ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-THPOST020

VISUALISATION OF PARETO OPTIMAL SPACES AND OPTIMISATION SOLUTION SELECTION USING PARALLEL COORDINATE PLOTS*

S. Smith†, M. Southerby, S. Setiniyaz, R. Apsimon, G. Burt
Lancaster University, Cockcroft Institute, LA1 4YR Lancaster, United Kingdom

Abstract

In this paper, we build on previous work where multi-objective genetic algorithms were used to optimise rf cavities using non-uniform rational basis splines (NURBS) to improve the cavity geometries and reduce peak fields. [1] These optimisations can produce thousands of Pareto optimal solutions, from which a final cavity solution must be selected based on design criteria, such as accelerating gradient and power requirements. As all points are considered equally optimal, this can prove difficult without further analysis. Here we focus on the visualisation of the Pareto optimal points and the final solution selection process. We have found that the use of clustering algorithms and parallel coordinate plots (PCP's) provide the best way to represent the data and perform the necessary trade-offs between the peak fields and shunt impedance required to pick a final design.

INTRODUCTION

Multi-objective optimisation (MO) methods are employed when an optimisation must be performed that involves two or more conflicting objectives that must be optimized simultaneously, which is often the case in rf cavity design. Mathematically these types of problems can be summarized as follows:

$$\text{min/max } f_m(x), \quad m = 1, 2, ...M, \tag{1}$$

$$\text{subject to: } \quad g_j(x) \geq 0, j = 1, 2, ...J, \tag{2}$$

$$x_{ilb} \leq x_i \leq x_{iub}, \quad i = 1, 2, ..., n \tag{3}$$

Where f_m are the objectives, g_j are the constraints and $x_{ilb/iub}$ are the bounds on the input parameters. As an example for cavity design, these may be the peak fields, a required operating frequency, and limits on some of the cavity dimensions, respectively. A single, unique, optimal solution to these MO problems does not exist, and instead a set is found that defines the trade-off between all the objectives. For complex optimisations with many inputs and objectives, these sets can contain thousands of solutions that are all considered equally optimal by the algorithm. These points can also have objective values that are very similar to each other, either because the algorithm has converged to a near continuous set, or in the case of rf design, due to numerical inaccuracies caused by meshing. This presents a difficulty to the designer who must then select a 'best' solution from this set to use for a final design based on some criteria. This paper explores methods that can be applied to these large sets in order to aid in the final solution selection process.

SINGLE CELL OPTIMISATION RESULTS

An optimisation was previously performed for a generic X-band (12 GHz), $\beta = 1$, $2\pi/3$ travelling wave cavity and the reader is referred to [1] for details. This optimisation produced a large Pareto optimal set of 5000 points which are shown in Fig. 1, with 2 objectives on the axis; shunt impedance per unit length (Z), and the peak electric field (E_{pk}/E_{acc}). The points are also colour coded with a third objective, the modified Poynting vector ($\sqrt{S_c}/E_{acc}$) [2].

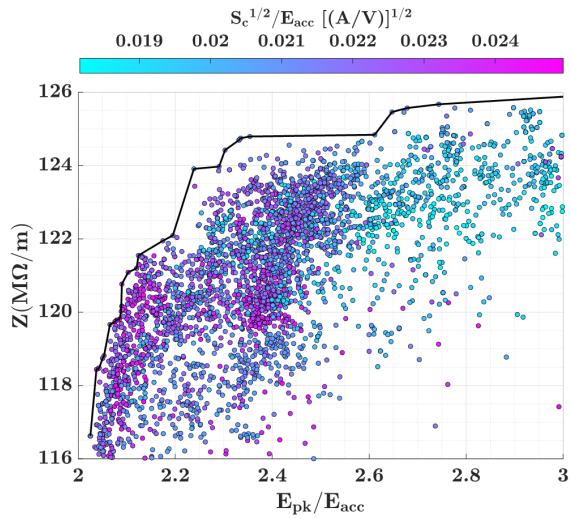

Figure 1: 2D scatter plot showing Pareto optimal set in the Z vs. E_{pk}/E_{acc} objective space, with $\sqrt{S_c}/E_{acc}$ in colour.

PARALLEL COORDINATE PLOTS

If there are only two objectives then selection can be relatively straightforward as the points can be plotted to create a 2D Pareto front as shown in black in Fig. 1. The designer can then easily select a point based on the trade off between the two objectives. If there are three objectives, then colour can be used to provide information about the third objective or a 3D plot could be used, however, as the number of objectives increases, it becomes significantly more difficult to visualise the data sets and select an individual final solution. A comprehensive list of appropriate visualisation techniques for Pareto optimal sets is given in [3] but we have found that the parallel coordinate plot (PCP) provides the best method for representing these sets, especially as the number of dimensions grows larger than four. PCP's display each solution as a piece-wise line that crosses equally

* Supported by Research Grant: ST/P002056/1
† s.smith26@lancaster.ac.uk

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-THPOST020

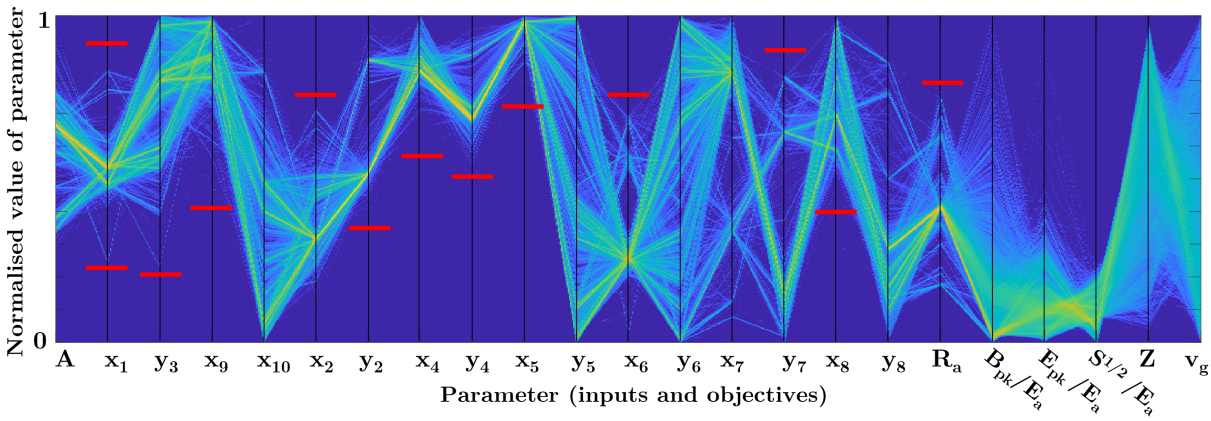

Figure 2: Density PCP for X-band travelling wave cavity, showing input and output spaces. Red lines show where input space could potentially be reduced for improved convergence speed. [3]

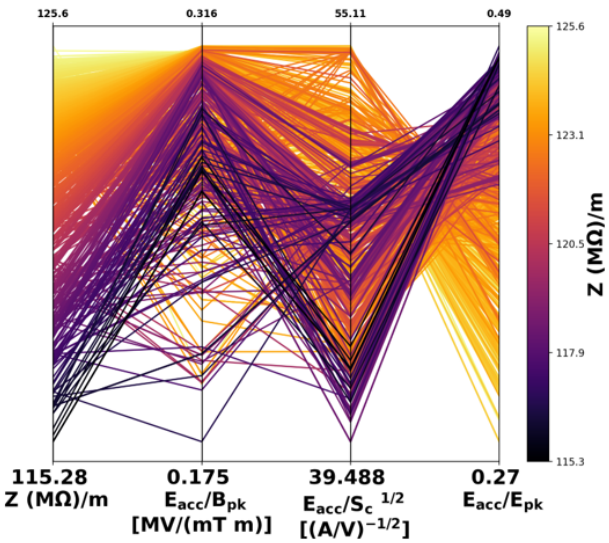

Figure 3: PCP showing Pareto optimal set for an optimized TW structure with four objectives (E_{pk}/E_{acc}, B_{pk}/E_{acc}, $\sqrt{S_c}/E_{acc}$ and Z). The inverse of three objectives are used so that all objectives require maximising. Z is also shown in colour for improved readability.

spaced y-axes, each of which represents a scale for a single objective. The point where the line crosses each axis gives the individual objective value, meaning that a large number of objectives and therefore dimensions can be visualised on a single plot. An example of a PCP with an entire Pareto optimal set is shown in Fig. 3. Four objectives are shown, with each quantity being normalized to E_{acc} and colour also representing the shunt impedance to make identification of the solutions with a large value easier. Trends in the set are easily distinguished using this method, for example it clearly shows that points that have a large Z also have a large E_{pk}/E_{acc} and that points with smaller values of E_{pk}/E_{acc} tend to have a larger value for $\sqrt{S_c}/E_{acc}$.

PCP's can also be used to visualise the entire input and objective space as shown in Fig. 2. Here, the colour of the lines represent the density of the solutions in that region. This can be useful for gaining insight into which inputs effect which objectives and where potential constraints could be placed to reduce the input space and increase the convergence speed for the algorithm on future optimisation runs. For example a constraint could be placed on x_4 to reduce the search space by half, as no optimal solutions have a value of x_4 less than half of the maximum. Although these figures are useful for discerning trends and providing information about trade-offs between the objectives, it would be impossible to select a single solution from the set as the 'best' as the space is almost continuous. Therefore before PCP's can be used to select an individual final solution, the Pareto optimal space must be reduced.

CLUSTERING THE N-DIMENSIONAL PARETO OPTIMAL OBJECTIVE SPACE

Although numerous mulit-dimensional visualisation techniques exist, many of them suffer from the same problem in that as the number of solutions to be visualised increases past a certain point, it becomes more and more difficult to distinguish the individual solutions and almost impossible to select one individual point as the final design. In order to deal with this issue, a number of methods exist that can be used to reduce these large data sets [4]. Most of these methods rely on some form of dimensional reduction (DR) of the solutions based on subjective decisions about the objectives by the designer. These methods can be useful for dealing with large data sets but information about the individual points can be lost when DR is performed. Ideally, as much information about the original solutions must be preserved so that any selection that happens after this process still leads to as close to an optimum design as possible.

We have found that the use of clustering algorithms such as K-means [5] or the Agglomerative Hierarchical Clustering

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-THPOST020

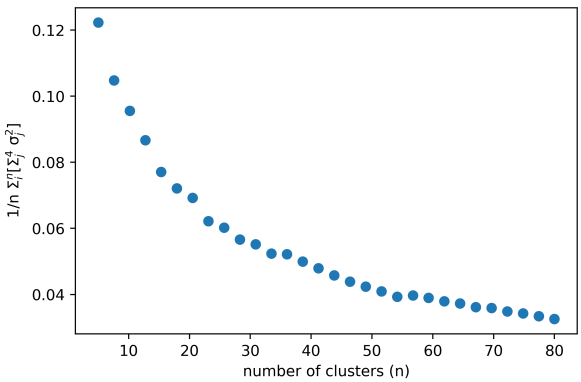

Figure 4: Mean standard deviation of objective cluster as a function of number of clusters. [3]

algorithm [6] are the most appropriate tools for this job. These methods can be used to group solutions based on their input or objective values, and then a single representative point of each cluster is chosen to be used in combination with a multi-dimensional visualisation technique in order to select a final design. In this way, thousands of points can be reduced to hundreds or tens, making the set more manageable.

A choice must be made about the number of clusters to use for a given data set. If too few clusters are used, then there will be significant overlap between the clusters, making a selection between them meaningless. If too many clusters are chosen then there is diminishing returns on the process compared to using the raw data set. An investigation was made into the variation of the standard deviation of each cluster with an increasing number of clusters with Fig. 4 showing the results of this investigation. The standard deviation drops rapidly until around 20 clusters, where it starts to decrease at a reduced rate. After 35 clusters this rate is again reduced making a choice of 25-35 clusters reasonable for this data set. After the clusters are created, a representative point must be selected to use on the PCP. The representative point is found by giving each solution within a cluster a weighted value. The equation used to determine the individual solution weight is as follows:

$$W_{soln} = \sum_{n}^{4} [obj \times range]_n,$$

where obj_n is the n^{th} objective normalised value, and $range_n$ is the normalised range for that objective. This weighting provides an accurate way to represent the individual cluster, which can then be used for visualisation.

SOLUTION SELECTION EXAMPLE

Fig. 5 shows a PCP of four objectives and 24 clusters (clustered using Agglomerative Hierarchical Clustering) for the Pareto optimal set. The inverse of the peak fields is used so that all objectives require maximising, making it easier to compare solutions. If for example, the cavity design was limited only by E_{pk}/E_{acc}, then cluster 1 could be chosen to select the solution from, assuming that no limits are imposed on the other objectives. This cluster is then opened to view all the solutions and is shown in Fig. 6. As there is little variation between the solutions within the cluster the final solution could be picked based on minor preferences or requirements for the other 3 objectives.

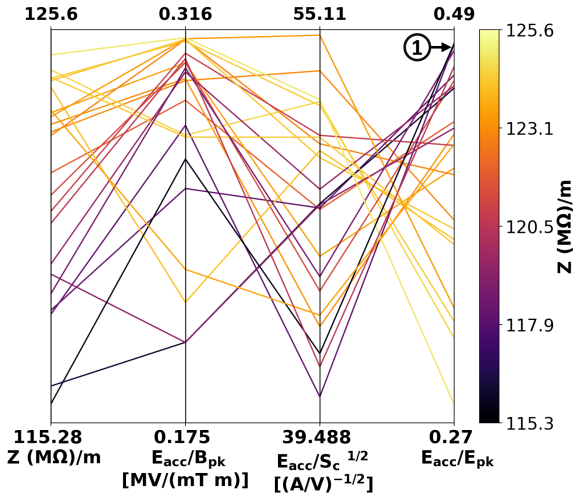

Figure 5: TW cavity objective space with four objectives showing 24 clusters. [3]

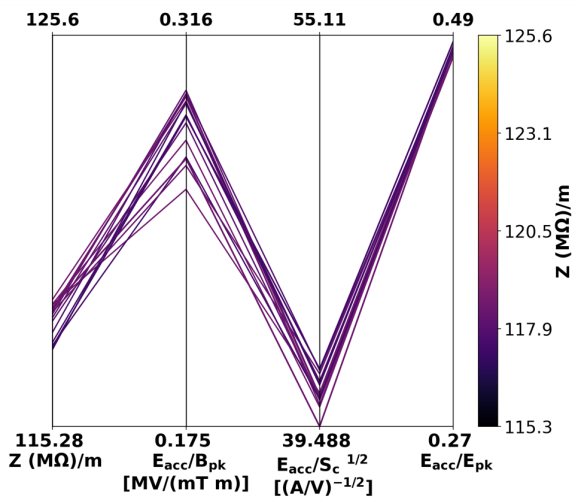

Figure 6: Opened cluster from Fig 5 for a design limited by E_{pk}/E_{acc} [3]

CONCLUSION

Multi-objective genetic algorithms have been applied to a 12 GHz TW single cell cavity design, generating a large Pareto optimal set. Once the Pareto optimal set has been found, Agglomerative Hierarchical Clustering has been used to cluster the solutions to reduce the large data set. Finally, PCP's have been used to visualise the data making the process of selecting of a final optimised design significantly easier.

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-THPOST020

REFERENCES

[1] S. Smith, R. Apsimon, G. Burt, S. Setiniyaz, and M. Southerby, "Multi-Objective Optimization of RF Structures," in *Proc. IPAC'21*, Campinas, SP, Brazil, 2021, pp. 1103–1106, doi:10.18942/JACoW-IPAC2021-MOPAB355

[2] A. Grudiev, S. Calatroni, and W. Wuensch, "New local field quantity describing the high gradient limit of accelerating structures," *Physical Review Special Topics-Accelerators and Beams*, vol. 12, no. 10, p. 102 001, 2009.

[3] S. Smith, M. Southerby, S. Setiniyaz, R. Apsimon, and G. Burt, "Multiobjective optimization and pareto front visualization techniques applied to normal conducting rf accelerating structures," *Phys. Rev. Accel. Beams*, Accepted May 2022, In press, 2022.

[4] B. Filipič and T. Tušar, "Visualization in multiobjective optimization," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2019, pp. 951–974.

[5] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern recognition*, vol. 36, no. 2, pp. 451–461, 2003.

[6] W. H. Day and H. Edelsbrunner, "Efficient algorithms for agglomerative hierarchical clustering methods," *Journal of classification*, vol. 1, no. 1, pp. 7–24, 1984.