Xue, L., Liu, B., Yu, Y., Cheng, Q. S., Imran, M. and Qiao, T. (2022) An unsupervised microwave filter design optimization method based on a hybrid surrogate model-assisted evolutionary algorithm. IEEE Transactions on Microwave Theory and Techniques, (doi: 10.1109/TMTT.2022.3219072).

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

https://eprints.gla.ac.uk/281507/

Deposited on: 30 September 2022

Enlighten – Research publications by members of the University of Glasgow
https://eprints.gla.ac.uk

An Unsupervised Microwave Filter Design Optimization Method Based on A Hybrid Surrogate Model-Assisted Evolutionary Algorithm

Liyuan Xue, *Student Member, IEEE*, Bo Liu, *Senior Member, IEEE*, Yang Yu, *Member, IEEE*, Qingsha S. Cheng, *Senior Member, IEEE*, Muhammad Imran, *Senior Member, IEEE*, Tianrui Qiao

***Abstract*—In resonator-coupled bandpass filter 3D design, it is a routine that the filter optimization methods are guided/supervised by designers' experience to carry out an iterative design optimization process. To realize automated or unsupervised filter 3D design optimization, a new method, called hybrid surrogate model-assisted evolutionary algorithm for filter optimization (H-SMEAFO), is proposed. H-SMEAFO aims to automatically obtain a highly optimal filter 3D design without designers' interaction (i.e., unsupervised) and is also not restricted to certain kinds of filter structures. In H-SMEAFO, the key innovations include a hybrid response feature-based objective function and a hybrid surrogate model-assisted global optimization algorithm; both are designed bespoke for filter design landscape characteristics. The performance of H-SMEAFO is demonstrated by an 8th order dual-band waveguide filter with 4 transmission zeros and a 6th order waveguide filter with 2 transmission zeros, for which, unsupervised design optimization does not appear to be possible using existing methods. Numerical results show the effectiveness and advantages of H-SMEAFO.**

***Index Terms*—Design optimization, Differential evolution, Filter response features, Microwave filters, Objective function, Surrogate modeling**

I. Introduction

Microwave filter 3D electromagnetic (EM) simulation-based design optimization is attracting much attention [1]. Given a filter structure and an initial design, filter design optimization aims to obtain an optimal 3D design (i.e., geometric parameter values) that satisfies a set of predefined performance specifications. Filter design landscapes are often highly multimodal (i.e., have numerous local optima) [2] and are challenging for most optimization algorithms. Therefore, filter 3D design optimization research focuses on (1) methods to obtain a high-quality initial design (e.g., the coupling matrix (CM) method [3], [4]), and (2) methods to perform optimization from the initial design. The paper focuses on the latter.

Thanks to the rich filter analysis and design methodology research [5], effective experience-guided design procedures are proposed. In these procedures, off-the-shelf local optimizers are iteratively employed. For example, in the step tuning procedure, the filter is optimized by successively adding one resonator at a time and carrying out a corresponding optimization run in each iteration [6], [7]. In advanced port tuning procedure [8], [9], auxiliary ports and tunable elements are added for providing circuit-based surrogates in each step of optimization. In these methods, because the designer's tuning procedure and experience play a major role, we call them "supervised" design optimization, which is the current routine.

In recent years, novel successful intelligent filter design optimization methods are proposed. For example, space mapping methods [1], [10], [11] use low-fidelity models (e.g., equivalent circuits) to reduce the necessary number of computationally expensive high-fidelity EM simulations. Cognition-driven optimization methods [12], [13] integrate designers' intuition of firstly optimizing the frequency features and then optimizing the ripple heights. The homotopy method [14] constructs a series of intermediate optimization problems from the initial design to the optimal design and has strength when the initial design is not of high quality. Machine learning techniques are employed in the above methods for speed improvement. Compared to off-the-shelf local optimizers, these methods obtain optimal filter designs with much higher quality more efficiently. Hence, optimization is playing an equally important role compared to designers' guidance. However, local optimizers are employed, and in many cases, designers' interaction is still necessary to jump out of local optima. We therefore call them "semi-supervised" design optimization.

This paper aims to propose "unsupervised" filter design optimization. "Unsupervised" refers to two characteristics: (1) It can satisfy stringent design specifications by pressing one button, and designers' interaction is not needed, and (2) It is general enough and not restricted to certain kinds of filter structures. The benefits include: (1) Much of designers' time (i.e., cost) is saved because they are transformed into computing time as no interaction is needed. This is especially useful considering nowadays's largely increased availability and largely decreased financial cost of computing resources. (2) Because designers' experience does not play a role, unsupervised design optimization applies to average engineers with less design experience.

Although unsupervised design optimization has already been realized for antennas, it is still very challenging for filters due to their landscape characteristics [2]. To achieve

Manuscript received August 23th, 2022. L. Xue, B. Liu, M. Imran are with James Watt School of Engineering, University of Glasgow, Glasgow, Scotland. G12 8QQ. (e-mail: l.xue.1@research.gla.ac.uk, Bo.Liu@glasgow.ac.uk, Muhammad.Imran@glasgow.ac.uk)

Y. Yu is with Key Laboratory of Microwave Remote Sensing, National Space Science Center, Chinese Academy of Sciences, Beijing, 100190, China. (emails: Issacyu@live.cn).

Q. S. Cheng, and T. Qiao are with Dept. of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, 518055, P.R.China. (emails: chengqs@sustech.edu.cn).

This work was supported by Glasgow Industrial Doctoral Project in collaboration with Nokia. Corresponding authors: Bo Liu, Qingsha S. Cheng

this goal, the following elements are essential: (1) Global optimization algorithms bespoke for filters. Unlike the above-mentioned successful local optimizers for filters, a global optimizer is essential considering the highly multimodal filter design landscape in one-off optimization. To the best of our knowledge, SMEAFO [2] is the only global optimizer designed for filters. However, it targets not complex problems and the convergence speed needs improvement. (2) Methods seeking an initial design that is as close to the final optimal design as possible. Several successful methods are available [15] and the method in [16] is used in this work. (3) Proper objective functions simplifying the filter design landscape. Except objective functions based on the magnitude of S-parameters (e.g., $\max(|S_{11}|)$), which is straightforward and very widely used, several promising methods are proposed [13], [14], [17], but this area is far from mature. Unsupervised filter design optimization is still in its infancy.

One thing that must be mentioned is the optimization time. For a single filter, unsupervised design is often slower than supervised design due to the use of global optimization (i.e., no-free-lunch). Hence, it is important to improve the optimization speed for unsupervised design, which is one of the goals of this research. However, due to no interaction being needed, many filters can be optimized in parallel, which alleviates this drawback. Hence, we assume a few days' time consumption for unsupervised design is reasonable compared to the testing and tuning time, which could be months.

In this paper, a new method, called hybrid surrogate model assisted evolutionary algorithm for filter optimization (H-SMEAFO), is proposed. Two key innovations include: (1) A new hybrid objective function that adaptively combines various features of filter response is proposed. (2) A new surrogate model-assisted hybrid optimization algorithm is proposed, which iteratively carries out Gaussian process surrogate model-assisted differential evolution and Nelder-Mead simplex algorithms self-adaptively. Two typical filters, including an 8th order symmetric dual-band waveguide filter with 4 transmission zeros and a 6th order waveguide filter with 2 transmission zeros, are used to test H-SMEAFO. Experiments and comparisons verify the effectiveness of H-SMEAFO to realize unsupervised filter design optimization reasonably efficiently for the first time (to the best of our knowledge) as well as its advantages compared to state-of-the-art methods.

The remainder of the paper is organized as follows. Section II presents the background knowledge. Section III elaborates on the H-SMEAFO method, including the algorithm structure, the new hybrid objective function, and the new hybrid optimization algorithm. Section IV presents the performance and advantages of H-SMEAFO using two real-world filter examples, respectively. The concluding remarks are provided in Section V.

II. Background Knowledge

A. Filter Design Process and Unsupervised Design Optimization

A typical filter design process is shown in Fig. 1. In this paper, resonator-coupled bandpass filters are targeted. They

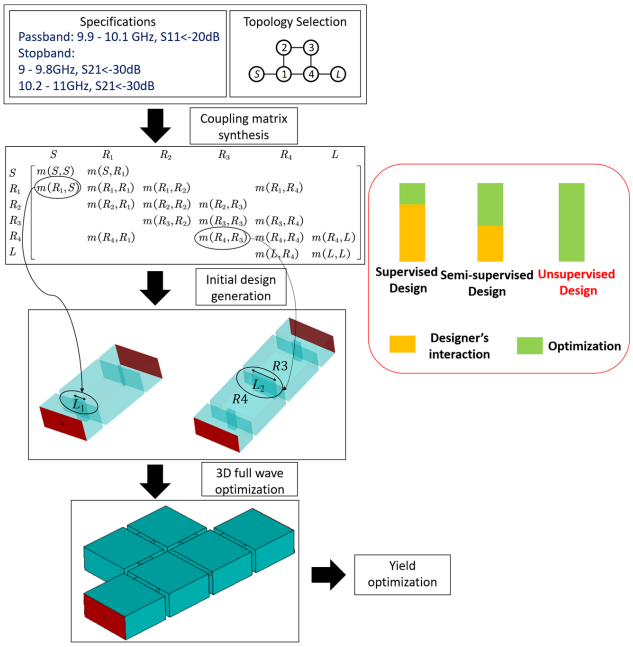

Figure 1. A typical filter design process

are most widely used in communication systems and are the focus of most filter design optimization research works. From the design specifications, the designer first needs to decide on a viable fabrication technique and filter structure. Then, the design and prototyping start. Note that unlike many other microwave devices (e.g., antennas, couplers), the blessing of filter design is that there are systematic ways to obtain an initial design with reasonable quality, and the procedure is called synthesis [18] and physical dimensioning [19].

Synthesis aims to generate element values corresponding to particular filtering performance in the normalized frequency domain, i.e., a CM [18] or lumped-element equivalent circuit [20]. The obtained element values often have a one-to-one correspondence to the geometrical dimensions of the physical elements, e.g., resonators or coupling structures. Obtaining the initial 3D design from the element values is called physical dimensioning. There are various physical dimensioning methods with different qualities and simplicity.

Although the initial 3D design provides much information and is often not very far from the optimal 3D design satisfying the specifications, their response is often poor and a consecutive 3D optimization using EM simulations is needed. This is because the initial design is generated based on the CM, which is the ideal condition and with much approximation. Also, due to the highly multimodal filter design landscape characteristics, the optimization problem is challenging for most optimization methods. Hence, the supervised, semi-supervised and unsupervised filter design optimization mentioned in Section I are for this step. As described in Section I, in contrast with supervised and semi-supervised design, unsupervised design should not rely on designers' experience-based decisions in the design optimization process and can obtain the optimal design satisfying the specifications on its own from the initial

design. Then, using the obtained optimal design from this step, the yield of the filter is optimized considering fabrication error [21], which is out of the scope of this paper.

Note that supervised, semi-supervised, and unsupervised filter design optimization complement each other and target different kinds of users. Supervised design procedures mainly target experienced designers, semi-supervised optimization methods target both experienced designers and average engineers, and unsupervised design optimization methods mainly target average engineers, although experienced designers may also use them for automated design.

As said in Section I, a global optimizer is essential for unsupervised filter design optimization and the SMEAFO algorithm [2] is a published global optimizer for filters. In this paper, the proposed H-SMEAFO inherits the machine learning method (i.e., Gaussian process (GP)) and the global search technique (i.e., differential evolution (DE)) from SMEAFO, which are briefly introduced as follows.

B. Gaussian Process

GP is a widely used machine learning method in engineering optimization, whose strengths include its strong learning ability (i.e., high surrogate model quality) and the ability to provide a statistically grounded prediction uncertainty for each candidate design. Given a set of observations $\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\}$ with the corresponding function values $\boldsymbol{y} = (y_1, y_2, ..., y_n)^T$, GP model treats the values of $\boldsymbol{y}$ as samples of a multivariate Gaussian distribution. Thus, the likelihood function can be expressed in terms of samples $\boldsymbol{y}$ as:

$$L_h = \frac{1}{(2\pi\sigma^2)^{n/2}\,|\boldsymbol{R}|^{1/2}} \exp\left[-\frac{(\boldsymbol{y}-\mathbf{1}\mu)^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{y}-\mathbf{1}\mu)}{2\sigma^2}\right] \tag{1}$$

where $\mathbf{1}$ is a $n \times 1$ vector of ones, $\boldsymbol{R}$ is the covariance matrix defined by the correlation function

$$R_{i,j} = \mathrm{Corr}\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = \exp\left(-\sum_{l=1}^{d} \theta_l \left|x_i^l - x_j^l\right|^{p_l}\right), \quad \theta_l > 0, 1 \le p_l \le 2 \tag{2}$$

where d is the dimension of $\boldsymbol{x}$, $\boldsymbol{\theta}$ and $\boldsymbol{p}$ are hyperparameters describing how fast the correlation decreases on the l-th variable and the corresponding function smoothness, respectively. Assuming the hyperparameter values are known, the μ and σ in (1) can be obtained in a closed form, where

$$\begin{aligned} \hat{\mu} &= \mathbf{1}^T\boldsymbol{R}^{-1}\boldsymbol{y}(\mathbf{1}^T\boldsymbol{R}^{-1}\mathbf{1})^{-1} \\ \hat{\sigma} &= (\boldsymbol{y}-\mathbf{1}\hat{\mu})^T\boldsymbol{R}^{-1}(\boldsymbol{y}-\mathbf{1}\hat{\mu})n^{-1} \end{aligned} \tag{3}$$

Substituting (3) into (1), the likelihood function can be maximized numerically, obtaining the optimal hyperparameter values. Given a new design $\boldsymbol{x}^*$, the value of $\hat{y}(\boldsymbol{x}^*)$ and the uncertainty $\hat{s}(\boldsymbol{x}^*)$ can be obtained by best linear unbiased estimation and mean square error:

$$\begin{aligned} \hat{y}(\boldsymbol{x}^*) &= \hat{\mu} + \boldsymbol{r}^T\boldsymbol{R}(\boldsymbol{y}-\mathbf{1}\hat{\mu}) \\ \hat{s}^2(\boldsymbol{x}^*) &= \hat{\sigma}^2\left[1 - \boldsymbol{r}^T\boldsymbol{R}^{-1}\boldsymbol{r} + \frac{\left(1-\mathbf{1}^T\boldsymbol{R}^{-1}\boldsymbol{r}\right)^2}{(\mathbf{1}^T\boldsymbol{R}^{-1}\mathbf{1})}\right] \end{aligned} \tag{4}$$

where $\boldsymbol{r} = [\mathrm{Corr}(\boldsymbol{x}^*, \boldsymbol{x}_1), \mathrm{Corr}(\boldsymbol{x}^*, \boldsymbol{x}_2), ..., \mathrm{Corr}(\boldsymbol{x}^*, \boldsymbol{x}_n)]^T$ describing the correlation between $\boldsymbol{x}^*$ and all sample designs.

The lower confidence bound prescreening [22] is employed to balance exploration and exploitation when using GP in optimization, which is defined by:

$$y_{lcb}(\boldsymbol{x}^*) = \hat{y}(\boldsymbol{x}^*) - \omega\hat{s}(\boldsymbol{x}) \tag{5}$$

where $\omega \in [0, 3]$ is a constant, which is set to 2 in our algorithm. With this prescreening, the quality of x is considered from not only the predicted value but also the prediction uncertainty when judging the fitness of a candidate design, which enhances the ability to jump out of local optima. The above GP model is implemented by the ooDACE toolbox [23].

C. Differential Evolution

DE algorithm is a popular method in engineering optimization. In H-SMEAFO, DE is selected as the global search engine and its mutation and crossover operators are used. Considering a population denoted by P containing n candidate designs, in which each design is $\boldsymbol{x}_i \in \mathbb{R}^d, i = 1, ..., n$, the mutation strategy (DE/current-to-best/1) is:

$$\boldsymbol{v}_i = \boldsymbol{x}_i + F \cdot (\boldsymbol{x}_{best} - \boldsymbol{x}_i) + F \cdot (\boldsymbol{x}_{r_1} - \boldsymbol{x}_{r_2}) \tag{6}$$

where $\boldsymbol{v}_i$ is the i-th solution after mutation, $\boldsymbol{x}_{best}$ is the best candidate in the current population P, $\boldsymbol{x}_{r_1}$ and $\boldsymbol{x}_{r_2}$ are two randomly selected exclusive candidate solutions, and $F \in [0, 1]$ is the scaling factor. The DE/current-to-best/1 strategy introduces an appropriate amount of population diversity to avoid being trapped in local optima for filter design landscape [2]. Then, the crossover operator is carried out with a given crossover rate, CR. The offspring $\boldsymbol{u}_i$ is generated by:

$$u_i^l = \begin{cases} v_i^l, & \text{if (rand(0,1)} \le CR) \mid j = \text{rand}(0, d) \\ x_i^l, & \text{otherwise} \end{cases} \tag{7}$$

where $l = 1, 2, ..., d$ indicates the l-th design variable.

D. Nelder-Mead Optimization

The Nelder-Mead (NM) simplex method is a widely used derivative-free local optimization algorithm that is suitable for non-smooth or even discontinuous landscapes. Hence, in H-SMEAFO, it is selected as the local optimizer to collaborate with DE.

NM starts from a set of $d+1$ points $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_d$ that do not lie in the same hyperplane (the so-called nondegenerate working simplex). Given an initial solution $\boldsymbol{x}_0$, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d$ can be generated by adding a small value $\boldsymbol{\varepsilon}_i$ to each component of $\boldsymbol{x}_0$, that is

$$\boldsymbol{x}_i = \boldsymbol{x}_0 + \boldsymbol{\varepsilon}_i\boldsymbol{\delta}_i, \quad i = 1, 2, ..., d \tag{8}$$

where $\boldsymbol{\delta}_i$ is a $d \times 1$ one-hot vector with only value one on i-th component. By ordering the initial set according to $y(\boldsymbol{x})$ in ascending order, denoted by $\boldsymbol{x}_0, \boldsymbol{x}_2, ...\boldsymbol{x}_d$, reflection, expansion, contraction, and shrinking are performed iteratively according to different conditions. Reflection, expansion, and

contraction are used to explore possible search directions for improvement adaptively.

For example, the reflection generates the reflected point $\boldsymbol{x}_r$ by:

$$\boldsymbol{x}_r = 2\sum_{i=0}^{d-1} \boldsymbol{x}_i/d - \boldsymbol{x}_d \tag{9}$$

Then, $y(\boldsymbol{x}_r)$ is calculated to determine whether accept $\boldsymbol{x}_r$ or carry out expansion or contraction. At least a new solution from expansion or contraction is calculated and accepted if appropriate. In the worst case, the current simplex is shrunk and calculated for the next iteration. More details of NM simplex are in [24].

III. The H-SMEAFO Method

A. The Structure of H-SMEAFO

As said in Section I, H-SMEAFO aims to realize unsupervised filter design optimization and improve the optimization speed under this condition. The reason why unsupervised design optimization is challenging is the filter's highly multimodal design landscape [2]. Therefore, the solution includes two directions: (1) simplify the problem, and (2) improve the optimizer.

To simplify the problem, a natural idea is to find a high-quality initial design using filter theory. Quite a few research works are focusing on this topic and there are several successful methods. Although the employed initial design generation method [16] is out of the scope of this paper, it also does not need interactive decisions from the designer and can co-work with H-SMEAFO. For example, CST Microwave Studio already has such a one-button initial design generation tool from the 2022 release, and [16] improves its solution quality. Another way to simplify the problem is to study the objective function because the design landscape characteristics are affected by the objective function used. Hence, a new objective function is proposed, so as to release pressure on the global optimization algorithm. With the simplified design landscape as the target, a new bespoke global optimization method hybridizing DE (assisted by GP surrogate model) and NM simplex algorithms adaptively is proposed, aiming to improve the global optimization speed and success rate.

The framework of H-SMEAFO is shown in Fig. 2, which works as follows.

Step 1: Sample n candidate designs and let them form the initial database. (All the designs visited by H-SMEAFO are considered candidate designs.)

Step 2: Judge whether the stopping criterion is met (e.g., reach the maximum allowed number of EM simulations or satisfy all specifications. See Section IV (A) and (B) for examples.). If yes, output the best design; Otherwise, go to Step 3.

Step 3: Decide the objective function to be used (Section III (B)).

Step 4: Select α best designs from the database based on ranking the objective function values to form a population P.

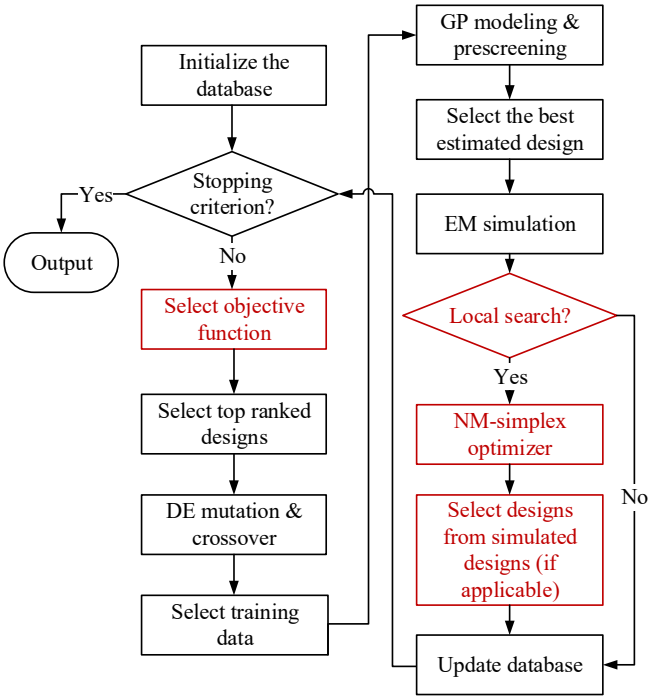

Figure 2. The flow diagram of H-SMEAFO

Step 5: Apply the DE/current-to-best/1 mutation (6) and the crossover operator (7) on P to generate α child solutions.

Step 6: For each child solution, select τ training data points and construct a local GP surrogate model.

Step 7: Prescreen the α child solutions generated in Step 5 using the lower confidence bound method (5), where GP models from Step 6 are used. Estimate the best child solution based on the lower confidence bound values. Carry out an EM simulation to it.

Step 8: Judge if NM local optimization needs to be triggered. If yes, carry out a complete NM simplex optimization run until the stopping criterion is met. The GP surrogate model is not used in this step. The condition to trigger the local optimization and its stopping criterion are explained in Section III (C).

Step 9: The candidate designs visited by NM local optimization in Step 8 are ranked and selected if it improves the current best design. The method is in Section III (C).

Step 10: Update the database by adding the estimated best design in Step 7 and the selected designs in Step 9 (if applicable) and its/their performance (EM simulation results). Go back to Step 2.

It can be seen that GP modeling, DE search, and some model management operators are inherited from SMEAFO. The benefits of them for filter design landscapes are described in [2]. The two main innovations include the hybrid objective function (Step 3) and the iterative adaptive optimization method hybridizing DE and NM simplex in an unsupervised filter optimization process (Step 8, 9 and the overall algorithm structure). They are shown as red blocks and are detailed in

the following subsections.

Besides the above innovations, a clear difference compared to SMEAFO is the initialization (Step 1). SMEAFO does not use an initial design and the advantage of employing filter design knowledge through the initial design is lost. In H-SMEAFO, with the initial design generated by the method in [16], Gaussian distributed random numbers with zero mean and variances of σ_f are added to it to form the initial samples. (Note that unlike the initial design, which is a single design obtained from the CM (Fig. 1), initial samples are a set of designs obtained from the initial design for the optimization algorithm to use.)

The value of σ_f is important to make use of the knowledge in the initial design. When a large random number is added to the initial design parameters, the pattern of the initial design can be overwritten, and thus cannot provide much assistance to the global optimizer. In contrast, a small random number can keep the pattern of the initial design but may make the initial samples clustered in a narrow region, preventing the search from jumping out of local optima.

In H-SMAEFO, the design parameters are divided into resonance-related and coupling-related, covering most of the filter design parameters. Resonance-related design parameters mainly control the center frequency and bandwidth and directly influence the frequency features. Coupling-related parameters, on the other hand, mainly control the ripple height feature. Empirical analysis shows that for our hybrid objective function in Section III (B), the former is more sensitive. Hence, σ_f has a different value for each category.

σ_f can be considered as a perturbation of initial geometric parameters. Given a resonator for which the physical length L is proportional to the guided wavelength λ_g at the resonant frequency f, the perturbation of L and the corresponding frequency has the following relation:

$$\Delta L \propto \frac{\Delta f}{f}\lambda_g \tag{10}$$

Hence, ΔL is varied with $FBW \times \lambda_g$, where FBW is the fractional bandwidth of the filter. In H-SMEAFO, for resonating-related design parameters, $\sigma_{fr} = 0.25 \times FBW \times \lambda_{gc}$ and for coupling-related parameters, $\sigma_c = FBW \times \lambda_{gc}$, where λ_{gc} is the guided wavelength of the central frequency. This empirical rule is also applicable to other global optimizers for filters.

B. The Hybrid Objective Function

Objective functions are representations of particular features extracted from filter response. An appropriate objective function for unsupervised design optimization should have the following requirements: (1) It is in line with the design specifications, i.e., a better objective function value refers to a candidate design that better meets the specifications, and (2) The design landscape resulted by the objective function is as smooth as possible, so as to relieve the pressure on the optimizer, i.e., avoiding being trapped in local optima and failing to satisfy the specifications.

Various objective functions have been proposed. Directly using the specifications, i.e., $\min(\max(|S_{11}|))$, is straightforward and very widely used. This objective function satisfies the former requirement but not the latter, and the resulted design landscape is often complex [2]. A recent promising objective function uses the difference between the extracted CM of a candidate design and ideal CM [25]. This objective function makes the design landscape very simple. However, the drawback is that the CM extraction often fails for candidate designs with poor responses [25], which often happens in unsupervised filter 3D optimization.

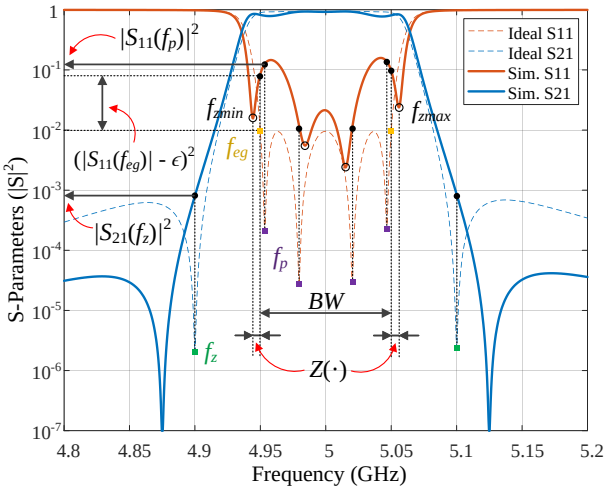

Figure 3. Key features considered in $F1$

Another recent promising objective function is the cognition-driven objective function [12], [13]. It uses poles and return loss as features, whose weights are adjusted at different optimization stages, making the design landscape much smoother. However, experiments in Section IV show that for unsupervised design optimization of filters with higher order and/or with transmission zeros, this objective function may not be exactly in line with the design specifications (i.e., the former requirement).

Therefore, a new hybrid objective function for minimization is proposed in H-SMEAFO targeting unsupervised design, which has two phases. At the beginning phase, many points around the initial design may have poor responses. Hence, the goal of the first phase is to obtain the general shape of the desired response with the highest speed. The objective function is based on the zero, pole, and edge (ZPE) objective function [26] with an added term Z to restrict the bandwidth, which is as follows (Fig. 3).

$$\begin{aligned} F_1(f_z, f_p, f_{eg}) = &\sum_i |S_{21}(f_{z_i})|^2 + \sum_j |S_{11}(f_{p_j})|^2 \\ &+ \sum_k (|S_{11}(f_{eg_k})| - \epsilon)^2 + Z(BW) \end{aligned} \tag{11}$$

where ϵ is the magnitude of S_{11} at the edge of the passband in the ideal condition. By considering the magnitude of S-parameters at the ideal frequencies of zeros, poles, and edges (denoted by f_z, f_p and f_{eg}, and indexed by i, j and k, respectively), the desired response shape can be formed efficiently. This ZPE objective function was originally proposed for optimization-based filter CM synthesis and shows high

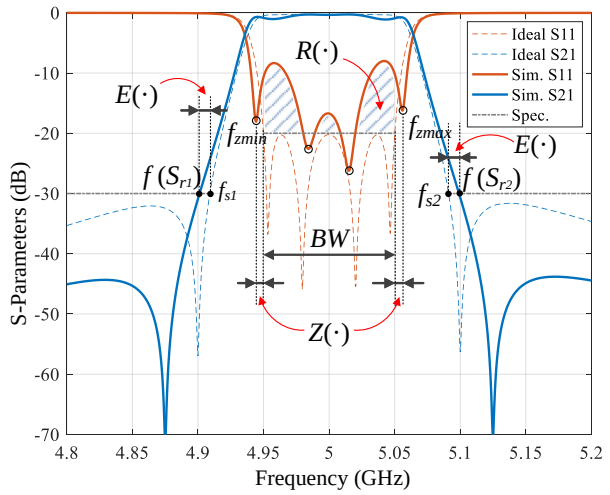

Figure 4. Key features considered in $F2$

efficiency [26]. Our pilot experiments show that it is much more efficient to obtain the general shape of the desired response than using $\min(\max(|S_{11}|))$ in unsupervised 3D optimization.

The added term $Z(BW)$ is defined as:

$$Z(BW) = (f_{z_{max}}(S_{11}) - f_{z_{min}}(S_{11}) - BW)/BW \quad (12)$$

where $f_{z_{max}}(S_{11})$ and $f_{z_{min}}(S_{11})$ are the maximum and minimum frequency of extracted reflection zeros, which are extracted by vector fitting [27], [28], [29]. Vector fitting can identify reflection zeros for candidate designs with poor responses, which is verified by existing research [13], [30]. The use of divided by BW is for normalization purpose. The purpose of the Z term is to penalize a typical kind of wrong response shape, for which, some of the reflection zeros are out of the passband, despite that the $\max(|S_{11}|)$ may even satisfy the specification in the passband.

However, this objective function cannot be used throughout the 3D optimization. The reason is the deviation between the ideal response and the real response obtained from physical design. In most cases, when the filter is fully optimized using F_1, the specifications are not met, i.e., F_1 is not always in line with the design specifications after the general response shape is formed.

Therefore, after phase 1, a new objective function for phase 2 follows up. The goal is to consider all the necessary features and satisfy the specifications. The phase 2 objective function F_2 include three terms, which are the inband ripple $R(\cdot)$, stopband edge frequency $E(\cdot)$, and bandwidth restriction $Z(\cdot)$. All of them are to be minimized.

The inband ripple term (Fig. 4) is defined as:

$$R(BW, S_r) = \frac{1}{S_r BW} \int_{BW} \max(|S_{11}(f)| - S_r, 0) \mathrm{d}f \quad (13)$$

where BW is the bandwidth, S_r is the specification (e.g., -20 dB for inband $|S_{11}|$). $\frac{1}{S_r BW}$ is the normalization term. The reason why using integration is that it calculates the average violation of the S_{11} specification, making the design landscape much smoother than using $\max(|S_{11}|)$ in the passband, which leads to a rugged landscape [2]. Our pilot experiments show that by using the same optimizer, not only the success rate is largely improved, but the convergence speed is also much faster compared to using $\max(|S_{11}|)$.

The stopband edge frequency term (Fig. 4) is defined as:

$$E(f_s, S_r) = \sum_i \max(f(S_{r_i}) - f_{s_i}, 0)/BW \quad (14)$$

where $f(S_r)$ is the first frequency whose $|S_{21}|$ meets the given stopband specification S_r (e.g., $|S_{21}|$ reaches -30 dB at the frequency $f(-30\text{dB})$), and f_s is the frequency specification (e.g., $|S_{21}|$ should be under -30 dB at f_s and below, both shown in Fig. 4), and i indicates the number of stopband specifications. This applies to each stopband. The use of divided by BW is for normalization purpose. Note that even for some design cases without stopband specifications, this term can be added using a reasonable estimation of f_s. Our pilot experiments show that using this term can improve the optimization speed in collaboration with the previous term.

The passband reflection zero term (Fig. 4), $Z(BW)$, is the same as that in F_1 except that the small difference (ΔB) between the distance of the maximum and minimum frequency of reflection zeros and the real bandwidth using -20 dB as the threshold, is also considered as a term in the equation. The value of ΔB can be estimated from the ideal response with the optimal CM. The final objective function for phase 2 is $F_2 = R(BW, S_r) + E(f_s, S_r) + Z(BW)$. No weight is needed because the terms are normalized.

In summary, the new objective function for unsupervised 3D design optimization works as follows. F_1 is firstly employed. By targeting the responses at critical frequencies, the general expected response shape is formed efficiently. When the average improvement of the F_1 value is less than 2% in recent consecutive 100 iterations, F_2 is employed. 2% is an empirical value and is not sensitive. F_2 considers exactly satisfying the specifications while at the same time, smoothening the design landscape to relieve the pressure of the global optimizer and promote the search speed. F_2 is used till the end of the unsupervised design optimization.

C. The Hybrid Surrogate Model-assisted Optimization Algorithm

As said in Section I, a filter design landscape characteristics-oriented global optimization algorithm is essential for unsupervised design even with the objective function in Section III (B). To the best of our knowledge, SMEAFO [2] is the only published global optimizer targeting filter design landscape characteristics, which shows good optimization ability for filters and diplexers [2], [30]. However, the drawback is the speed. It often needs a few thousand EM simulations to obtain the optimal design for not simple structures. As said in Section I, although the above time consumption is bearable considering unsupervised design, improving it is important. Hence, the goal of the hybrid surrogate model-assisted optimization algorithm in this subsection is to improve the speed of SMEAFO while maintaining its global optimization ability.

The flow diagram of the hybrid optimization algorithm is shown in Fig. 5. It can be seen that the GP model-assisted

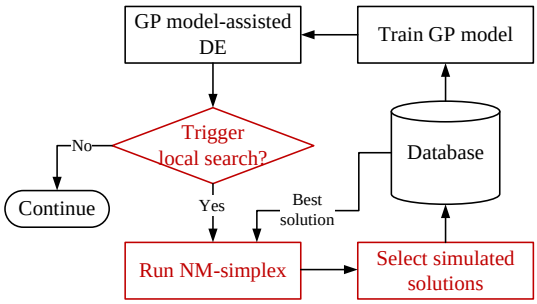

Figure 5. The working principle of the hybrid optimization method

DE part is inherited from SMEAFO. SMEAFO mainly relies on these operators to jump out of local optima for a filter design landscape. The difference compared to SMEAFO are as follows: (1) A Nelder-Mead local optimizer is introduced to the current best design obtained by GP-assisted DE, and (2) An iterative use of the global and local optimizers is introduced, as well as the data exchange method between them.

The reason why SMEAFO has a relatively slow speed is its exploration ability (i.e., a relatively large diversity of candidate designs that are visited). Note that this ability must be kept to jump out of local optima. However, due to no-free-lunch, the cost is the lower convergence speed. Without compromising the diversity of the population of candidate designs, a possible way to improve the convergence speed is to improve the current best design by a local search. For the current best design, local search algorithms may improve the objective function value much faster than GP-assisted DE although the exploration ability is lost. Not only the current best design is improved, but the useful candidate designs visited in local search can also complement the global search. Particularly, the updated current best design can benefit the whole population according to DE/current-to-best/1 mutation (6) when using GP-based DE again.

Hence, instead of using a Gaussian local search-based exploitation phase (only run for a single time in the near convergence space) as SMEAFO, the global and local optimizers are iteratively used. Once GP-based DE cannot find a better candidate design in 100 consecutive iterations, the local optimizer is triggered. Once the local optimizer reaches a local optimum and cannot improve, the population of GP-based DE is updated by introducing selected useful visited candidate designs from the local optimizer and restarted. The two optimizers complement each other using the above way to combine their strength.

NM simplex method is selected as the local optimizer. NM simplex is a derivative-free search method and is suitable for rugged highly multimodal landscape [24], while filter design landscape shares similar properties. NM simplex is carried out from the current best design from GP-based DE. In the local optimization process, no surrogate model is used. This is because local search requires a sufficiently accurate surrogate model, but building it using a limited number of EM simulations is often difficult for filter design landscape, while an inaccurate surrogate model may misguide the local search. Our pilot experiments using real-world filters show that without using a surrogate model converges even faster than using a GP model in NM simplex-based optimization.

A critical problem is the data feedback from the NM simplex optimization to the GP-based DE optimization. The candidate designs visited by NM simplex must be carefully selected for inclusion in the database considering both performance and the introduced diversity to avoid GP-based DE being trapped in local optima, i.e., the current best α solutions in the database lack diversity (Step 4 in Section III (A)). The scoring strategy in [31] is used to rank and select α simulated candidate designs to be included in the database when the current best design is improved by NM simplex. When NM simplex does not improve the current best design, it will be terminated when using more than 200 EM simulations or the objective function does not decrease for 20 consecutive iterations.

D. Parameter Settings

Compared to SMEAFO, H-SMEAFO does not introduce new algorithm parameters although there are a few empirical parameters for initialization in Section III (A), and those to switch the objective functions and optimizers in Section III (B) and (C), respectively. These parameters are based on experiments using various kinds of filters. Once set, they do not change. For the common algorithm parameters of SMEAFO and H-SMEAFO, the SMEAFO parameter setting rules are still applicable to H-SMEAFO. More details are in [2]. For simplicity, $n = 5 \times d$, $\alpha = 5 \times d$, $\tau = 5 \times d$, $F = 0.8$, and $CR = 0.8$, which are used in Section IV.

IV. Numerical Results and Comparisons

In this section, the performance of H-SMEAFO is verified by two real-world examples, including an 8th order dual-band waveguide filter with 4 transmission zeros [32] and a 6th order waveguide filter with 2 transmission zeros [33]. For both filters, the initial design is obtained by the method in [16]. As verified by experiments in the following subsections, several popular existing filter optimization methods are not able to achieve unsupervised design for them starting from the given initial designs. All the experiments are run on a workstation with Intel 3.2 GHz Core (TM) i7 CPU and 8 GB RAM under Windows operating system. CST Microwave Studio is used as the simulation tool. No parallel computing is considered and the time consumption is wall clock time.

To demonstrate the advantages of the proposed hybrid objective function, using the same hybrid surrogate model-assisted optimization algorithm in Section III (C) and the same initial design, the reference objective functions include the ZPE objective function [26], only F_2 in Section III (B), the objective function based on the difference between the extracted CM of a candidate design and ideal CM (it is called the CM difference method in the following) [25], the cognition-driven multi-feature objective function [13], and the widely used magnitude of S-parameters-based objective function (e.g., $\max(|S_{11}|)$ in the passband).

For each objective function, 5 independent runs are carried out, and the results are compared statistically. Due to the

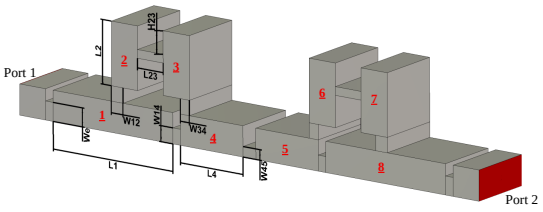

Figure 6. The structure of the X-band filter

computationally expensive EM simulations, more runs are not affordable. In all the comparisons, we consider a run to be successful when its obtained optimal design has a less than 0.1 overall specification violation (i.e., sum together) within 2500 EM simulations, although H-SMEAFO costs much fewer EM simulations than this budget.

To demonstrate the advantages of the proposed hybrid surrogate model-assisted optimization algorithm, using the same hybrid objective function in Section III (B), the reference method is SMEAFO because it is the only published global optimizer for filters to the best of our knowledge. Local optimizers are not included because they are not suitable for unsupervised design.

Since SMEAFO and H-SMEAFO are stochastic algorithms, random seeds affect the necessary number of EM simulations to satisfy the specifications. To avoid this effect and to focus the comparison on different search mechanisms, we divide the comparison into 5 groups. 5 sets of initial populations are thus randomly generated from the initial design (Step 1 in Section III (A)) and each group use one of them. In each group, the same initial population and the same random seed are used for 5 independent runs using SMEAFO and H-SMEAFO, and the convergence speed is compared statistically. The 5 groups show the same conclusion. Hence, the results from a typical initial population are displayed in the following subsections.

A. Example 1

The first example is an X-band symmetric 8th order dual-band filter with 4 transmission zeros, which is shown in Fig. 6. This filter is designed to operate at the center frequency of 10 GHz with two passbands symmetrically located at 9.35-9.70 GHz and 10.30-10.65 GHz. Four of eight cavity resonators are independent, and the corresponding ones have the same dimensions. The four resonators (i.e., resonators 1-4 or 5-8) form a cascaded quadruplet which generates two transmission zeros at 9.88 GHz and 10.12 GHz, respectively. Since the dual-band filter has a symmetric structure, there are four transmission zeros in total. With the transmission zeros, a higher stopband rejection between two passbands is realized. This filter has 10 design variables, in which, $[L1, L2, L4]$ ($L2 = L3$) target the resonant frequency, and $[W12, W34, W14, W45, We, H23, L23]$ target the coupling between resonators. The filter is modeled in CST Microwave Studio with about 12000 meshes, and each EM simulation costs 1 to 1.5 minutes.

The design specifications are shown Table I and Fig. 7. Note that the central stopband is formed by 4 transmission zeros: 2 transmission zeros at 9.88 and the other 2 at 10.12 GHz are introduced considering higher stopband rejection.

Table I
DESIGN SPECIFICATIONS FOR EXAMPLE 1

Notation	Item	Frequency Range (GHz)	Specification (dB)
PB_1	Passband 1 Reflection Coefficient (S_{11})	9.35 - 9.70	-20
PB_2	Passband 2 Reflection Coefficient (S_{11})	10.30 - 10.65	-20
SB	Stopband Transmission Coefficient (S_{21})	9.85 - 10.15	-40
SB_l	Stopband Left Edge Transmission Coefficient (S_{21})	≤ 8.8	-20
SB_r	Stopband Right Edge Transmission Coefficient (S_{21})	≥ 11.2	-20

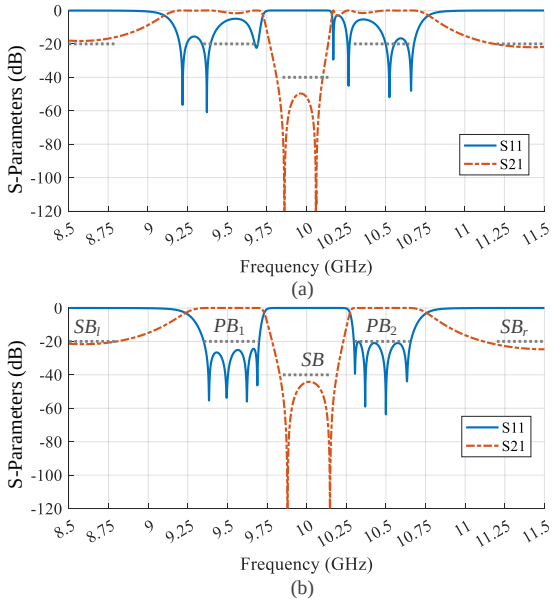

Figure 7. Response of the X-band filter: (a) The initial design. (b) An optimized design by H-SMEAFO. (The dotted lines show the specifications)

In all 5 runs, H-SMEAFO satisfies the specifications. The initial design, a typical optimal design and its response are shown in Table II and Fig. 7. For the 5 runs, an average of 678 EM simulations are used, costing 13 hours. It can be seen that the time consumption is reasonable, especially considering unsupervised design.

The 6 reference objective functions are firstly compared.

Table II
THE INITIAL DESIGN AND A TYPICAL OPTIMIZED DESIGN (ALL SIZES IN MM) (EXAMPLE 1)

Variable names	$W12$	$W34$	$W14$	$W45$	We
Initial values	4.504	3.464	5.210	3.417	6.186
Optimized values	4.234	3.489	4.192	2.713	6.894
Variable names	$H23$	$L23$	$L1$	$L2$	$L4$
Initial values	7.645	9.940	46.374	20.983	24.193
Optimized values	6.766	8.962	46.035	20.673	23.359

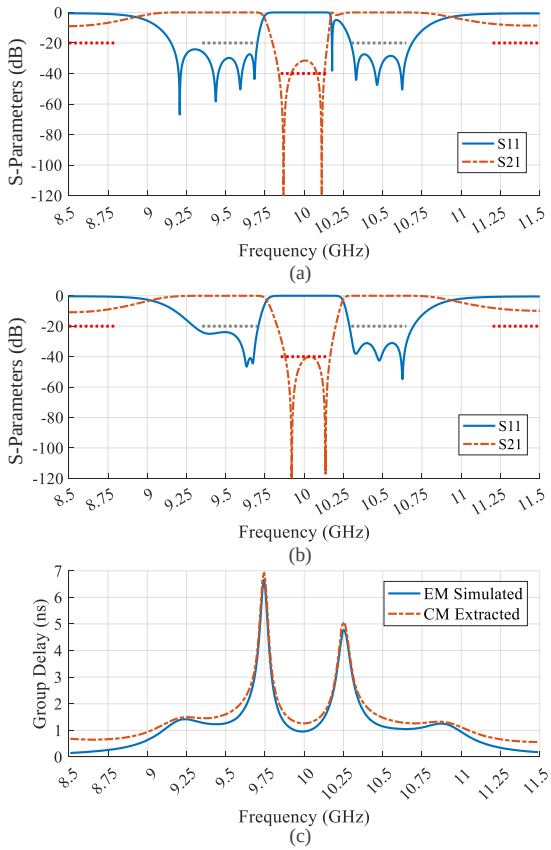

Figure 8. Response of the dual-band filter. (a) A typical response using the ZPE objective function. (b) A typical response using the cognition-driven multi-feature objective function. (c) A typical group delay comparison using the CM difference-based objective function.

In all 5 runs, the ZPE objective function, the CM difference objective function, and the cognition-driven multi-feature objective function do not succeed for this example. A typical response of the optimal design obtained by the ZPE objective function is shown in Fig. 8(a). This result verifies the statement in Section III (B): The ZPE objective function is not in line with the design specifications after the general response shape is formed. For a similar reason, the cognition-driven multi-feature objective function concentrates on the appropriate positions for the zeros and poles and the ripple height of the passband, as shown in Fig. 8(b). Therefore, it does not satisfy the specifications of the central stopband and the bandwidth requirement of the passbands, although the response shape is formed correctly.

For the CM difference-based objective function, the result verifies the statements in Section III (B). The main challenge is that the CM extraction (in this case, by group delay) is difficult to be accurate as shown in Fig. 8(c) for many candidate designs, which misleads the optimization. Note that unsupervised design is considered, while the effectiveness of the above three functions for semi-supervised design has been shown in the literature.

The magnitude of S-parameters, F_2, and the hybrid objective functions show success in the 5 runs but have different performances. For the magnitude of S-parameters-based objective

Table III
STATISTICAL RESULTS FOR DIFFERENT OBJECTIVE FUNCTIONS

	$\max(\|S_{11}\|)$	only F_2	Hybrid
Min. number of EM simulations	1179	526	646
Max. number of EM simulations	1430	1298	740
Ave. number of EM simulations	1286	955	678
Standard deviation	129	390	36

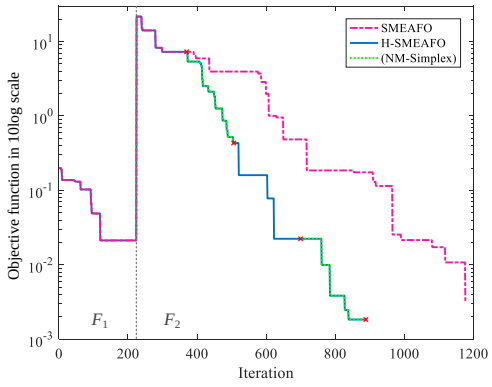

Figure 9. Typical convergence trends of the hybrid optimizer and the SMEAFO optimizer (example 1)

function (simplified as $\max(|S_{11}|)$ in the following although S_{21} is also considered), 3 out of 5 runs obtain successful results. For F_2 and the hybrid objective function, all 5 runs are successful. The comparison result is shown in Table III. Only successful runs using $\max(|S_{11}|)$ are considered in Table III.

It can be observed from Table III that the hybrid objective function improves the efficiency significantly (30% to 50% compared to the reference methods using average values). Moreover, the standard deviation is much smaller than the reference methods, showing its more stable performance.

Then, the comparisons between different optimizers using the same hybrid objective function are carried out. As said above, using the same initial population and random seed, the convergence trends of the hybrid optimization algorithm and SMEAFO differ when NM simplex is triggered. At this point, the two algorithms have the same current best design and training data points. After that, the two search mechanisms are carried out separately. The results of a typical initial population out of the 5 are as follows considering all the runs show the same conclusion.

Using the overall constraint violation threshold of 0.1 as mentioned before, H-SMEAFO takes 654 EM simulations on average to converge using this typical initial population, while SMEAFO takes an average of 1015 EM simulations over 5 runs. Therefore, the new hybrid optimizer reduces about 30% of the total EM simulations. The corresponding convergence trend is shown in Fig. 9. The effectiveness of NM simplex optimization and the iterative global and local search mechanism for filter design landscape can be observed.

B. Example 2

The second example is a C-band 6th order waveguide filter with 2 transmission zeros, which is shown in Fig. 10. The

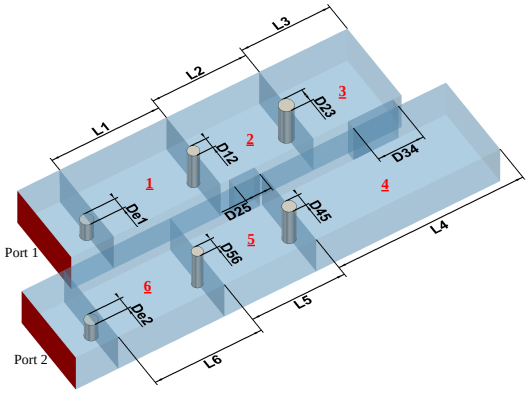

Figure 10. The structure of the C-band filter

Table IV
DESIGN SPECIFICATIONS FOR EXAMPLE 2

Notation	Item	Frequency Range (GHz)	Specification (dB)
PB	Passband Reflection Coefficient (S_{11})	4.9 - 5.1	-20
SB_l	Stopband Left Edge Transmission Coefficient (S_{21})	$\leq$ 4.87	-30
SB_r	Stopband Right Edge Transmission Coefficient (S_{21})	$\geq$ 5.15	-30

working frequency is 4.9 to 5.1 GHz. Resonators 2-5 form a cascaded quadruplet. In this filter, two transmission zeros are generated at the upper and lower stopbands, obtaining higher stopband rejections. The resonators are coupled via inductive posts or coupling irises. This is because the frequency range of the passband is too narrow to be realized by capacitive coupling irises. Therefore, the filter is designed using an E-plane cut structure. Resonator 4 uses a TE102 mode to realize a negative coupling between resonators 2 and 5. There are 14 design variables, in which $[L1, L2, L3, L4, L5, L6]$ target resonant frequency and $[D12, D23, D34, D45, D56, D25, De1, De2]$ target the coupling between resonators. The filter is modeled in CST Microwave Studio with about 12000 meshes, and each EM simulation costs about 1 to 1.5 minutes. The design specifications are shown in Table IV and Fig. 11 (b).

In all 5 runs, H-SMEAFO satisfies the specifications. The initial design and a typical optimal design are shown in Table V, with the response in Fig. 11. For the 5 runs, an average of 776 EM simulations are used, costing 16 hours. The time consumption is reasonable, especially considering unsupervised design.

As for the comparison with 6 reference objective functions, the ZPE objective function, the CM difference objective function, and the cognition-driven multi-feature objective functions do not succeed due to the same reason mentioned in example 1. The $\max(|S_{11}|)$ objective function also fails to find a design that satisfies all specifications. It is shown that the optimization is trapped in a local optimum with a maximum passband $|S_{11}|$ of -17.9 dB. A potential reason is that, with the increasing number of orders of the filter, the edges of the passband

Table V
THE INITIAL DESIGN AND A TYPICAL OPTIMIZED DESIGN (ALL SIZES IN MM) (EXAMPLE 2)

	$D12$	$D23$	$D34$	$D45$	$D56$
Initial value	2.616	3.233	18.830	3.233	2.616
Optimized value	2.509	3.246	21.230	2.919	2.325
	$D25$	$L1$	$L2$	$L3$	$L4$
Initial value	12.484	50.004	43.615	41.343	86.963
Optimized value	12.101	50.006	43.310	41.320	86.568
	$L5$	$L6$	$De1$	$De2$	
Initial value	42.976	49.971	2.918	2.919	
Optimized value	43.023	50	2.481	2.492	

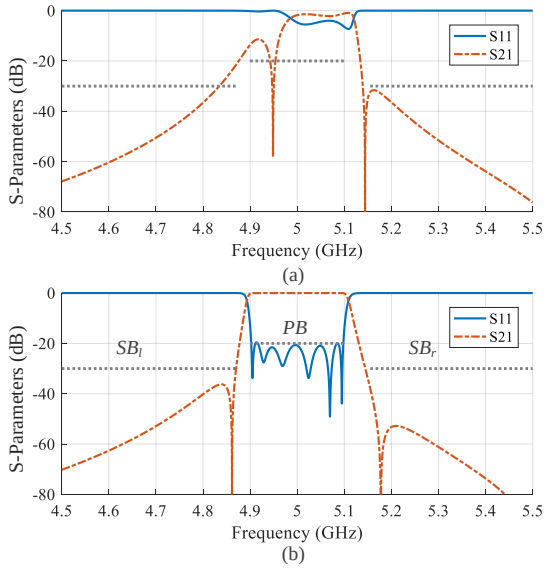

Figure 11. Response of the C-band filter: (a) The initial design. (b) An optimized design by H-SMEAFO. (The dotted lines show the specifications)

response become steeper and the $\max(|S_{11}|)$-based objective function leads to a more complex design landscape [2], which causes the search to fail.

3 out of 5 runs succeed using only the F_2 objective function. The comparison result is shown in Table VI, in which only successful runs are considered. It can be observed that the hybrid objective function improves the efficiency significantly (60% compared to only using F_2). Moreover, the standard deviation is much smaller than only using F_2, showing the advantages of F_1 and the hybrid work of F_1 and F_2.

The comparisons between different optimizers show the same observation as example 1. Again, all 5 initial populations show similar comparison results. Using a typical population among the 5 initial populations, the new hybrid optimizer reduces about 30% of the total EM simulations compared to

Table VI
STATISTICAL RESULTS FOR DIFFERENT OBJECTIVE FUNCTIONS

	only F_2	Hybrid
Min. number of EM simulations	1460	511
Max. number of EM simulations	2350	925
Ave. number of EM simulations	1916	776
Standard deviation	445	168

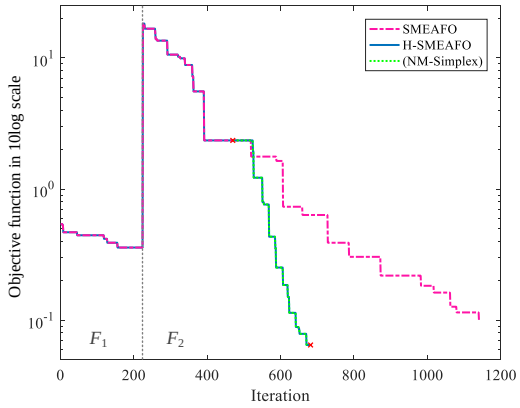

Figure 12. Typical convergence trends of the hybrid optimizer and the SMEAFO optimizer (example 2)

SMEAFO (average of 5 runs). The corresponding convergence trend is shown in Fig. 12, in which H-SMEAFO takes 712 EM simulations, whereas SMEAFO takes 1211 EM simulations. In this case, NM simplex optimization in H-SMEAFO directly converges to the final optimal design when it is triggered. The effectiveness of NM simplex optimization and the iterative global and local search mechanism for filter design landscape is demonstrated again.

V. Conclusions

In this paper, the hybrid surrogate model-assisted evolutionary algorithm for filter optimization (H-SMEAFO) has been proposed. Real-world filters, for which, unsupervised design does not appear to be possible using existing filter optimization methods, show that H-SMEAFO realizes unsupervised filter design optimization with reasonably good efficiency (e.g., about half a day using a normal desktop computer). The significant benefits mentioned in Section I are therefore realized. The effectiveness of H-SMEAFO comes from the new hybrid objective function and the hybrid optimization method bespoke for filter design landscapes. Both techniques are applicable to research works aiming at unsupervised filter design or filter global optimization. Moreover, due to its unsupervised nature, the designer's guidance or decision is not needed when using H-SMEAFO. Hence, H-SMEAFO can be programmed into a software tool, and the users only need to press a button and let H-SMEAFO obtain the optimal design satisfying the specifications on its own. Future works include continuously improving the speed for H-SMEAFO, aiming to realize unsupervised design with an experienced designer's speed.

Acknowledgement

The authors would like to thank Dr. Yi Wang, University of Birmingham, UK, for valuable discussions.

References

[1] J. E. Rayas-Sanchez, S. Koziel, and J. W. Bandler, "Advanced rf and microwave design optimization: A journey and a vision of future trends," *IEEE Journal of Microwaves*, vol. 1, no. 1, pp. 481–493, 2021.

[2] B. Liu, H. Yang, and M. J. Lancaster, "Global optimization of microwave filters based on a surrogate model-assisted evolutionary algorithm," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 6, pp. 1976–1985, 2017.

[3] R. J. Cameron, "Advanced filter synthesis," *IEEE Microwave Magazine*, vol. 12, no. 6, pp. 42–61, 2011.

[4] Y. Yu, B. Liu, Y. Wang, M. J. Lancaster, and Q. S. Cheng, "A general coupling matrix synthesis method for all-resonator diplexers and multiplexers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 3, pp. 987–999, 2020.

[5] R. Cameron, "Advanced coupling matrix synthesis techniques for microwave filters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 1, pp. 1–10, 2003.

[6] M. Guglielmi, "Simple cad procedure for microwave filters and multiplexers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, no. 7, pp. 1347–1352, 1994.

[7] X. Shang, W. Xia, and M. J. Lancaster, "The design of waveguide filters based on cross-coupled resonators," *Microwave and Optical Technology Letters*, vol. 56, no. 1, pp. 3–8, 2014.

[8] D. Swanson, "Narrow-band microwave filter design," *IEEE Microwave Magazine*, vol. 8, pp. 105–114, 2007.

[9] Q. S. Cheng, J. W. Bandler, and S. Koziel, "Space mapping design framework exploiting tuning elements," *IEEE transactions on microwave theory and techniques*, vol. 58, no. 1, pp. 136–144, 2009.

[10] J. W. Bandler, Q. S. Cheng, S. A. Dakroury, A. S. Mohamed, M. H. Bakr, K. Madsen, and J. Sondergaard, "Space mapping: the state of the art," *IEEE Transactions on Microwave theory and techniques*, vol. 52, no. 1, pp. 337–361, 2004.

[11] S. Koziel, J. W. Bandler, and K. Madsen, "A space-mapping framework for engineering optimization theory and implementation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 10, pp. 3721–3730, 2006.

[12] C. Zhang, F. Feng, V.-M.-R. Gongal-Reddy, Q. J. Zhang, and J. W. Bandler, "Cognition-driven formulation of space mapping for equal-ripple optimization of microwave filters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 7, pp. 2154–2165, 2015.

[13] F. Feng, W. Na, W. Liu, S. Yan, L. Zhu, J. Ma, and Q.-J. Zhang, "Multifeature-assisted neuro-transfer function surrogate-based em optimization exploiting trust-region algorithms for microwave filter design," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 2, pp. 531–542, 2020.

[14] P. Zhao and K. Wu, "Homotopy optimization of microwave and millimeter-wave filters based on neural network model," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 4, pp. 1390–1400, 2020.

[15] R. J. Cameron, C. M. Kudsia, and R. R. Mansour, *Microwave filters for communication systems: fundamentals, design, and applications*. John Wiley & Sons, 2018.

[16] Y. Yu, "Knowledge-guided microwave diplexers and multiplexers design based on computational intelligent techniques," Ph.D. dissertation, University of Birmingham, Feb. 2022.

[17] M. S. Sorkherizi and A. A. Kishk, "Use of group delay of sub-circuits in optimization of wideband large-scale bandpass filters and diplexers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 8, pp. 2893–2905, 2017.

[18] R. J. Cameron, C. M. Kudsia, and R. R. Mansour, *Synthesis of a General Class of the Chebyshev Filter Function*, 2018.

[19] ——, *Waveguide Realization of Single and Dual Mode Resonator Filters*, 2018.

[20] J.-S. G. Hong and M. J. Lancaster, *Microstrip filters for RF/microwave applications*. John Wiley & Sons, 2004.

[21] Z. Zhang, B. Liu, Y. Yu, and Q. S. Cheng, "A microwave filter yield optimization method based on off-line surrogate model-assisted evolutionary algorithm," *IEEE Transactions on Microwave Theory and Techniques*, 2022.

[22] M. Emmerich, K. Giannakoglou, and B. Naujoks, "Single-and multi-objective evolutionary optimization assisted by Gaussian random field metamodels," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 4, pp. 421–439, 2006.

[23] I. Couckuyt, A. Forrester, D. Gorissen, F. De Turck, and T. Dhaene, "Blind kriging: Implementation and performance analysis," *Advances in Engineering Software*, vol. 49, pp. 1–13, 2012.

[24] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the nelder–mead simplex method in low dimensions," *SIAM Journal on optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[25] M. Sharifi Sorkherizi and A. A. Kishk, "Use of group delay of subcircuits in optimization of wideband large-scale bandpass filters and diplexers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 8, pp. 2893–2905, 2017.
[26] S. Amari, "Synthesis of cross-coupled resonator filters using an analytical gradient-based optimization technique," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, no. 9, pp. 1559–1564, 2000.
[27] D. Deschrijver and T. Dhaene, "A note on the multiplicity of poles in the vector fitting macromodeling method," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 4, pp. 736–741, 2007.
[28] B. Gustavsen and A. Semlyen, "Rational approximation of frequency domain responses by vector fitting," *IEEE Transactions on Power Delivery*, vol. 14, no. 3, pp. 1052–1061, 1999.
[29] P. Zhao and K.-L. Wu, "Circuit model extraction of parallel-connected dual-passband coupled-resonator filters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 2, pp. 822–830, 2018.
[30] Y. Yu, B. Liu, Y. Wang, and Q. S. Cheng, "Automated diplexer design with key performance indicator-based objectives," *IEEE Microwave and Wireless Components Letters*, pp. 1–4, 2022.
[31] B. Liu, V. Grout, and A. Nikolaeva, "Efficient global optimization of actuator based on a surrogate model assisted hybrid algorithm," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5712–5721, 2017.
[32] X. Shang, Y. Wang, G. Nicholson, and M. Lancaster, "Design of multiple-passband filters using coupling matrix optimisation," *IET microwaves, antennas & propagation*, vol. 6, no. 1, pp. 24–30, 2012.
[33] E. Ofli, R. Vahldieck, and S. Amari, "Novel e-plane filters and diplexers with elliptic response for millimeter-wave applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 3, pp. 843–851, 2005.