Chapter 5

Conclusion

This section will present a brief discussion of the results obtained in Chapter 4 and their significance. Some suggestions about how this work can be extended are also given.

The verification of the accuracy of the calculations and models implemented is considered in Section 5.1. The effects of the algorithm's parameters and options are discussed in Section 5.2. Section 5.3 reviews the results obtained, and Section 5.3.1 considers the implications of the comparisons between these results and published results. Finally, some suggestions for future development of this work are given in Section 5.4, followed by a brief conclusion in Section 5.5.

5.1 Accuracy Verification

The tests conducted in Section 4.2 to verify the accuracy of the calculations and models used in this dissertation are briefly considered here.

The accuracy of the ideal components (inductors, capacitors, and transmission lines) is very high because these ideal components have exact models that characterise their performance. The microstrip line, open-end effect, and via inductance models used here are also in excellent agreement with those provided with EEsof. The accuracy of the width step is acceptable if the ratio of the two lines' widths is kept low, and the frequency is low. The T-junction model is only accurate at very low frequencies, and when the line widths and relative dielectric constants are small. The cross model is not expected to perform well because it is a modified T-junction model is used rather than a true cross model.

The results with full circuits are very impressive except for the microstrip results for Problems 2 to 4. These problems all have the same substrate which has a relative dielectric constant of 1 and a height of 3 mm. These two factors mean that there will be large fringing fields leading to dramatic discontinuity effects. The disagreement between the current algorithm and EEsof is understandable under these circumstances. Problem 5 uses the same substrate, but does not have large errors because the maximum frequency in this case is much lower than the maximum frequency for Problems 2 to 4. The agreement with EEsof results is thus seen to be good as long as the substrate height is small compared to the wavelength in the substrate, and fringing fields are not excessive.

5.2 Options and Parameter Values

The algorithm developed has a large number of options and parameters that affect its performance. The tests summarised in Section 4.3 will be considered here.

The effect of the Pareto-like optimality is remarkable. The results obtained using Pareto-like optimality to design networks with one to six elements are comparable to, and often better than, the results where the algorithm is applied to only one length at a time. This is despite the fact that the algorithm's total power is concentrated on one network length in the non-Pareto case, whereas the Pareto-like case divides the algorithm's power between a number of network lengths. This causes the processing time required in the non-Pareto case to be much higher than that required in the Pareto-like case. Some reasons for this remarkable result are the fact that different length solutions interact, and diversity is maintained. The various length solutions in a population work together in the Pareto-like case to produce better results in all network lengths simultaneously. For example, a good two element network usually forms the basis for a good three element network. Diversity is essential in a genetic algorithm, and simply means that the entire population does not converge to one part of the search space. Diversity is maintained in the Pareto-like case because the solutions of different lengths, while similar as noted above, are nevertheless different.

Messy genetic algorithms also allow multiple length solutions to coexist, but the approach used here is different. A messy genetic algorithm starts with a large number of short solutions and slowly constructs longer solutions with the sole objective of obtaining the best maximal length solution. The approach used here is to allow solutions of all lengths to coexist at all times with the objective of simultaneously obtaining the best solutions for every length.

The addition of a local optimiser to the genetic algorithm leads to improvements to the results in all cases and dramatic improvements to long individuals. Genetic algorithms are known to have limitations when trying to find highly accurate results and will thus often miss the small range of component values with very good results, despite getting close to these good ranges. The addition of a local optimiser overcomes this problem because the local optimiser will take values close to a good result and move them to that good result (the local optimum). Long individuals are affected more than short individuals because they represent a more difficult problem for the algorithm because there are more possible combinations of parameters.

The default parameters for the genetic algorithm operator parameters produce very good results. The combined probability of arithmetic and binary crossover is 0.7 which is very close to the value of 0.6 recommended by Goldberg [17] for binary crossover in a binary genetic algorithm. When the crossover rates become too high the performance of the algorithm deteriorates because good individuals only have a very small probability of surviving unaltered to the next generation. Increasing mutation probability causes the algorithm performance to decrease because mutation tends to disrupt good results. However mutation is still necessary to ensure that the entire search space is considered and premature convergence does not occur. Unlike the other mutation operators, increasing the non-uniform mutation probability is seen to improve the algorithm performance because non-uniform mutation is similar to simulated annealing which is a good optimisation algorithm in its own right.

The tournament size results show that a compromise must be reached between large tournaments which lead to fast local search, and small tournaments which cover the whole search space. Section 4.3.4 shows that a value of three or four produces the best results in this case. These values close to the value of two that Bäck [43] states is commonly used.

An increase in the number of optimisation steps is expected to decrease the final error because more local search steps will isolate local minima faster. This reasoning is borne out by the results in Section 4.3.5. The problem is that the number of optimisation steps has a large influence on the time required to run the algorithm so large values are undesirable. Furthermore, if the number of local search steps becomes too large (larger than the values considered here), the algorithm will converge on local minima too fast, neglecting the global search and decreasing the algorithm's performance.

The performance and time requirements of the algorithm depend heavily on the population size and number of generations used. The results in Section 4.3.6 show that the performance of the algorithm initially improves rapidly when population size and number of generations are increased and then tapers off. This effect is due to the fact that once the best result is obtained with a high probability, further increases in population size and number of generations will only have a very small effect.

The algorithm takes a comparatively long time to run. The main factors that cause this are the extremely wide ranges of component values used, the time required to calculate fitness, and the number of function evaluations required to obtain a good result.

The wide range of component values was chosen to give worst-case results for the algorithm and better results can be expected with smaller component ranges.

The fitness evaluation comprises the vast majority of the processing time required by the algorithm and this time is very difficult to reduce. This effect could be minimised by reducing the number of function evaluations required, but the current algorithm is comparatively expensive in this regard.

This algorithm requires a comparatively large number of function evaluations to obtain a result because genetic algorithms are not the most efficient optimisation algorithms. The versatility of the current algorithm is largely due to the use of a genetic algorithm, and slower convergence is the price paid for this versatility. The convergence of the genetic algorithm was improved by using a local optimiser, but this requires gradient information which was calculated using finite differences. This means that the gradient is calculated by computing one extra fitness value for each variable thereby increasing the number of fitness evaluations required. Some possibilities for reducing the time and number of function evaluations required are given in Section 5.4.

5.3 Results

The results obtained when the algorithm is applied to the ten test problems are considered here.

While the worst results are often very poor, the best and median results are usually very good, showing that the algorithm finds good results in the vast majority of cases. This effect is probably due to the stochastic nature of the algorithm and is a known limitation of genetic algorithms. Another possibility is that the algorithm is being misled by the way solutions of differing lengths interact. This could be a problem when the best solution with three elements is not the basis of a good solution with four elements, for example.

Longer networks usually have a much higher variation in errors than shorter networks because longer networks represent a more challenging problem for the algorithm. The extra elements in longer networks mean that there are more possible combinations of elements that must be evaluated than for shorter networks.

The distributed solutions are generally worse than the lumped solutions, but these results do not mean that distributed solutions obtained from lumped prototypes will produce better results than those given here. This is because the distributed parameter ranges were limited to ensure that the distributed elements are realisable, while the lumped parameter ranges have no such limitations.

The fact that a number of transmission lines with the same characteristic impedance or line width are cascaded in the distributed solutions to Problems 4 and 7 simply means that a transmission line longer than the maximum specified length is required. Fewer elements could be used by allowing longer transmission lines, but in general, this is not a good approach. The algorithm could be modified to avoid this effect by ensuring that there is some minimum difference between the impedances of connected series transmission lines.

The mixed solutions usually obtain results that are at least as good as the better of the lumped and distributed solutions. This is because the lumped and distributed cases are simply instances of the mixed case. The mixed solutions are sometimes slightly worse than the lumped or distributed solutions because the mixed problem has significantly more possible combinations of elements than the other cases and is thus a much more challenging problem.

Some of the microstrip results violated the maximum allowable line width ratios at a discontinuity. This is expected because a penalty function was used and solutions that violate the constraints by a small amount are known to occur when penalty functions are used.

The microstrip results are similar to the distributed results in most cases, but the microstrip results are generally not quite as good as the distributed results. This is caused by the presence of discontinuities and the limitation placed on the microstrip line widths at a discontinuity. The importance of compensating for discontinuities is clearly shown by this observation.

5.3.1 Comparison to Published Results

The present algorithm is compared to other published results in Section 4.5.1 and the implications of these results are considered below.

The results obtained are impressive. The current algorithm performs at a level comparable to the best published results in all cases except Problem 5. The only algorithms that obtain similar performance are those developed by Abrie [2] and Dedieu *et al.* [82], but this is offset by the fact that these algorithms are more limited than the present one. The algorithm proposed by Dedieu *et al.* can only determine component values and must be supplied with a circuit configuration. The main limitation of the published version of Abrie's algorithm is that it is not able to vary both the impedance and length of transmission lines simultaneously. (However, this limitation has recently been removed, although results have not yet been published.) The consideration of mixed lumped-distributed networks in the new algorithm presented here is much more general than any other algorithm. Additionally, all published algorithms only consider one matching network size at a time whereas the current algorithm considers all lengths from one to a specified maximum. This means that this algorithm is able to achieve similar results to the best published techniques while being more versatile.

Problem 5 is the one case where the current algorithm does not perform as well as the published results. The stochastic nature of genetic algorithms means that there will be some variation in the results. Additionally the local optimiser used is poor, and a better local optimiser could lead to better results. It is significant that Abrie's [2] algorithm, which performs very well on this problem, uses a good local optimiser (a quasi-Newton method). Another possibility is that the current algorithm is being misled by the way solutions of differing lengths interact.

The main disadvantage of the current algorithm is that it is not as efficient as the published algorithms in terms of the number of function evaluations required. This problem has already been considered in Section 5.2, but the important points are repeated here. This algorithm is more versatile than other algorithms implying a more complex problem which is more difficult to solve. The advantages of this versatility include the fact that resonant sections are allowed, multiple network lengths are considered simultaneously, and both the length and impedance of transmission lines can be varied.

5.4 Extensions to This Work

While this work has led to the development of a very powerful optimisation and impedance matching algorithm, a number of further developments are possible. Some possibilities for such further development are considered below.

The current work is limited to the design of ladder networks, but this is not a fundamental limitation of the approach used here. Other structures such as those with component loops could be considered by an extended algorithm.

The natural next step after developing an impedance matching network design algorithm is to move towards a complete amplifier design algorithm. The algorithm described here has the potential to be developed into a complete amplifier design package and could even be extended to automatically select an active device.

Other problems such as filter and antenna design could also be considered because the advantages obtained here would also be useful in these problems. Applying the algorithm to other problems would involve little more than using a different fitness function.

The optimisation algorithm developed here has a number of unique properties and these could be developed further. The use of Pareto-like optimality to allow solutions with varying lengths has tremendous potential and will be considered further. The combination of binary and floating point operators could be extended to consider other genetic operators such as the fuzzy operators suggested by Klir and Yuan [111]. The integration of a genetic algorithm and a local optimiser could form the basis for the integration of further optimisation algorithms into a genetic algorithm.

The time required by this algorithm is its main limitation and could be reduced in a number of ways. A better local optimiser could be used to accelerate convergence, but this would entail major changes to the current algorithm. Eliminating the finite difference calculations by using an optimisation algorithm that does not require gradients could lead to an additional reduction in the number of function evaluations. The only drawback is that local optimisation algorithms that do not use gradients tend to converge more slowly than local optimisation algorithms that use gradients, so this change will not necessarily produce an improvement. Better initialisation of the algorithm could also be used to ensure that the algorithm has a good starting point, thereby reducing the number of function evaluations required to get near a good solution. This initialisation could, for example, be done using analytic techniques to design a matching network for a single frequency and allowing the algorithm to optimise these narrowband results to obtain a good broadband solution.

5.5 Summary

A new impedance matching algorithm based on a hybrid genetic algorithm has been developed. The consequences of the tests performed in Chapter 4 were presented above and are summarised below.

The results achieved with this new impedance matching algorithm compare very well with published results. This algorithm has a number of advantages over published techniques including the simultaneous synthesis of networks with multiple lengths, a very general consideration of mixed lumped-distributed networks, the simultaneous modification of both the length and width of transmission lines, and the inclusion of dispersion and discontinuity effects.

The hybrid genetic algorithm also has a number of unique features including the use of a Pareto-like optimality to obtain variable length solutions, and the simultaneous application of both binary and floating-point genetic operators. The variable length solution capability of the algorithm is particularly notable and warrants further investigation.

While the results achieved so far have been impressive, the algorithm can still be developed further and some suggestions for such development are given. These suggestions mainly relate to ways of overcoming the limitations of this algorithm, and other applications of the new techniques that have been developed.