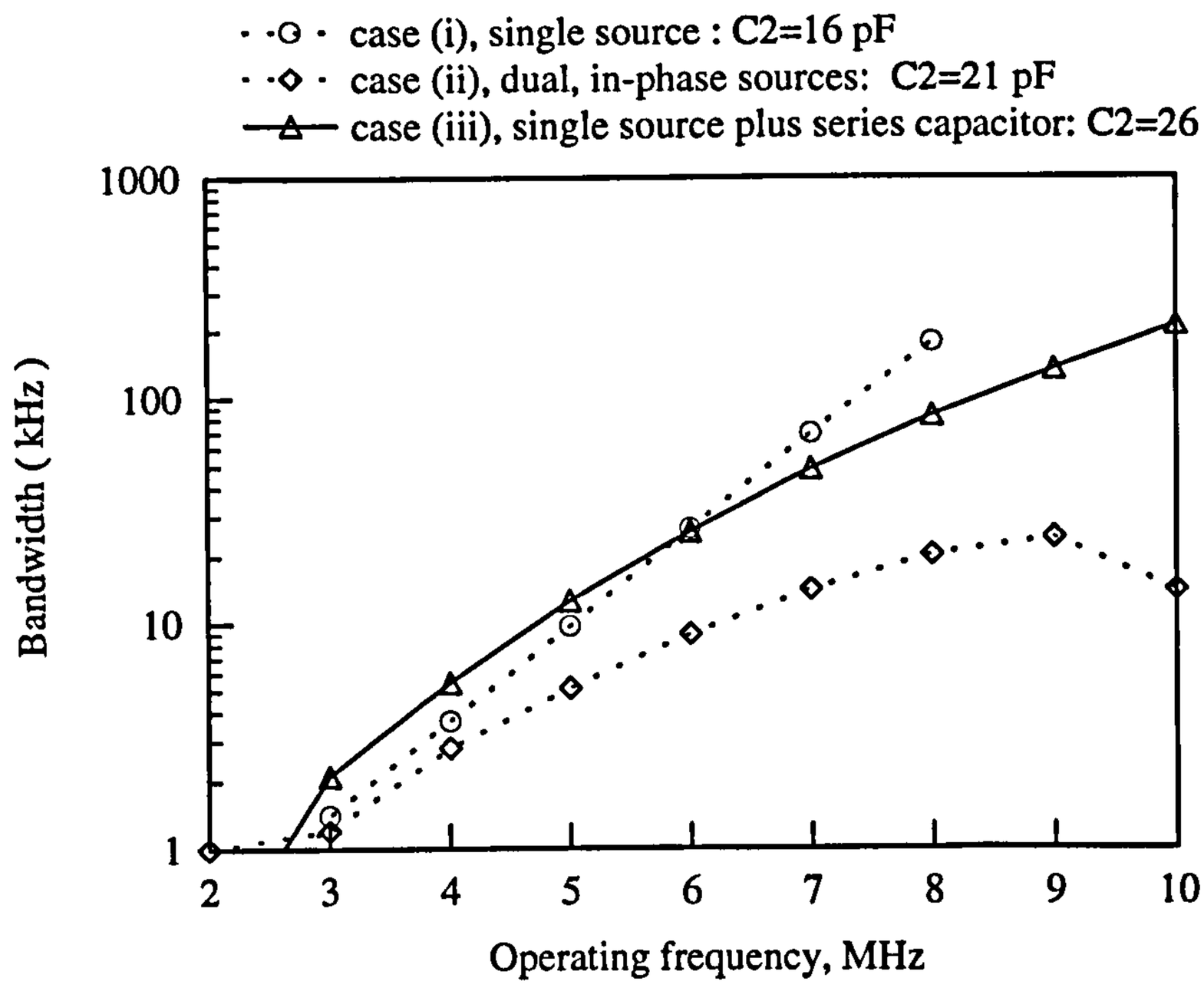


Figure 6.25 Bandwidth for VSWR of 2:1 from the network (b) used for cases (i), (ii) and (iii) when C_2 was fixed.

The two configurations of network (a) and (b) were also examined using the GA for use in cases (i) and (ii). As the results of case (iii), the network (b) was again more suitable to the two cases than the network (a). Therefore, the component values of this circuit were manipulated by the GA to produce the appropriate optimal matching conditions for case (i) and case (ii). The predicted bandwidth of these two systems were both also shown in figure 6.25. It should be noted that the vehicular loop fed by a single source can not be matched to a transmission line using this type of matching

circuit at 9 and 10 MHz as at such frequencies the antenna input impedance was changed from inductive to capacitive. Therefore, the results shown in figure 6.25 for the single source case are only over the frequency range from 2 to 8 MHz. At frequencies below 5 MHz all these antenna systems have extremely narrow bandwidth while beyond this frequency the bandwidth was increased for cases (i) and (iii). Especially, the bandwidth of case (i) is better than case (iii). However, there is hardly any improvement of bandwidth on case (ii).

6.6 Summary

This chapter has shown the application of GA to optimise a number of vehicular antenna configurations for NVIS propagation and also to design the matching networks for these antennas.

Firstly, section 6.2 provided the implementation of the wire-grid model of a Land Rover vehicle used for this investigation.

In section 6.3 the commonly used vehicular whip antenna was examined as the first instance. This was facilitated by manipulating the tilt and azimuth angles θ and ϕ of an electrically short whip antenna sited on the vehicle using the GA to achieve optimal percentage of power radiated towards zenith as required for NVIS propagation. The GA has identified that bending the whip from the vehicle can produce optimal NVIS performance. However the evaluated maximum G_{45} is less than 10%. The next example was the loop configuration. From this investigation the GA successfully produced a geometrical configuration of this antenna with optimal NVIS performance. The loop antenna provides better performance than a whip. A further example is the loop fed by dual sources. The optimal results for this design have been

found to occur when the extra source was positioned at the other end of the loop and its phase was tuned to the same as the original driving source. Compared to the loop fed by a single source, a significant improvement in efficiency was achieved from this dual, in-phase fed loop at frequencies beyond 4 MHz.

Section 6.4 further developed a capacitively loaded vehicular loop with the identical performance to the dual-feed case. This antenna configuration was developed iteratively by examining various loading networks with a simplified loop and then with the real vehicle-loop system. The position and component values of the load were all searched by the GA. Finally, the loop driven by a single sources and in series with an appropriate capacitor was found to be the optimal configuration, both electrically and mechanically.

Finally, in section 6.5 the GA was also used to design the matching network for the developed loop antenna systems. The results show that the GA can correctly and easily produce a matching network. In addition, the transfer bandwidths for these antenna systems have also been evaluated and compared in this section. The results indicated that these antennas have the expected narrow bandwidths which must be considered when implementing the NVIS system.

CHAPTER 7

TRANSMISSION-LINE MODEL OF THE VEHICLE-MOUNTED LOOP ANTENNA

7.1 Introduction

The analysis in the previous chapter has shown that the current distribution on the loop is the main factor affecting the antenna's NVIS performance. Once the maximum current is concentrated on the top horizontal section of the loop the optimal NVIS performance is obtained. In that iterative design procedure the GA was used to optimise the value of the capacitive load and then the moment method code-NEC2 simulated the electromagnetic properties of this complex antenna-vehicle system. This combined method is particularly powerful. However, the theoretical analysis of this case is somewhat cumbersome because the platform on which the antenna is mounted is complex. This chapter therefore describes the development of a simple but applicable transmission-line model which gives a good approximation to the current distribution of the analysed system.

In the past, transmission-line models have been used successfully to predict the input impedance of various types of antennas. These include the folded dipole (Thiele et al. 1980), the biconical antenna (Collin 1985 p87) and electrically loaded dipole

(Austin 1989). Based on these applications, the model developed here is expected to possesses the ability to accurately predict the current distribution on the loop.

In section 7.2 based on the basic transmission-line theory, the model for this work will be derived. Meanwhile, the principle of current moment is also included in this section for the prediction of the optimum capacitive load.

Sections 7.3 and 7.4 will apply this model to predict the current distribution and the optimal capacitive loads for a capacitively loaded loop and the vehicular NVIS antenna system, respectively. The comparisons of the resulting results between the transmission-line model and the numerical computation using the MoM will also be shown in these sections.

7.2 Model Formulation

The optimal NVIS vehicle-antenna system developed in the previous chapter is formed by a rectangular loop antenna loaded with a capacitor. The antenna and the conductive surface of vehicle body formed a closed loop. Therefore, when the transmission-line model is applied to approximate this system, the vehicle section can be considered as part of the transmission line. It can be simulated by a section of transmission line with a different radius from the antenna section. The configuration of the transmission-line model for this antenna-vehicle system is shown in figure 7.1, where points A and B, as before, are the positions of antenna feed and capacitive load, respectively and Xc is the reactance of the load. The combination essentially forms an asymmetrical transmission line.

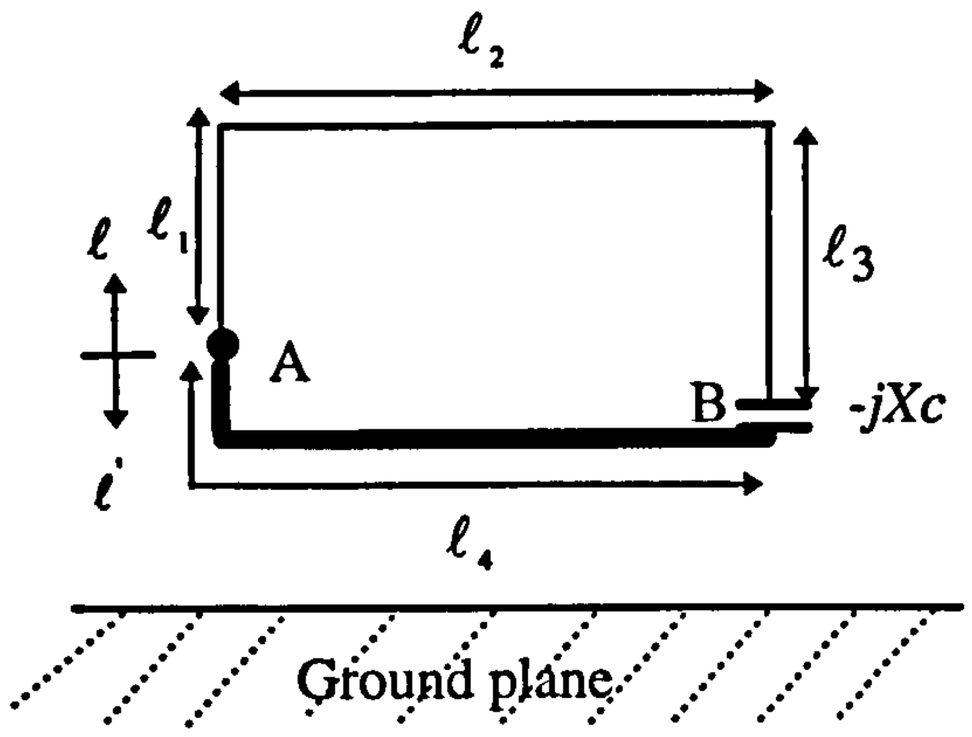


Figure 7.1 Transmission-line model of the vehicle-mounted loop antenna with a capacitive load.

As in the loaded dipole antenna discussed in previous section, the current standing wave on the lines may be assumed to be sinusoidal. If the current functions for the section of antenna and that of the vehicle are symbolised by I_a and I_v, respectively, they can be assumed to be

$$I_a(\ell) = I_0 \sin[k_0(a-\ell)] \qquad 0 \le \ell \le \ell_1 + \ell_2 + \ell_3 \tag{7.1}$$

$$I_v(\ell') = I_0' \cos[k_0(b-\ell')] + I_1' \qquad 0 \le \ell' \le \ell_4 \tag{7.2}$$

Where, I_0 is the current amplitude; ℓ and ℓ' are the distances from the feed along the antenna section and vehicle section, respectively. The unknowns a, b, I_0' and I_1' can be therefore determined by referring to the boundary conditions which are: (i) the current at the points A and B must be continuous; (ii) transmission line conditions must be satisfied at the feed and load. Thus we can write

$$I_a(\ell = 0) = I_v(\ell' = 0) \tag{7.3}$$

$$I_a(\ell = \ell_1 + \ell_2 + \ell_3) = I_v(\ell' = \ell_4) \tag{7.4}$$

$$I_a(\ell = \ell_1 + \ell_2 + \ell_3) = C\frac{\partial\ V_a(\ell = \ell_1 + \ell_2 + \ell_3)}{\partial\ t} \tag{7.5}$$

$$I_v(\ell' = \ell_4) = C\frac{\partial\ V_v(\ell' = \ell_4)}{\partial\ t} \tag{7.6}$$

Substituting equations 7.1 and 7.2 into equations 7.3-6 the unknowns a, b, $I_0^{'}$ and $I_1^{'}$ are therefore obtained. The solutions are given by

$$a = \ell_1 + \ell_2 + \ell_3 + \frac{1}{k_0}\tan^{-1}\frac{Z_0}{X_c} \tag{7.7}$$

$$b = \frac{1}{k_0}\cos^{-1}\left(\frac{1}{\sqrt{d^2+1}}\right) \tag{7.8}$$

$$I_0^{'} = I_0 \sin(k_0 a) / [\cos(k_0 b) - \cos(k_0 b - k_0 \ell_4) - \frac{Z_0}{X_c}\sin(k_0 b - k_0 \ell_4)] \tag{7.9}$$

$$I_1^{'} = I_0 \sin(k_0 b) - I_0 \sin(k_0 a)\cos(k_0 b) / [\cos(k_0 b) - \cos(k_0 b - k_0 \ell_4) - \frac{Z_0}{X_c}\sin(k_0 b - k_0 \ell_4)] \tag{7.10}$$

where, d is given by

$$d = \frac{-1 + \cos(k_0 \ell_4) + p \cdot \sin(k_0 \ell_4)}{p \cdot \cos(k_0 \ell_4) - \sin(k_0 \ell_4)} \tag{7.11}$$

and p by

$$p = \frac{\sin[k_0(\ell_1 + \ell_2 + \ell_3) + \tan^{-1}\frac{Z_0}{X_c}] - \sin(\tan^{-1}\frac{Z_0}{X_c})}{\cos(\tan^{-1}\frac{Z_0}{X_c})} \tag{7.12}$$

From equations 7.7-12 once the configuration of the loaded loop, the value of the load and the characteristic impedance for this model are determined the current distribution on the loop can be approximated from 7.7-12. Additionally, it should be emphasised that the "skin effect" will be used to determined the equivalent electromagnetic radius of the transmission line used to represent the modelled

structure. This, of course, will vary with the operating frequency and so yields a transmission line of variable characteristic impedance. The higher the operating frequency, the thinner the skin depth (Collin 1992 p54) resulting in a larger characteristic impedance.

In addition, if the loop is located over a perfectly conducting ground, as presented in the real state of affairs, and the distance of the lower horizontal section of this loop (i.e. vehicle section) from the ground plane is electrically short, the effect of this section on the far-field radiation can be neglected. This is because the loop current and its image are of equal amplitude but opposite phase. The mechanism is shown in figure 7.2. Thus, when the model is used to predict the value of the capacitive load for optimal NVIS performance, only the current moment on the top horizontal section should be maximised by loading with the appropriate capacitor. In the mathematical expression, the solution of C for the maximum current moment on the top horizontal section of the transmission line can be obtained by using gradient method which is shown as follows

$$\frac{\partial \int_{\ell_1}^{\ell_1+\ell_2} I_a(\ell)\, d\ell}{\partial C} = 0 \tag{7.13}$$

where ℓ_1 and $\ell_1 + \ell_2$ refer to the ends of the horizontal top section (see figure 7.1).

The following sections will apply this model to predict the current distribution and the optimal capacitive loads for two antenna examples. The configuration of the first antenna is a simple capacitively loaded loop and then, in the second case, the real antenna-vehicle system is modelled. All the predicted results from the model will be

compared to those from the computation of the Moment Method or from the prediction of GA/NEC2.

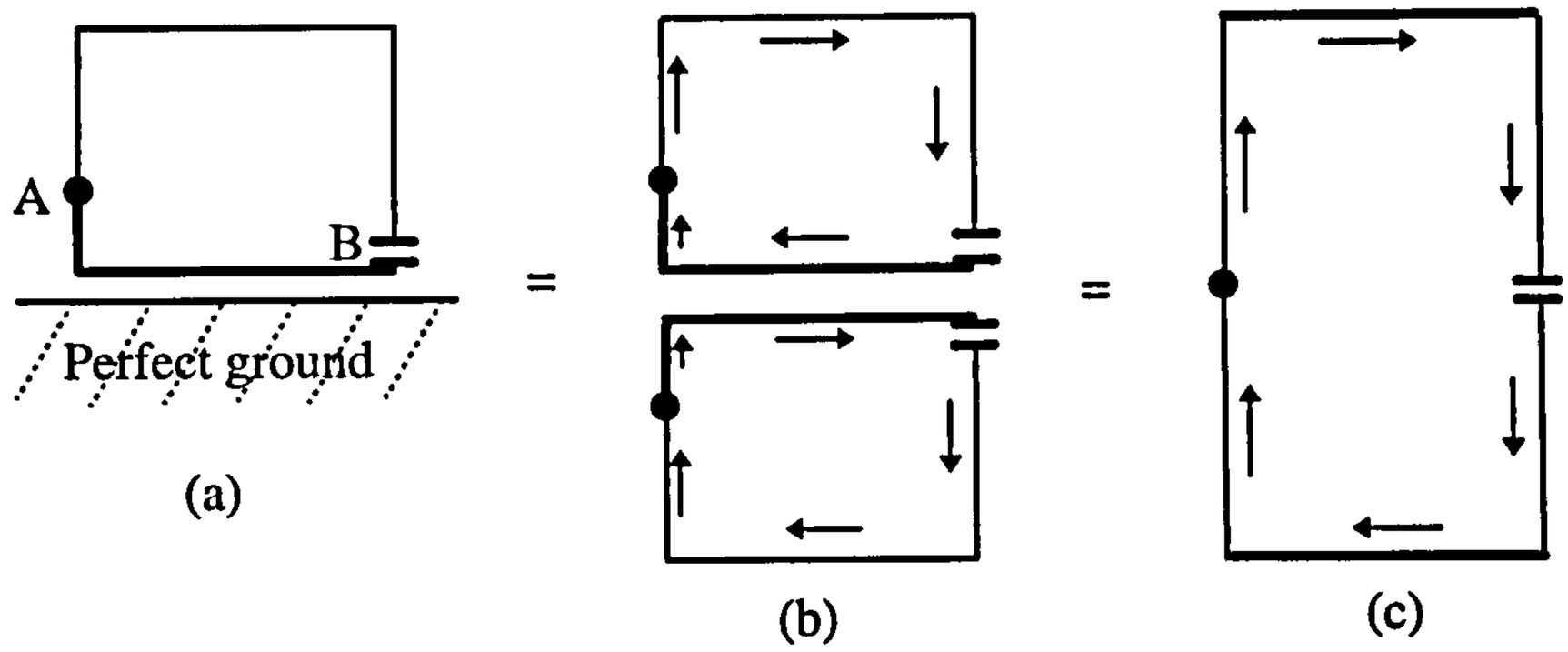


Figure 7.2 Transmission-line model of a loaded loop antenna mounted on the vehicle.
(a) Physical model.
(b) and (c) Equivalent model using image theory.

7.3 Prediction of a Capacitively Loaded Loop Antenna

Figure 7.3 is a capacitively loaded loop antenna as used in section 6.4.1. The dimensions of this antenna are deduced from the physical antenna-vehicle system

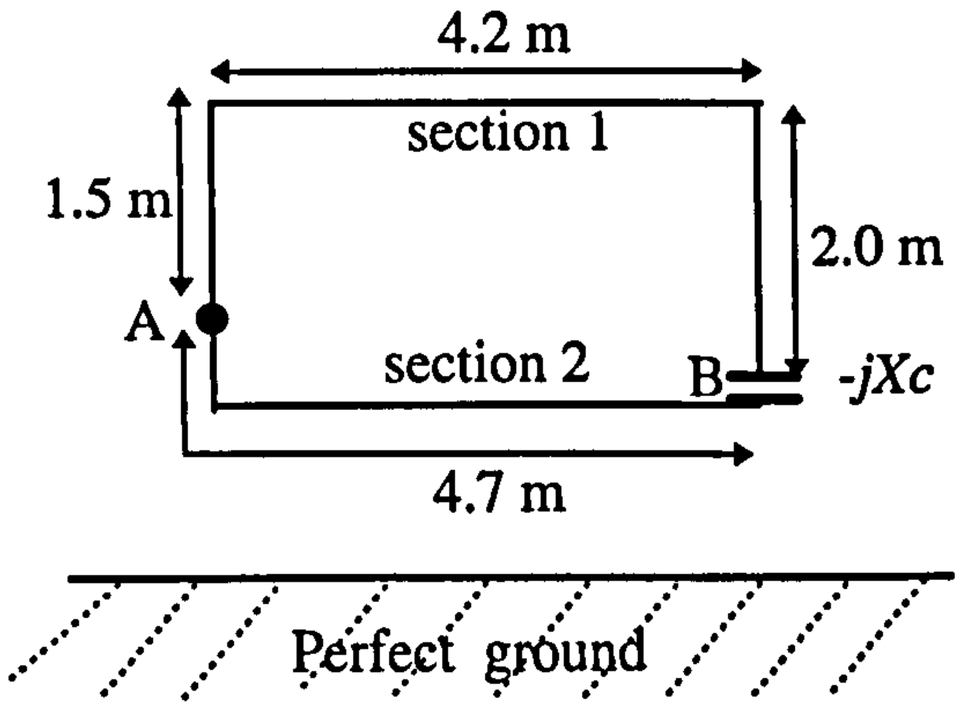


Figure 7.3 Dimensions of the transmission-line model for a capacitively loaded loop antenna.

investigated in chapter 6. Based on the optimised results from the GA/NEC2 method, the position of the capacitor was fixed at point B and the associated capacitor values used here over the NVIS band from 2 to 10 MHz were selected as those listed in table 6.5. To completely illustrate the transmission-line model for this loaded loop antenna, another criterion - characteristic impedance - was therefore obtained by selecting the same driving current at the feed point of both this model and the real loop antenna.

Figures 7.4 (a)-(e) present the predicted normalised current distributions along the loop at the frequencies of interest from the two different methods. The solid line is the solution from the transmission-line model and the dotted line is for that from the NEC2 code. In these plots, the current predicted by the transmission-line model was normalised by I_0; while that predicted by NEC2 was normalised by the maximum current. Clearly, the agreement of current distribution from the two methods is good. The appropriate characteristic impedance Z_c used for the model at each frequency is shown in figure 7.5. The results show that the impedance varies with frequency within the range from 428 to 441 Ω.

The predicted capacitor values against the operating frequency for both the transmission-line model and the GA/NEC2 method for optimal NVIS performance are illustrated in figure 7.6. The loads can be predicted accurately by the transmission-line model with a deviation of less than 10% from that predicted by the GA/NEC2. The deviation of the predicted capacitors against the operating frequency is also shown in figure 7.6.

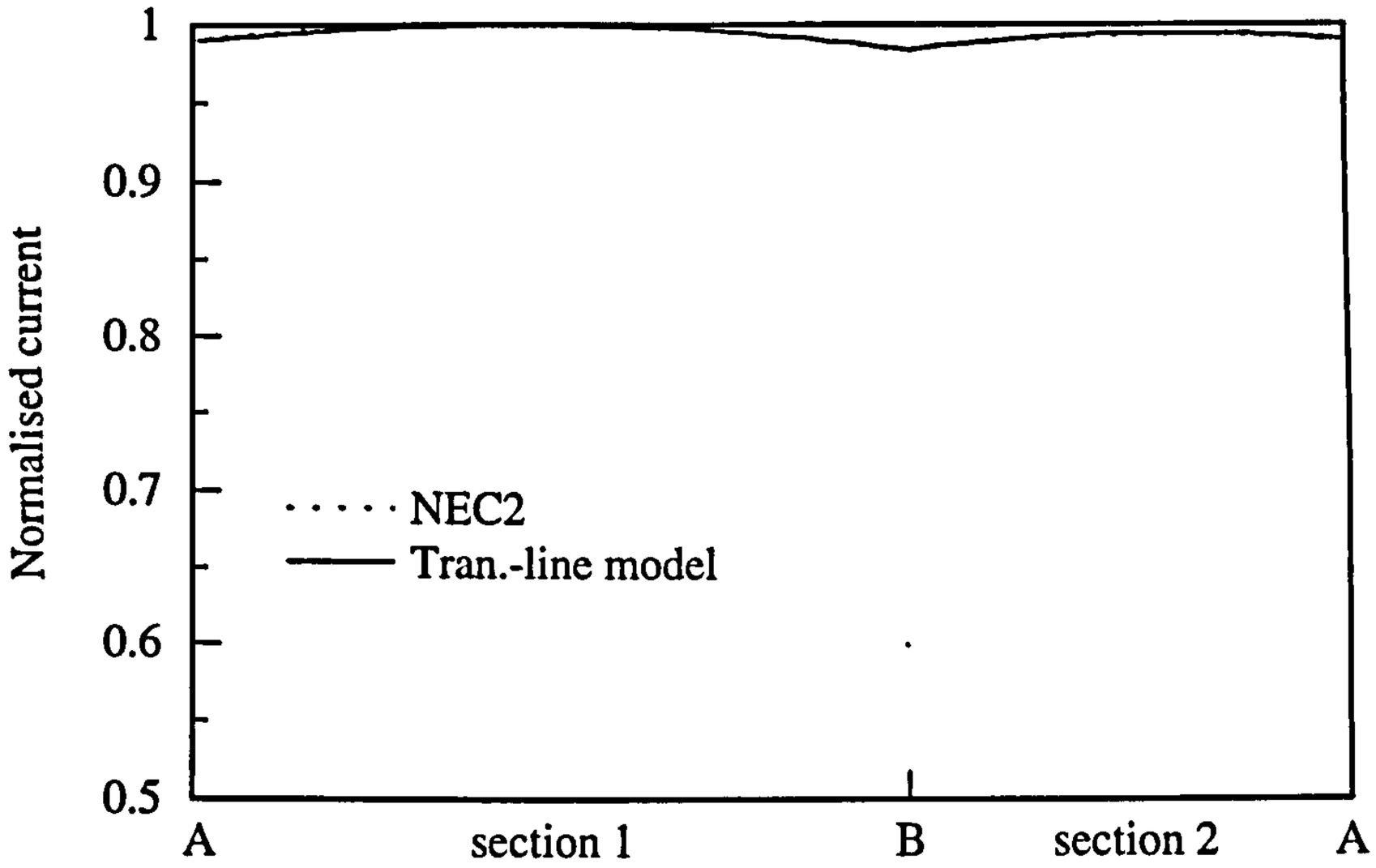


Figure 7.4(a) Comparison of the normalised predicted current distribution along a capacitively loaded loop antenna between using NEC2 and the transmission-line model at 2 MHz.

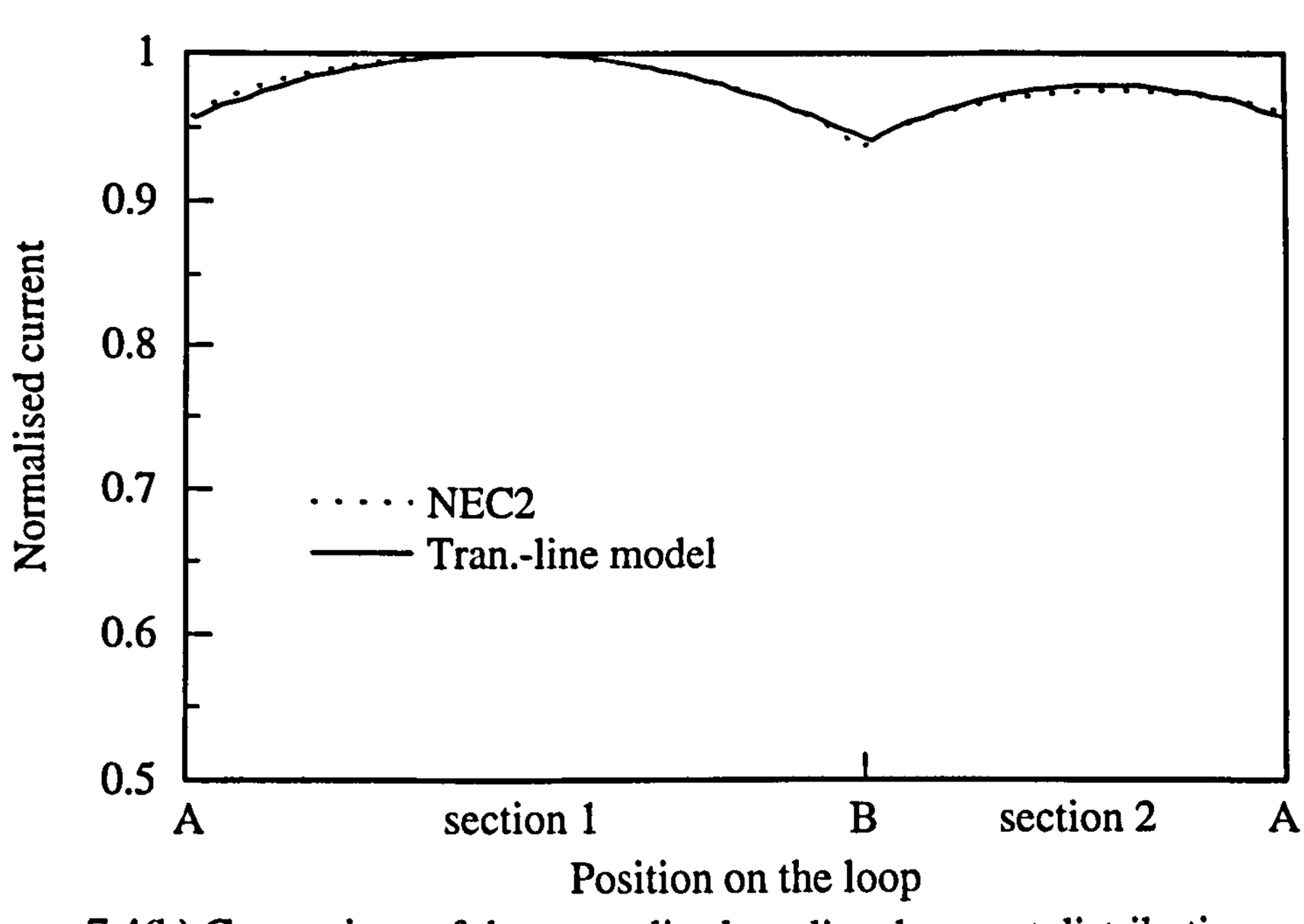


Figure 7.4(b) Comparison of the normalised predicted current distribution along a capacitively loaded loop antenna between using NEC2 and the transmission-line model at 4 MHz.

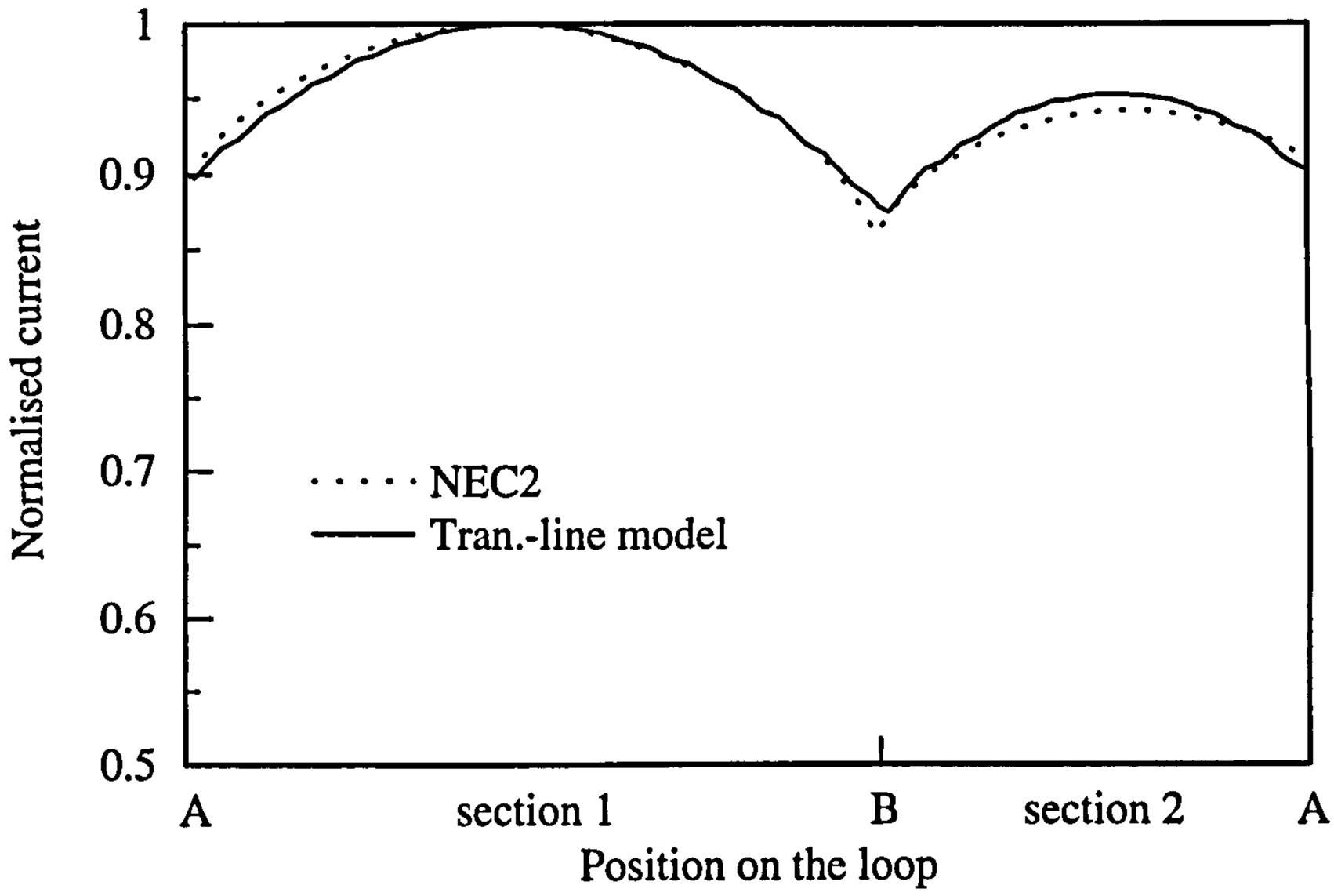


Figure 7.4(c) Comparison of the normalised predicted current distribution along a capacitively loaded loop antenna between using NEC2 and the transmission-line model at 6 MHz.

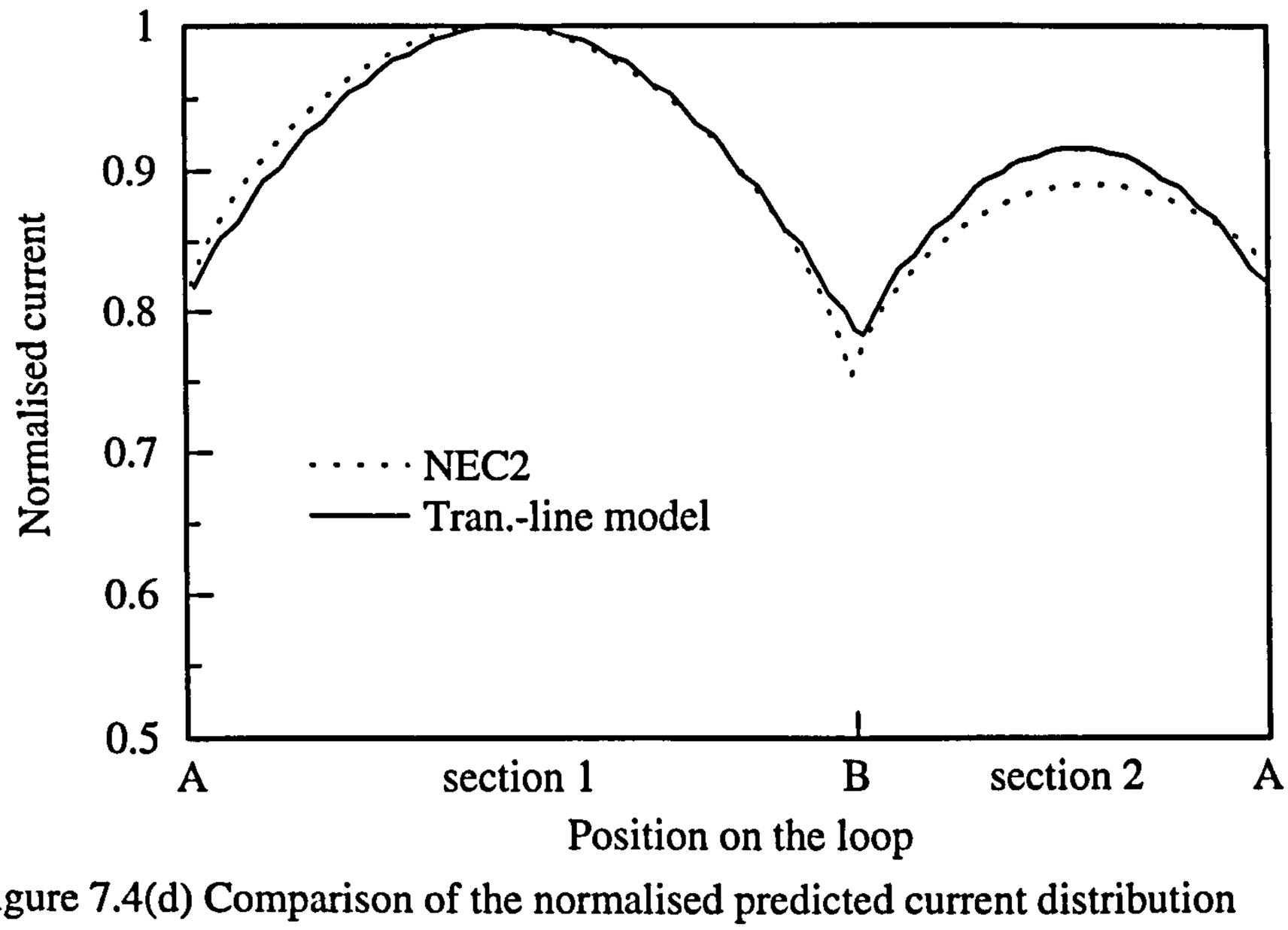


Figure 7.4(d) Comparison of the normalised predicted current distribution along a capacitively loaded loop antenna between using NEC2 and the transmission-line model at 8 MHz.

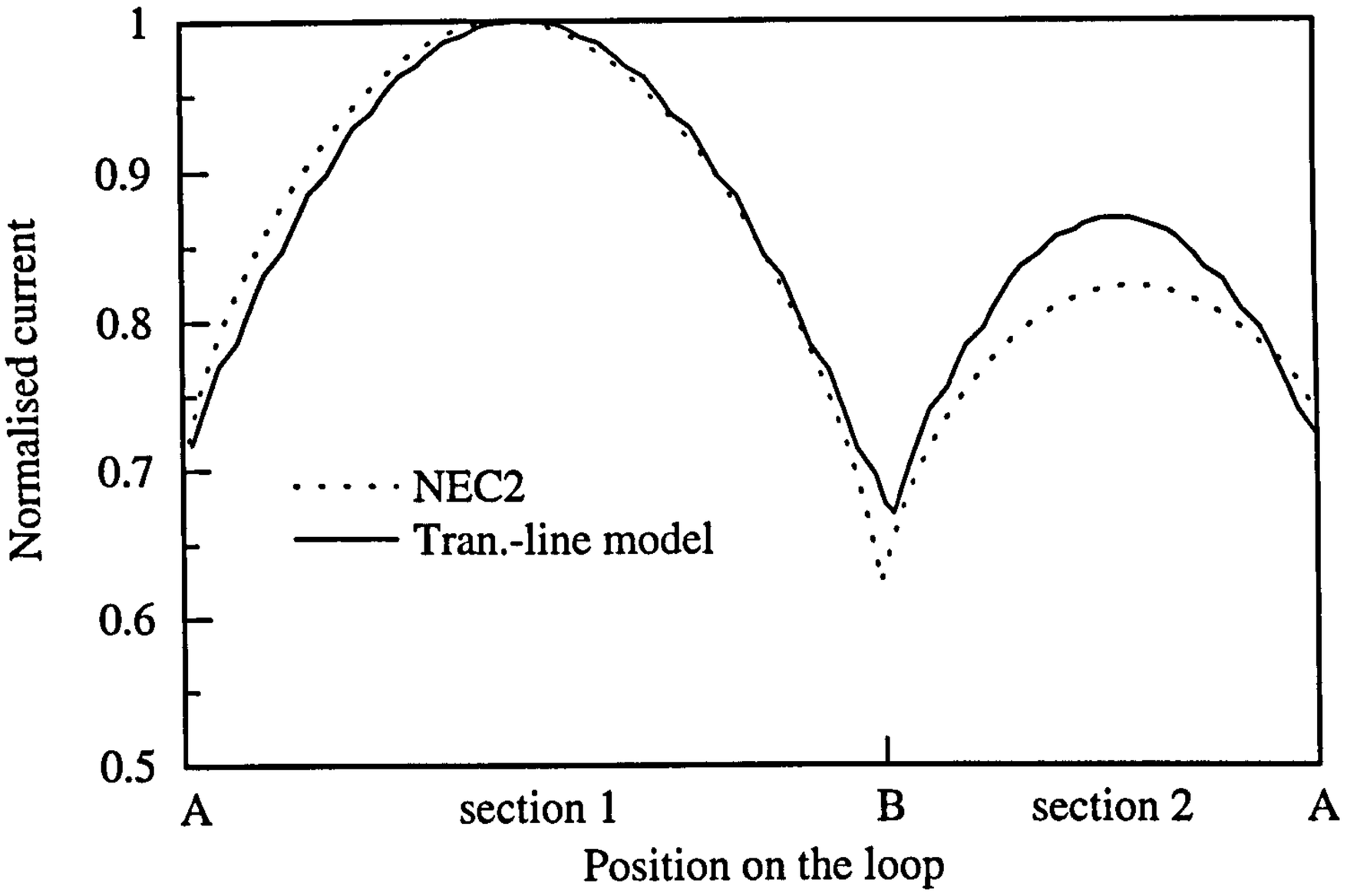


Figure 7.4(e) Comparison of the normalised predicted current distribution along a capacitively loaded loop antenna between using NEC2 and the transmission-line model at 10 MHz.

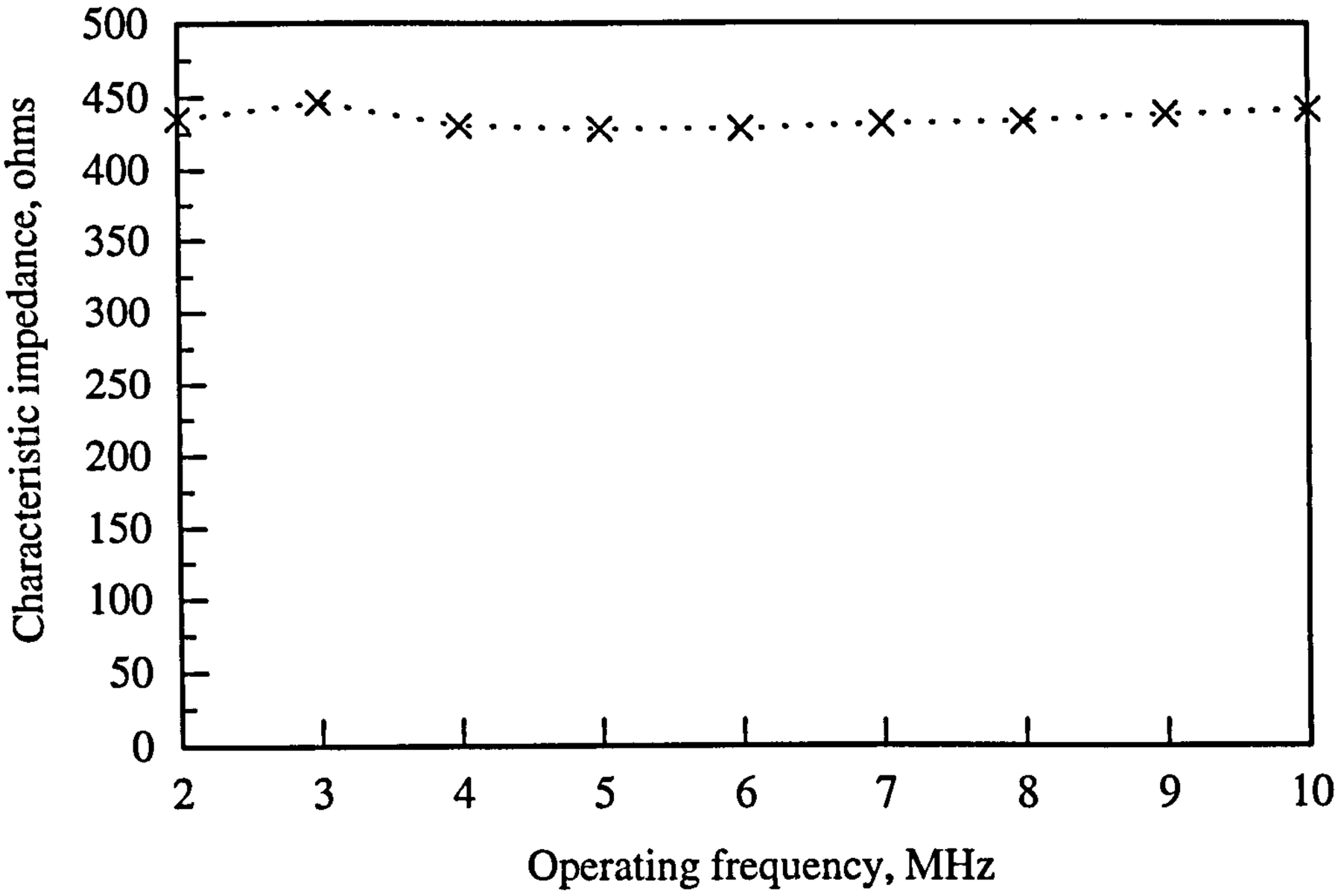


Figure 7.5 Variation in characteristic impedance of the transmission line with operating frequency for the capacitively loaded loop antenna.

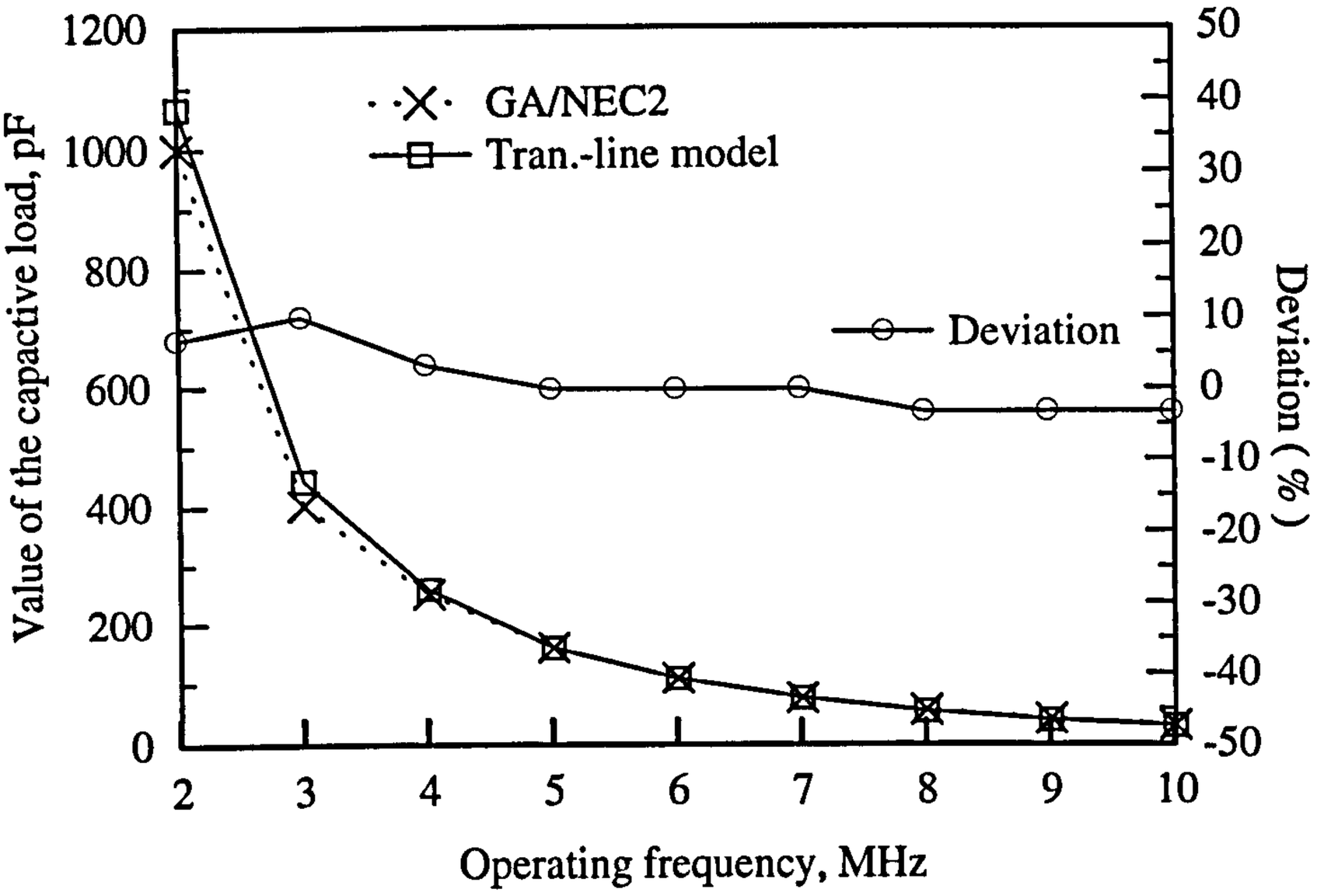


Figure 7.6 Variation in value of the capacitive loads and the deviation with operating frequency for the capacitively loaded loop antenna.

7.4 Prediction of the Vehicular NVIS Antenna System

In this section the transmission-line model will be applied to the real vehicle-mounted loaded loop antenna system which was obtained in section 6.4.2. Figure 7.7 shows the model of this system. As in the previous case, the capacitor was positioned at point B and its value was selected from figure 6.10. Again the characteristic impedance used for this model over the frequency band was obtained by feeding the transmission line with the same current source as that evaluated by NEC2 from the real configuration. As shown in figure 7.8 the resulting characteristic impedance varied with the frequency within the range from 285 to 350 Ω.

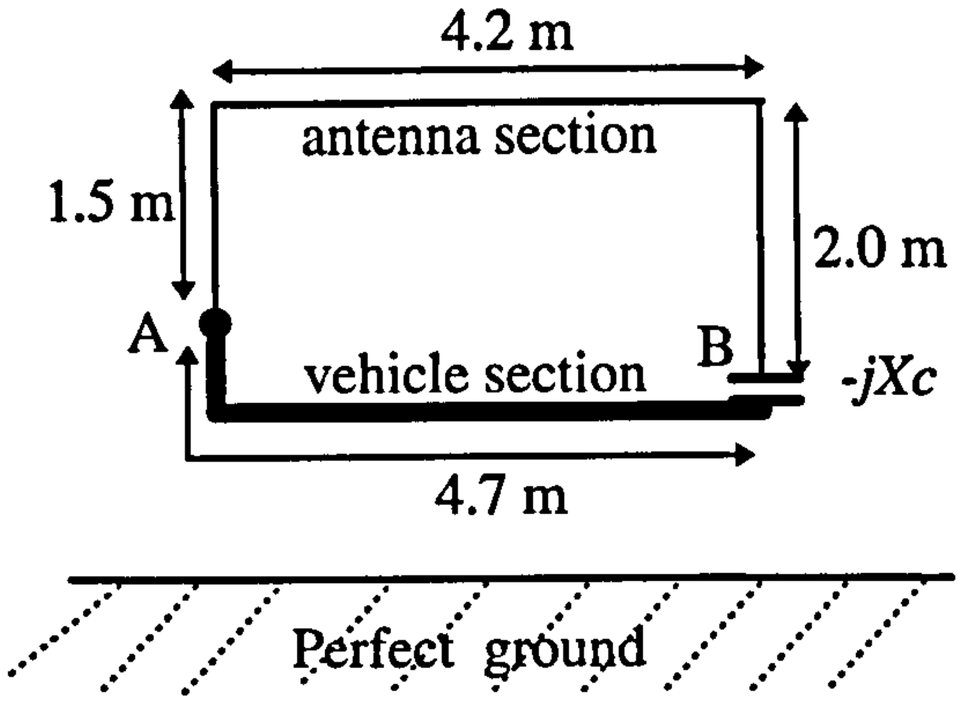


Figure 7.7 Dimensions of the transmission-line model for a loaded loop.

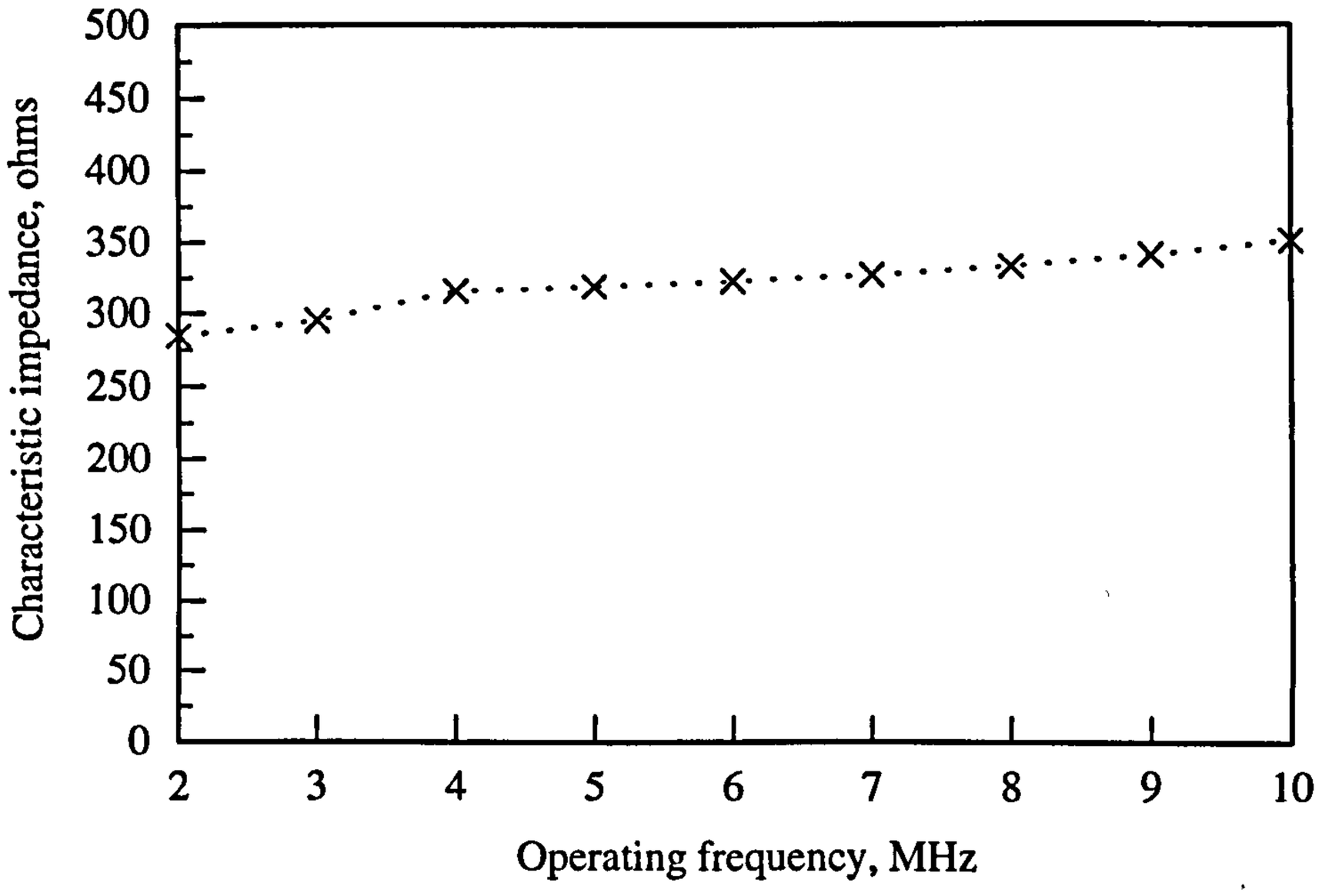


Figure 7.8 Variation in characteristic impedance of the transmission line with operating frequency for the antenna-vehicle system.

The predicted normalised current distributions are illustrated in figures 7.9 (a)-(e). In these plots because the current distribution on the vehicle section of the real configuration is difficult to be decided, therefore only that on the loop antenna section was shown and compared with that from the transmission-line model. Clearly, a good agreement has obtained between both.

The predicted capacitor values by maximising the current moment on the upper horizontal section of the transmission-line model were shown in figure 7.10. The model predicted the values of loads with deviations all less 10% from the values optimised by the GA/NEC2.

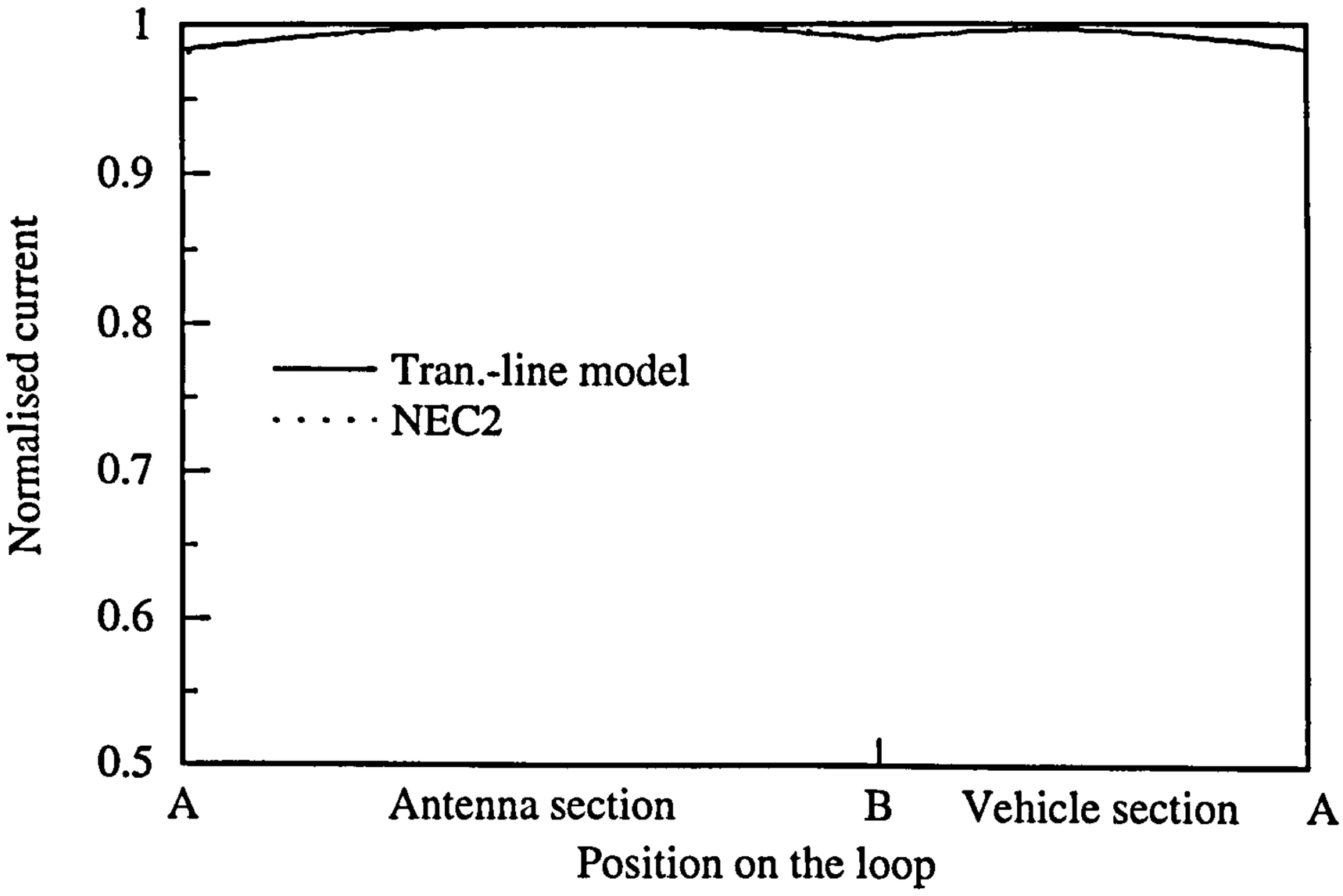


Figure 7.9 (a) Comparison of normalised current distribution on the loop antenna between the real configuration and the transmission-line model, frequency = 2 MHz

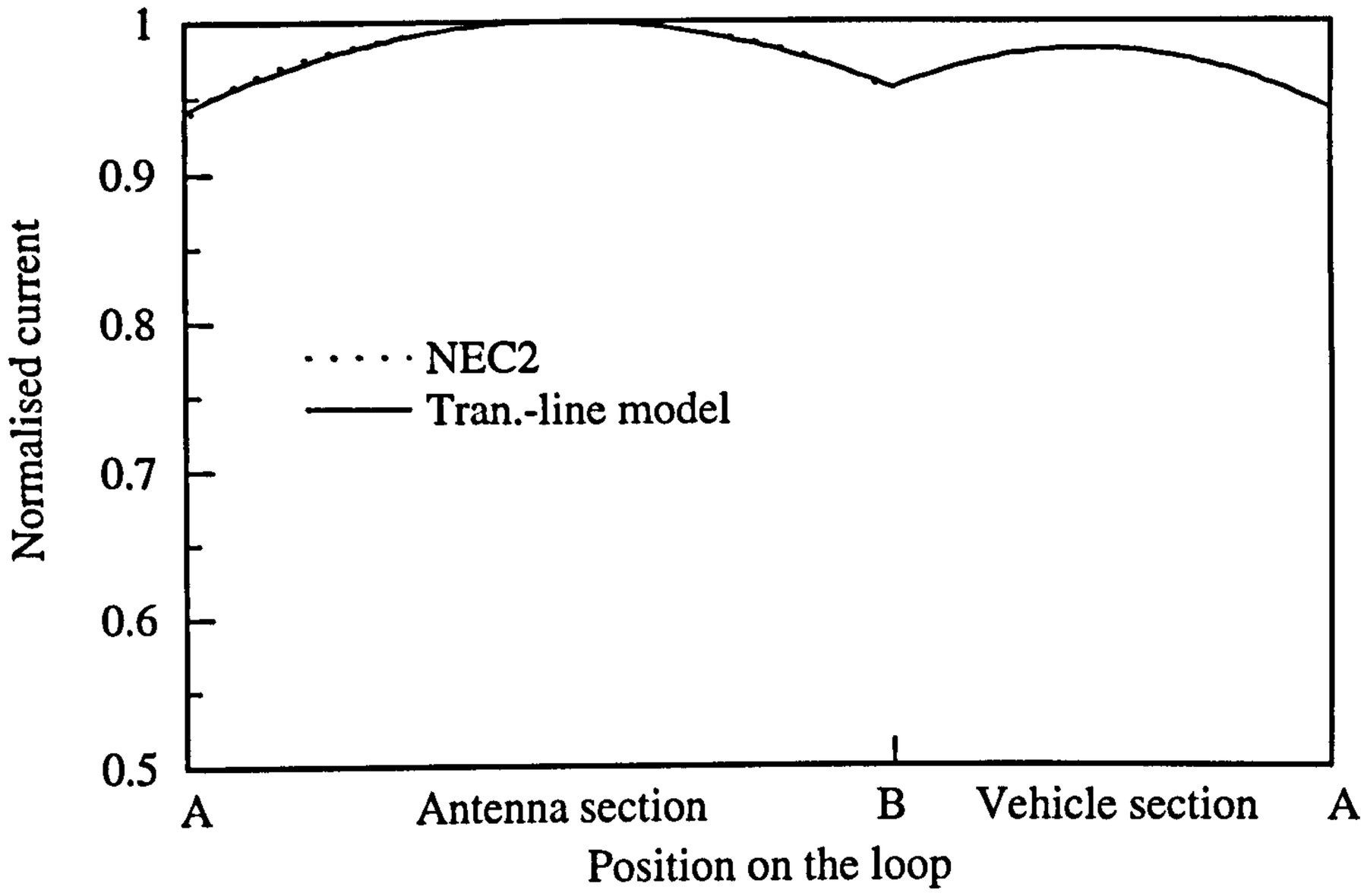


Figure 7.9 (b) Comparison of normalised current distribution on the loop antenna between the real configuration and the transmission-line model, frequency = 4 MHz

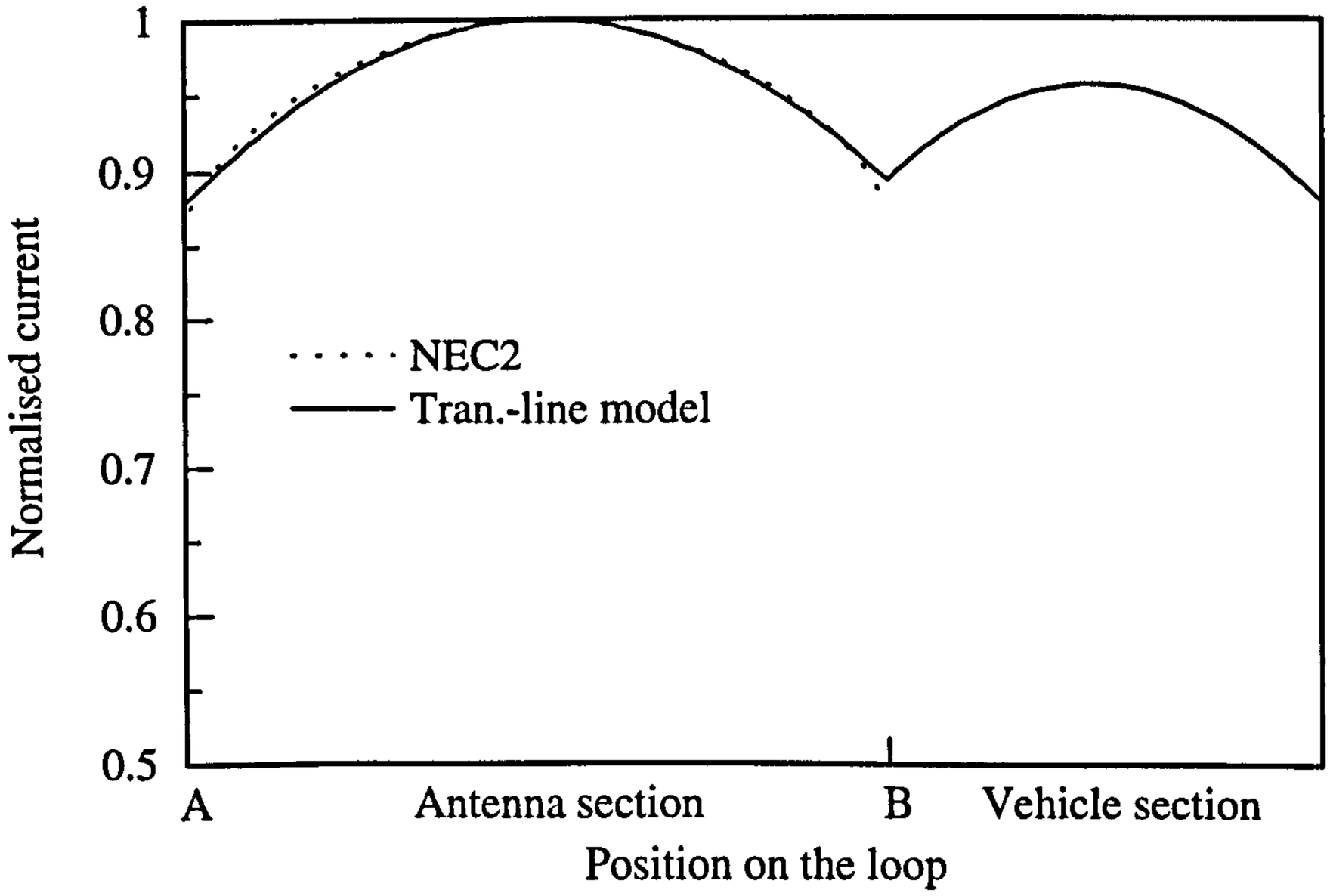


Figure 7.9 (c) Comparison of normalised current distribution on the loop antenna between the real configuration and the transmission-line model, frequency = 6 MHz

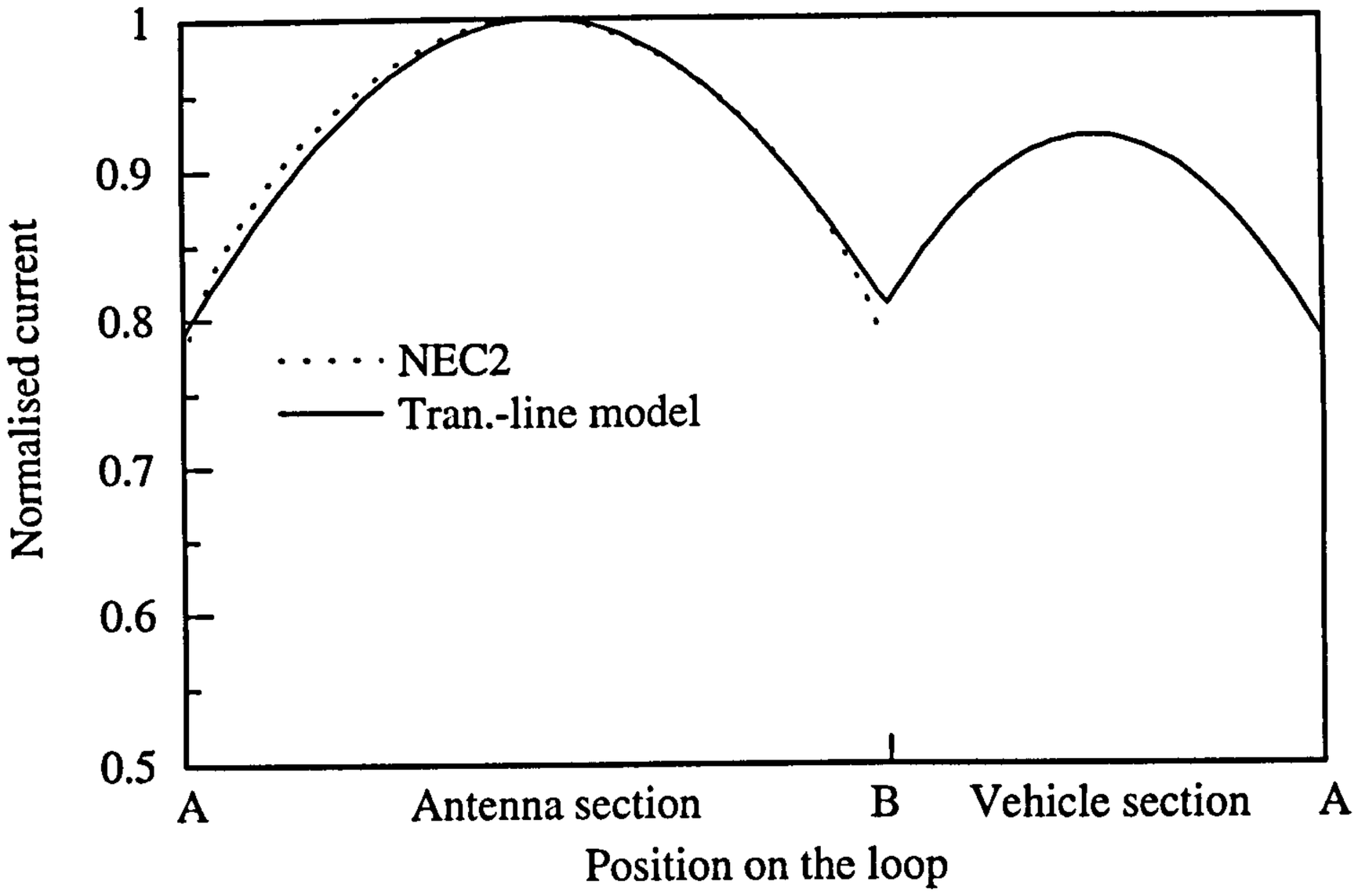


Figure 7.9 (d) Comparison of normalised current distribution on the loop antenna between the real configuration and the transmission-line model, frequency = 8 MHz

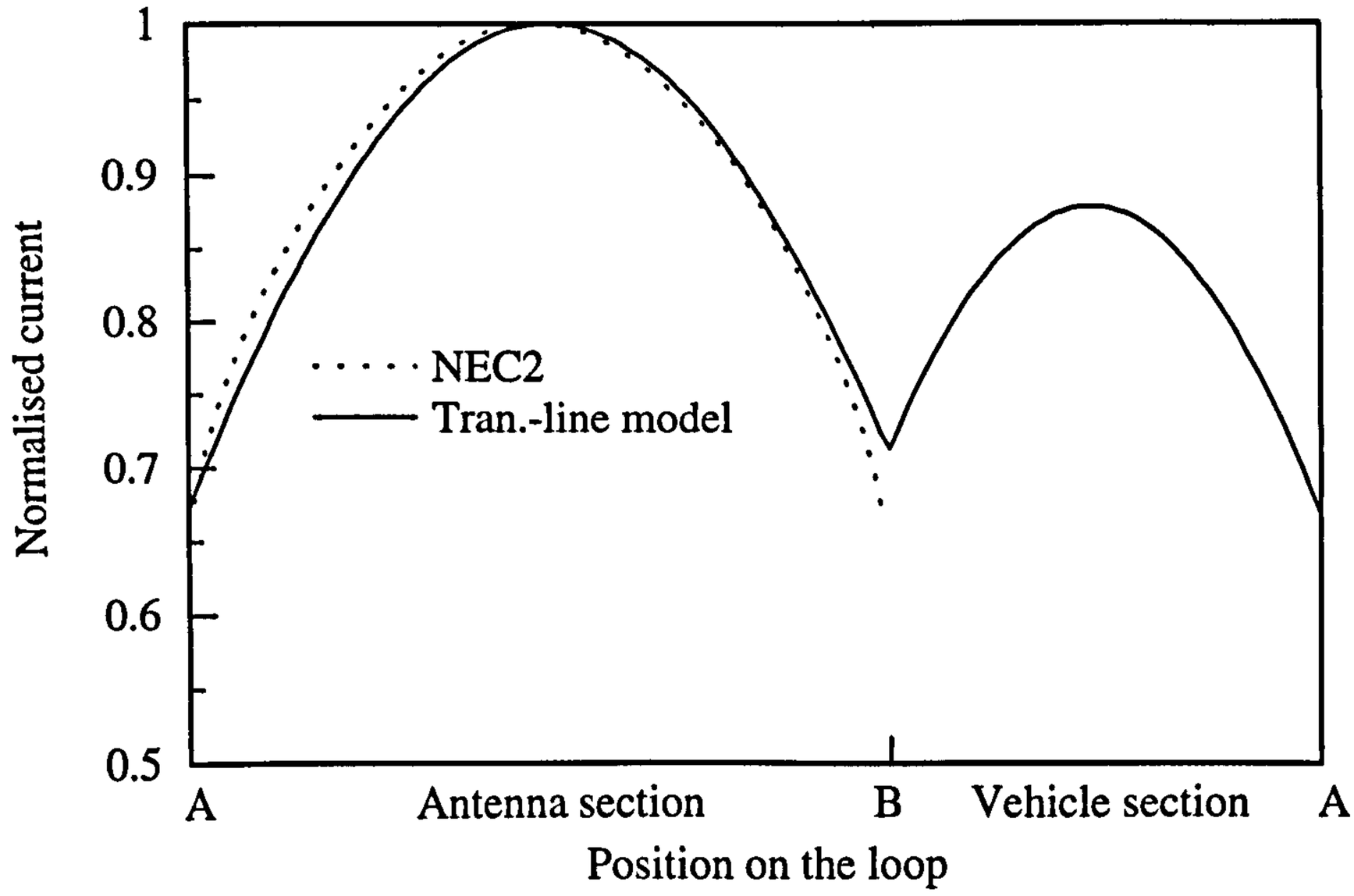


Figure 7.9 (e) Comparison of normalised current distribution on the loop antenna between the real configuration and the transmission-line model, frequency = 10 MHz

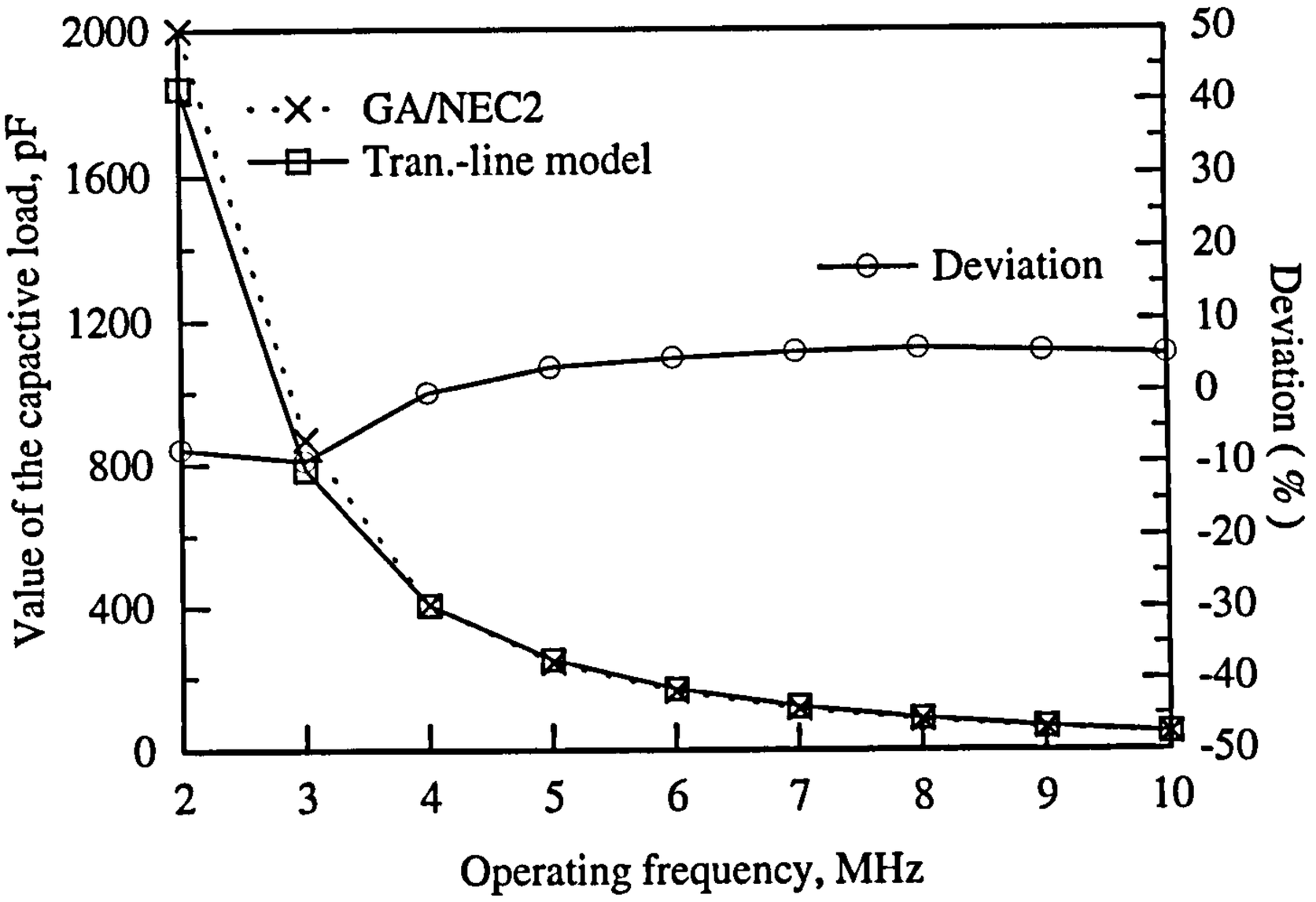


Figure 7.10 Variation in value of the capacitive loads and the deviation with operating frequency for the antenna-vehicle system.

7.5 Summary

A transmission-line model developed in this chapter has been described that forms an accurate tool for analysing the current distribution on the investigated vehicular loaded loop antenna with optimal NVIS performance. Moreover, this model can also accurately predict the capacitive load that must be added to the loop to produce maximum power radiated towards the zenith. Due to its simple theory, this model is extremely useful for predicting the performance of such a loaded loop mounted on a complex structure and in determining the values of the loading capacitance for particular geometrical configurations.

CHAPTER 8

CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER WORK

This research has applied a novel design procedure to the optimisation of antenna systems for specified objectives. This was achieved by using the deductive method of the Genetic Algorithm (GA) in conjunction with the Method of Moments as implemented in the numerical electromagnetics code - NEC2. The growing and successful use of this new GA technique for antenna optimisation was reviewed in chapter 2, and the software for this combined calculation procedure was developed based on the well-known methodology of both the GA strategy and the Moment Method. Before applying this technique practically to cope with the antenna problems under investigation here the accuracy of both the GA procedure and the NEC2 code were firstly validated using typical examples in the literature. The former was obtained by successfully searching for a global solution from a typical function with both a global and many local maxima while the latter was validated by accurately predicting the input impedance and radiation patterns of a vertical folded loop antenna mounted on a conducting box which was used to simulate a small hand-held portable telephone after representing it with a wire-grid model. Since this geometrical configuration most closely approximates the vehicular antenna in this work it was an appropriate validation procedure. These details were all discussed in chapters 3 and 4 of this

thesis. This work showed that the process of this program's development was both accurate and efficient.

An important finding of this study was the proof of the usefulness of the combined GA/NEC2 approach for designing an antenna with a specified performance which would be too difficult to obtain using conventional theoretical and experimental methods. In chapter 5 the GA's iterative procedure based on "natural evolution" produced an improved, V-shaped, parasitic array with linear elements and antenna performance of a high gain. It had been assumed previously that such performance could only be achieved by elaborate curvature of the elements, or by the use of considerably more elements in a conventional, parallel element structure. This work has produced a much simpler antenna – and hence one that is cheaper to construct – that can yield the same performance.

Another important application of this GA/NEC2 method was the investigation of vehicular NVIS antenna systems. This was illustrated in chapter 6. Initially in this chapter the GA confirmed the previously predicted result of the optimal configuration of a whip or a loop antenna with a single- or dual-feed system mounted on the vehicle for NVIS propagation from experiment or Characteristic Mode analysis. However, comparing the complexity between these methods the GA is obviously easier to implement than the others.

A result of considerable importance was the use of an appropriate capacitor at the opposite end to the feeding position of an optimised vertically oriented loop antenna can so modify the current distribution along the upper horizontal section of the loop to yield optimal NVIS performance. Previously, such performance could only be achieved by feeding the loop with two in-phase voltage sources at those points: a

cumbersome procedure. The loading capacitor could be made motor-driven so making frequency agility a fairly simple matter.

The GA also enabled an ideal impedance design matching circuit for the capacitively loaded NVIS antenna system to be designed.

Finally, to assist in predicting the radiation characteristics of this NVIS antenna, a transmission-line model has been proposed to approximate the loop-vehicle system. This model can precisely simulate the current distribution along the loaded loop and, furthermore, it accurately predicted the appropriate value of capacitive load for optimal NVIS performance by maximising the current moment along the top horizontal section of the loop.

The results presented in this thesis have been based on the numerical experiments of the antenna systems using the combined GA/NEC2 method. The ability of this method to explore an antenna configuration by being able to produce an appropriate specified objective was proved. Clearly this novel technique has the potential to be used for many other antenna designs. Although this study has been primarily concerned with an assessment of the GA/NEC2 approach for a vehicle-antenna system intended for optimal NVIS performance there are numerous other related applications. As was discussed in this thesis this mode of propagation is particularly necessary for relatively short-haul communications beyond ground-wave range, and between terminals that are either beyond line-of-sight, or that are severely obstructed by large land masses. However, such mobile terminals are not limited to just motor vehicles. They could be aircraft, ships or other land vehicles. From the results achieved here it is felt that the method has considerable potential and could usefully be employed to design antenna systems for a variety of other applications.

REFERENCES

1. Altman, Z., Mittra, R. and Boag, A., 1997, "New designs of ultra wide-band communication antennas using a genetic algorithm," IEEE Trans. on Antennas and Propagat., Vol.45, No.10, pp.1494-1501.

2. Altshuler, E.E., 1993 "A monopole antenna loaded with a modified folded dipole," IEEE Trans. on Antennas and Propagat., Vol.41, No.7, pp.871-876.

3. Altshuler, E.E., 1996a "A monopole loaded with a loop antenna," IEEE Trans. on Antennas and Propagat., Vol.44, No.6, pp.787-791.

4. Altshuler, E.E., 1996b, "Hemispherical coverage using a double-folded monopole," IEEE Antennas and Propagat., Vol.44, No.8, pp.1112-1119.

5. Altshuler, E.E. and Linden, D.S., 1997, "Wire-antenna designs using genetic algorithms," IEEE Antennas and Propagat. Magazine, Vol.39, No.2, pp.33-43.

6. Ares, F., Rengarajan, S.R., Villanueva, E., Skochinski, E. and Moreno, E, 1996, "Application of genetic algorithms and simulated annealing technique in optimising the aperture distributions of antenna array patterns," Electronics Letters, Vol.32, No.3, pp.148-149.

7. Austin, B.A., 1989, "A simulation exercise in antenna analysis using MININEC," Int. J. Elect. Enging Educ., Vol. 26, pp.355-366.

8. Austin, B.A. and Murray, K.P., 1991, "The generation of antenna characteristic modes from the impedance matrix using the Moment Method," Proceedings of the Seventh International Conference on Antennas and Propagation (ICAP-91), IEE Conference Publication No.333, Vol.2, pp.713-716.

9. Austin, B.A. and Murray, K.P., 1998, "The application of characteristic-mode techniques to vehicle-mounted NVIS antennas," IEEE Antennas and Propagat. Magazine, Vol.40, No.1, pp.7-21.

10. Baker, D.C., 1991, "On the functioning of a helicopter-borne HF loop antenna," Applied Computational Electromagnetics Society Journal," Vol.6, No.2, pp.53-73.

11. Balanis, C.A., 1997, "Antenna theory: analysis and design," , New York: Wiley, p.767.

12. Boag, A., Boag, A., Michielssen, E. and Mittra, R., 1996, "Design of electrically loaded wire antennas using genetic algorithms," IEEE Trans. on Antennas and propagat., Vol.44, No.5, pp.687-695.

13. Bojsen, J.H., Shaer-Jacobsen, H., Nilsson, E. and Andersen, J.B., 1971,

"Maximum gain of Yagi-Uda arrays," Electronics Letters, Vol.7, pp.531-532.

14. Burberry, R.A., 1982, "HF loop antennas for air, land and sea mobiles," IEE International Conference on HF Radio Systems and Techniques, IEE Pub. No.206, pp.18-22.

15. Burke, G.J. and Poggio, A.J., 1981a, "Numerical Electromagnetics Code - Method of Moments, Part I: Theory," Lawrence Livermore National Laboratory, Technical Document UCID-18834, Calif., U.S.A.

16. Burke, G.J. and Poggio, A.J., 1981b, "Numerical Electromagnetics Code - Method of Moments, Part III: User's Guide," Lawrence Livermore National Laboratory, Technical Document UCID-18834, Calif., U.S.A.

17. Chen, C.A. and Cheng, D.K., 1975, "Optimum element lengths for Yagi-Uda arrays," IEEE Trans. on Antennas and Propagat., Vol.AP-23, No.1, pp.8-15.

18. Cheng, D.K. and Chen, C.A., 1973, "Optimum element spacings for Yagi-Uda arrays," IEEE Antennas and Propagat., Vol.AP-21, No.5, pp.615-623.

19. Cheng, D.K. and Chen, C.A., 1991, "Gain optimization for Yagi-Uda arrays," IEEE Antennas and Propagat. Magazine, Vol.33, No.3, pp.42-45.

20. Cheng, D.K. and Liang, C.H., 1982, "Shaped wire antennas with maximum directivity," Electronics Letters, Vol.18, No.19, pp.816-818.

21. Chu, A.W.C., Long, S.A. and Wilton, D.R., "The radiation pattern of a monopole antenna attached to a conducting box," IEEE Trans. on Antennas and Propagat., Vol.38, No.12, pp.1907-1911.

22. Collin, R.E., 1985, "Antenna and radiowave propagation," New York: McGraw-Hill.

23. Collin, R.E., 1992, "Funfations for microwave engineering," New York: McGraw-Hill.

24. Cox, J.W.R., 1991, "Comparison of predicted aircraft wire antenna terminal impedance (using NEC) with measurement in the HF band," International Conf. On Antennas and Propagation, ICAP-91, York, U.K., IEE Pub. No.333, Vol.2, pp.717-720.

25. Cox, J.W.R. and Vongas, 1991, "Calculated and measured radiation characteristics of an HF loop antenna mounted upon a helicopter," IEE International Conference on HF Radio Systems and Techniques, IEE Pub. No.206, Vol.2, pp.73-76.

26. Davies, K., 1990, "Ionospheric Radio," London: Peter Peregrinus Ltd.

27. Ebine, Y. and Yamada, Y., 1991, "A vehicular-mounted vertical space diversity antenna for a land mobile radio," IEEE Trans. on Vehicular Technology, Vol.VT-40, No.2, pp.420-424.

28. Ehrenspeck, H.W. and Poehler, H., 1959, "A new method for obtaining maximum gain from Yagi antennas," IRE Trans. Antennas Propagat., Vol.AP-7, pp.379-386.

29. Eley, A.S., Barr, E.E. and Choi, J.H., 1991, "Near vertical incidence skywave antenna examination," Proceeding of the 5th International Conf. on HF Systems and Techniques, IEE Pub. No.339, pp.79-83.

30. Fanni, A., Marchesi M., Serri, A. and Usai, M., 1997, "A greedy genetic algorithm for continuos variables electromagnetic optimization problems," IEEE Trans. on Magnetics, Vol.33, No.2, pp.1900-1903.

31. Ferguson, T.R. and Balestri, R.J., 1988, "Efficient solution of large moments problems: Wire grid modelling criteria and conversion to surface currents," Applied Computational Electromagnetics Society Journal, Vol.3, No.1, pp.55-81.

32. Fiedler, D.M. and Farmer, E.J., 1996, "Near vertical incidence skywave communication - theory, techniques and validation," Sacramento: Worldradio Books.

33. Finlayson, B.A., 1972, "The method of weighted residuals and variational principles," New York: Academic Press.

34. Goldberg, D.E., 1989, "Genetic algorithms in search, optimization and machine learning," New York: Wiley.

35. Goldberg, D.E., Deb, K. and Clark, J.H., 1991, "Genetic algorithms, noise, and the sizing of populations," Complex Syst., Vol.6, No. 3, pp.333-362.

36. Goodman, J.M., 1992, "HF communications science and technology," New York: Van Nostrand Reinhold.

37. Gotthard O. and Landstorfer, F.M., 1979, "Shaped wire antennas for circular polarisation," IEE International Conference on Antennas and Propagat., IEE Pub. No.248, pp.166-170.

38. Hagn, G.H., 1973, "On the relative response and resolute gain towards the zenith of HF field-experiment antennas measured with an ionospheric sounder," IEEE Trans. on Antennas and Propagat., Vol.AP-21, pp.571-574.

39. Hagn, G.H. and Van Der Laan, J.E., 1970, "Measured relative responses toward the zenith of short whip antennas on vehicles at high frequency," IEEE Trans. on Vehicular Technology, Vol.VT-19, No.3, pp.230-236.

40. Hagn, G.H. and Vincent, W.R., 1974, "Comments on the performance of selected low-power HF radio sets in the Tropics," IEEE Trans. on Vehicular Technology, Vol.VT-23, pp.55-58.

41. Hallen, E., 1938, "Theoretical investigations into the transmitting and receiving qualities of antennae," Nova Acta Regiae Soc. Sci. Upsaliensis, Ser. IV, 11, No.4, pp.1-44.

42. Hansen, R.C., 1975, "Optimum inductive loading of short whip antennas," IEEE Trans. on Vehicular Technology, Vol.VT-24, No.2, pp.21-29.

43. Harrington, R.F., 1961, "Time-harmonic electromagnetic fields," New York: McGraw-Hill Press.

44. Harrington, R.F., 1968, "Field computation by moment methods," New York: Macmillan Press.

45. Haupt, R.L., 1994, "Thinned arrays using genetic algorithms," IEEE Trans. on Antennas and Propagat., Vol.42, No.7, pp.993-999.

46. Haupt, R.L., 1995a, "Comparison between genetic and gradient-based optimization algorithms for solving electromagnetics problems," IEEE Trans. on Magnetics., Vol.31, No.3, pp.1932-1935.

47. Haupt, R.L., 1995b, "Optimum quantised low sidelobe phase tapers for arrays," Electronics Letters, Vol.31, No.14, pp.1117-1118.

48. Haupt, R.L., 1995c, "An introduction to genetic algorithms for electromagnetics," IEEE Antennas and Magazine, Vol.37, No.2, pp.7-15.

49. Holland, J.H., 1973, "Genetic algorithms and the optimal allocations," *SIAM* Journal of Computing, Vol.2, No.2, pp.88-105.

50. Holland, J.H., 1992, "Genetic Algorithm," Scientific American, July, pp.44-50.

51. Jesch, R.L., 1985, "Measured vehicular antenna performance," IEEE Trans. on Vehicular Technology, Vol.VT-34, No.2, pp.97-107.

52. Johnson, J.M. and Rahmat-Samii, Y., 1997, "Genetic algorithms in engineering electromagnetics," IEEE Antennas and Propagat Magazine., Vol.39, No.4, pp.7-25.

53. Jones, E.A. and Joines, W.T., 1997, "Design of Yagi-Uda Antennas using genetic algorithms," IEEE Trans. on Antennas and Propagat., Vol.45, No.9, pp.1386-1392.

54. Katsibas, K.D., Balanis, C.A., Tirkas, P.A. and Birtcher, C.R., 1998, "Folded

loop antenna for mobile hand-held units," IEEE Trans. on Antennas and Propagat., Vol.46, No.2, pp.260-266.

55. King, R.W.P., 1969, "The loop antenna for transmission and reception," in Collin, R.E. and Zucker, F.J., "Antenna theory, Part I," McGraw-Hill, Chap.11, pp.458-482.

56. Landstorfer, F.M., 1976, "A new type of directional antenna," in Antennas Propagat., Soc. Int. Symp. Dig., Vol.11-15, pp.169-172.

57. Landstorfer, F.M., 1977, "Optimization of the directivity of linear and horn-antennas," IEEE Conference on CAD of Electronic and Microwave Circuits and System, Hull, U.K., pp.96-101.

58. Landstorfer, F.M., 1979, "New development in VHF/UHF antennas," IEE Conference Pub. No.169, pp.132-141.

59. Lee,K.S.H., Martin, L. and Castillo, J.P., 1976, "Limitations of wire-grid modelling of a close surface," IEEE Trans. on EMC, Vol.EMC-18, No.3, pp.123-129.

60. Li, Q., Rothwell. E.J., Chen, K.M. and Nyquist, D.P., 1996, "Scattering center analysis of radar targets using fitting scheme and genetic algorithm," IEEE Trans. on Antennas and Propagat., Vol.44, No.2, pp.198-207.

61. Liang, C.H. and Cheng, D.K., 1983, "Directivity optimization for Yagi-Uda arrays of shaped dipoles," IEEE Trans. on Antennas and Propagat., Vol.AP-31, No.3, pp.522-525.

62. Lin, Y.T. and Richmond, J.H., 1975, "EM modeling of aircraft at low frequencies," IEEE Trans. on Antennas and Propagat., Vol.AP-23, No.1, pp.53-56.

63. Linden, D.S. and Altshuler, E.E., 1996, "Automating wire antenna design using genetic algorithms," Microwave Journal, Vol.39, No.3, pp.74-86.

64. Lo, Y.T. and Lee, S.W., 1966, "A study of space tapered arrays," IEEE Trans. on Antennas and Propagat., Vol.AP-14, pp.22-30.

65. Ludwig, A.C., 1987, "Wire-grid modelling of surfaces," IEEE Trans. on Antennas and Propagat., Vol.AP-35, No.9, pp.1045-1048.

66. Maslin, N.M., 1987, "HF communications - a systems approach," London: Pitman.

67. Michielssen, E., Sajer.,J.M., Ranjithan, S. and Mittra, R., 1993, "Design of lightweight, broad-band microwave absorbers using genetic algorithms," IEEE

Trans. on Microwave Theory and Techniques, Vol.41, No.6/7, pp.1024-1031.

68. Miller, E.K. and Burke, G.J., 1992, " Low frequency computational electromagnetics for antenna analysis," Proceedings of the IEEE, Vol.80, No.1, pp.24-43.

69. Mitchell, M., 1997, "An introduction to genetic algorithms," London: MIT Press.

70. Mitchell, R.J., Chambers, B. and Anderson, A.P., 1996, "Array pattern synthesis in the complex plane optimised by a genetic algorithm," Electronics Letters, Vol.32, No.20, pp.1843-1845.

71. Moore, J. and Pizer, R., 1986, "Moment methods in electromagnetics," Letchworth: Research Studies Press.

72. Murray, K.P., 1993, "The design of antenna systems on complex structures using characteristic modes," PhD Thesis, University of Liverpool, U.K.

73. Newman, E.H. and Kingsley, K., 1991, "An introduction to the method of moments," Computer Physics Communications, Vol.68, pp.1-18, North-Holland.

74. Nishikawa, K., 1984, "Effect of automobile body and earth on radiation patterns of antennas for FM radio," Trans. IECE of Japan, Vol.E67, No.10, pp.555-562.

75. Nishikawa, K. and Asano Y., 1986, "Vertical Radiation patterns of mobile antenna in UHF band," IEEE Trans. on Vehicular Technology, Vol.VT-35, No.2, pp.57-62.

76. Nyquist, D.P, and Chen, K.M., 1968, "The traveling-wave linear antenna with nondissipative loading," IEEE Trans. on Antennas and Propagat., Vol.AP-16, No.1, pp.21-31.

77. Perez, R., 1998, " Wireless communications design handbook," Vol.3, New York: Academic Press, p.201.

78. Pocklington, H.C., 1897, "Electrical oscillations in wire," Cambridge Philos. Proc., Vol.9, pp.324-332.

79. Poggio, A.J. and Miller, E.K., 1973, "Integral equation solutions of three dimensional scattering problems," in Computational Techniques for Electromagnetics, R. Mittra, Ed. Ch.4, Pergamon, Elmsford, New York.

80. Quagliarella, D., Periaux, J., Poloni, C. and Winter, G., 1998, "Genetic algorithms and evolution strategy in engineering and computer science," New York: John Wiley & Sons.

81. Rao, B.L.J., Ferris, J.E. and Zimmerman W.E., 1969, "Broadband characteristics

of cylindrical antennas with exponentially tapered capacitive loading," IEEE Trans. on Antennas and Propagat., Vol.AP-17, No.17, pp.145-151.

82. Reilly, M.H., 1991, "Upgrades for efficient three-dimensional ionospheric ray tracing: Investigation of HF near vertical incidence skywave effects," Radio Science, Vol.26, No.4, pp.971-980.

83. Richmond, J.H., 1966, "A wire-grid model for scattering by conducting bodies," IEEE Trans. on Antennas and Propagat., Vol.AP-14, pp.782-786.

84. Richmond, J.H., 1974, "Radiation and scattering by thin-wire structures in complex frequency domain," NASA Report CR-2396, Contract NGL36-008-138, Hampton, Virginia.

85. Riolo, R.L., 1992, "Survival of the fittest bits," Science American, July, pp.89-91.

86. Sarabandi, K. and Li, E.S., 1997, "Characterization of optimum polarization for multiple target discrimination using genetic algorithms," IEEE Trans. on Antennas and Propagat., Vol.45 No.12, pp.1810-1817.

87. Smith, P.H., 1942, "L-type impedance transformation circuits," Electronics, Vol.15, pp.130-133.

88. Stutzman, W.L. and Thiele, G.A., 1981, "Antenna theory and design," New York: Wiley, p.226.

89. Taboada, J.M., Rodriguez, J.L., Landesa, L. and Obelleiro, F., 1999, " Automatic wire-grid generation for electromagnetic analysis of arbitrary-shaped conducting bodies by NEC," Computer Applications in Engineering Education, Vol.7, No.1, pp.31-43.

90. Taflove, A., 1995, "Computational electromagnetics: The finite-Difference Time-Domain Method," Boston, MA: Artec.

91. Taniguchi, Y., Hirano, M. and Arai, A., 1985, " The application of numerical methods to vehicle antennas," International Symposium on Antennas and Propagation, ISAP-85, Kyoto, Japan. Conf. Proceedings, pp.441-444.

92. Tennant, A., Dawoud, M.M. and Anderson, A.P., 1994, "Array pattern nulling by element position perturbations using a genetic algorithm," Electronics Letters, Vol.30, No.3, pp.174-176.

93. Thiele, G.A., 1969, "Analysis of Yagi-Uda type antennas," IEEE Trans. on Antennas and Propagat., Vol.AP-17, pp.24-31.

94. Thiele, G.A. and Ekelman, E.P., 1980, "Design formulas for vee dipoles," IEEE

Trans. on Antennas and Propagat., Vol.AP-28, No.4, pp.588-590.

95. Thollon, F. and Burais, N., 1995, "Geometrical optimization of sensors for eddy currents non destructive testing and evaluation," IEEE Trans. on Magnetics, Vol.31, No.3, pp.2026-2031.

96. Tirkas, P.A. and Balanis, C.A., 1992, "Finite difference time domain method for antenna radiation," IEEE Trans. on Antennas and Propagat., Vol.40, No.3 pp.334-340.

97. Trueman, C.W. and Kubina, S.J., 1991, "Automated radius calculation for wire-grid models," ACES 7th Annual Review of Progress, Monterey, U.S.A..

98. Uler, G.F., Mohammed, O.S. and Koh, C.S., 1994, "Utilizing genetic algorithm for the optimal design of electromagnetic devices," IEEE Trans. on Magnetics, Vol.30, No.6, pp.4296-4298.

99. Virani, H.R., 1988, "Electrically small antennas," Journal of the Institution of Electronic and Radio Engineers, Vol.58, No.6, pp.266-274.

100. Weile, D.S. and Michielessen, E., 1997, "Genetic algorithm optimization applied to electromagnetics: a review", IEEE Trans. on Antennas and Propagat., Vol.45, No3, pp.343-353.

101. Weile, D.S., Michielessen, E. and Boag, A., 1996, "Optimization of broad-band loaded wire antennas in real environments using genetic algorithms", Proc. 12th Ann. Rev. Progress Appl. Computat. Electromagn., Monterey, CA, pp.726-733.

102. Weile, D.S., Michielessen, E. and Goldberg D.E., 1996, "Genetic algorithm design of Pareto optimal broadband microwave absorbers", IEEE Trans. on Electromagnetic Compatibility, Vol.38, No3, pp.518-525.

103. Wheeler, H.A., 1975, "Small Antennas," IEEE Trans. on Antennas and Propagat., Vol.AP-23, No.4, pp.462-469.

104. Whiteley, L.D., 1993, "Foundations of genetic algorithms • 2," California: Morgan Kaufmann Publishers.

105. Willey, R.E., 1962, "Spacing tapering of linear and planar arrays," IRE Antennas Propagat., Vol.AP-10, pp.369-377.

106. Yagi, H., 1928, "Beam transmission of ultra-short waves," Proceedings of IRE, Vol.16, p.715.

107. Yan, Keen-Keong and Lu, Yilong, 1997, "Sidelobe reduction in array-pattern synthesis using genetic algorithm," IEEE Trans. on Antennas and Propagat., Vol.45, No7, pp.1117-1121.

108. Yeh, Y.S. and Mei, K.K., 1967, "Theory of conical equiangular spiral antennas," Part I - Numerical Techniques, IEEE Trans. on Antennas and Propagat., AP-15, No.5, pp.634-639.

Appendix A: Additional Reported Results

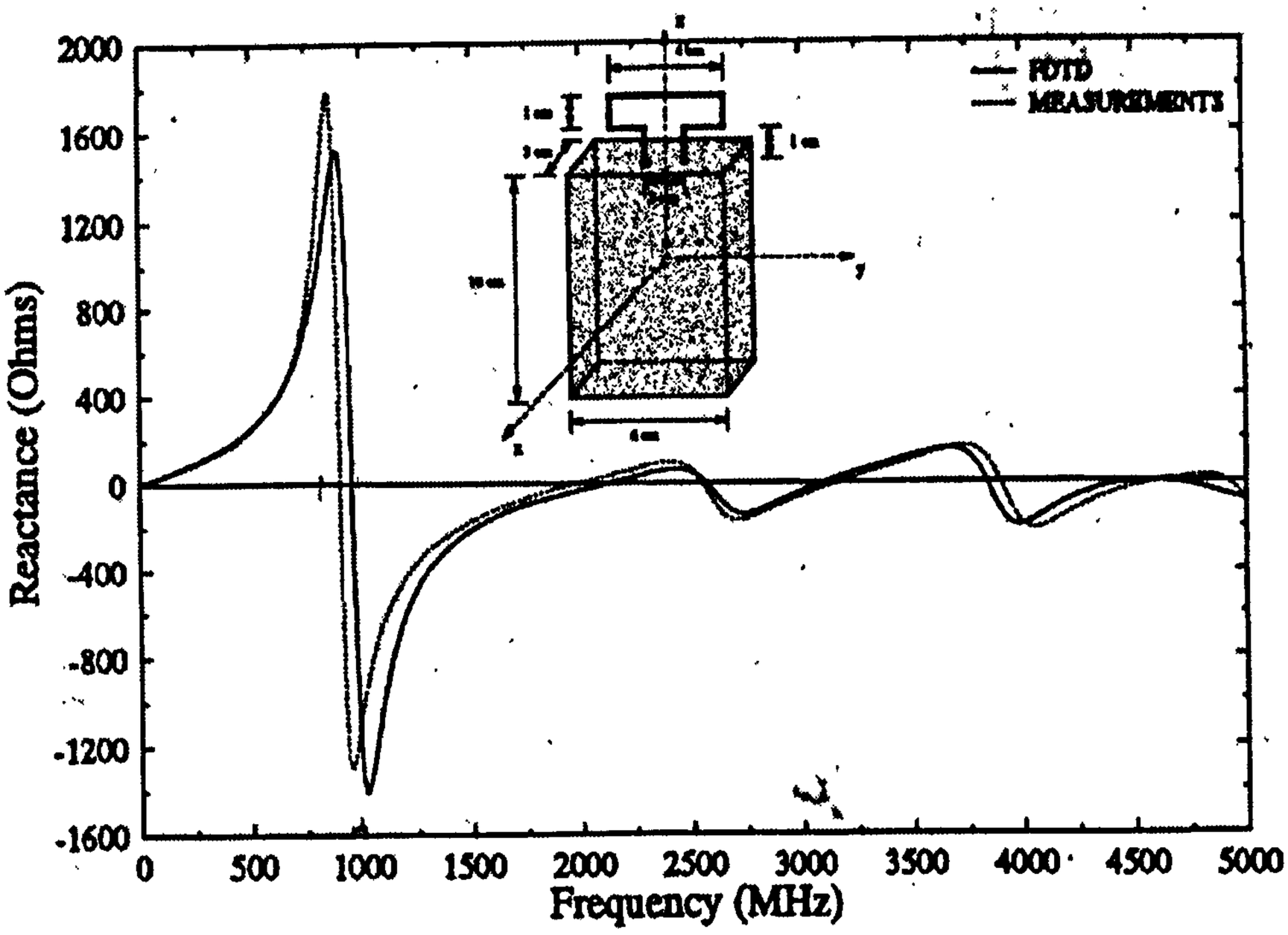


Figure A.1 Input reactance of wire folded loop antenna mounted on a conducting box. (Source: Katsibas et al. 1998)

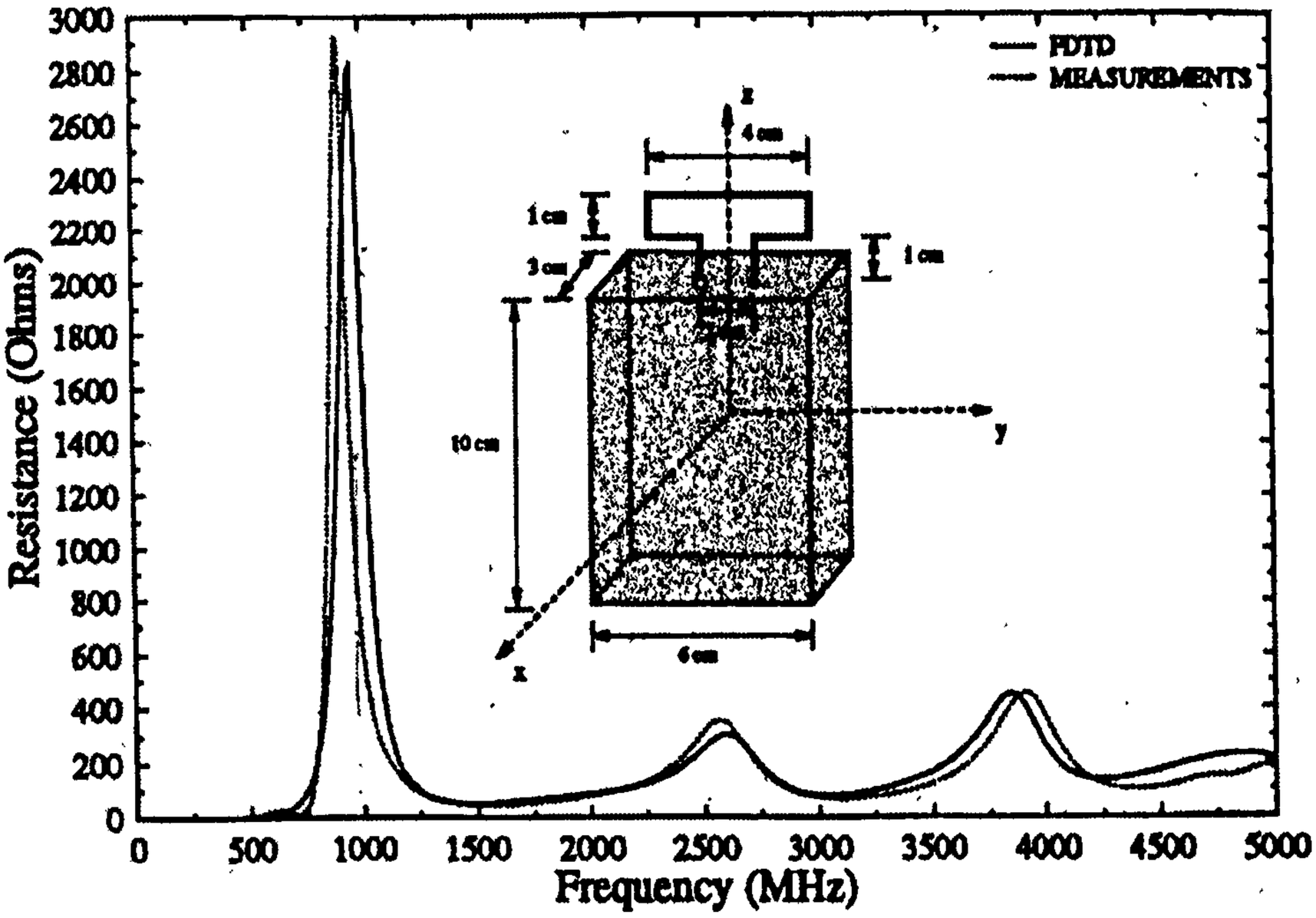


Figure A.2 Input resistance of wire folded loop antenna mounted on a conducting box. (Source: Katsibas et al. 1998)

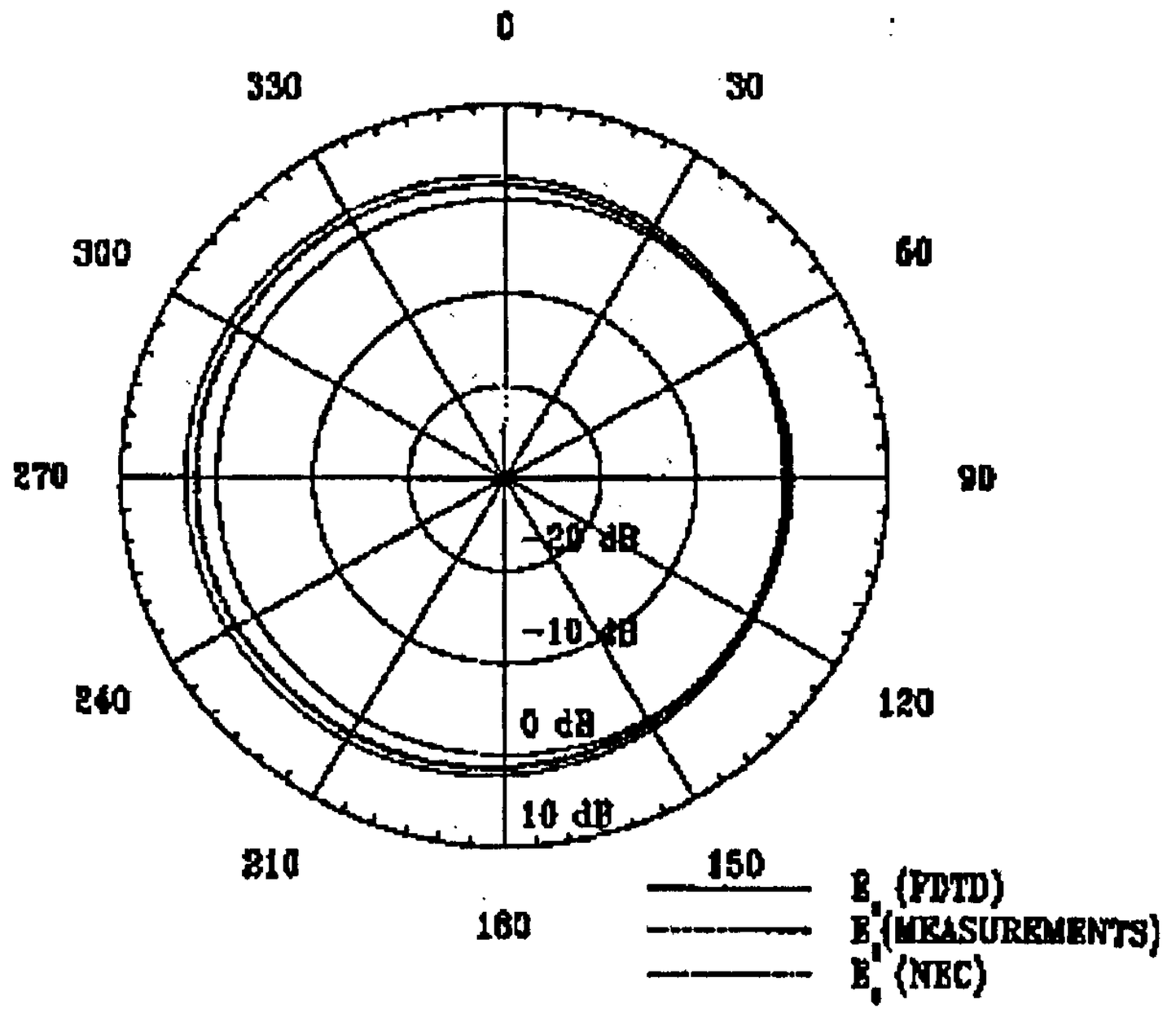


Figure A.3 Azimuth plane patterns on xy-plane of wire folded loop antenna mounted on a conducting box.
(Source: Katsibas et al. 1998)

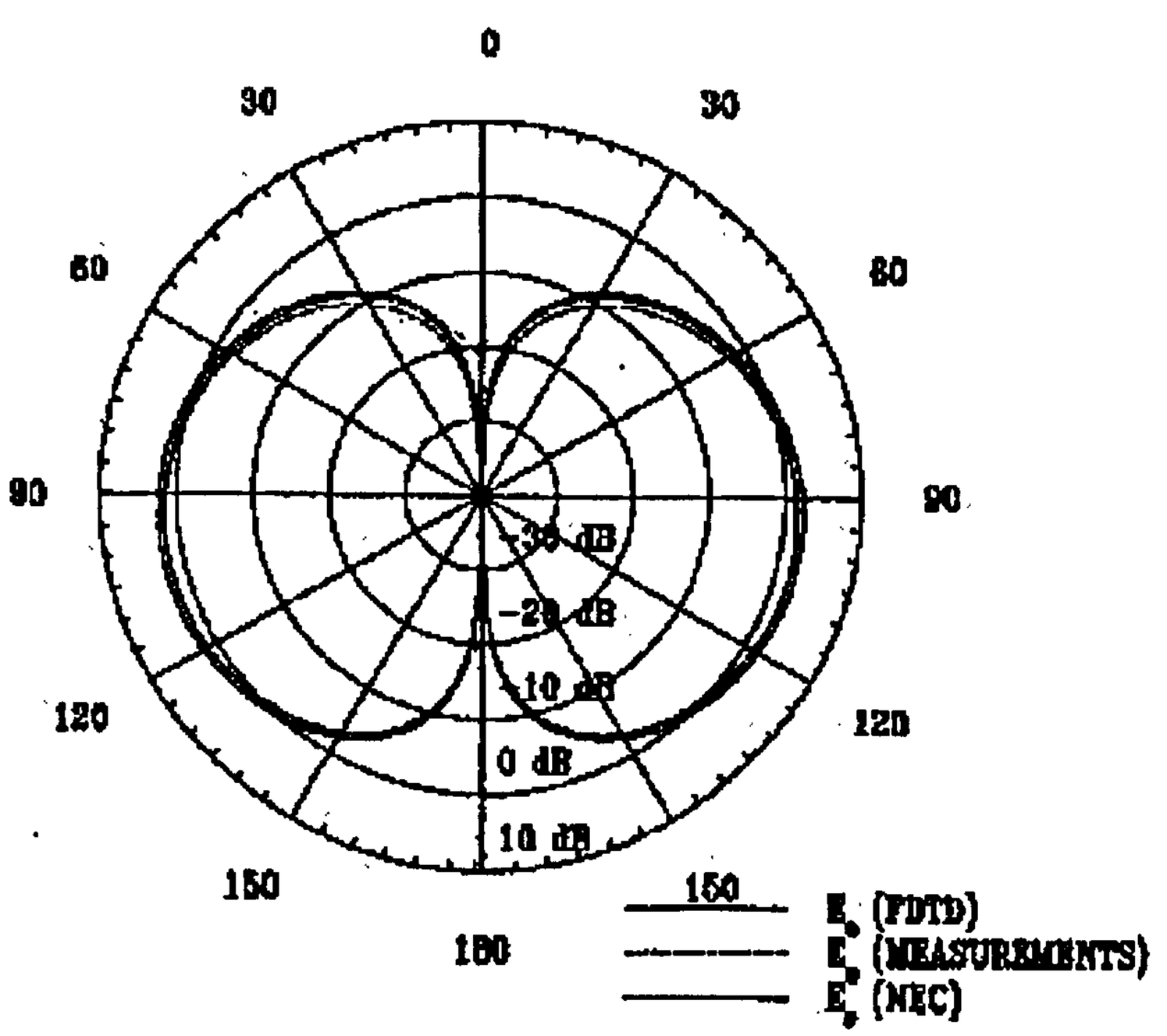


Figure A.4 Elevation plane patterns on xz-plane of wire folded loop antenna mounted on a conducting box.
(Source: Katsibas et al. 1998)

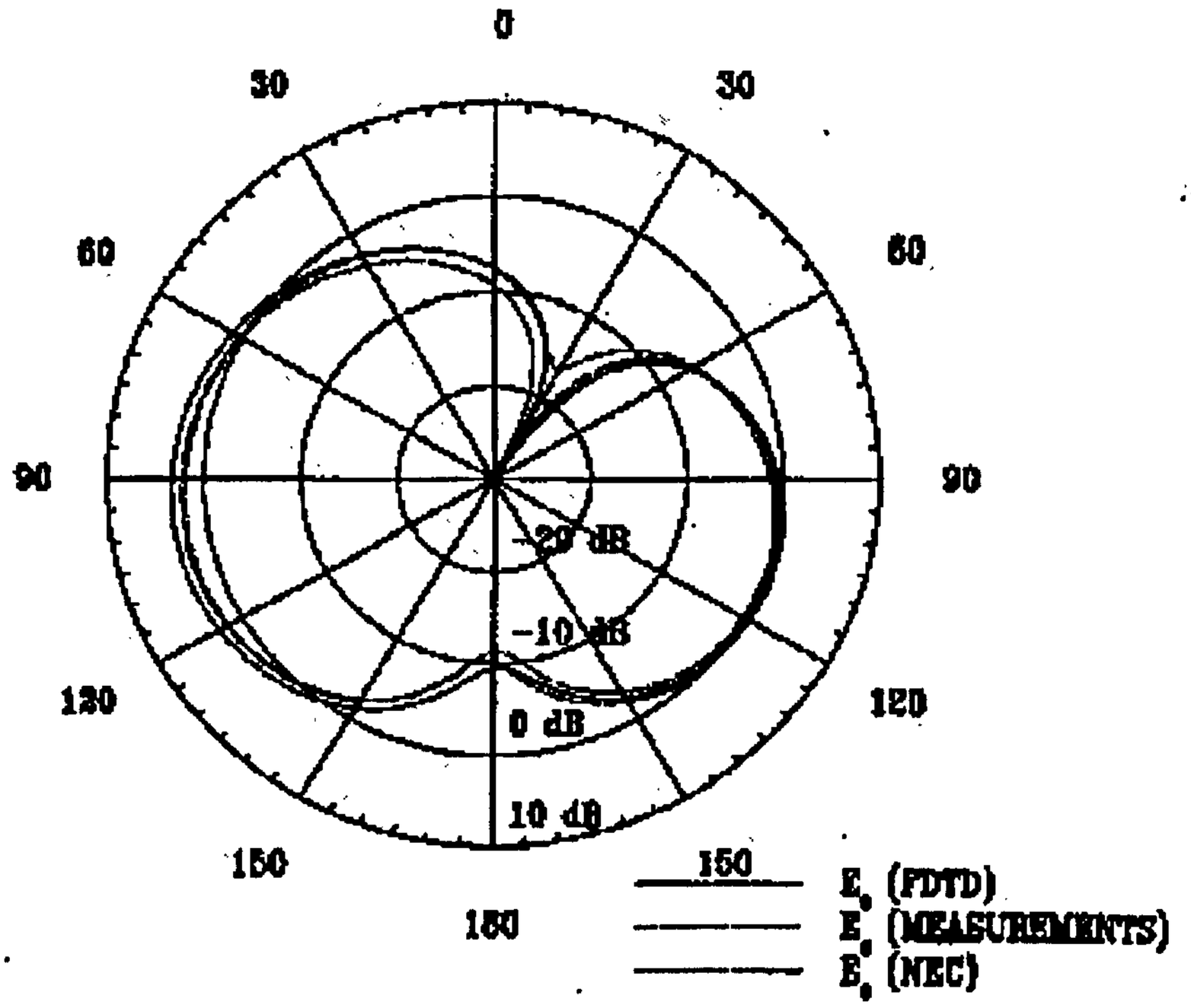


Figure A.5 Elevation plane patterns on yz-plane of wire folded loop antenna mounted on a conducting box.
(Source: Katsibas et al. 1998)

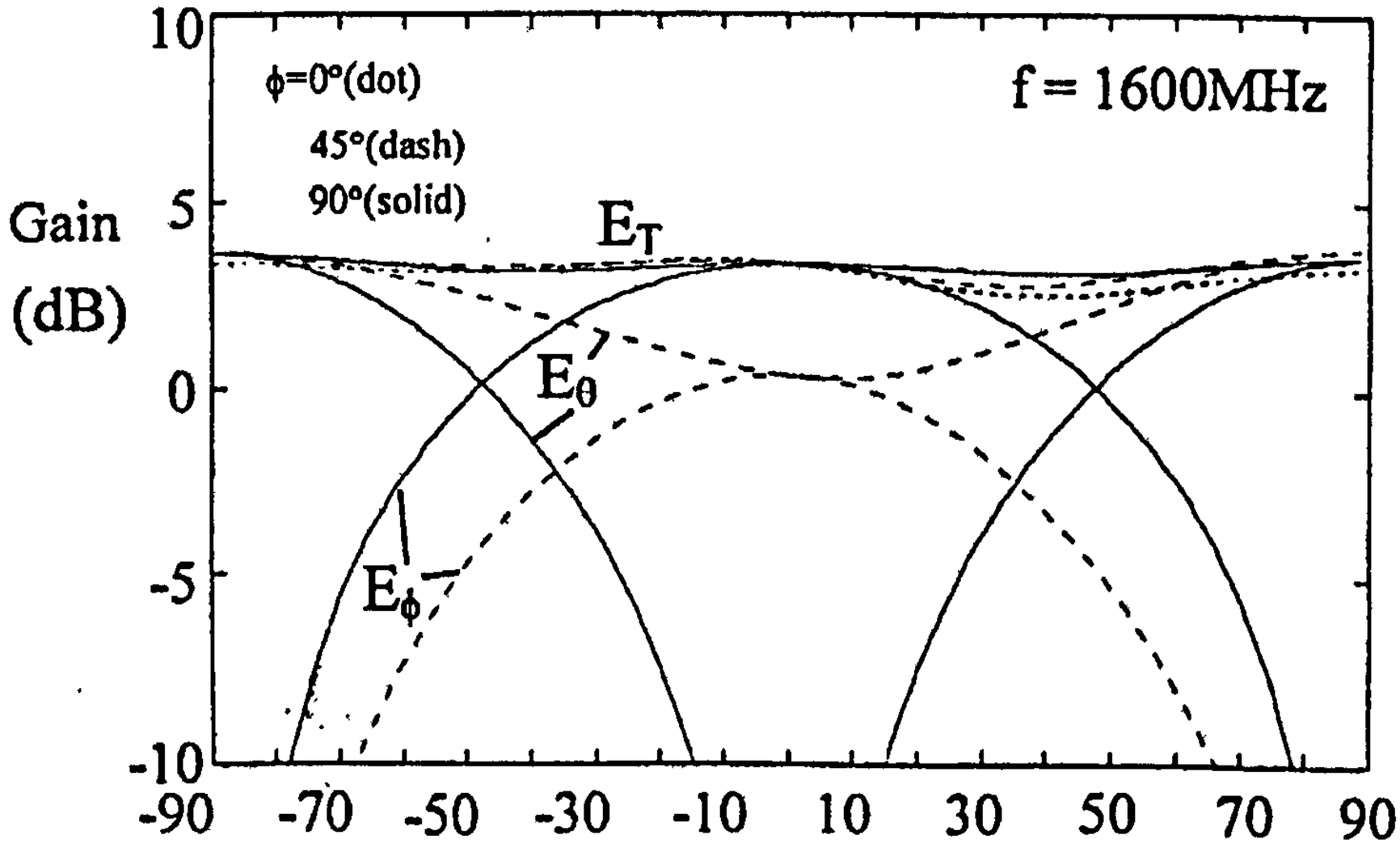


Figure A.6 The computed E_θ, E_ϕ, and E_T fields in the θ plane for φ=0°, 45°, and 90°, at 1600 MHz. (Source: Altshuler and Linden 1997)

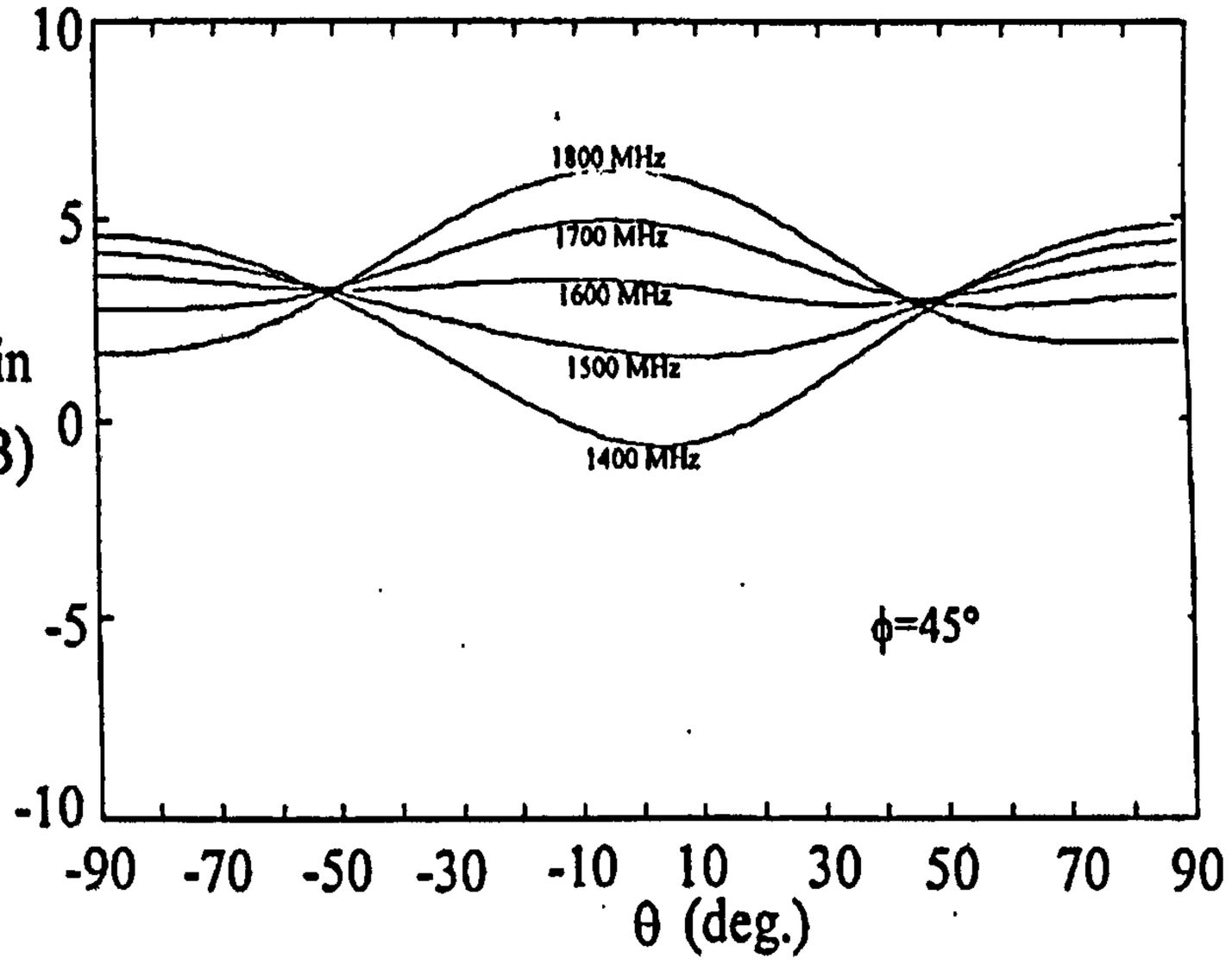


Figure A.7 The computed E_T fields in the θ plane for ϕ=45°, at frequencies from 1400 to 1800 MHz. (Source: Altshuler and Linden 1997)

Appendix B: List of Publications

1. W.C. Liu and B.A. Austin, "Genetic algorithms in antenna design", The 13th National Radio Science Colloquium, Leicester, U.K., 1997.

2. B.A. Austin and Wen-Chung Liu, "Genetic algorithm optimization of vehicle-mounted loop antenna for NVIS applications", Electronics Letters, Vol.35, No.4, pp.252-253, 1999.

3. B.A. Austin and Wen-Chung Liu, "An optimized shaped Yagi-Uda array using the genetic algorithm", National Conference on Antennas and Propag., York, U.K., March, 1999, IEE Pub. No. 461, pp.245-248.

4. Wen-Chung Liu and B.A. Austin, "An optimized shaped parasitic array using the genetic algorithm", IEE Proceedings Microwaves, Antennas and Propagation", 1999, accepted.

Genetic Algorithms in Antenna Design

W.C.Liu and B.A.Austin
Department of Electrical Engineering and Electronics
University of Liverpool
Liverpool, L69 3BX
United Kingdom

Genetic Algorithms (GAs), used in conjunction with the Numerical Electromagnetics Code (NEC), are a revolutionary new approach to wire antenna design. Until now traditional techniques used the method of moments (MoM) to determine the current distributions on each wire of a selected antenna configuration and, from these, the radiation characteristics are calculated and compared with a target specification. The procedure is iterative and both experience and some appropriate candidate configuration are usually required to ensure success. By contrast, the GA is based on the processes of natural selection and the "survival of the fittest" in an evolutionary system which leads iteratively to that desired specification with minimal foresight or pre-conditioning on the part of the designer [1]. In using the GA all design parameters are encoded into a string of binary bits known, unsurprisingly, as a chromosome, while each parameter is represented by several bits - the genes. The genes are the basic building blocks of the GA and map directly to the associated real, physical properties of the antenna. The immense power of the technique is its ability to satisfy a performance criterion without any *a priori* knowledge of candidate configurations, and the facility for finding the global optimum result.

This presentation will describe the philosophy of the GA in general and will then use it to design a form of Yagi Uda array where maximum gain with a fixed number of elements is the performance criterion. To reduce the search-space and hence computing time the antenna size, element spacing and diameter are preselected while element shape is left as a design variable available for random manipulation by the GA. In addition the maximum tolerable beamwidth is set at + and - 15° with maximum radiation *in the* $\phi = 0°$ direction. The effects of such GA features as the objective function, pairing type, crossover *point and mutation probability* will be discussed. Finally the GA-predicted results are compared with those of Liang [2] and Landstorfer [3], for the same problem, and are shown to be in excellent agreement with them; though with markedly different geometrical configurations.

The GA is remarkably robust and is able to design optimised antennas, of unexpected shapes, almost automatically.

References

[1] Goldberg, D.E., *Genetic Algorithms in Search, Optimization and Machine Learing*, New York, Addison-Wesley, 1989.

[2] Liang, C.H. and Cheng, D.K. *IEEE Trans. Antennas and Propagation*, **AP-31**, 3, May 1983, 522-525.

[3] Landstorfer, F.M. *IEEE Conference on CAD of Electronic and Microwave Circuits and System*, Hull, U.K., 1977, 96-101.

Genetic Algorithm optimisation of a vehicle-mounted loop antenna for NVIS applications

B.A. Austin and Wen-Chung Liu

The performance of a vehicle-mounted loop antenna for near-vertical incidence skywave applications is optimised by using a Genetic Algorithm in conjunction with the Method of Moments. Optimum loop current distribution is crucial to performance and this is achieved by appropriate reactive loading.

Introduction: We describe in this letter a new vehicle-mounted loop antenna configuration that yields optimum performance when radiating towards the zenith within the 2-10 MHz range typical of most near-vertical incidence skywave (NVIS) applications [1]. Previous work [2] showed that such performance could be achieved by driving a vertically-mounted loop with two in-phase sources positioned at the bases of the two vertical members of a rectangular loop antenna, as shown in Fig.1, in contrast to that from the usual single-source loop fed at A. However, such a dual-source method of feeding an antenna is cumbersome and presents many practical problems. We have therefore used a Genetic Algorithm (GA) in conjunction with the Method of Moments code NEC2 to design an optimised antenna system [3] that requires just a single source and a suitably located, lumped capacitive load to achieve the same performance.

Implementation and results: The target of the G.A. was the radiation of maximum power within a cone, centred on the zenith, with a half-angle of 45°, as previously defined in [2]. No restriction was placed in the GA on either the sign or the position of the reactive load that was intended to change the current distribution on the antenna so that maximum power would always be radiated within that target cone. A steady-state GA with a total of 60 chromosomes in each generation was used. Furthermore, in each chromosome, 6 and 16 bits, respectively, were used to represent the position and value of the load, and the GA was linked to the NEC2 code to simulate the performance of the rectangular loop of 5cm diameter aluminium tube mounted vertically along the longest diagonal dimension (4.2m) of a typical long-wheelbase vehicle.

The GA/NEC2 combination produced an optimum result when the load was a capacitor positioned at point B, the point previously occupied by the second source. The GA also determined the value of this capacitor at all frequencies within the NVIS range to ensure maximum zenithal radiation. Fig.2 shows the results obtained for three cases: (i) an unloaded loop driven by a single source at A (the conventional approach); (ii) dual, in-phase sources at A and B, and (iii) a single source at A plus a series capacitor at position B. It will be noted that the very small electrical dimensions of the loop at the low-frequency end of the NVIS spectrum is the factor limiting performance below approximately 4 MHz. However, both the dual-source loop and the capacitively-loaded loop produce markedly improved performance (better than 3.4 dB at 10 MHz, in fact) above that frequency, with there being no essential difference between them. This result is most significant because it shows that a single source at one end of such a rectangular loop, and an appropriate capacitor at the other, will produce the optimum current distribution in the loop for NVIS applications. Fig. 3 gives the capacitor values calculated for the loop of the dimensions indicated in Fig.1 between 2 and 10 MHz.

Mechanism of operation: In essence, below 4 MHz, the current is uniformly distributed around the loop because it is electrically very small. Above that frequency however, the amplitude of the current is no longer constant but varies with position around the loop as the frequency increases towards 10 MHz at which point the loop perimeter is 0.4λ, which no longer satisfies the small loop criterion of constant current [4]. The consequence of this is that instead of maximum power being radiated upwards it begins to shift away from the zenith which leads to the performance shown for case (i) in Fig. 2. The inclusion of the appropriate capacitor at its optimum position (B), so modifies the current distribution that

maximum radiation always occurs towards the zenith at all frequencies within the NVIS range. This capacitor could be made motor-driven making frequency agility a fairly simple matter.

Conclusions: A practical vehicle-mounted loop antenna, which can produce a marked performance improvement compared to the conventional, single-source configuration for the NVIS applications, is achieved by a GA/MM technique.

References

1. REILLY, M.H.: 'Upgrades for efficient three-dimensional ionospheric ray tracing: Investigation of HF near vertical incidence sky wave effects', *Radio Science*, July-August 1991, **26**, (4), pp. 971-980.

2. AUSTIN, B.A., and MURRAY, K.P.: 'The application of Characteristic-Mode Techniques to Vehicle-Mounted NVIS Antennas', *IEEE Antennas and Propag. Magazine*, 1998, **40**, (1) pp. 7-21

3. WEILE, D.S., and MICHIELSSEN, E.: 'Genetic algorithm optimization applied to electromagnetics: a review', *IEEE Trans. Antennas Propag.*,1997, **45**, (3) pp. 343-353

4. KING, R.W.P.: 'The Loop Antenna for Transmission and Reception', Chap. 11 in Collin and Zucker "*Antenna Theory-Part I* "(Press, Place, 1969), pp. 458-482.

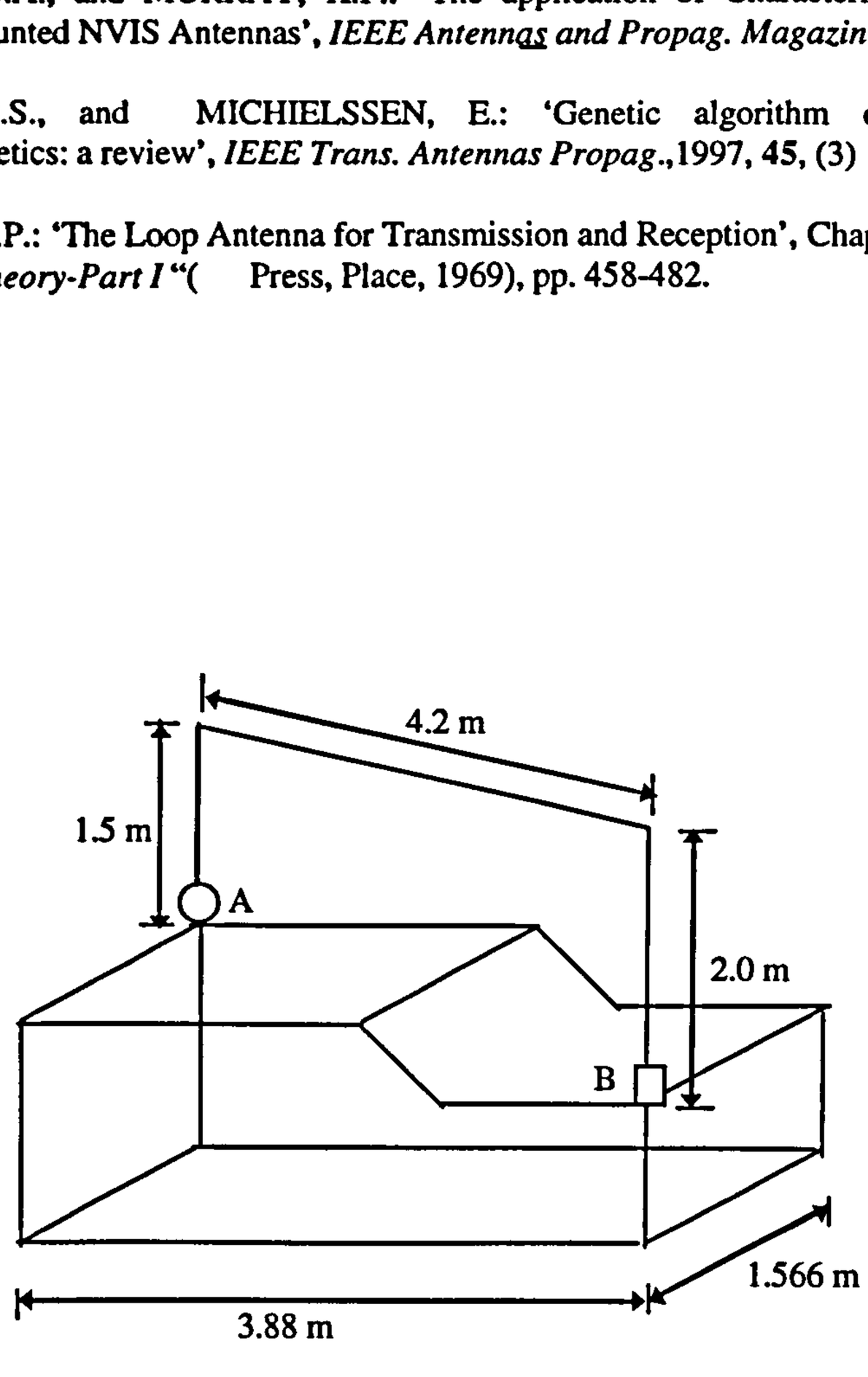


Fig. 1 Configuration of the vehicle-mounted loop antenna

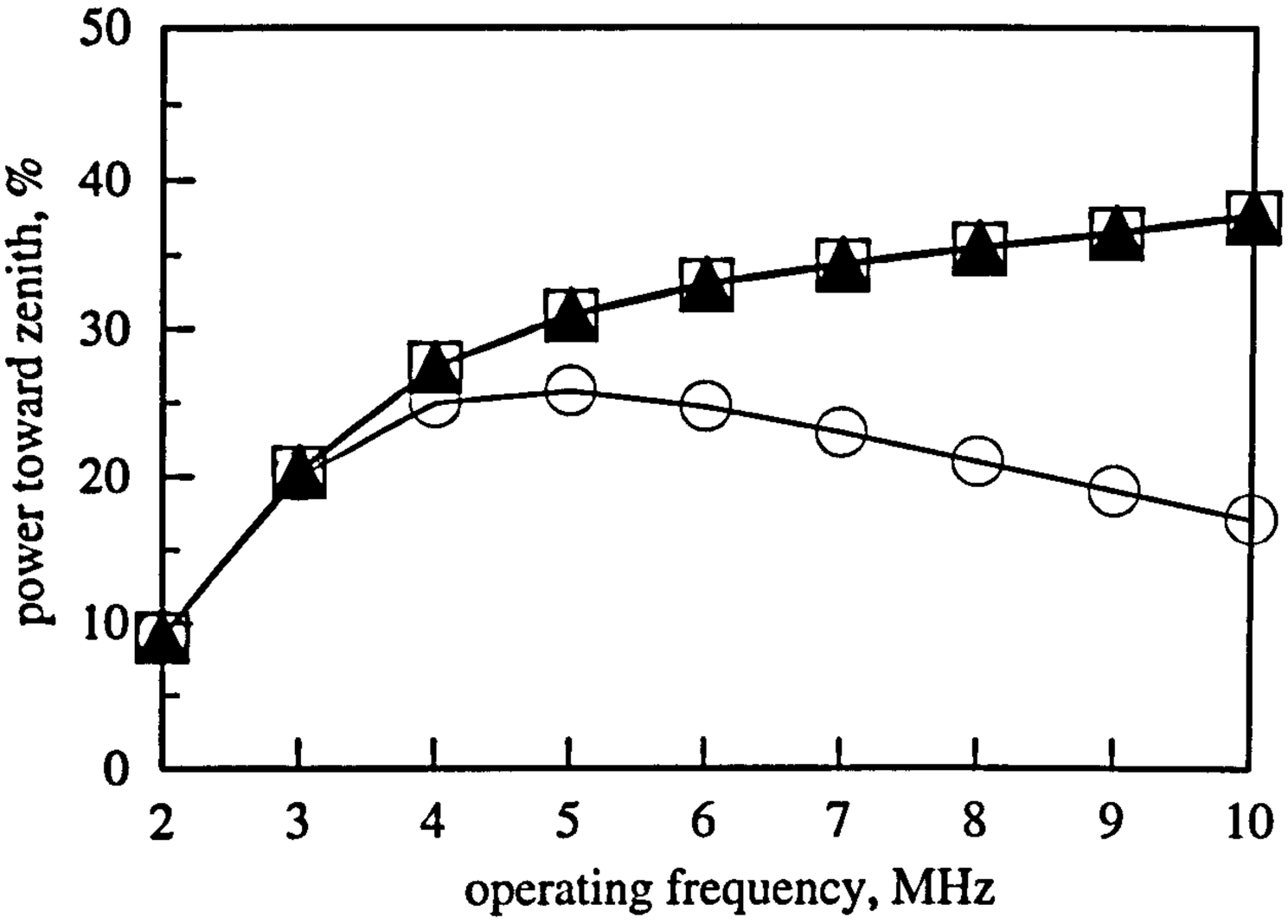


Fig. 2 Percentage of power radiated toward the zenith

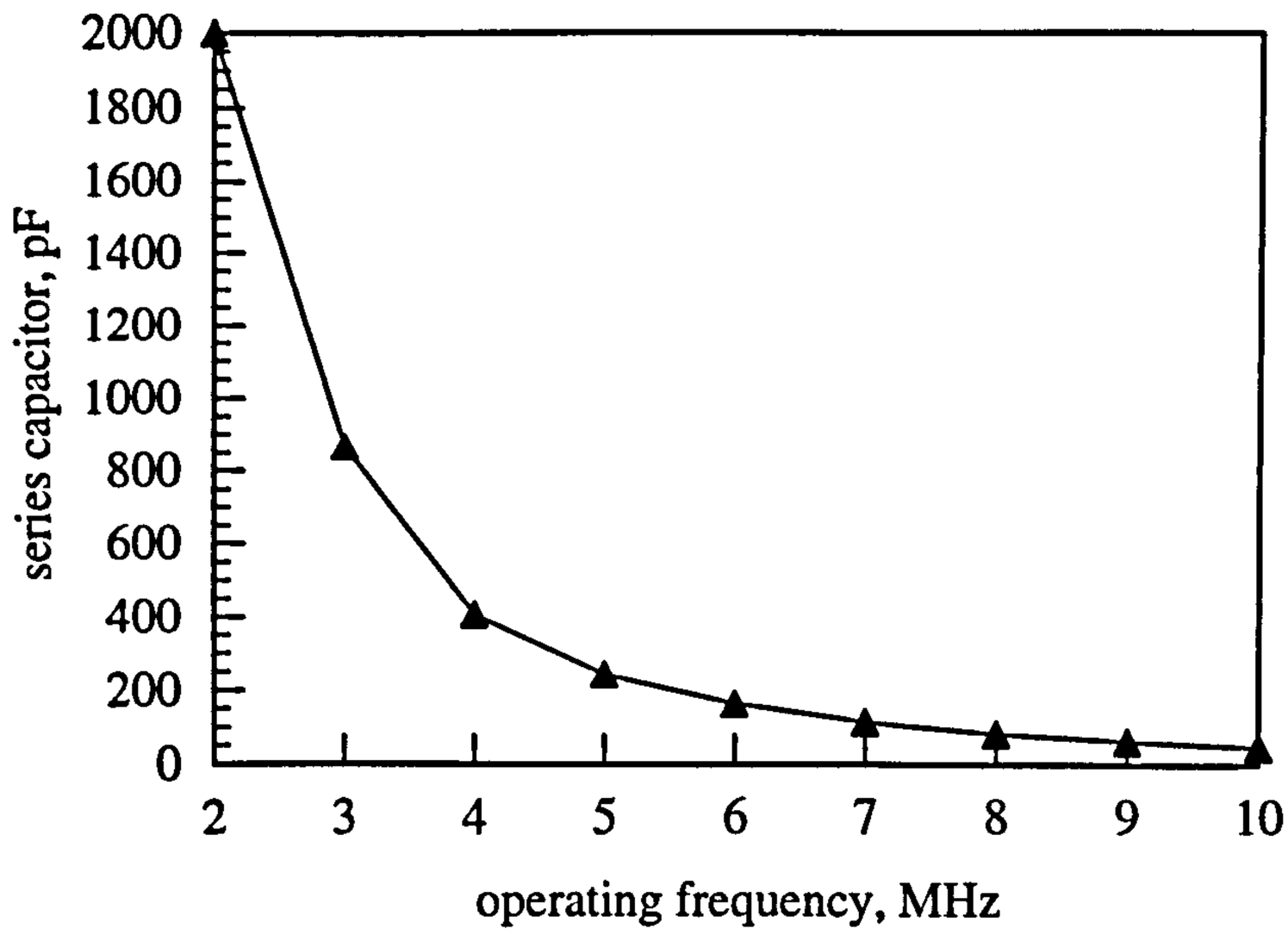


Fig. 3 Capacitor values for optimum radiation toward the zenith

AN OPTIMIZED SHAPED YAGI-UDA ARRAY USING THE GENETIC ALGORITHM

B A Austin and Wen-Chung Liu

Abstract: The performance of a shaped, three-element Yagi-Uda antenna is improved by using a genetic algorithm(GA) in conjunction with the method of moments code NEC2 for maximum gain. To carry out the scheme of this algorithm, the length, wire angles and spacing of the elements are all included in the design parameters and encoded by a string of binary bits. The results show that a simple, Vee-shape, Yagi-Uda array can achieve markedly increased directivity without the need for elaborate element curvature.

INTRODUCTION

Genetic Algorithms (GAs), used in conjunction with the Numerical Electromagnetics Code (NEC), are a revolutionary new approach to wire antenna design and optimization (1,2). Until recently traditional techniques used the method of moments (MoM) to determine the current distributions on each wire of a selected antenna configuration and, from these, the radiation characteristics were calculated and compared with a target specification. The procedure is iterative and both experience and some appropriate candidate configuration are usually required to ensure success. By contrast, the GA is based on the processes of natural selection and the "survival of the fittest" in an evolutionary system which leads iteratively to that desired specification with minimal foresight or pre-conditioning on the part of the designer (3). In using the GA all design parameters are encoded into a string of binary bits known, unsurprisingly, as a *chromosome*, while each parameter is represented by several bits – the *genes*. The *genes* are the basic building blocks of the GA and map directly to the associated real, physical properties of the antenna. The immense power of the technique is its ability to satisfy a performance criterion without any *a priori* knowledge of candidate configurations, and the facility for finding the global optimum result.

The objective of this work is to use the GA to synthesize a shaped, three-element Yagi-Uda array with optimal gain. Previous work (4,5) has produced maximum gain by fairly complex shaping of the three elements of the array. Essentially vee-shaped elements were shown to be optimum; however the requirement for peculiar curvature adds to manufacturing costs. In this work the vee-shape array is pre-selected and both the gain and front-to-back ratio (F/B) are considered in the objective function to ensure good antenna performance. During the iterations the method of moments code NEC2 was used to calculate the gain, F/B, input impedance, beamwidth etc. of each antenna designed by the GA.

APPROACH

Figure 1 shows the shape of the objective antenna. It is symmetrical and the feed point is placed at the origin on the x-y plane. The radius of each element is 0.0025λ. Thus the lengths, spacing and included angle of the elements are left as the design variables for random manipulation by the GA. The range of ℓ_i was allowed to vary between 0.3λ and 1.1λ; d_i between 0.1λ and 0.5λ; and the included half-angle, ϕ_i, between 10° and 90°. Figure 2 is a flow chart of the GA used in this work. It begins by encoding the design parameters into a string of binary bits called a *chromosome*. Six, five and seven bits respectively were chosen for the variables: element length, element spacing and included angle. Each antenna therefore requires a *chromosome* with $(6\times3)+(5\times2)+(7\times3)=49$ bits (or *genes*) to store the real configuration. After the parameter encoding an initial population of 90 candidate antennas were created by simple randomisation. The randomness makes it possible to explore a broad population of possible solutions in the entire search space. Generally, the good initial *chromosomes* will help the algorithm to achieve an optimum result quickly. On commencement of the iterations a mathematical function was applied to evaluate the individual 'cost' of each created antenna. This function is always set according to the search objective and is called the objective function or the cost function. Clearly the associated cost

of each antenna will decide its fitness for survival into the next generation. The goal here is to obtain both high gain and F/B so the objective function must force all new configurations towards this target. It therefore consists of three components: gain, F/B and the maximum tolerable beamwidth set at + and - 15° with the maximum radiation in the $\phi = 0°$ direction. In view of the wide variation of initial values obtained, each of the three parameters was normalized by an appropriate factor (6). We selected Gain = 13 dBi and F/B = 20 dB as goals since they exceeded typical results reported for similar configurations reported in the literature, while the radiated power in the angle ($-15° \le \phi \le 15°$), is normalized by the total radiated power. Numerical experiments showed that the three normalized terms made very different contributions to the value of the objective function. To avoid any of them from dominating the iteration process, each was weighted by an associated constant (6). In this example the improvement in gain and the radiated power concentrated within the specified angle changed much more slowly than did the F/B so they must be more heavily weighted. Thus values of 24, 18, 1 respectively were selected from a number of preliminary runs. The objective function was then:

$$\text{Objective function} = 24\frac{\text{Gain}}{13} + 18\frac{\sum_{-15°\le\phi\le15°} P(\phi)}{\sum_{-180°\le\phi\le180°} P(\phi)} + \frac{F/B}{20}$$

A stop criterion was set in the next step. It is always chosen according to the specific features of the problem. In this case, it was set arbitrarily to operate when the objective function is asymptotic to its maximum value for, at least, thirty iterations. If the procedure does not stop, the *chromosomes* are then ranked according to their respective costs. The higher the cost the better the chance of survival to generate the new *chromosomes*. Based on the philosophy of natural selection or "survival of the fittest", the top one-third (30 *chromosomes*) of the ranking list are kept – the rest are discarded. Two distinct operations applied in the mating process – select two survivors for mating or pairing, and then exchange the *genes* between the two paired *chromosomes*. The first operation can, of course, be achieved by many different schemes – the roulette wheel selection is the famous one. In order to generate diverse offspring every two alternative *chromosomes* of the survivors were mated as shown in Figure 3a. The second operation is to select the loci (or crossover points) to swap the *genes* of the parents. Figure 3b shows the two-point crossover process we used. Two random points divided the gene string of each parent into three sections or two sections if the two points were the same. After swapping *genes*, each of the offspring will inherit either both ends or occasionally only one end of the gene string of its parent. This is unlike the general single-point crossover in which only one end of the gene string of each parent will be left with the offspring. Each pairing produced two offspring resulting in a total of 60 after the mating process is complete. This is a so-called steady state GA because the number of new generation parents and offspring is retained unchanged. The probability of the algorithm converging on a local optimum is avoided by selecting 5% of the *chromosomes* randomly from each generation and then inverting two bits in each *chromosome*. Thus all the processes are manipulated and the evaluation of each associate cost of the new generation begins again. The algorithm runs through these processes iteratively until it converges.

RESULTS

As expected there is considerable fluctuation initially until steady state is reached. This is due to the large difference between the initial candidate antennas which were generated randomly. This process ensures that the best *chromosome* of the new generation is very different to that of the previous one. After sufficient runs, the relationship between parents is very close so the changes in the new paired *chromosomes*, except for those which underwent mutation, will be small. Table 1 shows the final results after 102 iterations. Finally, we compare the geometry of this GA - optimized antenna with the antennas which resulted from the optimizations of Liang (4) and Landstorfer (5). Firstly, the length of the elements were all close to 1.5λ as in (4) and (5). Secondly, the general vee-shape of the array is another common characteristic.

However, the GA technique produced an optimized array of three, linear elements with included angles of the order of 100° in marked contrast to the elaborate element curvature that was found to be necessary in (4) and (5).

Liang (4) devised an antenna with maximum gain of 11.8 dBi while Landstorfer (5) obtained 11.5 dBi. Figure 4 is a plot of the gain against the azimuth angle for our optimized array. The maximum gain of 12.1 dBi exceeds that of both while the half power beamwidth meets the specified 30°. The results also shows the F/B = 16.6 dB and front-to-side ratio (F/S) of 15.9 dB. The computed input impedance is Z_{in}=20.4+j19.9 Ω.

CONCLUSION

The GA in conjunction with NEC is a powerful new method of antenna design. As an example, an improved Vee-shape Yagi-Uda array with linear elements was designed. It had been assumed previously that such performance could only be achieved by elaborate curvature of the elements or by the use of considerably more elements in a conventional, parallel element array.

REFERENCES

1. Weile D S and Michielssen E, 1997, "Genetic algorithm optimization applied to electromagnetics: a review", IEEE Trans. Antennas Propag., 45, (3), 343-353

2. Altshuler E E and Linden D S, 1997, "Wire-antenna designs using genetic algorithms", IEEE Antennas and Propag. Magazine, 39, (2), 33-43

3. Goldberg D E, 1989, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley Press, New York, USA

4. Liang C H and Cheng D K, 1983, "Directivity optimization for Yagi-Uda arrays of shaped dipoles", IEEE Trans. Antennas Propag., AP-31, (3), 522-525

5. Landstorfer F M, 1977, "Optimisation of the directivity of linear and horn-antennas", IEEE Conference on CAD of Electronic and Microwave Circuits and System, 96-101, Hull, U.K.

6. Jones E A and Joines W T, 1997, "Design of Yagi-Uda antennas using the genetic algorithm", IEEE Trans. Antennas Propag., 45, (9), 1386-1392

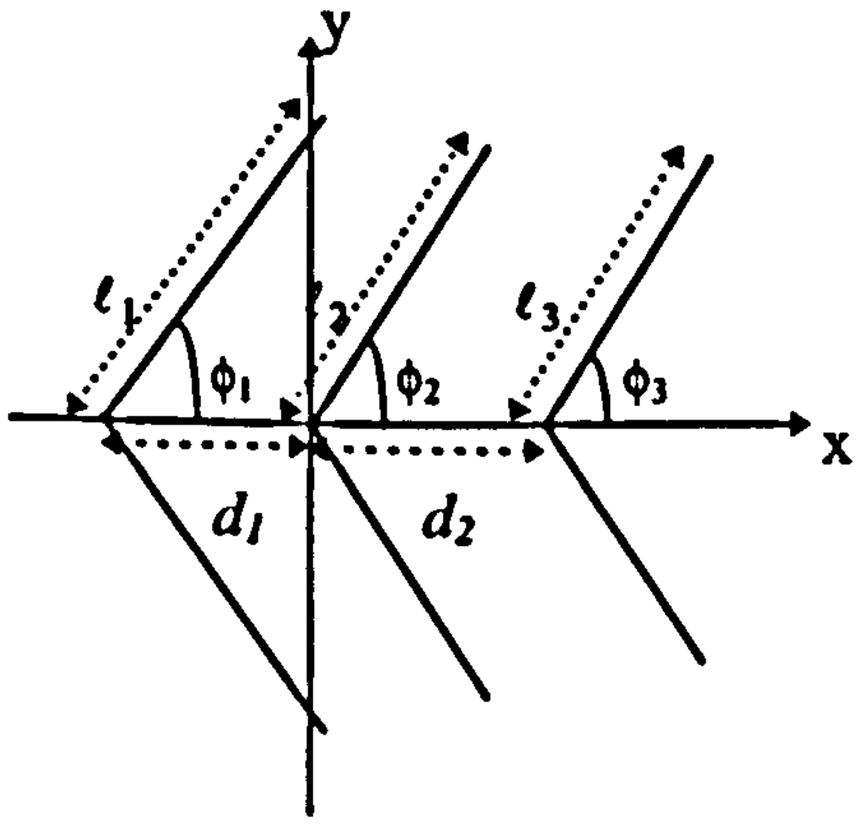


Figure 1: The geometry of a three-element Vee-shape Yagi-Uda array

parent 1 + parent 3 = child 31 +child 32
parent 2 + parent 4 = child 33 +child 34
parent 3 + parent 5 = child 35 +child 36
parent 4 + parent 6 = child 37 +child 38
parent 5 + parent 7 = child 39 +child 40
.....
.....
.....
parent 28 + parent 30 = child 85 +child 86
parent 29 + parent 1 = child 87 +child 88
parent 30 + parent 2 = child 89 +child 90

Figure 3a: Alternative pairing scheme

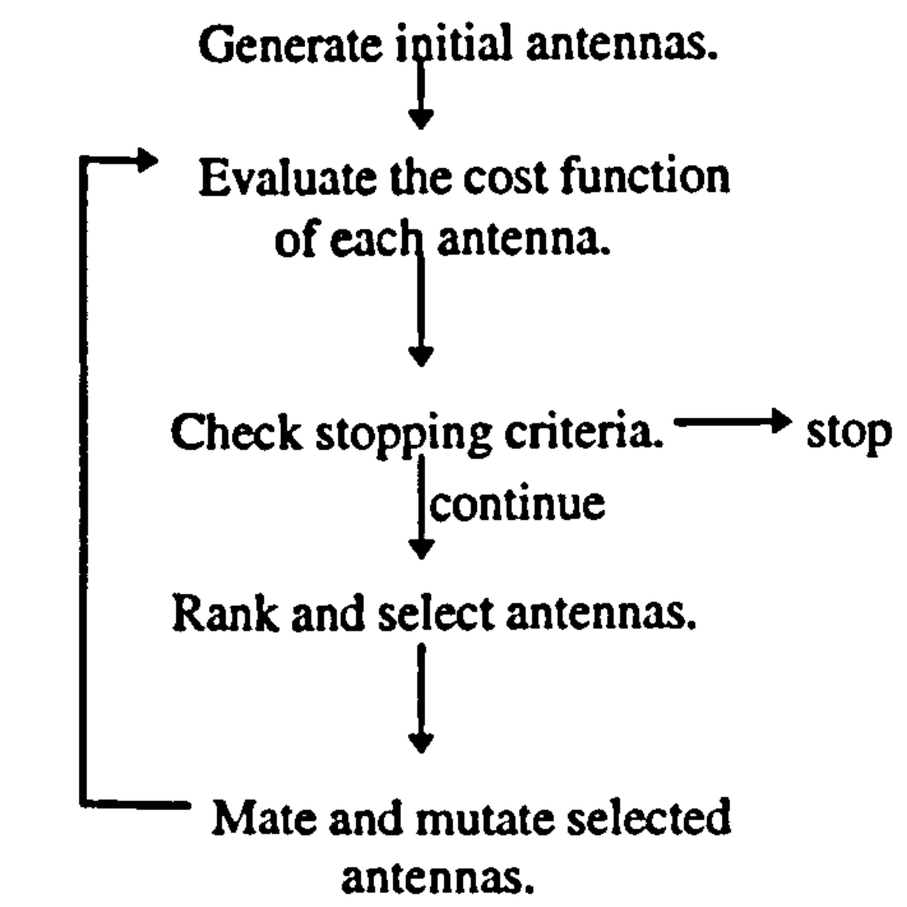


Figure 2: A flow chart of the genetic algorithm used to design a Vee-shape Yagi-Uda array

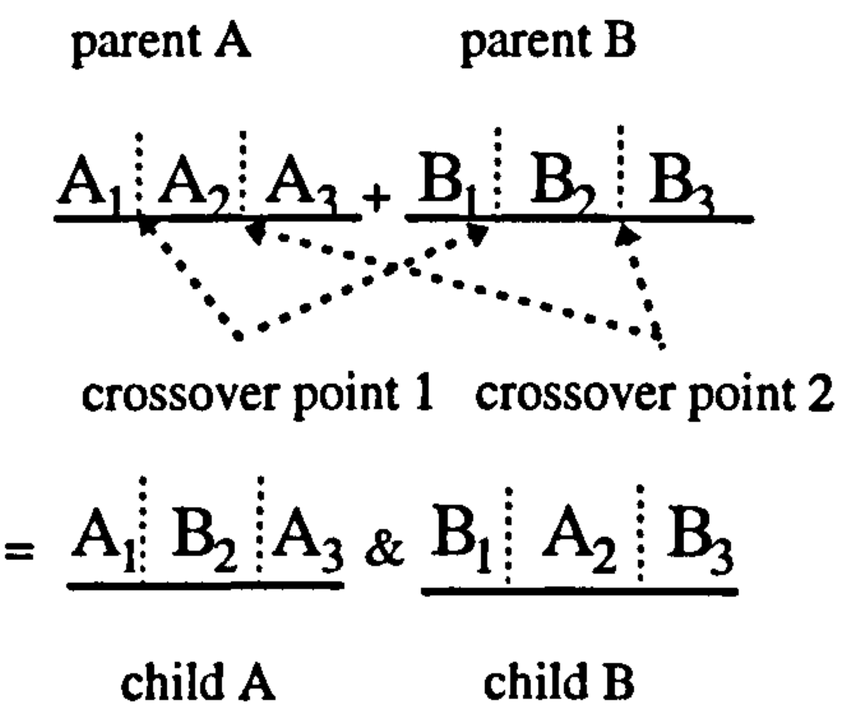


Figure 3b: Two-point crossover scheme

Table 1 - Geometry parameters for the optimum three-element Vee-shape Yagi-Uda array

ℓ_1	ℓ_2	ℓ_3	ϕ_1	ϕ_2	ϕ_3	d_1	d_2
0.756λ	0.745λ	0.722λ	50°	50°	54°	0.165λ	0.255λ

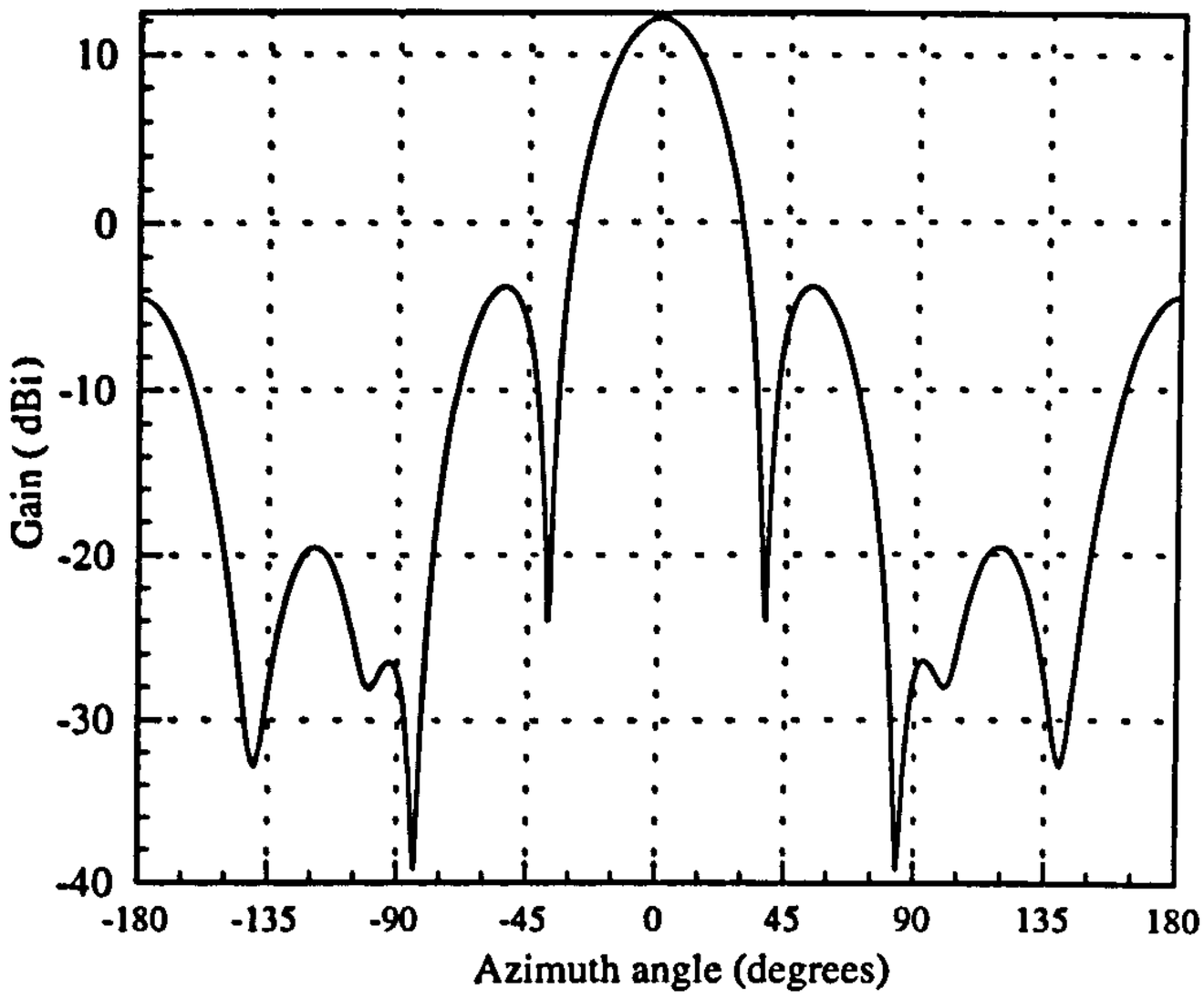


Figure 4: Radiation pattern of the optimized Vee-shaped Yagi-Uda array

An optimized shaped parasitic array using the genetic algorithm

Wen-Chung Liu and B A Austin

Indexing terms: Parasitic arrays, Shaped arrays, Genetic Algorithms

Abstract: The performance of a shaped, three-element parasitic antenna is improved by the use of a genetic algorithm (GA) in conjunction with the method of moments (MoM). The GA enabled the shaping and spacing of the elements to be changed at will as the optimum gain and front-to-back ratio (F/B) were sought. Appropriate interfacing to the NEC code made rapid testing possible and the combination thus yields a powerful antenna-synthesis tool. The results show that a simple, $3\lambda/2$, V-shape, parasitic array can achieve markedly increased gain over its straight, parallel-element counterpart but without the need for the elaborate element curvature, as was previously reported elsewhere.

1. Introduction

Genetic Algorithms (GAs), used in conjunction with the Numerical Electromagnetics Code (NEC), are a revolutionary new approach to wire antenna design and optimization [1, 2]. Until recently traditional techniques for antenna synthesis used the method of moments (MoM) to determine the current distributions on each wire of a selected antenna configuration and, from these, the radiation characteristics were calculated and compared with a target specification. The procedure is iterative and both experience and some appropriate candidate configurations are usually required to ensure success. By contrast, the GA is an optimisation procedure which is based on the processes of natural selection and the "survival of the fittest" in an evolutionary system which leads ultimately to the closest approximation to the antenna specification with minimal foresight or pre-conditioning on the part of the designer [3]. In using the GA all antenna design parameters are encoded into a string of binary bits known as a chromosome, while each parameter is represented by several bits – the genes. The genes are the basic building blocks of the GA and map directly to the associated real, physical properties of the antenna. The immense power of the technique is its ability to satisfy a performance criterion without any *a priori* knowledge of candidate configurations, and the facility for finding the global optimum result.

The objective of this work was to use the GA to synthesize a shaped, three-element parasitic array with optimal gain. Previous work [4,5] had produced maximum gain by fairly complex shaping of the three, approximately 1.5λ, elements in the array. Essentially V-shaped elements were shown to be optimum; however the requirement for peculiar curvature adds to manufacturing costs. In this work the V-shape, with straight elements, of the array emerged as a candidate configuration after some preliminary runs of the GA and so the antenna's geometry was then constrained to this particular shape. Both the gain (G) and front-to-back ratio (F/B) were considered in the objective function to ensure good antenna performance. The method of moments code NEC2 was used to determine the gain, F/B, input impedance and beamwidth of each antenna designed by the GA.

2. Synthesis procedure

Fig. 1 shows the general shape of the candidate antenna. It is symmetrical and the feed point is placed at the origin on the x-y plane. The radius of each element was set as 0.01λ to allow comparison with the results in [5]. Thus the lengths, spacing and included angles of the elements are left as the design variables for random manipulation by the GA. The range of ℓ_i was allowed to vary between 0.3λ and 1.1λ; d_i between 0.1λ and 0.5λ; and the included half-angle, ϕ_i, between 10° and 90°. Fig.2 is a flow chart of the GA used in this work. It begins by encoding the design parameters into a string of binary bits: the chromosome. Seven, six and eight bits respectively were chosen for the variables: element length, element spacing and included angle. Each antenna therefore requires a chromosome with $(7\times 3)+(6\times 2)+(8\times 3)=57$ bits (or genes) to store the real configuration. After the parameters were encoded an initial population of 90 candidate antennas was created by simple randomisation. The

randomness makes it possible to explore a broad population of possible solutions in the entire search space. Generally, the 'good' initial chromosomes will help the algorithm to achieve an optimum result quickly. On commencement of the iterations a mathematical function was applied to evaluate the individual 'cost' of each created antenna. This function is always set according to the search objective and is the required objective function or cost function. Clearly the associated cost of each antenna will decide its fitness for survival into the next generation. The goal here is to obtain both high gain and F/B so the objective function must force all new configurations towards that target with the maximum tolerable beamwidth set at + and - 15° and with maximum radiation in the $\phi = 0°$ direction. In view of the wide variation of initial values obtained, each of the three parameters was normalized by an appropriate factor: a standard procedure for such problems, aptly described in [6]. We selected G = 13 dBi and F/B = 20 dB as goals since they exceeded typical results for similar configurations reported in the literature (see [4,5]), while the radiated power within the E-plane ($-15° \le \phi \le 15°$) is normalized by the total radiated power. As expected in such a multi-dimensional problem numerical experiments showed that the three normalized terms made very different contributions to the value of the objective function. To prevent any of them from dominating the iteration process, each was weighted by an associated constant; again as in [6]. In this example the improvement in gain and the radiated power concentrated within the specified angle changed much more slowly than did the F/B – a feature of such parasitic arrays – so they must be more heavily weighted. Thus values of 24, 18, 1 respectively were selected after a number of preliminary runs. The resulting objective function was then:

$$\text{Objective function} = 24\frac{\text{Gain}}{13} + 18\frac{\sum_{-15°\le\phi\le15°} P(\phi)}{\sum_{-180°\le\phi\le180°} P(\phi)} + \frac{F/B}{20} \qquad (1)$$

A stop criterion, chosen according to the specific features of the problem, was set in the next step. In this case, it was set arbitrarily to operate once the objective function became asymptotic to its maximum value and remained so for at least thirty iterations. If the procedure does not stop, the chromosomes are then ranked according to their respective costs. The higher the cost the better the chance of survival to generate the new chromosomes. Based on the philosophy of natural selection or "survival of the fittest", the top one-third (30 chromosomes) of the ranking list are kept – the rest are discarded. Two distinct operations then applied in the mating process which then followed: select two survivors for mating or pairing, and then exchange the genes between the two paired chromosomes. The first operation can, of course, be achieved by many different schemes – the so-called roulette wheel selection is the famous one. In order to generate diverse offspring every two alternative chromosomes of the survivors were mated. The second operation is to select the loci (or crossover points) to swap the genes of the parents. Fig. 3 shows the two-point crossover process used. Two random points divided the gene string of each parent into three sections, or two sections if the two points were the same. After swapping genes, each of the offspring will inherit either both ends or occasionally only one end of the gene string of its parent. This is unlike the general single-point crossover in which only one end of the gene string of each parent will be left with the offspring. Each pairing produced two offspring resulting in a total of 60 after the mating process was complete. This is a so-called steady state GA because the number of new-generation parents and offspring remains unchanged [8]. The probability of the algorithm converging to a local optimum is avoided by selecting 5% of the chromosomes randomly from each generation and then inverting two bits in each. Thus all the processes are manipulated and the evaluation of each associate cost of the new generation begins again. The algorithm runs through these processes iteratively until it converges.

3. Results

Table 1 and Fig. 1 show the geometrical features of the improved antenna reached after 102 iterations. It is of interest to compare the geometry of this GA-optimized antenna with the antennas which resulted from the optimizations of Landstorfer [4] and Liang [5]. In all three cases the lengths of the elements were close to 1.5λ – a feature which makes the shaped antenna unlike the conventional Yagi-Uda array which has nominally $\lambda/2$ elements. The general V-shape of the array is another common characteristic

and this, coupled with the increased lengths of the elements, places the antenna in the category of the so-called V-dipole which achieves greater gain than its $\lambda/2$ counterpart [7]. The inclusion of parasitic elements, as in the Yagi-Uda array, enhances this yet further and, as shown here, does not require elaborate element-shaping to achieve optimal gain for a given number of elements – unlike the approach adopted in [4] and [5] where such element curvature was found to be necessary to achieve that same performance. This is undoubtedly an advantage from a manufacturing point of view.

Landstorfer [4], by an experimental procedure, devised an array with G = 11.5 dBi, while Liang [5] achieved 11.8 dBi. Fig. 4 shows the gain plotted against the azimuth angle ϕ for our optimized array and compares it with the performance of that in [5]. The maximum gain of the GA-antenna is 11.8 dBi with a half power beamwidth that meets the specified 30°. The computed F/B is 28 dB while the input impedance is Z_{in}=10.2 - j8.8 Ω. The performance of the V-shaped array is as good as any achieved with complex curvature of the elements.

4. Conclusion

The GA used in conjunction with the moment method such as the Numerical Electromagnetic Code (NEC) is a powerful new technique of antenna design and optimisation. To illustrate its capability in maximising antenna gain and front-to-back ratio an improved, V-shaped, parasitic array with linear elements was developed. Whereas it had been assumed previously that such performance could only be achieved by elaborate curvature of the elements, or by the use of considerably more elements in a conventional, parallel element Yagi-Uda structure, this work has shown that a much simpler antenna – and hence one that is cheaper to construct – can yield the same performance.

5. References

1. Weile D. S. and Michielssen E.: 'Genetic algorithm optimization applied to electromagnetics: a review', *IEEE Trans. Antennas Propag.*, 1997, **45**, (3), pp. 343-353

2. Altshuler E. E. and Linden D. S.: 'Wire-antenna designs using genetic algorithms', *IEEE Antennas and Propag. Magazine*, 1997, **39**, (2), pp. 33-43

3. Goldberg D. E.: 'Genetic Algorithms in Search, Optimization and Machine Learning' (Addison-Wesley Press, New York, 1989)

4. Landstorfer F. M.: 'Optimisation of the directivity of linear and horn-antennas', *IEEE Conference on CAD of Electronic and Microwave Circuits and Systems,* Hull, U.K., 1977, pp.96-101

5. Liang C. H. and Cheng D. K.: 'Directivity optimization for Yagi-Uda arrays of shaped dipoles', *IEEE Trans. Antennas Propag.*, 1983, **AP-31**, (3), pp. 522-525

6. Jones E. A. and Joines W. T.: 'Design of Yagi-Uda antennas using genetic algorithms', *IEEE Trans. Antennas Propag.*, 1997, **45**, (9), pp. 1386-1392

7. Stutzman W. L. and Thiele G. A.: 'Antenna Theory and Design' (J. Wiley & Sons, New York, 1981), p.204

8. Johnson J. M. and Rahmat-Samii Y.: 'Genetic Algorithms in Engineering Electromagnetics', *IEEE Antennas and Propag. Magazine*, 1997, **39**, (4), pp. 7-25

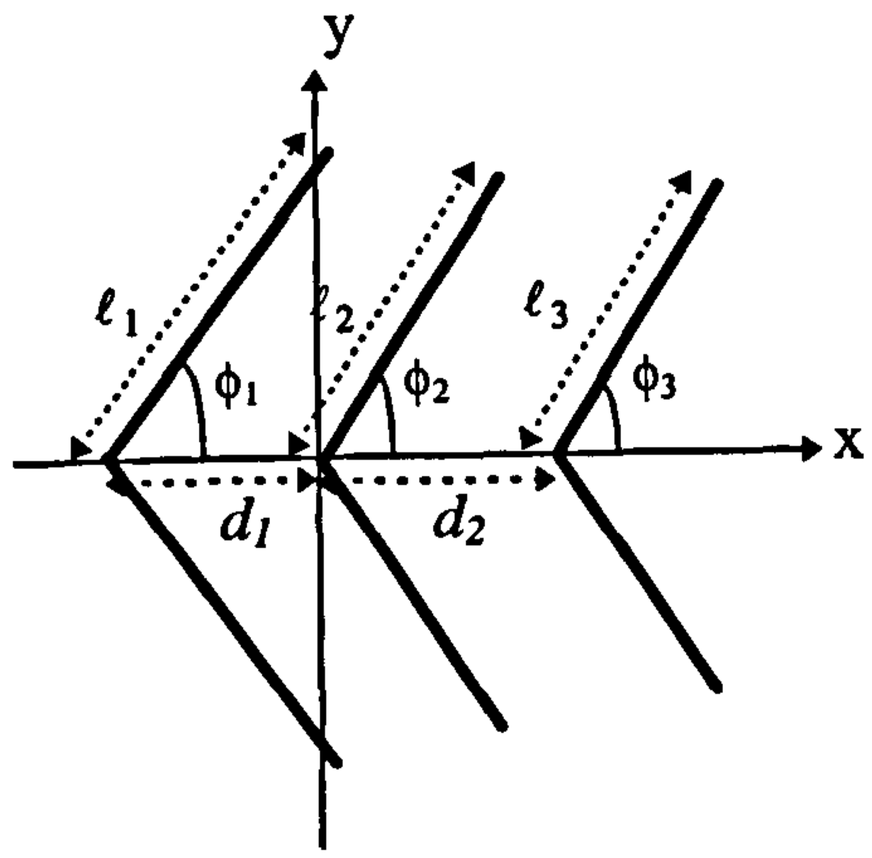


Fig. 1 The geometry of a three-element V-shaped parasitic array

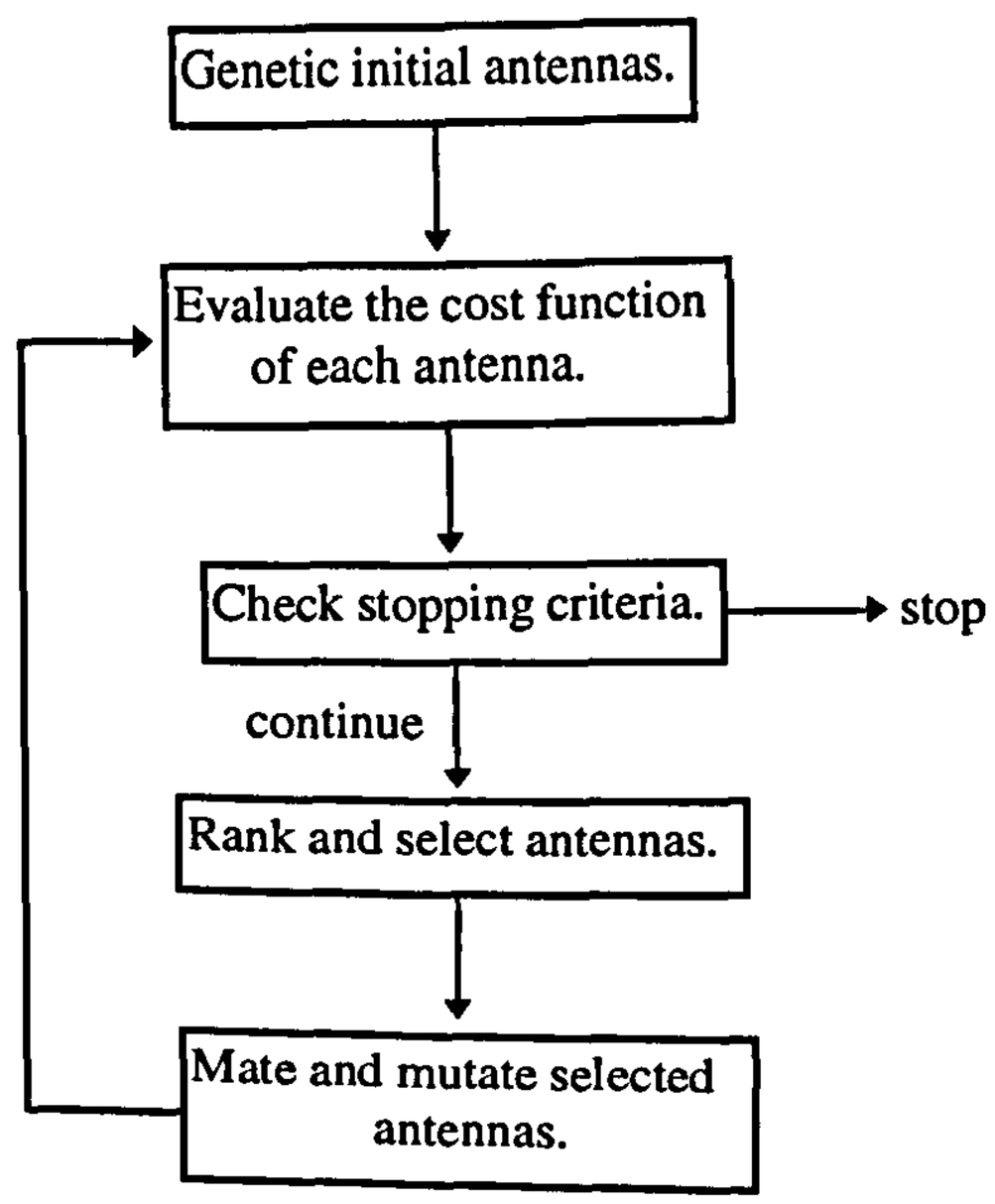


Fig. 2 A flow chart of the genetic algorithm used to design a V-shaped parasitic array

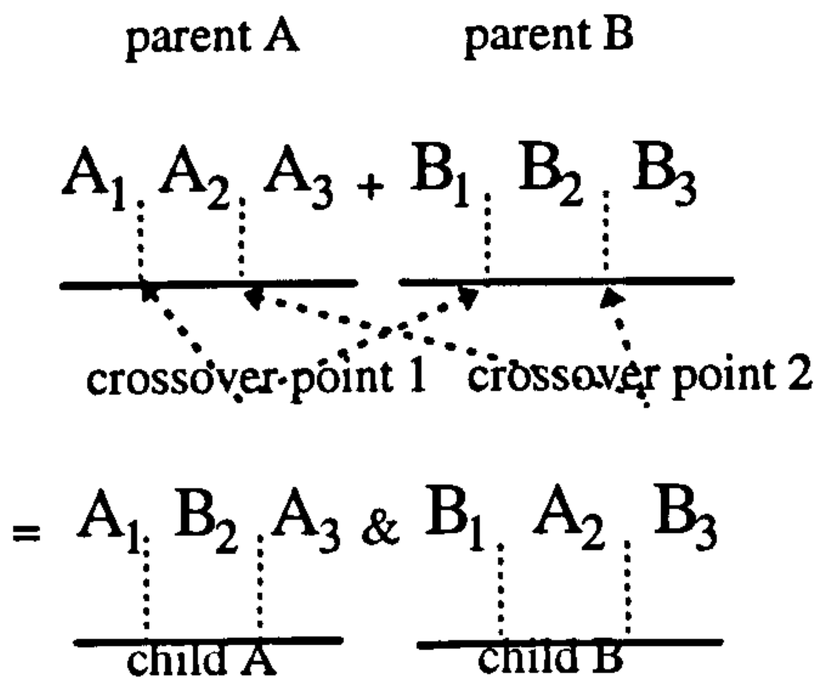


Fig. 3 Two-point crossover scheme

Table 1 Geometry of the optimum three-element V-shaped parasitic array

ℓ_1	ℓ_2	ℓ_3	ϕ_1	ϕ_2	ϕ_3	d_1	d_2
0.747λ	0.728λ	0.697λ	50°	50°	50°	0.113λ	0.233λ

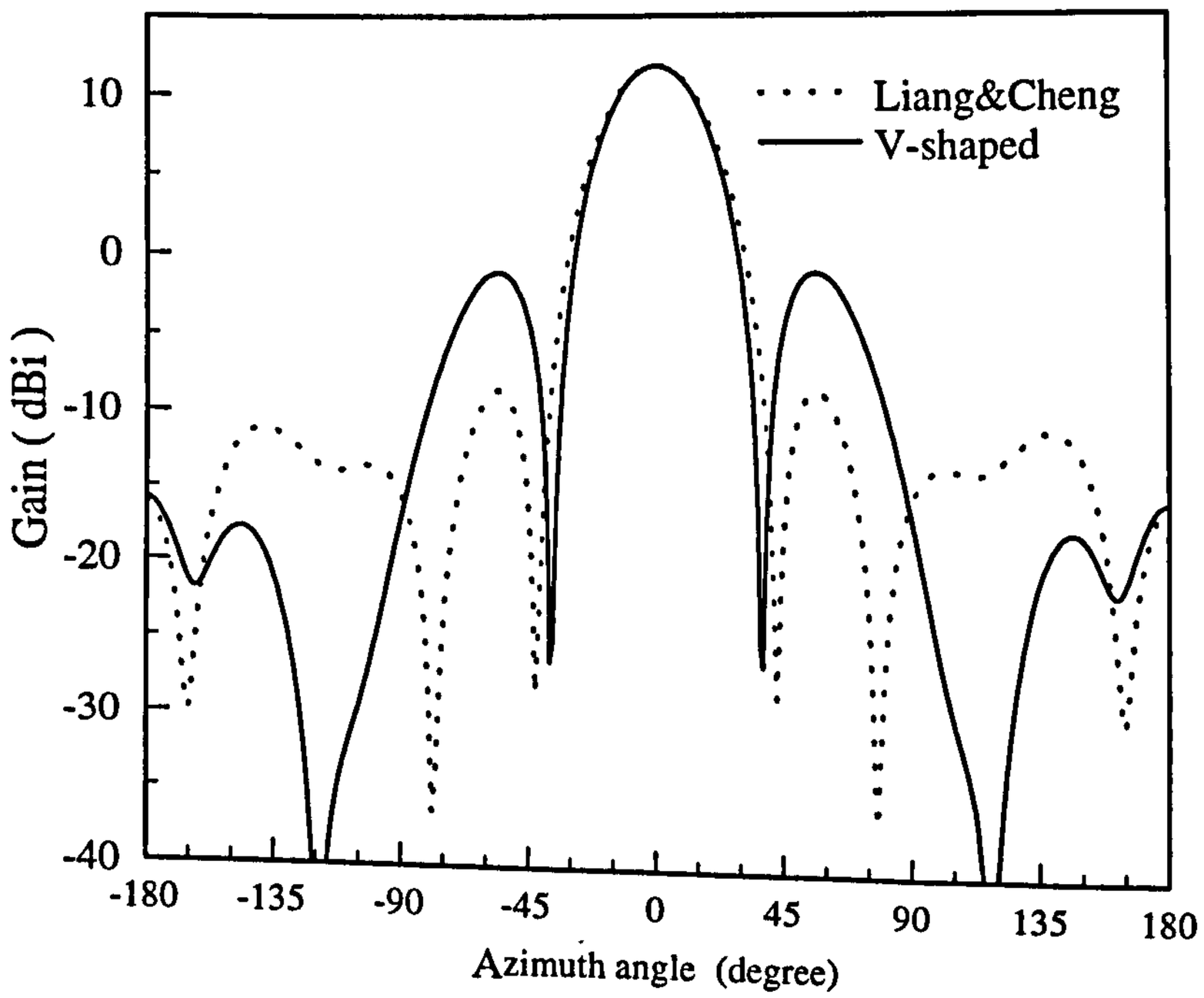


Fig. 4 Comparison of the radiation pattern for the V-shaped and the curved antennas.