Multiobjective optimization and Pareto front visualization techniques applied to normal conducting rf accelerating structures

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

(Received 8 March 2022; accepted 20 May 2022; published 14 June 2022)

There has been a renewed interest in applying multiobjective (MO) optimization methods to a number of problems in the physical sciences, including to rf structure design. The results of these optimizations generate large datasets, which makes visualizing the data and selecting individual solutions difficult. Using the generated results, Pareto fronts can be found giving the trade-off between different objectives, allowing one to utilize this key information in design decisions. Although various visualization techniques exist, it can be difficult to know which technique is appropriate and how to apply them successfully to the problem at hand. First, we present the setup and execution of MO optimizations of one standing wave and one traveling wave accelerating cavity, including constraint handling and an algorithm comparison. In order to understand the generated Pareto frontiers, we discuss several visualization techniques, applying them to the problem, and give the benefits and drawbacks of each. We found that the best techniques involve clustering the resulting data first to narrow down the possible choices and then using multidimensional visualization methods such as parallel coordinate plots and decision maps to view the clustered results and select individual solutions. Finally, we give some examples of the application of these methods and the cavities selected based on arbitrary design requirements.

DOI: 10.1103/PhysRevAccelBeams.25.062002

I. INTRODUCTION

The ideas of multicriteria decision making in economics were first introduced in the mid-1800s by Edgeworth and Pareto, and since then they have steadily developed and are now ubiquitous in many fields including engineering and the life sciences [1]. The key idea is that real world scenarios often exist where one faces several conflicting criteria and an optimum balance needs to be found between them. Multiobjective optimization methods can be employed in these cases to explore the trade-offs between the desired objectives.

In these types of problems, a unique global optimum solution does not exist, and instead it is necessary to find a set of points that define the trade-off between the objectives. One way that an optimum point can be defined is with the idea of Pareto optimality. A point is considered Pareto optimal if there are no further changes of the inputs that can improve at least one objective function without detriment to at least one other function. When this set of solutions is found and mapped from the decision space to the objective space, they make up an N-dimensional Pareto frontier, as shown in Fig. 1 for the 2D case where the Pareto front can be discrete or continuous. The decision space represents the N-dimensional space containing all values of the inputs to the optimization, and the objective space is the N-dimensional space containing the outputs or objective values that the optimization produces.

Although a variety of techniques exist to solve multiobjective (MO) optimization problems and estimate the Pareto frontiers, two of the most popular methods are to

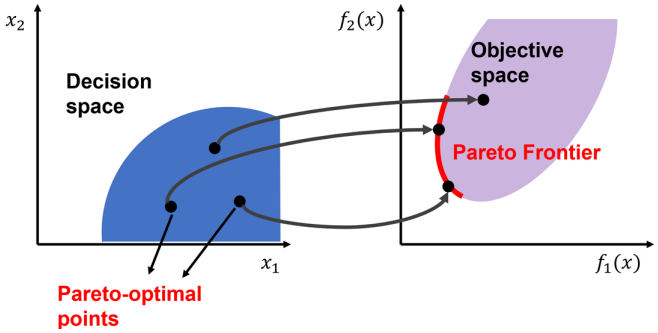

FIG. 1. Mapping decision space to objective space and Pareto optimal points to the Pareto front in 2D.

*s.smith26@lancaster.ac.uk
†m.southerby@lancaster.ac.uk

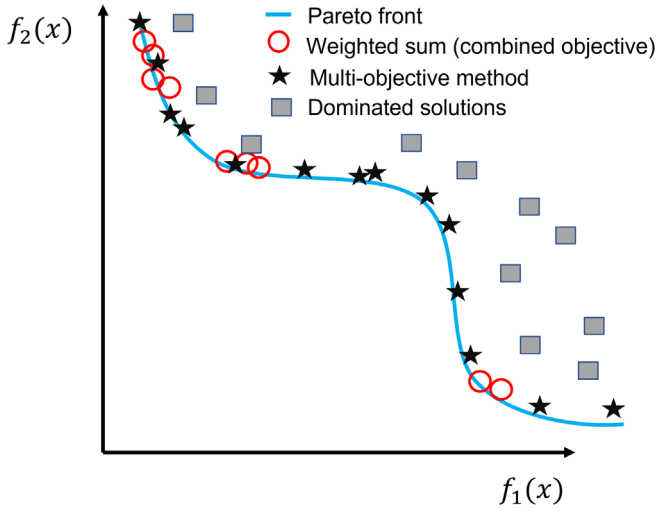

FIG. 2. Pareto front for two objectives showing the benefits of MO methods over weighted sum optimization methods. Adapted from Ref. [4].

combine individual objective functions together using weightings and then use single-objective optimization methods [2] or to solve the problems directly using genetic algorithms [3]. The weighted sum approach involves selecting scalar weights w_i and minimizing a combined function given by

$$U = \sum_{i=1}^{k} w_i f_i(x), \tag{1}$$

where k is the number of objective functions and f_i are the individual objective functions [2]. Although single-objective methods can be effective if the information about trade-offs is not needed, they require multiple optimization runs in order to obtain an approximation of the Pareto-optimal set, which can lead to a large computational complexity. The weighted sum approach utilizes a convex combination of objectives, and, therefore, nonconvex regions cannot be detected as shown in Fig. 2. This is because the sum of the weights is constant and negative weights are not allowed. Choosing the correct weightings may also not be immediately obvious and may take time to determine, again increasing the complexity of the optimization. Therefore, for large search spaces and complex objectives, genetic algorithms are preferred, as they can find multiple objective trade-offs in a single optimization run.

A. Optimization of rf structures

The requirements for any rf accelerating structure are, typically, a high shunt impedance, a high accelerating gradient, and low peak electromagnetic fields to avoid rf breakdown and heating effects. These objectives are normally conflicting, and, therefore, MO methods are ideal for exploring the trade-offs between different cavity geometries to find a set of optimal solutions. Traditionally, rf accelerating structures have been optimized by sweeping the various cavity parameters while holding others constant and generating parameter curves or by combining individual objectives into a single-objective function and then using single-objective optimization methods to minimize the combined function [5]. While these methods can work well, there are a number of drawbacks. First, when using a combined goal function, trade-offs where a small change in the weighting of one of the objectives may lead to large improvements in the others. Also, often the requirements for a design can change, and in this case the optimized solution is likely no longer optimal, and an optimization would need to be performed again. An advantage of estimating the Pareto frontier and obtaining a set of Pareto-optimal solutions is that a designer can quickly select an optimum cavity design based on the new requirements while also being able to explore the entire design space with the optimization potentially leading to cavity geometries with improved performance that may not be considered otherwise. It also allows the designer to easily reconsider the limits of the design and use this information in a more holistic analysis of the system.

Within the accelerator field, MO methods have previously been used for a wide range of optimization problems including dc photoinjectors [6,7], heat load and trip rates [8], beam dynamics [9], and resonator design under geometric uncertainties [10]. Recently, genetic algorithms have also been used for a wide variety of rf cavity designs [11–15], and, with improvements in parallelization and integrated commercial optimization tools, this trend is likely to continue. After estimating the Pareto frontier using a multiobjective genetic algorithm (MOGA), the designer must select the "best" design from a pool of appropriate candidates. In 2D, this selection is relatively straightforward, as the trade-off is easy to visualize, but, as the number of objectives is increased, this becomes more difficult. An array of dimension reduction methods and visualization techniques have been proposed to assist the decision maker in this process (see Sec. V), and these are explored and applied in this work to the problem of rf cavity design.

The rest of this paper is organized as follows. The cavity modeling, optimization objectives, and constraints used are described in Sec. II. The optimization methodology including algorithms, frequency handling, and convergence is discussed in Sec. III. A comparison between multiobjective and single-objective methods is given in Sec. IV. Visualization methods are introduced in Sec. V along with the clustering methods used for the analysis. Section VI explores the application of the visualization methods to the results for a generic standing wave and traveling wave cavity optimization. Section VII gives two examples of how the best methods can be applied to select a final design. Finally, Sec. VIII discusses the findings and draws conclusions from the rest of the paper.

II. APPLYING MOGAS TO RF CAVITIES

Mathematically, MO problems can be summarized as follows:

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

$$\text{subject to } g_j(x) \geq 0, \quad j = 1, 2, \ldots, J, \tag{3}$$

$$x_{ilb} \leq x_i \leq x_{iub}, \quad i = 1, 2, \ldots, n, \tag{4}$$

where f_m are the objectives, g_j are the constraints, and $x_{ilb/iub}$ are the upper and lower bounds of the input parameters. In order to formulate the optimization problem, the objectives, constraints, and geometrical limits need to be determined. The objectives define the targets of the optimization and can be either to maximize, minimize, or target a specific value. The constraints, on the other hand, impose hard limitations on the objectives of optimization and define a feasible design. The optimization may find a solution with the best objective value, but if it does not satisfy the constraints, then it is not considered feasible.

A. Objectives

rf cavities broadly fall into two categories: standing wave and traveling wave, each having a slightly different design process. In either case, for the highest efficiency, the particle energy gain per unit power dissipation must be maximized, a quantity known as the shunt impedance, which is given by

$$r_s = \frac{V_0^2}{P},$$
$$Z \equiv \frac{r_s}{L} = \frac{E_0^2}{P/L}, \tag{5}$$

where r_s is the shunt impedance, P is the power dissipated in the cavity walls, E_0 is the axial electric field, and Z is the quantity to be maximized, the shunt impedance per unit length. V_0 is the axial cavity voltage given by

$$V_0 = \int_{-L/2}^{L/2} E(0, z)dz, \tag{6}$$

where L is the cavity length. Z can be increased with the addition of nose cones as shown in Fig. 3 that increase the axial electric field relative to the total stored energy [16]. By doing this, though, the peak surface electric fields are increased, which can eventually lead to rf breakdown if they exceed a certain limit; therefore, the quantity E_{pk} must be minimized to avoid this, leading to the first set of conflicting objectives. Another quantity that has recently been found to give a good indication of rf breakdown is the modified Poynting vector [17] given by

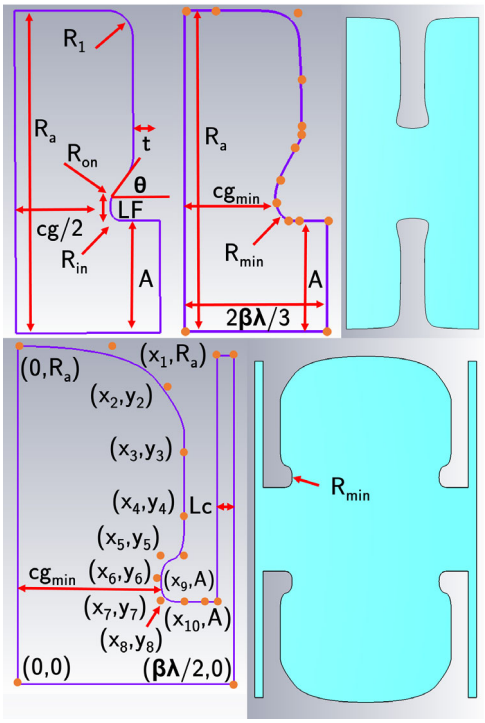

FIG. 3. Standard and spline geometries for the X-band traveling wave cavity (top) and the C-band standing wave cavity (bottom), showing input parameters.

$$S_c = ||\mathrm{Re}(\bar{S})|| + g_c \cdot ||\mathrm{Im}(\bar{S})||, \tag{7}$$

where $\bar{S}$ is the complex Poynting vector and g_c is a constant obtained from fitting to experimental data, normally taken as $1/6$. This is the second quantity to be minimized. Finally, high surface magnetic fields cause rf pulsed heating, which can eventually lead to cracks and surface roughening, essentially limiting the accelerating gradient [18]. This needs to be avoided, and, therefore, the final objective to be minimized is the peak magnetic field (B_{pk}). Normalizing the peak fields to the accelerating gradient and putting them into appropriate units, the four final objectives are shown in Table I.

B. Constraints

For both the standing wave and traveling wave cavities, the frequency of the fundamental accelerating mode (TM_{010}) must be kept constant. In reality, it is difficult

TABLE I. Optimization objectives.

Quantity	Min/Max	Units
$E_{\mathrm{pk}}/E_{\mathrm{acc}}$	Min	Unitless
$B_{\mathrm{pk}}/E_{\mathrm{acc}}$	Min	mT/MV m
$\sqrt{S_c}/E_{\mathrm{acc}}$	Min	$\sqrt{\frac{A}{V}}$
Z	Max	$\frac{M\Omega}{m}$

to keep exactly constant, and, therefore, a small value of ϵ is chosen for simulation purposes such that

$$f_{\mathrm{TM}_{010}} = f_{\mathrm{desired}} \pm \epsilon \tag{8}$$

with typical values for ϵ of 5–30 MHz. The frequency can later be tuned to the exact value with little effect on the other objectives. For the traveling wave cavity, another important parameter that must be considered is the group velocity v_g. The group velocity is given by

$$v_g = \frac{d\omega}{dk}, \tag{9}$$

and it defines the velocity of the energy flow through the structure, shown by the rf power flow:

$$P_{\mathrm{rf}} = v_g U, \tag{10}$$

where ω is the angular frequency, k is the wave number, and U is the cavity stored energy per unit length. v_g can either vary along the length of a linac or remain constant, and it is normally chosen to balance the electromagnetic fill time and the field amplitude or power dissipation requirements. A number of methods for dealing with this constraint in the optimization are explored in the next section.

C. Modeling methods

Two types of cavity modeling methods were considered, one using standard shapes such as circular radii and angles to define the nose cone geometry (referred to as the standard parameters model henceforth) and the other a model generated using nonuniform rational basis splines (NURBS). This method uses a series of control points with a spline interpolated between them to define the cavity geometry. This allows for much greater control of the nose cone and outer cavity wall shapes and has been shown to increase the shunt impedance by as much as 13% [19]. The two generic cavities that have been used in this optimization are shown in Fig. 3. The first (top) is an X-band (12 GHz), $\beta = 1$, $2\pi/3$ traveling wave cavity, and the second (bottom) is a C-band (5.712 GHz), $\beta = 1$, $\pi/2$ biperiodic standing wave cavity. The positions of the control points were related back to the standard cavity parameters by defining a minimum nose cone gap and a minimum nose cone radius, chosen by estimating the minimum values that could feasibly be manufactured. In both cases, the control point positions were set in such a way that concave nose cone shapes were not possible in order to reduce the search space.

III. OPTIMIZATION

To calculate the electromagnetic fields in the cavities, the eigenmode solver in the commercial code CST Studio Suite® [20] was used. The optimizations were performed using Isight® [21], a software package for design optimization and exploration. The two main MOGAs used for the cavity optimization were the well-known fast nondominated sorting genetic algorithm (NSGA-II) [22] and the archive-based micro genetic algorithm (AMGA), which offers an improvement over NSGA-II by using an archive to maintain its search history [23].

A. Algorithms

An example of a generic MOGA flow chart is shown in Fig. 4. These methods are inspired by the process of evolution in biology. An initial population is sampled from the design space and then progressed using selection, crossover, and mutation to form offspring. This process is then repeated, improving the population characteristics until a stop criterion is achieved.

The baseline algorithm parameters shown in Table II were selected based on the number of inputs with the parameters kept as close as possible for a fairer comparison of the algorithms. The crossover probability controls the probability of whether parent solutions are recombined to generate the offspring, and the mutation probability controls the probability that an offspring is mutated; this is left as the optimum choice of 1/n, where n is the number of design variables. Both the crossover distribution index and mutation distribution index are inversely proportional to the amount of perturbation in the design variables; therefore, smaller values improve the resilience to premature converge but reduce the focus of the search. The number of generations defines the number of times the processes of mutation and crossover are completed, and the number

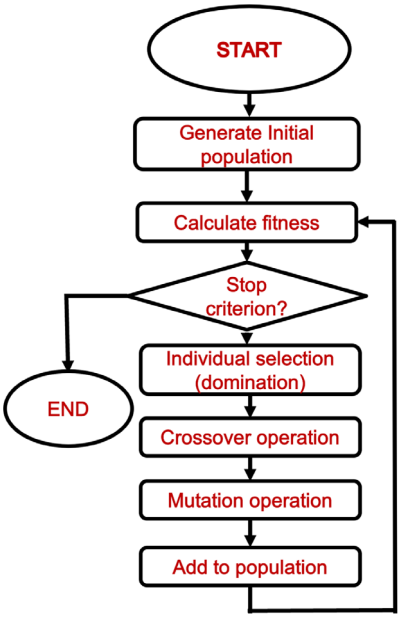

FIG. 4. Generic MOGA flow chart showing how the algorithm progresses, improving the population through successive generations.

TABLE II. Baseline algorithm parameter selections.

Parameter	NSGA-II	AMGA
Initial size	60	500
Population size	60	40
No. function evaluations	⋯	12 000
Archive size limit	⋯	500
Crossover probability	0.9	0.9
Mutation probability	1/n	1/n
Crossover distribution index	9	9
Mutation distribution index	18	18
No. generations	200	⋯

of total evaluations can be calculated by multiplying the population size and the number of generations. The simulations can be initialized in various ways, including initializing the population with random points within the defined parameter bounds; seeding the simulation with a single point, from which it generates an exponentially decaying cloud of points around the initial point to explore similar designs; and seeding with a starting population from an earlier simulation which then acts as the first generation of the simulation.

B. Pareto front convergence

An example of a generated 2D Pareto front is shown in Fig. 5, showing the trade-off between Z and $E_{\mathrm{pk}}/E_{\mathrm{acc}}$. Although the "true" Pareto frontier is not known, information about the convergence characteristics can be obtained by plotting the Pareto front vs the generation number for a particular simulation as shown in the figure. Between the values of $E_{\mathrm{pk}}/E_{\mathrm{acc}} = 2.5$–$5$, the front has already converged after approximately 125 generations, while future generations increase the spread of the front at the extreme ends. The difference between the Pareto fronts for 200 and 250 generations is negligible with most of the points falling on the same curve, meaning that, for this particular set of simulation parameters and constraints, there is little benefit to running for more generations.

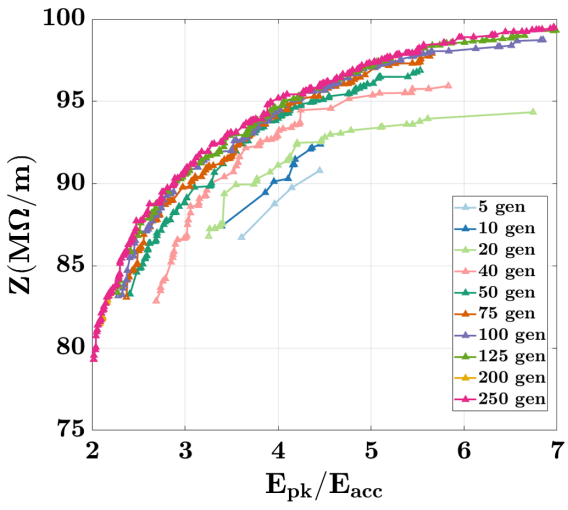

FIG. 5. Z vs $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ Pareto front convergence example showing how the Pareto front converges after 250 generations for the C-band standing wave (SW) structure.

C. Group velocity condition

Figure 6 shows some of the results of an investigation into the algorithms and various methods for dealing with the v_g condition for the traveling wave cavity. One of the main parameters that affects v_g is the aperture radius, and, therefore, this had to be taken into consideration in the simulations. The options considered for handling the v_g requirements included having v_g as a constraint, similar to the frequency, and allowing the aperture radius to vary; having v_g as an objective with an appropriate choice of the aperture radius upper and lower bounds constraint; choosing a fixed value of the aperture radius based on estimates and allowing v_g to vary between constrained bounds, e.g., 0.95%–1.05% c; and having v_g as both an objective and a constraint.

With v_g as a constraint (±0.05% c) only, the algorithm convergence speed was slow, with a maximum Z of 104 MΩ/m achieved after 12 000 evaluations. With the large range of possible aperture radii, the algorithm struggled to converge on values that would provide a large shunt impedance as it was not actively searching. When using v_g as an objective, both the NSGA-II and AMGA algorithms were able to find more appropriate values for the aperture radius, but many of the obtained values for v_g fell outside of the ±0.05% c range, meaning they were unfeasible. Better performance was obtained by estimating a good value of the aperture radius and keeping it fixed.

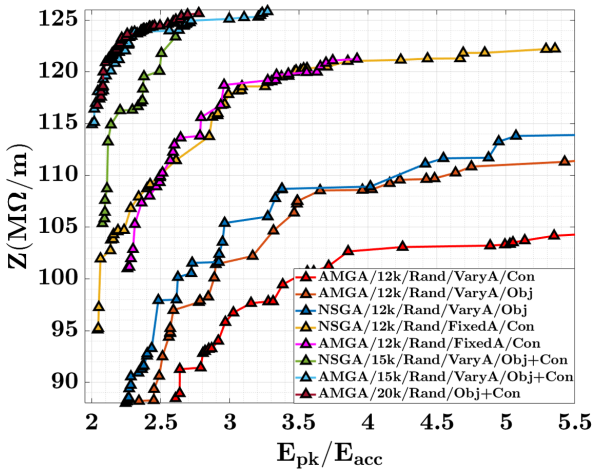

FIG. 6. Algorithm and constraint handling method comparison for the traveling wave (TW) X-band structure, showing the Z vs $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ Pareto fronts. The legend follows the format algorithm/number of individual evaluations/initialization mode/fixed or varying aperture radius/constraint handling technique.

The estimate was made by using the standard parameter cavity and changing the nose cone gap to achieve a large value of Z while keeping v_g within the acceptable bounds. Both algorithms converged to around 120 MΩ/m, which is a significant improvement, but Z is dependent on the aperture radius, so, by fixing it, the maximum value of Z also becomes limited. The best results were obtained by having v_g as an objective while also constraining it to $\pm 0.05\%$ c values. This meant that the algorithms were able to search for the best values of the aperture radius while also remaining within the feasible limits. Figure 6 also shows that the AMGA algorithm performance was significantly better for the objective and constraint case, although the opposite is true when the aperture radius took a fixed value with NSGA-II outperforming AMGA slightly. Based on this, for the final optimizations, AMGA was used for the traveling wave cavity and NSGA-II for the standing wave. The optimization results shown here are for a single TW cell, although typically v_g will vary along the length of the linac. Performing an individual optimization for each individual cell would be computationally expensive; therefore, we would propose that one optimization would be performed for a group of 5–10 TW cells, with the aperture radius being adjusted for each to obtain the required v_g. This would then be repeated for the next group of cells along the linac until the total desired number of cells is reached.

D. Frequency handling

Figure 7 shows the feasible points in the Z vs $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ 2D objective space with a total frequency constraint of 60 MHz. By reducing ϵ from 30 to 1 MHz and finding the Pareto frontier each time, the relative effect of the frequency constraint on the final estimated Pareto frontier can be found. At higher values of Z, there is a more noticeable shift in the Pareto frontier, most likely because of the dependence of Z on the cavity radius and frequency. As ϵ is reduced to 1 MHz, a larger shift in the frontier is seen with increased noise due to the lower number of feasible solutions. A second simulation with a reduced ϵ that is seeded with individuals from the first run could be used to converge to a more accurate frontier, although this would be at the expense of a longer total simulation time. Reducing ϵ by 10 MHz approximately halves the number of feasible points and, therefore, doubles the time (6000 to 12 000 evaluations) to a similar level of convergence. As the time to convergence is directly related to the ϵ value, the use of $\epsilon = 20$ MHz is justified as a trade-off between simulation time and accuracy. Individual solutions can also eventually be tuned to the correct frequency within the bounds with minimal effect on the other objective parameters.

E. Geometry results

Figure 8 shows the results of a comparison between the converged SW cavities, one modeled with the NURBS and one with standard parameters. We can see that the spline performance at lower values of $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ is ≈10% better with the difference between the two decreasing as $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ increases, confirming the findings in Ref. [19]. This difference can be explained by observing the geometries of the spline model at low $E_{\mathrm{pk}}/E_{\mathrm{acc}}$. These geometries are difficult to realize with only the standard parameters, whereas as the nose cone increases in size and length it becomes increasingly easy to recreate a similar geometry with the standard parameters. It is noted that it could also be the case that the true Pareto front has been missed and that the differences exist only for $E_{\mathrm{pk}}/E_{\mathrm{acc}} < 3$, although further investigation will be required.

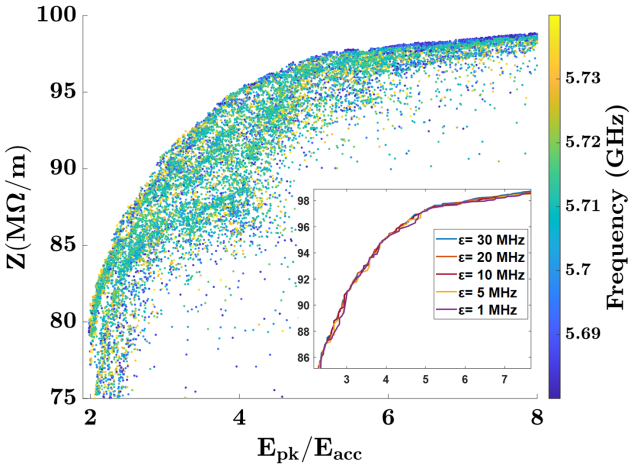

FIG. 7. Frequency handling investigation results, showing all optimization points and the variation in the Z vs $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ Pareto fronts for different values of ϵ (the feasible difference from the ideal frequency).

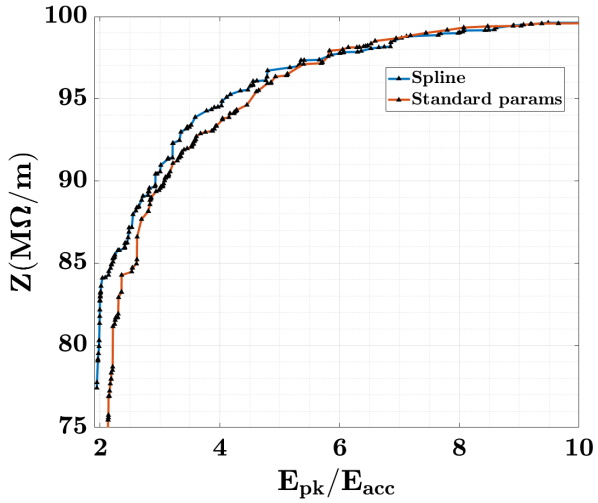

FIG. 8. Comparison of Z vs $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ Pareto fronts for spline model vs standard parameters for the C-band cavity, showing an improvement when splines are used.

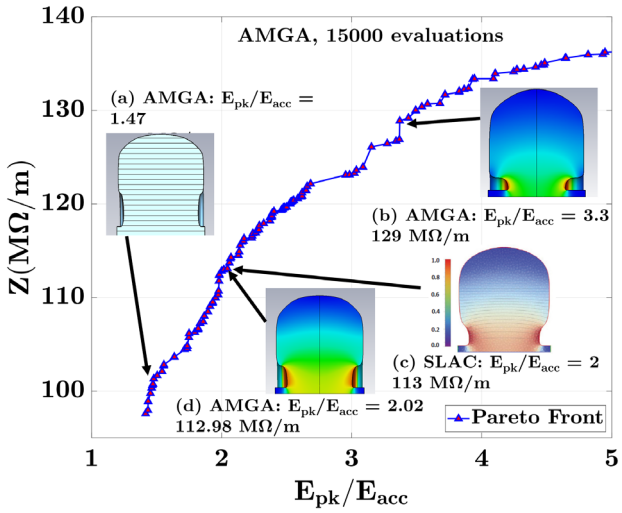

FIG. 9. Z vs $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ Pareto front for an S-band SW cavity optimization. SLAC image reproduced from Ref. [19].

IV. SINGLE- VS MULTIOBJECTIVE OPTIMIZATION

In order to illustrate the power of this approach, an example is provided in Fig. 9. The cavity used in this example is described in Ref. [19] and is an S-band SW structure that was optimized with a combined optimization function of

$$\text{optimization function} = \frac{Z}{\mathrm{Max}[\frac{E_{\mathrm{pk}}}{E_{\mathrm{acc}}}, 2]}. \tag{11}$$

This was used along with a NURBS geometry and evolutionary algorithm to find a cavity with $E_{\mathrm{pk}}/E_{\mathrm{acc}} = 2$ and $Z = 113\ \mathrm{M\Omega/m}$ within 20 000 evaluations. This cavity is shown in Fig. 9(c). In order to compare with this result, a similar geometry was set up and ran with the AMGA algorithm for 15 000 evaluations with the weightings chosen to explore the region where the original cavity lies. The algorithm manages to find a very similar point to the original [Fig. 9(d)], and the Pareto front provides new information about the objective space. For example, we now know that we can reduce the peak fields by ≈25% [Fig. 9(a)] for only an 11% reduction in Z, possibly allowing for a higher gradient before breakdown. We could also move up to $E_{\mathrm{pk}}/E_{\mathrm{acc}} = 3.3$ [Fig. 9(b)] for an improved Z of 129 MΩ/m at the expense of $E_{\mathrm{pk}}/E_{\mathrm{acc}}$.

V. VISUALIZING THE N-DIMENSIONAL PARETO OPTIMAL OBJECTIVE SPACE

After the chosen MOGA is applied to the single-cell design space, the output solutions construct the objective space. A subset of the objective space is the Pareto optimal subspace that constructs the Pareto frontier. Within the Pareto optimal subspace, by definition, no further solutions can be removed, as no single solution is dominated by all other solutions. Certain solutions within this subset will have more desirable objective values, however to the detriment of other objective values.

Considering two objectives, we can construct a 2D Pareto front, showing the set of optimal solutions. Often, a 2D Pareto front makes a convex or concave shape, and the trade-off between the two objectives is easily visualized. Although 2D fronts do not provide information about all the objectives, as a first check the 2D fronts can be plotted to immediately explore the trade-offs between two objectives. For example, Fig. 10(a) shows the Pareto front for Z vs $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ while also showing the values of $\sqrt{S_c}/E_{\mathrm{acc}}$ and $B_{\mathrm{pk}}/E_{\mathrm{acc}}$ in color. By including the geometry at select points, the trade-off between $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ and Z can be understood—a longer nose cone increases Z but also causes the peak electric field to increase, as expected. Both $\sqrt{S_c}/E_{\mathrm{acc}}$ and $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ increase with increasing Z, although a Pareto front between them does exist, with slight differences in the geometry favoring one over the other. This has been confirmed in other work [24], with the Pareto frontier in Fig. 10(b) allowing this transition to be visualized for the traveling wave cavity case with solutions to the left minimizing $\sqrt{S_c}/E_{\mathrm{acc}}$ and solutions on the right minimizing $E_{\mathrm{pk}}/E_{\mathrm{acc}}$.

Although these 2D plots can be useful for exploring the trade-offs to some extent, if the number of objectives increases from two to N, the Pareto front is now a (N − 1)-dimensional hypersurface. There is no guarantee that points that lie on one Pareto front are close to the others. This is demonstrated in Fig. 11, where the points from each 2D Pareto front are plotted on the same axes. The Pareto front of Z vs $\sqrt{S_c}/E_{\mathrm{acc}}$ is shown in green, along with the points that fall on the other two Pareto fronts and points that fall on more than one front. It should be noted that no points fall on all three Pareto fronts and very few fall on two, meaning that selecting candidates using 2D plots alone is not feasible. In light of this, more intricate methods of visualization must be implemented when N ≥ 3 in order to assist in the selection process.

In this study, we wish to select a single-cell design based upon how it performs with respect to the four objectives discussed in Sec. II. The MOGA provides a single set of solutions which puts limits on the visualization techniques that can be implemented. As each MOGA solution refers to an independent cell geometry, they must be visualized independently, which leaves the techniques shown in the red solid and dashed box in Fig. 12.

A. Issues associated with dimensionality reduction

While the techniques shown in the dashed red box in Fig. 12 provide methods to visualize independent solutions of a single set objective space, the original objective values are transformed and displayed. This is undesirable, as the

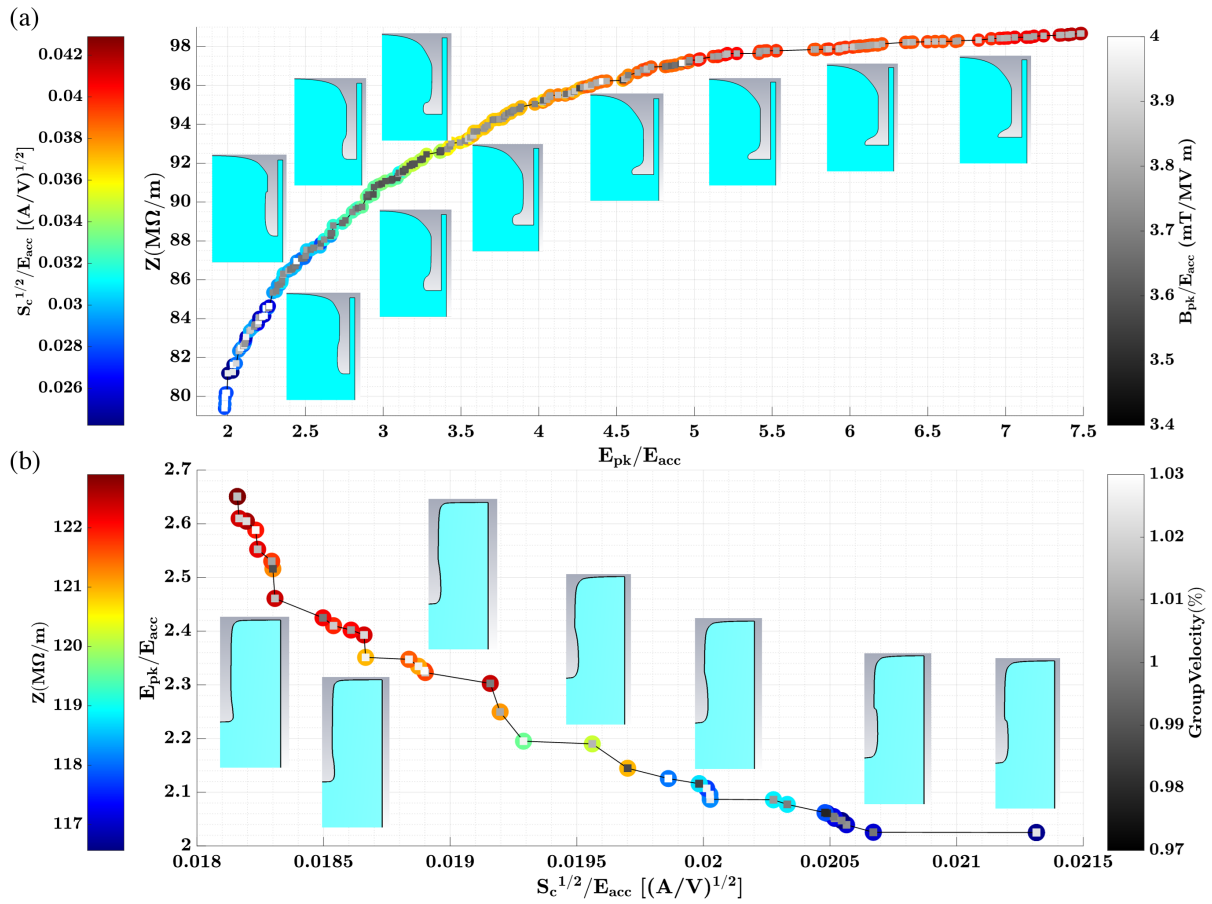

FIG. 10. Pareto fronts showing cavity geometry. (a) Z vs $E_{\mathrm{pk}}/E_{\mathrm{acc}}$, SW, with $B_{\mathrm{pk}}/E_{\mathrm{acc}}$ and $\sqrt{S_c}/E_{\mathrm{acc}}$ in color; (b) $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ vs $\sqrt{S_c}/E_{\mathrm{acc}}$, TW, with Z and v_g in color.

aim is to provide a set of plots that allow a user to understand the trade-offs between the original single-cell objectives. Many visualization techniques rely on dimensionality reduction (DR) or goal functions. While these methods can be appropriate, they can lead to the loss of individual objective information. Some common examples of DR are prosection [26] and 3D radial coordinate visualization (3dRadVis) [27]. Prosection is a DR technique that combines two objectives into one by using the Pythagoras theorem on a localized section of the Pareto optimal subspace. This method is highly effective at combining objectives; however, as this study requires the conservation of original objective values, it is not appropriate.

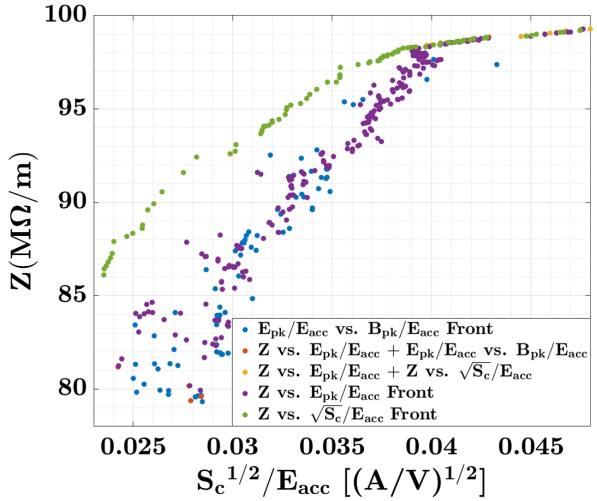

FIG. 11. Z vs $\sqrt{S_c}/E_{\mathrm{acc}}$ objective space. Points that fall on the Z vs $E_{\mathrm{pk}}/E_{\mathrm{acc}}$, $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ vs $B_{\mathrm{pk}}/E_{\mathrm{acc}}$, and multiple fronts are shown.

3dRadVis is a DR technique that performs a reduction ($N \rightarrow 2$) and plots the solution as a point within a unit circle. Objective nodes are placed evenly spaced around a unit circle, and solutions with higher objective values are plotted closer to the respective objective node. It can be thought that a solution is connected to each objective node via a spring with the spring constant proportional to the value of that objective (normalized to the maximum objective value). The solution is then plotted where the N springs would equilibrate. Figure 13 displays a 3dRadVis plot with three objective nodes (three peak fields), and the color dimension is Z. While the plot does show some insight into the objective space, such as the relative trade-offs between objectives for high and low Z, it does not give

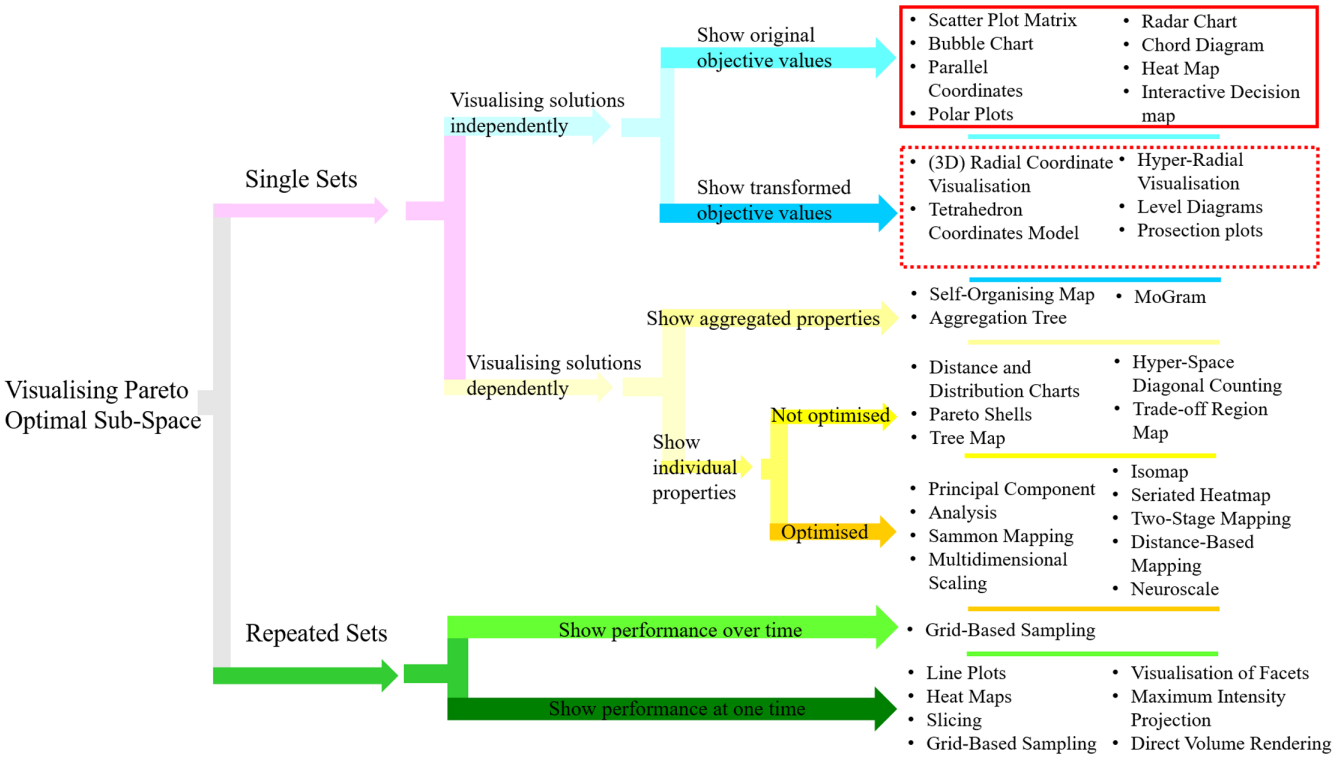

FIG. 12. Summary of methods that aid the visualization of a Pareto optimal subset for various types of dataset. Adapted from Ref. [25].

any indication of individual values. In addition, there is no information regarding whether points close to an objective node are large in that objective, very small in other objectives, or a mixture of the two. This leads us to search for techniques that show trade-offs between objective values while also preserving the individual objective values, so that a selection process can occur. Note that DR can be split into two sections, feature extraction (reducing the dimension of the dataset, what we mean when we say DR), and feature selection (reducing the dataset to a subset holding interesting features; the MOGA does this for us) [28].

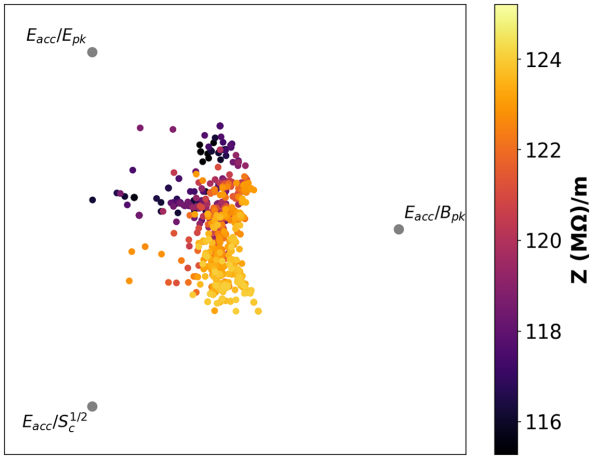

FIG. 13. 3dRadVis example showing three surface fields as objectives with color dimension showing Z.

Not all datasets retain all useful information when acted upon with a DR operator. This can be interpreted as follows; suppose we wish to reduce a solution of N dimensions to N − 1. The solution can be represented with an (N × 1) vector of objective values. If we now acted upon this vector with an (N − 1 × N) DR matrix, the resulting vector will be (N − 1 × 1):

$$\begin{pmatrix} A & B & C \\ D & E & F \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} AX+BY+CZ \\ DX+EY+FZ \end{pmatrix} = \begin{pmatrix} X' \\ Y' \end{pmatrix}, \quad (12)$$

where (X, Y, Z) is a vector of objectives and (X', Y') is the new vector of objectives after DR has been applied. It is clear that reducing the dimension by one has lost all information regarding the original objective values; instead, the remaining vector can be thought of as a goal function. As the dimension has been reduced, the DR matrix is nonsquare and does not have an inverse. This means, given a vector (X', Y'), the original values (X, Y, Z) cannot be determined.

As an example, consider a circular particle distribution with Gaussian particle density. The dataset can be reduced in dimension $(x, y \to r)$. While the original values of x, y are lost, the original circular distribution can be recreated as the lost information has little importance. This reflects the

fact that the dataset has rotational symmetry. If, instead, the particle distribution is elliptical, performing DR $(x, y \rightarrow r)$ will lose important information regarding the rotation angle of the ellipse. In the context of this study, combining $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ and $\sqrt{S_c}/E_{\mathrm{acc}}$ into one objective (as is done in prosection) will enable good points to be selected; however, all information of individual objectives, which is important in single-cell design, is forfeit. The aim is to provide visualization methods that can be repeated for all single-cell design studies and preserve all original objective values.

B. Clustering Pareto optimal subspace

From the definition of a nondominated solution, it is possible that a MOGA finds two similar solutions, with one slightly better than the other in one objective but slightly worse in another. Various clustering algorithms exist that group similar solutions based on objective values [29], that could be used to reduce the dataset, while still retaining the important information. The algorithm we have chosen to implement in this study is the agglomerative hierarchical clustering algorithm [30]. The term agglomerative refers to the merging of data from N clusters and N solutions into n clusters for N solutions, where $n \leq N$. The algorithm creates clusters based on the Euclidean distance between solutions in objective space. The user selects the number of clusters created. The question of the optimal number of clusters is not trivial, however. The objective space is, in theory, continuous. Clusters do have discrete boundaries but have a specific chosen cutoff, where a solution could be in either cluster 1 or cluster 2, based on the total numbers of clusters chosen. If the number of clusters is too few, dissimilar solutions are clustered together, rendering the clustering meaningless. If too many clusters are chosen, the visualisation can look overcrowded, reducing the effectiveness. As each cluster of solutions will be displayed using a visualization technique as a single solution, it must be as representative of the cluster as possible. The chosen solution is found by giving each solution within a cluster a weighted value. The equation used to determine the individual solution weight is as follows:

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

where obj_n is the nth objective normalized value and range_n is the normalized range for that objective. This weighting provides an accurate way to represent the individual cluster, which can then be used for visualization.

In order to investigate the effect of clustering on the points, parallel coordinate plots (PCPs) have been used. PCPs display a solution as a line crossing equally spaced axes. Each axis represents an objective, and the objective value is given by the point at which the solution crosses the axis. The production of a parallel coordinate possesses four degrees of freedom: (i) objective axis order, (ii) axis scale (logarithmic, linear, etc.), (iii) axis rotation (maximizing or minimizing), and (iv) plot order of solutions (important for larger datasets). PCPs are able to display as many objectives as the screen resolution will allow, making them a popular choice for multiobjective visualization problems. An example of the use of parallel coordinate plots is shown in Fig. 14(a), where the four objectives are shown on the four y axes, with the inverse of $E_{\mathrm{pk}}/E_{\mathrm{acc}}$, $B_{\mathrm{pk}}/E_{\mathrm{acc}}$, and $\sqrt{S_c}/E_{\mathrm{acc}}$ used so that all objectives require maximizing, making the PCP easier to read, as better solutions are those with the highest value for each objective.

Figure 14 also shows how the number of total clusters changes the intercluster structure. Figure 14(a) shows that ten clusters is too few, as a large variance can be observed within the cluster. When 30 clusters are used, as shown in Fig. 14(b), the variance is reduced, and all solutions have similar values for the objectives. Going from 30 clusters to 100 clusters in Fig. 14(c) does not improve the situation enough to allow for the increased number of solutions, which reduces the use of visualization techniques as the plots become harder to read. Another way to look at this is

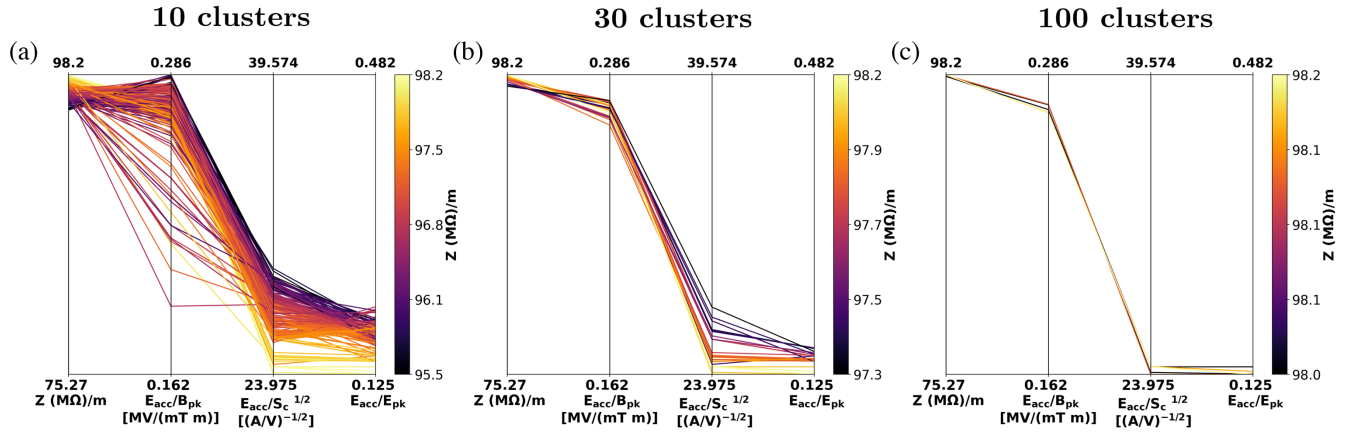

FIG. 14. Three PCPs showing the objective values inside of a single cluster for different numbers of total clusters: (a) 10, (b) 30, and (c) 100 clusters. The opened clusters are the ones with the highest value of Z.

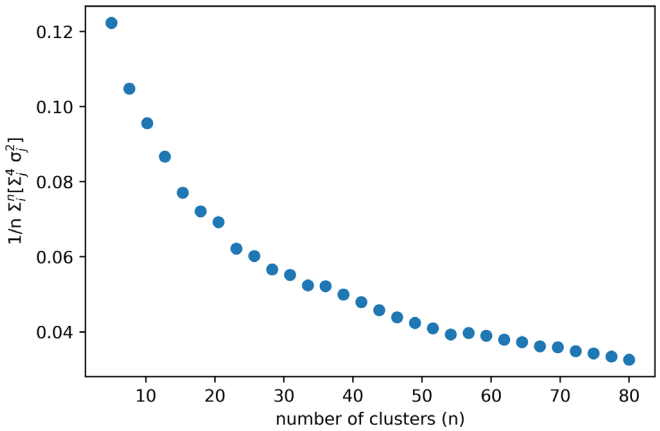

FIG. 15. Mean standard deviation of objective cluster as a function of the number of clusters.

to use the standard deviation of the clusters. The standard deviation of a cluster can be defined by squaring the standard deviation of each objective in a cluster and summing together. Figure 15 shows how the number of clusters affects the error associated with clustering. Increasing the number of clusters from five decreases the mean standard deviation quickly until around 20 clusters. As the number of clusters increases over 20, the mean standard deviation drops at a slower rate while also increasing the number of refined data points. A natural choice of clusters resides somewhere between 20 and 30.

C. Cluster variance

In order to visualize the variance within each cluster, PCP plots were created for cases with 12, 24, 36, and 48 clusters. The results are shown in Fig. 16, with only the first

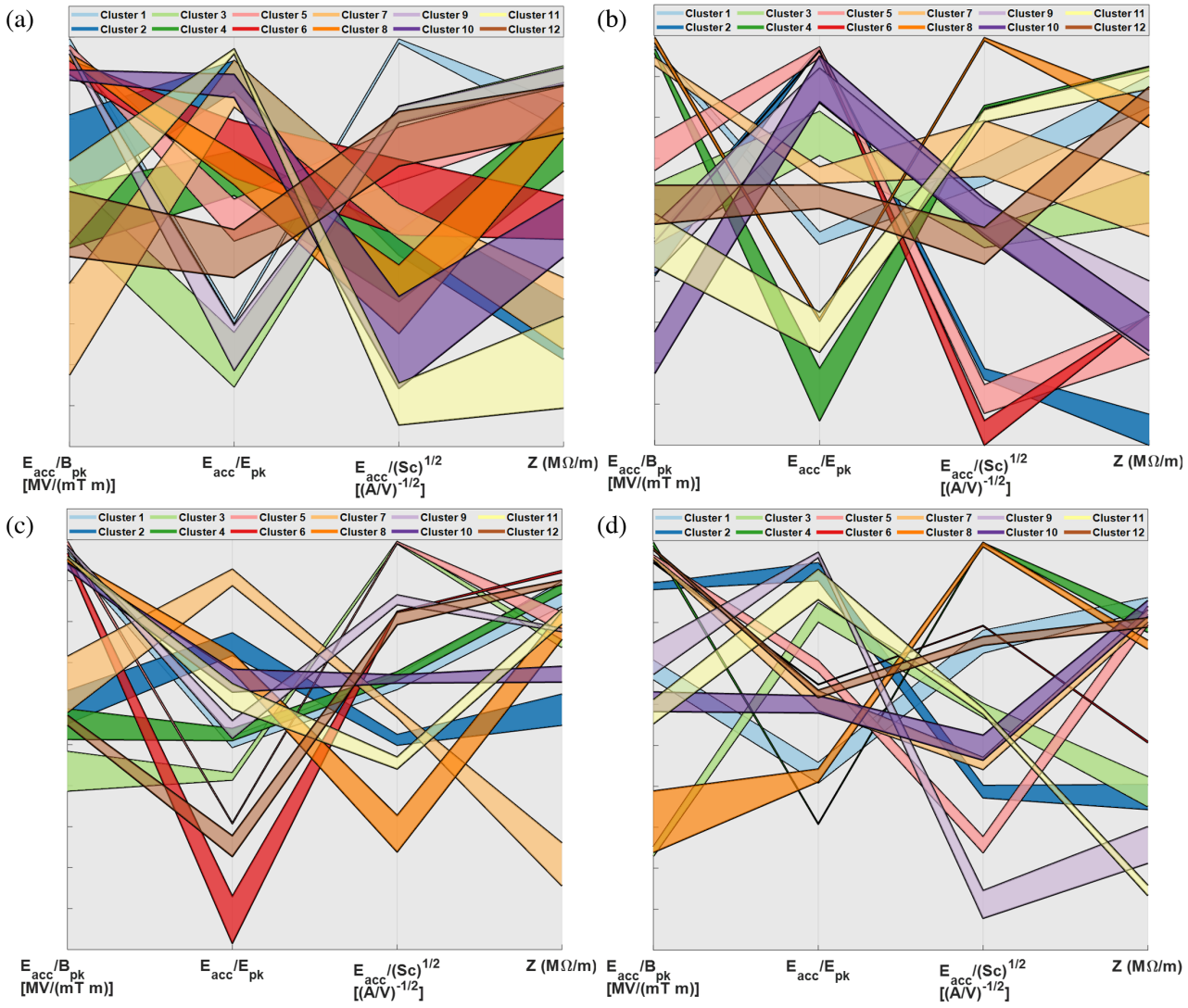

FIG. 16. PCP plots for four different clustering cases. Each line represents one cluster, and the width of the shaded regions represents the 25% and 75% percentiles within each cluster. (a) 12 clusters, (b) first 12 clusters from 24 clusters, (c) first 12 clusters from 36 clusters, and (d) first 12 clusters from 48 clusters.

12 clusters from each group being shown. The clusters are represented with a shaded region which is their 25% and 75% percentiles. In Fig. 16(a), with only 12 clusters there is a significant overlap between the clusters, meaning that, by selecting one for specific objective traits, other solutions that fall into the nearby clusters could be missed. Clearly, as the number of clusters is increased, the variance is reduced, and it becomes easier to select an individual cluster with the objective values required. As the number of clusters becomes too large (> 40), the PCP plots of the centroids become too difficult to read, defeating the purpose of clustering. For 24–36 clusters, the clusters are easily distinguishable and can be read on a single-line PCP plot, allowing individual clusters to be selected. Although some overlap still exists, this can be overcome by inspecting two or three clusters that are close together instead of one in order to select a final solution.

VI. VISUALIZATION TECHNIQUES APPLIED TO RF SINGLE CELLS

The visualization methods discussed here will explore their application to the results of the C-band SW single-cell and X-band TW single-cell optimization.

The methods discussed are as follows: parallel coordinate plots [31], radar charts [32], polar plots, decision maps (DMs) [33], heat maps [34], and 3D Pareto surface approximations.

We show how using the methods aids the selection process for two example single-cell designs. It is worth noting that the visualization techniques shown are used to aid a selection, not provide a final design. In reality, different designs possess different design specifications, and those design requirements coupled with the methods shown can lead to the final selection.

A. Parallel coordinate plots

Figure 17 shows a PCP for a C-band SW single cell. The plot is shown twice with different plot order, showcasing degree of freedom 4. The left plot prioritizes showing low-Z Pareto optimal solutions and the relationship between low-Z single cells and the surface field values. The right plot shows high-Z solutions more clearly. A selection process is not readily obvious due to the large number of solutions, hence the need for clustering. PCPs can also be used to explore the entire input and objective space; this is particularly useful if the plot is color coded with the density of lines as shown in Fig. 18. It is clear from this plot which variables have the largest effects on the objectives and can also inform the choice of geometrical constraints or values of the input parameters for future optimization runs.

A refined PCP is plotted in Fig. 19. The plot is more readable than Fig. 17, as all solutions can be traced fully. High-Z solutions also have low peak surface magnetic fields; however, they have high $S_c^{1/2}/E_{\mathrm{acc}}$ and $E_{\mathrm{pk}}/E_{\mathrm{acc}}$. Conversely, low-$Z$ solutions have low $S_c^{1/2}/E_{\mathrm{acc}}$ and $E_{\mathrm{pk}}/E_{\mathrm{acc}}$; however, they can adopt a large range of values for $B_{\mathrm{pk}}/E_{\mathrm{acc}}$. A particular solution can be selected by observing the plot. Very high values of Z result in very large surface peak electric and S_c fields. Similarly, low peak electric field returns very low Z. This essentially omits solutions at the two extremes in Z (the yellow and black lines in Fig. 19). Solutions that give small peak fields while not compromising massively on Z are the solutions in purple. Another way to arrive at this is to take the only solutions we have not discussed, orange solutions. Orange solutions offer high Z and peak surface magnetic fields; however, they have high S_c and surface electric field. This means we achieve highly in two objectives and low in the remaining two. Purple solutions allow for high values in

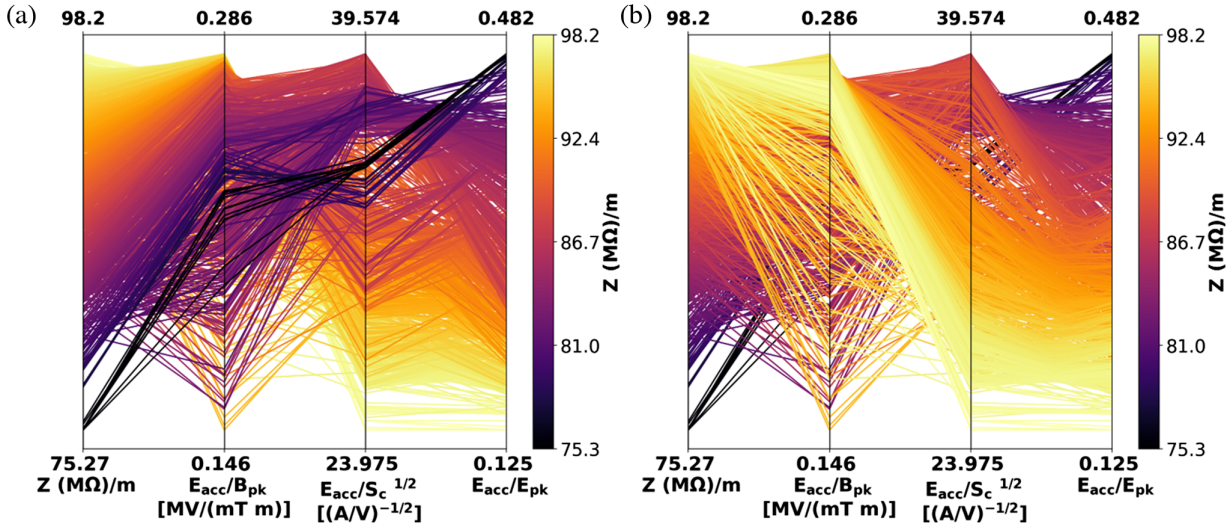

FIG. 17. PCPs showing the Pareto optimal space for a C-band SW single cell. The same plot is displayed twice with different plot order, where (b) is the reverse of (a). All objectives require maximising.

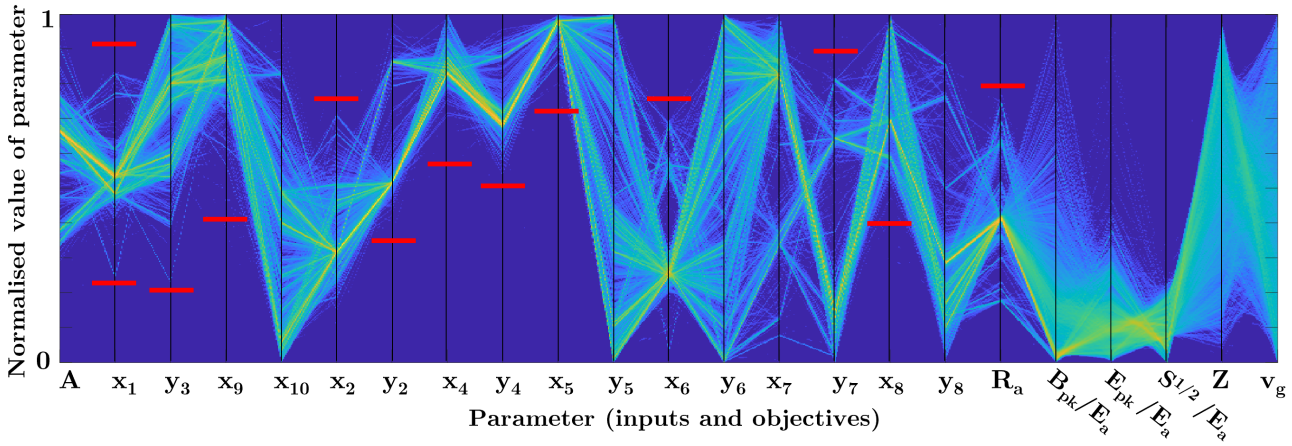

FIG. 18. Density PCP for X-band traveling wave cavity, showing input and output spaces. The density of the lines represents the number of solutions in a certain region, and red lines show where initial input space constraints could potentially be reduced for improved convergence speed [35].

three objectives while compromising only slightly on the remaining objective, Z. If shunt impedance per unit length was of particular importance, the pink band of solutions between the orange and purple solutions are possible. PCPs allow the user to understand how the range of objective values interact. This can be realized in Fig. 19, where the majority of clustered solutions pass through the same point on the $E_{\mathrm{acc}}/B_{\mathrm{pk}}$ axis.

B. Radar charts

Radar charts (RCs, also known as spider charts) are similar to PCPs, the difference being the solutions are plotted across axes equally spaced in θ, as opposed to the x or y Cartesian axis. RCs are visually more interesting than PCPs; however, comparing objectives for a single solution is more difficult, as the distance along each axis is harder to assess relative to parallel axes. When many solutions are plotted, RCs aid visualizing certain features of a dataset due to the periodicity of the plot. Each solution makes one complete turn and as a result provides an additional objective interaction relative to PCPs. An RC is shown in Fig. 20 for a C-band SW single cell with 20 clusters.

The RC shows essentially the same information as the PCP in Fig. 19, however with fewer clusters. The RC shows a symmetry between $E_{\mathrm{acc}}/E_{\mathrm{pk}}$ and $E_{\mathrm{acc}}/S_c^{1/2}$ in addition to a symmetry between $E_{\mathrm{acc}}/B_{\mathrm{pk}}$ and Z. Again, it is easy to see the solutions in purple offer good solutions that minimize the peak fields while not minimizing Z. The RC also shows the pink solution as the best trade-off between all four objectives, provided they are of equal

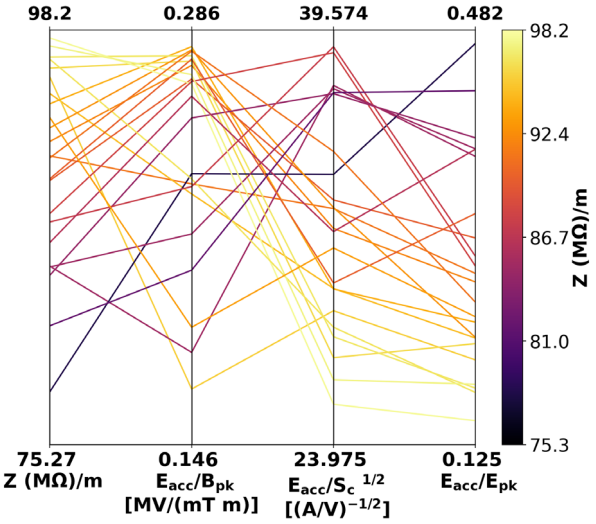

FIG. 19. Refined PCP for C-band SW single cell, showing four objectives with 30 clustered solutions.

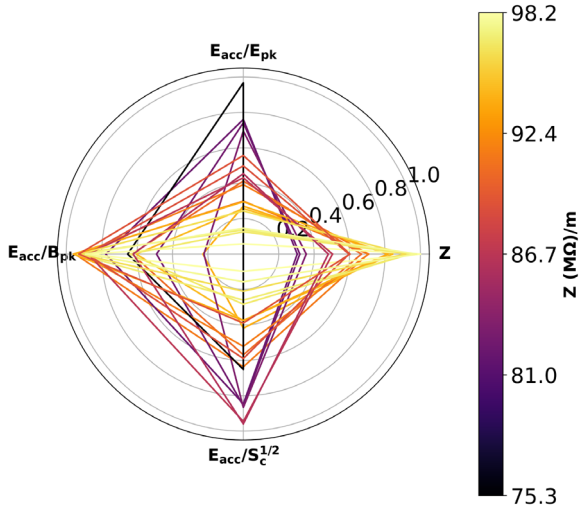

FIG. 20. Refined RC showing four objectives for the C-band SW single cell with 20 clustered solutions.

importance. The orange and yellow solutions show very strong values for Z and $B_{\mathrm{pk}}/E_{\mathrm{acc}}$; however, they require the user to tolerate high $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ and S_c.

PCPs and RCs allow datasets with a large number of objectives to be compared with a single plot, making them an attractive technique for multiobjective visualization. The addition of a color scale incorporates an additional level of clarity, which can allow objectives physically far away on a plot to be compared with more ease. While all visualization techniques display the data in a way that allows the user to draw connections between various objectives, PCPs and RCs have the ability to display this information in a very simple form. The relationship between Z and $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ is far more easily interpreted from a PCP than a DM. This is because PCPs display raw data. PCPs and RCs can also quickly show when there is not a clear correlation between objectives. An example of this is shown in Fig. 20, which demonstrates the lack of structure between $B_{\mathrm{pk}}/E_{\mathrm{acc}}$ and $S_c^{1/2}/E_{\mathrm{acc}}$. Relationships between nonadjacent objectives are easier to understand using an RC. RC solutions complete one full turn when plotted and, as a result, provide an additional objective connection than a PCP. RCs are also able to show subtle symmetry between objectives. This can be seen in Fig. 20, where the clustered data can be effectively separated further into two main clusters. The first main cluster creates a flat diamond shape (in yellow) and corresponds to all high-Z solutions. The other main cluster (purple and black) creates a lopsided kite that is maximized in all peak fields and has relatively low Z. The solutions that offer the best trade-off are the solutions that do not readily fit into either of these main clusters or have properties that can be linked with both main clusters (solution in pink).

PCPs and RCs, however, are useful for selecting a final design only if the user has compressed the original dataset into a small number of clusters. The PCP or RC becomes rapidly less useful as the number of plotted solutions increases and effectively works only when the number of solutions is already small. Clustered PCPs and RCs provide a wide array of clustered solutions from which a final design can chosen based upon design requirements.

C. Decision maps

Standard decision maps display biobjective Pareto fronts as a function of a third objective [33]. Recently, animated interactive DMs have been developed in order to display additional objectives. Figure 21 displays four objectives in the form of a DM; the 2D Pareto front of $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ vs $\sqrt{S_c}/E_{\mathrm{acc}}$ is shown as a function of binned Z and the surface magnetic field. A candidate can again be selected by omitting other candidates. Minimum-Z solutions with small peak surface fields can be omitted, as can maximum-Z solutions. In order to best compromise between objectives, the Pareto fronts that have Z in the range

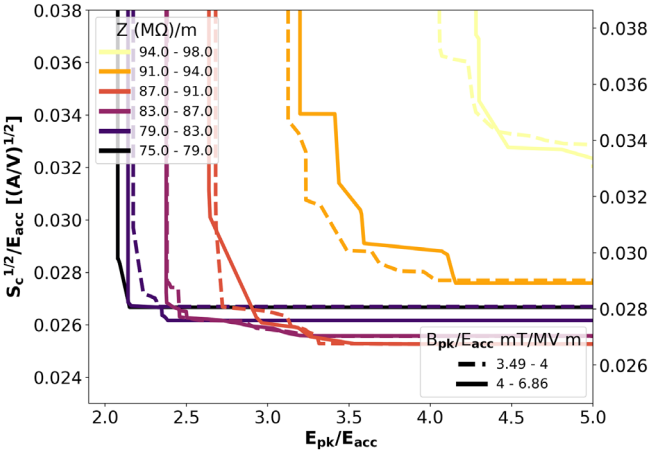

FIG. 21. DM for C-band SW cell. The Pareto fronts of $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ vs $S_c^{1/2}/E_{\mathrm{acc}}$ for each bin in Z are displayed as a function of high and low values of $B_{\mathrm{pk}}/E_{\mathrm{acc}}$.

86–91 MΩ/m possess solutions with the best trade-off between all four objectives. Some of the advantages of DMs are that a large number of initial points do not reduce the efficacy of the plot, as the data are binned during the process. Increasing the number of points will actually improve the accuracy of the individual points, meaning it is an attractive way of dealing with very large datasets. Another advantage of a DM is that the raw objective values can be inferred more easily than the PCP or RC. A disadvantage is that the number of objectives is limited to four, unless an interactive method is employed, where five objectives can be used with scroll bars. As a result, for a large dimensional set, PCPs and RCs provide a more appropriate choice of visualization technique.

D. Heat maps

Another approach is to use heat maps (HMs), which can be useful for visualizing the entire design and objective space while also allowing individual solutions to be explored. Heat maps display the interaction between two matrix elements as a color of varying intensity. The entire design and objective space can be visualized in this way, giving a complementary method to PCPs for visualizing entire datasets, although they can suffer from the same problem of overcrowding. For the single-cell study, the matrix rows show all feasible solutions, and the matrix columns display the inputs and/or objectives. The corresponding map provides a 2D visualization of how each objective interacts with a solution as well as how the objectives interact for any given solution. If the data are plotted in no particular order, heat maps are of little use, as it is hard to extract meaningful information due to the random order of the solutions. If the data are clustered or ranked by similarity, then it becomes easier to see parameter and objective interactions, as explained in Ref. [36].

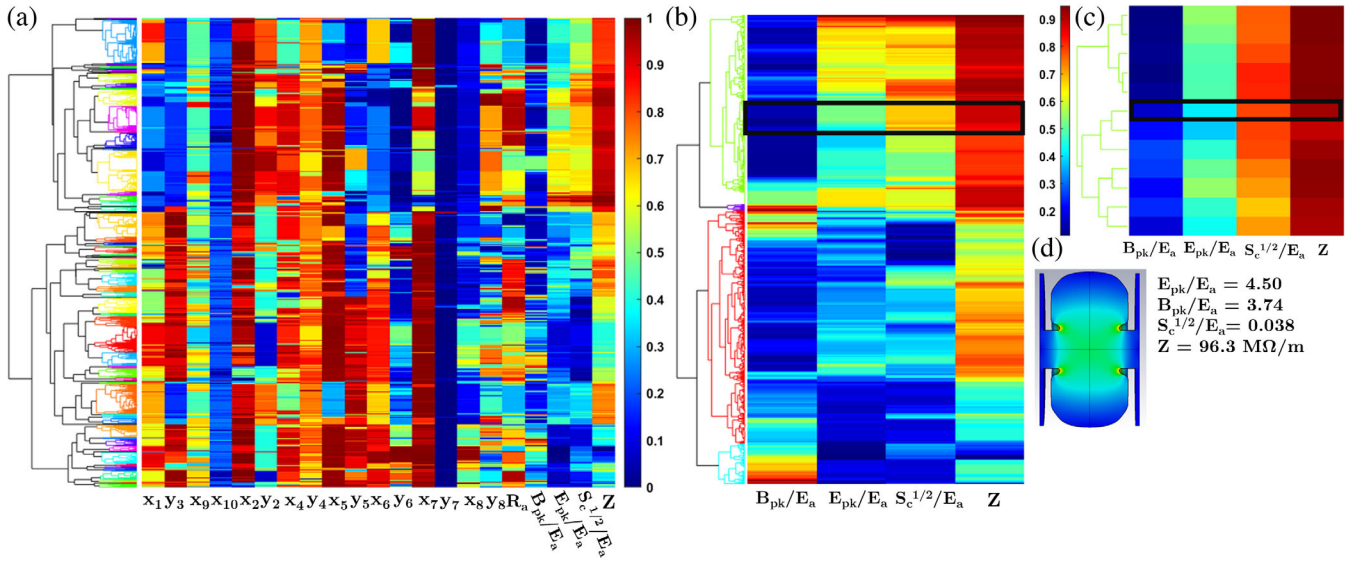

FIG. 22. (a) Heat map showing both design and objective space. (b) Heat map showing only objective space. (c) Enlarged section with low values of E_{pk}/E_{acc} and high Z. (d) Example of a selected cavity using the heat map method.

Figure 22 shows the heat map for the Pareto optimal points for the SW cavity with the frequency within the bounds 5.712 ± 5 MHz. On the left, the entire objective space is shown, normalized and clustered using hierarchical clustering but keeping all of the solutions. By knowing the maximum and minimum values of the objectives, regions of interest can be identified and the cluster expanded to zoom in. This can be repeated until a small subset of solutions is obtained, and then the individual geometries can be explored. One advantage of heat maps is that, by clustering the entire input and output space, the clusters can also provide insight into cavity shapes that perform similarly. An example of a solution selection using this method is shown in the figure. An area where E_{pk}/E_{acc} is ≈4–5 and high values of Z occur is expanded, and then a solution is selected. This method is not ideal for finding an optimal solution, as individuals with better performance can easily be missed, but a general feel for the design space can be easily obtained, and the interactive element makes exploring the space quite simple. Also, if large values of an objective exist in a set with smaller values, then this may skew the heat map, causing the full range of colors to not be used, meaning that certain datasets may need to be filtered beforehand.

E. 3D visualization

Another option for visualizing the optimization results would be to use a 3D scatter plot or 3D bubble plot. These were found to be difficult to read, and it becomes increasingly more difficult to select individuals as the number of Pareto optimal points increases. In 3D, for a continuous solution space the Pareto optimal points can be an approximation of a smooth 3D Pareto surface, and, therefore, another method explored was to interpolate between the points or fit a polynomial surface to construct an approximation of the 3D surface. Although this can give a good indication of what the surface may look like, the plots were not useful for selecting individual candidates. One way to plot the 3D surface and retain useful information about individuals is by binning the data in one of the variables and then constructing 2D Pareto fronts for each bin. An example of the application of this method is shown in Fig. 23 for the SW cavity. By projecting the fronts onto the three axes and using color to represent a fourth objective

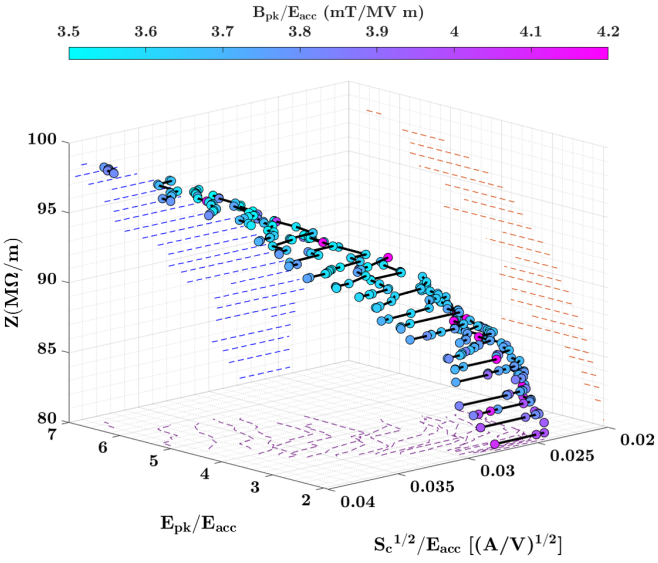

FIG. 23. Z vs E_{pk}/E_{acc} vs $\sqrt{S_c}/E_{acc}$ 3D scatter plot of Pareto optimal space for a C-band SW single cell. Individual 2D fronts are projected onto the planes of the plot for clarity.

(B_{pk}/E_{acc} in this case), one can clearly identify the 3D Pareto surface shape and potential regions of interest.

3D Pareto front approximations provide a simple method for visualizing four objectives and enable preservation of the entire 3D front shape, showing clearly how the objectives interact with each other. One can then hone in on regions to select individual solutions or repeat the plots with smaller bin sizes or a reduced objective space. The main drawbacks are that, for more than four objectives, multiple plots would be required in order to visualize all the Pareto fronts and that some information is lost when binning in one objective, as points that do not fall on the individual 2D Pareto fronts are not shown.

VII. SOLUTION SELECTION

In this section, the two most useful visualization techniques are explored and applied to select solutions for both the SW and TW cavities, each having different design goals.

A. Standing wave cavity selection

The DM in Fig. 24 shows the $\sqrt{S_c}/E_{acc}$ vs E_{pk}/E_{acc} Pareto fronts inside the 1 MΩ/m bins for the high-Z cases where $Z > 86$ MΩ/m. The points have been clustered using the method discussed in Sec. V, using 200 clusters as the individual points are small enough to see when plotted on the scatter plot, allowing for more clusters. It can be clearly seen that to get the lowest possible values of E_{pk}/E_{acc} and $\sqrt{S_c}/E_{acc}$ a higher value of B_{pk}/E_{acc} must be tolerated. An application of this method for solution selection is now explored assuming that the design is limited by the peak electric field ($E_{pk}/E_{acc} < 4.5$) and a high shunt impedance is required for maximum efficiency. First, the region of 4.4–4.6 is explored to find potential solutions that are close to $E_{pk}/E_{acc} = 4.5$, as solutions close to the limit will provide the highest values of Z, which can be seen in Fig. 24. Looking at the solutions in this region, three candidates are selected: SW1 with the lowest value of $\sqrt{S_c}/E_{acc}$ and a smaller value of B_{pk}/E_{acc} if this is required in the design, SW2 with the maximum value of Z, and SW3, a solution with a trade-off somewhere in the middle. The selected solutions are shown in the plot and also given in Table III.

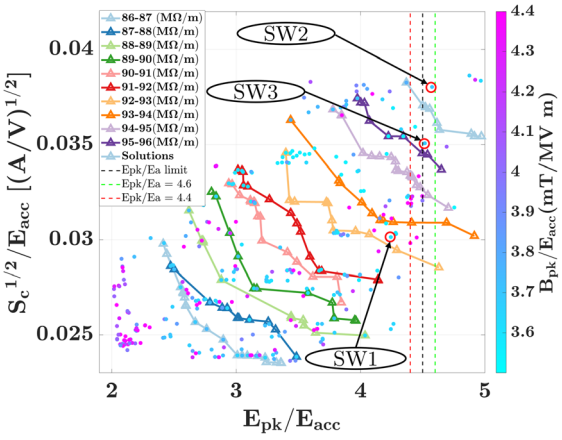

FIG. 24. 2D clustered scattered points with DM overlaid for the SW cavity to enable solution selection. Dashed lines show regions of interest if a limit of $E_{pk}/E_{acc} < 4.5$ is imposed. Circles show selected solutions based on other objective values.

TABLE III. SW and TW example selections.

Quantity	SW1	SW2	SW3	TW1	TW2
E_{pk}/E_{acc}	4.3	4.57	4.51	2.99	2.08
B_{pk}/E_{acc} (mT/MV m)	3.59	3.70	3.53	3.16	3.68
$\sqrt{S_c}/E_{acc}$ ($\sqrt{\frac{A}{V}}$)	0.0295	0.038	0.035	0.0182	0.0253
Z ($\frac{M\Omega}{m}$)	92.05	96.5	95.2	124.3	118.4
v_g (% c)	...	...	...	0.978	0.99

B. Traveling wave cavity selection

The TW cavity selection was performed using the clustering and PCP visualization technique. First, all solutions with $v_g = 0.97\%$–1.03% c were clustered into 24 clusters as shown in Fig. 25. Two design choices were considered, one where the design was E_{pk}/E_{acc} limited and the other limited by $\sqrt{S_c}/E_{acc}$. This led to the choice of cluster 1 for the first case and cluster 2 for the second. Although there was a cluster with a smaller value of E_{pk}/E_{acc}, it had a significantly lower value for the shunt impedance and was, therefore, not considered. The two selected clusters were opened to inspect the individual solutions and perform the final selection; the opened

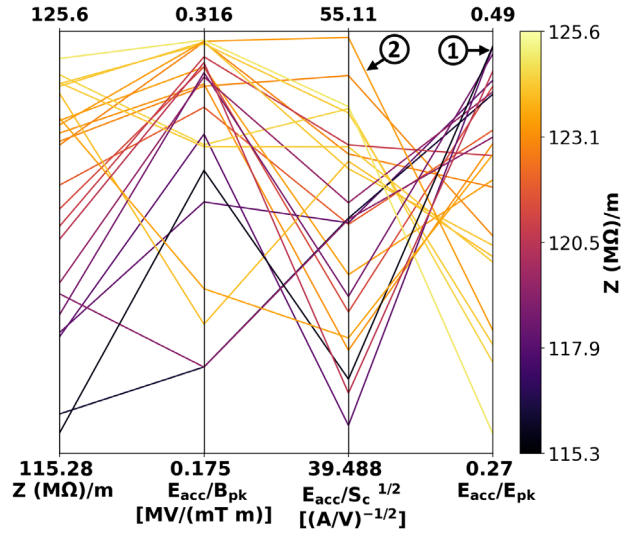

FIG. 25. TW cavity objective space with four objectives showing 24 clusters. Clusters 1 and 2 are the selected clusters for opening and picking a solution.

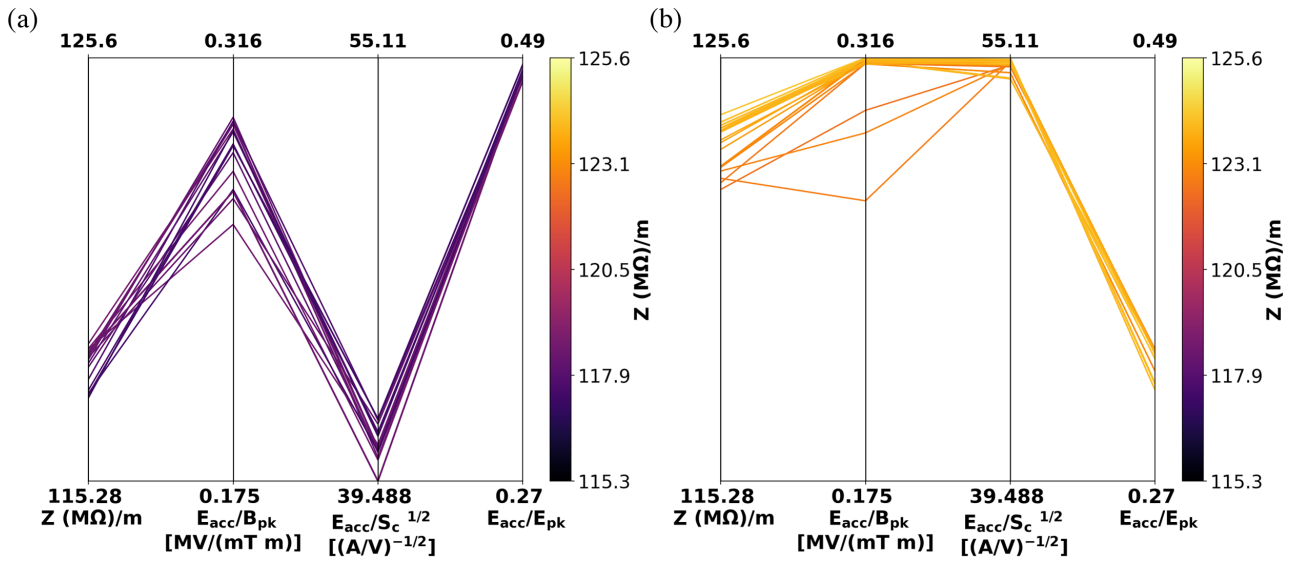

FIG. 26. Opened clusters from Fig. 25. (a) Cluster 1. (b) Cluster 2.

clusters are shown in Fig. 26. For cluster 1, there is little variance in the value of $E_{\mathrm{pk}}/E_{\mathrm{acc}}$, and, therefore, the choice was based on the best value for the other three objectives; the final chosen solution is shown in Fig. 27. For cluster 2, there is more variance in the value of $B_{\mathrm{pk}}/E_{\mathrm{acc}}$, with the outliers having a lower Z value and, therefore, not considered. The solutions with the smallest values of $\sqrt{S_c}/E_{\mathrm{acc}}$ also have a reduced Z, and, therefore, a solution with a larger Z was chosen, as the variance in $\sqrt{S_c}/E_{\mathrm{acc}}$ was small enough to justify taking the solution with larger Z. The final solutions are shown in Table III.

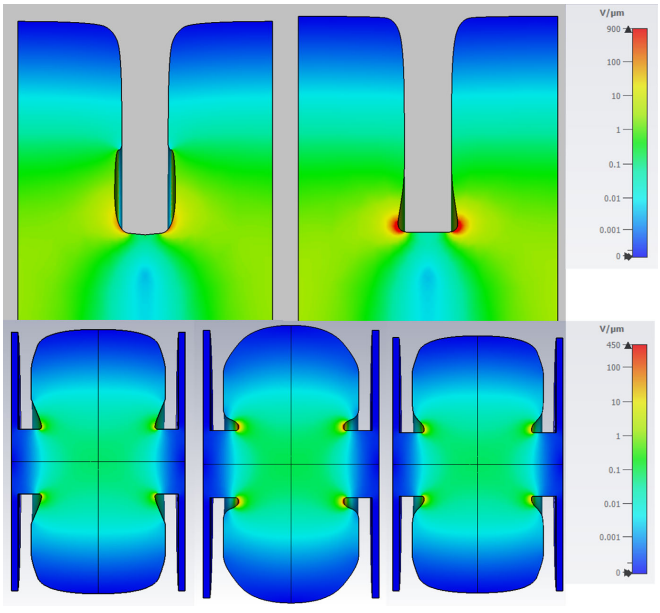

FIG. 27. Two potential final designs for an X-band TW single cell, chosen using clustering and the PCP visualization method [top from left to right: S1 (low $E_{\mathrm{pk}}/E_{\mathrm{acc}}$) and S2 (low $\sqrt{S_c}/E_{\mathrm{acc}}$)] and a C-band SW single cell, chosen using the DM method [bottom from left to right: S1 (low $\sqrt{S_c}/E_{\mathrm{acc}}$), S2 (high Z), and S3 (trade-off between Z and $\sqrt{S_c}/E_{\mathrm{acc}}$)]. Field shown is the absolute value of the electric field for 1 J stored energy.

VIII. DISCUSSION AND CONCLUSIONS

In this work, MOGAs have been applied to both X-band TW and C-band SW single-cell designs. The single-cell design was constructed using the NURBS method, allowing for more complexity in the shape, which has been shown to improve the performance in the low $E_{\mathrm{pk}}/E_{\mathrm{acc}}$ region by up to 10%. The benefits of MO optimization methods as opposed to single objective have been explored, showing that MO methods provide more insight into the design space and trade-offs with a comparable amount of computational effort. The best ways of dealing with the optimization constraints of the frequency and group velocity for rf cavity design have also been explored. The algorithms produced a set of individually unique solutions that are termed nondominated. If no subjective weight is applied to an objective, all solutions are deemed as optimal as each other in the Pareto optimal objective space. In order to select the final single-cell design, the trade-off between Pareto optimal solutions has been explored through the use of multiobjective visualization techniques. Various techniques that preserve initial objective values have been identified, and the application of these techniques to aid in the solution selection process has been demonstrated.

As the optimizations produce a large number of Pareto optimal points, clustering methods have been used to group very similar solutions to allow visualization methods to be used more effectively. This technique can be applied to any optimization results without the loss of important solutions,

as individuals with the best characteristics within the clusters are preserved. It was found that if the number of objectives is ≥ 4, then clustered parallel coordinate plots are the superior option for selecting individual solutions, as they are scalable to a large number of objectives and allow the designer to view the entire objective space easily and identify trade-offs. If the number of objectives is ≤ 4, then decision maps overlaid onto 2D scatter plots were found to be the best for selecting solutions, as all objectives and more solutions can be displayed on one plot. These methods can be used to aid designers, as the limits and requirements can easily be modified to select new designs. The results can also be incorporated into a systemwide design including the rf power parameters such as the pulse length, repetition rate, and power levels, for example.

ACKNOWLEDGMENTS

This work was supported by STFC Research Grant No. ST/P002056/1.

[1] O. L. De Weck, Multiobjective optimization: History and promise, in *Invited Keynote Paper, GL2-2, The Third China-Japan-Korea Joint Symposium on Optimization of Structural and Mechanical Systems, Kanazawa, Japan* (2004), Vol. 2, p. 34, http://strategic.mit.edu/docs/3_46_CJK-OSM3-Keynote.pdf.

[2] R. T. Marler and J. S. Arora, Struct. Multidiscip. Optim. **41**, 853 (2010).

[3] H. Tamaki, H. Kita, and S. Kobayashi, in *Proceedings of IEEE International Conference on Evolutionary Computation* (IEEE, New York, 1996), pp. 517–522.

[4] A. Van der Velden, P. Koch, Simulia, Isight design optimization methodologies, *ASM Handbook Volume 22B Application of Metal Processing Simulations* (2010), https://www.asminternational.org/search/-/journal_content/56/10192/05281G/PUBLICATION.

[5] E. Knapp, B. Knapp, and J. Potter, Rev. Sci. Instrum. **39**, 979 (1968).

[6] I. V. Bazarov and C. K. Sinclair, Phys. Rev. Accel. Beams **8**, 034202 (2005).

[7] T. Luo, H. Feng, D. Filippetto, M. Johnson, A. Lambert, D. Li, C. Mitchell, F. Sannibale, J. Staples, S. Virostek *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **940**, 12 (2019).

[8] B. Terzić, A. S. Hofler, C. J. Reeves, S. A. Khan, G. A. Krafft, J. Benesch, A. Freyberger, and D. Ranjan, Phys. Rev. Accel. Beams **17**, 101003 (2014).

[9] R. Bartolini, M. Apollonio, and I. Martin, Phys. Rev. Accel. Beams **15**, 030701 (2012).

[10] P. Putek, S. G. Zadeh, M. Wenskat, and U. van Rienen, Phys. Rev. Accel. Beams **25**, 012002 (2022).

[11] H. Feng, S. De Santis, K. Baptiste, W. Huang, C. Tang, and D. Li, Rev. Sci. Instrum. **91**, 014712 (2020).

[12] M. Kranjčević, A. Adelmann, P. Arbenz, A. Citterio, and L. Stingelin, Nucl. Instrum. Methods Phys. Res., Sect. A **920**, 106 (2019).

[13] M. Kranjčević, S. G. Zadeh, A. Adelmann, P. Arbenz, and U. Van Rienen, Phys. Rev. Accel. Beams **22**, 122001 (2019).

[14] M. Sawamura, R. Hajima, R. Nagai, and N. Nishimori, Design optimization of spoke cavity of energy-recovery linac for non-destructive assay research, in *Proceedings of SRF2011, Chicago, IL* (JACoW, 2011), MOPO036.

[15] Z. Tang, Y. Pei, and J. Pang, Nucl. Instrum. Methods Phys. Res., Sect. A **790**, 19 (2015).

[16] T. P. Wangler, *RF Linear Accelerators* (John Wiley & Sons, New York, 2008).

[17] A. Grudiev, S. Calatroni, and W. Wuensch, Phys. Rev. Accel. Beams **12**, 102001 (2009).

[18] D. P. Pritzkau and R. H. Siemann, Phys. Rev. Accel. Beams **5**, 112002 (2002).

[19] M. Nasr and S. Tantawi, New geometrical-optimization approach using splines for enhanced accelerator cavities' performance, in *Proceedings of the 9th International Particle Accelerator Conference* (JACoW, 2018).

[20] Dassault Systemes, CST STUDIO SUITE, R2021x, https://www.3ds.com/products-services/simulia/products/cst-studio-suite/, computer code.

[21] Dassault Systemes, ISIGHT & THE SIMULIA EXECUTION ENGINE, 2021x, https://www.3ds.com/products-services/simulia/products/isight-simulia-execution-engine/, computer code.

[22] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, IEEE Trans. Evol. Comput. **6**, 182 (2002).

[23] S. Tiwari, P. Koch, G. Fadel, and K. Deb, in *Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation* (Association for Computing Machinery, New York, NY, 2008), pp. 729–736, https://dl.acm.org/doi/abs/10.1145/1389095.1389235.

[24] K. N. Sjobak, A. Grudiev, and E. Adli, New Criterion for Shape Optimization of Normal-Conducting Accelerator Cells for High-Gradient Applications, in *Proceedings of the 27th International Linear Accelerator Conference* (JACoW, 2014).

[25] B. Filipič and T. Tušar, in *Proceedings of the Genetic and Evolutionary Computation Conference Companion* (Association for Computing Machinery, New York, NY, 2019), pp. 951–974.

[26] T. Tušar and B. Filipič, IEEE Trans. Evol. Comput. **19**, 225 (2015).

[27] A. Ibrahim, S. Rahnamayan, M. V. Martin, and K. Deb, in *Proceedings of the 2016 IEEE Congress on Evolutionary Computation (CEC)* (IEEE, New York, 2016), pp. 736–745.

[28] S. Khalid, T. Khalil, and S. Nasreen, A survey of feature selection and feature extraction techniques in machine learning, in *Proceedings of the 2014 Science and Information Conference* (IEEE, 2014), pp. 372–378.

[29] F. Nielsen, *Introduction to HPC with MPI for Data Science* (Springer, New York, 2016).

[30] K. Sasirekha and P. Baby, Int. J. Sci. Res. Publ. **83**, 83 (2013), https://www.ijsrp.org/research-paper-0313.php?rp=P15831.

[31] R. M. Edsall, Computational Statistics and Data Analysis **43**, 605 (2003).

[32] W.-Y. Liu, B.-W. Wang, J.-X. Yu, F. Li, S.-X. Wang, and W.-X. Hong, in *Proceedings of the 2008 International*

Conference on Machine Learning and Cybernetics, Kunming, China (IEEE, 2008), Vol. 2, pp. 857–862, https://ieeexplore.ieee.org/document/4620524.

[33] A. V. Lotov, V. A. Bushenkov, and G. K. Kamenev, *Interactive Decision Maps: Approximation and Visualization of Pareto Frontier* (Springer Science, New York, 2013), Vol. 89.

[34] A. Pryke, S. Mostaghim, and A. Nazemi, in *Proceedings of the International Conference on Evolutionary Multi-criterion Optimization* (Springer, New York, 2007), pp. 361–375.

[35] Barnaby, https://www.mathworks.com/matlabcentral/fileexchange/43611-parallel-coordinate-plots-gui-toolbox.

[36] D. J. Walker, R. Everson, and J. E. Fieldsend, IEEE Trans. Evol. Comput. **17**, 165 (2012).