Bayesian Optimisation of a Frequency Selective Surface Using a Regularised Objective Function

Kilian Bihannic, Jérémy Omer, Renaud Loison, Guillaume Reille

Bayesian Optimisation of a Frequency Selective Surface Using a Regularised Objective Function

Kilian Bihannic*, Jeremy Omer †, Renaud Loison‡, Guillaume Reille§
*IRMAR, Rennes, France, Kilian.Bihannic@insa-rennes.fr
†IRMAR, Rennes, France, Jeremy.Omer@insa-rennes.fr
‡IETR, Rennes, France, Renaud.Loison@insa-rennes.fr
§Thales DMS France, Elancourt, France, guillaume.reille@fr.thalesgroup.com

***Abstract*—This work introduces a penalty-based regularisation to optimise the geometric parameters of a periodic frequency selective surface. The equivalent circuit of a multi-layer bandpass filter is used to illustrate the benefits of the approach for wideband applications. The proposed regularisation is obtained by convolving the frequency response and the desired thresholds before measuring the infraction. We apply bayesian optimisation on the unconstrained optimisation problem and assess the performance of the strategy by the number of objective function evaluations. A faster convergence is observed with the proposed regularisation.**

***Index Terms*—Bayesian Optimisation, Electromagnetics, Frequency Selective Surface, Regularisation, Wideband Filtering**

Introduction

Periodic and quasi-periodic surfaces are planar arrangements of dielectric and metallic unitary elements [1]. They have been widely used in Frequency Selective Surfaces (FSS), reflectarrays and polarising surfaces. When impinging on a surface with geometry $\boldsymbol{x} \in \Omega_x \subset \mathbb{R}^d$, an incident electromagnetic wave of frequency $f \in \Omega_f \subset \mathbb{R}$ is reflected and transmitted with a given phase and magnitude. The S parameters are complex numbers characterising these properties of the surface. For filtering applications, the module of the reflexion S_{11} or transmission S_{12} coefficients must be bounded at operating frequency bandwidths. Computing precisely the S parameters at discrete frequencies requires expensive calls to an electromagnetic solver. Nevertheless, some FSS can be reasonably approximated by an equivalent circuit at normal incidence.

Bayesian Optimisation strategies [2] are iterative global optimisation algorithms designed to optimise expensive to evaluate objective functions. After evaluating the objective function on an initial Design of Experiment [3], a Gaussian Process [4] is trained on the observations and used as a surrogate model. The objective function is then evaluated at the point maximising an acquisition function. The model is then retrained with the new observation and the acquisition function is maximised once again. In practice both regularity and convexity of the objective function impact the performance of the optimisation strategy.

In this paper, we formulate the design of a second order multi-layer bandpass filter [5] as an unconstrained optimisation problem. Using an equivalent circuit [6], we study objective functions assessing the filtering properties of the FSS. A natural objective function is the norm of the infraction of thresholds at the operating frequencies. We noticed that this formulation presents an undesirable modelling property. Due to the resonant nature of the frequency response and the piece-wise constant nature of the bounding constraints, the infraction of the bounds might not capture how close we are from a realisable scattering profile. This observation motivated our investigations on reformulations of the optimisation problem. We propose to penalise the objective function by a necessary condition of optimality. This condition is obtained by applying an autoconvolution over frequency on the bounds and on the response before measuring the infraction.

We present the filter chosen as a case of study and formulate the optimisation problem in the next section. We recall the principles of bayesian optimisation in Section II. In Section III we present our main contribution, a penalisation adapted to wideband filtering applications. Numerical results are presented and discussed in Section IV.

I. Case of study

A. Bi-periodic Frequency Selective Surface

The multi-layer FSS introduced in [5] (Figure 1) is used as a case of study to formulate the optimisation problem. Operating at $f_0 = 10$ GHz, the surface presents two layers of dielelectric material covered by capacitive patches arranged with a sub-wavelength periodicity. An inductive grid is placed between the dielectric layers. This FSS is a Chebytchev filter [7].

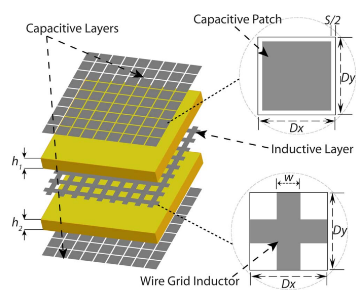

Fig. 1. Second order bandpass FSS from [5]

The optimisation is carried over two parameters. The first is the periodicity $D = D_x = D_y$, between $\lambda_0/10 = 3$ mm and $\lambda_0/5 = 6$ mm. The second parameter is the normalised width of the grid $w/D \in [0, 1]$. We stack the optimisation variables in the vector $\boldsymbol{x} \in \mathbb{R}^2$. Other parameters of the surface are fixed to the values proposed in [5]. The dielectric constants of the materials are $\epsilon_r = 3.4$ and $\mu_r = 1$. The two dielectric substrate thicknesses h_1 and h_2 are both equal to 0.59 mm. The space between two patches is fixed to $s = 0.15$ mm.

An equivalent circuit for the patches and the grid is well known [6]. This cheap analytical approximation is effective at normal incidence and used in this paper to investigate different formulations of the optimisation problem.

B. Problem Formulation

A frequency mask is defined to specify the optimal response. The response must be upper and lower bounded by user-defined thresholds. Here, we want the reflexion coefficient to achieve the following specifications (Figure 2):

- $20 \log_{10} |S_{11}|(\boldsymbol{x}, f) > -3$ dB from 7 to 8 GHz
- $20 \log_{10} |S_{11}|(\boldsymbol{x}, f) < -15$ dB from 9.33 to 10.66 GHz
- $20 \log_{10} |S_{11}|(\boldsymbol{x}, f) > -3$ dB from 12 to 13 GHz

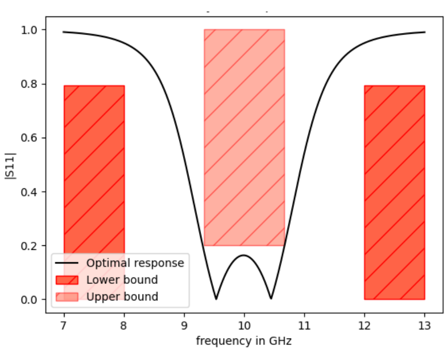

Fig. 2. Optimal response $\boldsymbol{s}(\boldsymbol{x}^{opt})$ optimal w.r.t. the mask.

We aim to solve the following realisability problem:

$$\begin{aligned} &\text{find } \boldsymbol{x}^* \in \Omega_x \subset \mathbb{R}^d \\ &\text{such that } \boldsymbol{l} \leq \boldsymbol{s}(\boldsymbol{x}^*) \leq \boldsymbol{u} \end{aligned} \tag{1}$$

where $\boldsymbol{s}(\boldsymbol{x}) = (|S_{11}|(\boldsymbol{x}, f_i))_{f_i \in \Omega_f}$ denotes the discrete linear response vector of length $F \in \mathbb{N}$. The bounds $\boldsymbol{l}$ and $\boldsymbol{u}$ are piecewise constant functions of frequency, characterised by the thresholds and bandwidth of the frequency mask.

II. Bayesian Optimisation

A. Principles

Unconstrained bayesian optimisation is an iterative method adapted to costly objective functions. In this paper, the objective function is assumed noiseless. A sparse initial design of experiment [3] $\boldsymbol{X} = [\boldsymbol{x}^1, \cdots, \boldsymbol{x}^n] \in \mathbb{R}^{n \times d}$ such as a latin hypercube sampling or a quasi-random sequence [8] is used to evaluate n initial values of the objective function $\boldsymbol{y} = [y^1, \cdots, y^n] \in \mathbb{R}^n$. At iteration k, two steps are repeated iteratively to select the next point where the objective is evaluated. First, a gaussian process is trained on $(\boldsymbol{X}, \boldsymbol{y})$, which provides a surrogate model for the objective function. An acquisition function $\alpha(\boldsymbol{x})$ is then maximised to obtain $\boldsymbol{x}^{n+1}$ where the objective is evaluated. The algorithm stops when a budget of evaluation is exhausted. The overall strategy is summarised in Algorithm 1.

Algorithm 1 Bayesian Optimisation

1: $\boldsymbol{X} \in \mathbb{R}^{n \times d} \leftarrow$ initial Design of Experiment of size n
2: $\boldsymbol{y} \in \mathbb{R}^n \leftarrow$ associated objective function values
3: $\epsilon \leftarrow$ precision
4: **while** $n <$ budget and $y^{best} > \epsilon$ **do**
5: $\quad \hat{\Theta}^{n+1} \leftarrow$ optimise GP hyperparameters on $(\boldsymbol{X}, \boldsymbol{y})$
6: $\quad \boldsymbol{x}^{n+1} \leftarrow$ optimise acquisition function $\alpha(\boldsymbol{x})$
7: $\quad y^{n+1} \leftarrow (\boldsymbol{x}^{n+1})$ evaluate objective function
8: $\quad y^{n+1}_{best} \leftarrow \min(y^{best}, y^{n+1})$ update best evaluation
9: $\quad (\boldsymbol{X}, \boldsymbol{y}) \leftarrow (\boldsymbol{X}, \boldsymbol{y}) \cup (\boldsymbol{x}^{n+1}, y^{n+1})$ update database
10: $\quad n \leftarrow n + 1$
11: **end while**

In typical Bayesian Optimisation settings, lines 2 and 7 corresponding to the evaluation of the objective function are the most time consuming. We dedicate next sections to line 5 and 6 corresponding to gaussian processes and acquisition functions, since both line require to solve an optimisation problem.

B. Gaussian Process

A gaussian process [4] $Y(\boldsymbol{x})$ is a collection of gaussianly distributed random variables. It is fully characterised by its mean $m(\boldsymbol{x})$ and a positive definite covariance function $k_\Theta(\boldsymbol{x}, \boldsymbol{x}')$ such as the exponential or matern kernel. With a gaussian prior, an evaluation of the objective is seen as an observation of $Y(\boldsymbol{x})$. The hyperparameters of the covariance kernel Θ are estimated by maximising the likelihood of the observations. The posterior distribution $\mathbb{P}_n\{y(\boldsymbol{x})|\boldsymbol{x}, \boldsymbol{X}, \boldsymbol{y}, \Theta\}$ is used to make predictions at $\boldsymbol{x} \in \mathcal{X}$. It is also a gaussian distribution with mean $\mu(\boldsymbol{x})$ and variance $\sigma^2(\boldsymbol{x})$:

$$\mu(\boldsymbol{x}) = \boldsymbol{k}_{\boldsymbol{x}}^T K_\Theta^{-1} \boldsymbol{y} \tag{2}$$

$$\sigma^2(\boldsymbol{x}) = k_\Theta(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{k}_{\boldsymbol{x}}^T K_\Theta^{-1} \boldsymbol{k}_{\boldsymbol{x}} \tag{3}$$

where $k_{\boldsymbol{x}}$ is the vector of correlations between $\boldsymbol{x}$ and the observed points in $\boldsymbol{X}$ and K is the correlation matrix. This class of surrogate models is flexible through the choice of the covariance function and provides not only a mean estimate for the objective function but also an estimation of the uncertainty of the model.

C. Acquisition Function

Minimising directly the surrogate model $\mu(\boldsymbol{x})$ to generate the next point that should be evaluated may converge in a local minimum. Instead, an acquisition function makes a compromise between minimising the model and exploring the design space in order to generate a dense sequence in the optimisation domain. The expected improvement acquisition function proposed in the algorithm EGO [2] is defined as:

$$\alpha_{EI}(\boldsymbol{x}) = \mathbb{E}[(y(\boldsymbol{x}) - y^n_{best})_+ | \boldsymbol{x}, \boldsymbol{X}, \boldsymbol{y}] \tag{4}$$

where $y^n_{best} = \min_{i=1,\cdots,n} y^i$ and $(z)_+ = \max(0, z)$. This acquisition function is chosen in the rest of this paper. The next section is dedicated to our contribution, the formulation of the optimisation problem as an unconstrained objective function.

III. Filtering Objective Functions

A. Necessary and sufficient conditions of realisability

A natural way to translate problem (1) as an unconstrained optimisation problem consists in summing the constraints [9] to quantify the realisability of a frequency response by a real number. The infraction vector $\boldsymbol{h}(\boldsymbol{x}) \in \mathbb{R}^F$ takes the form:

$$\boldsymbol{h}(\boldsymbol{x}) = \max\left(\boldsymbol{0}, \boldsymbol{l} - \boldsymbol{s}(\boldsymbol{x}), \boldsymbol{s}(\boldsymbol{x}) - \boldsymbol{u}\right) \tag{5}$$

The norm of that vector corresponds to the Riemann sum associated to the area intersecting the scattering parameter and the bounds. For any norm $||\cdot|| : \mathbb{R}^F \mapsto \mathbb{R}^+$ we have the following necessary and sufficient condition of realisability:

$$||\boldsymbol{h}(\boldsymbol{x})|| = 0 \iff \boldsymbol{l} \leq \boldsymbol{s}(\boldsymbol{x}) \leq \boldsymbol{u} \tag{6}$$

Objective functions of the form $||\boldsymbol{h}(\boldsymbol{x})||$ are often used in the litterature, the infinite norm corresponding to the maximum infraction of the bounds, the L1 norm corresponding to the average infraction over the bandwidth [10] and the L2 norm corresponding to the euclidean distance, often used in space mapping [11]. These functions share an undesirable modelling property, illustrated in Figure 3. A response $\boldsymbol{s_1} = \boldsymbol{s}(\boldsymbol{x}_1)$ resonating twice in the bandwidth is expected to be close from the global minimum, but is more penalised than a response $\boldsymbol{s_2} = \boldsymbol{s}(\boldsymbol{x}_2)$ not resonating at all. This observation motivated our investigations on reformulations of the objective function.

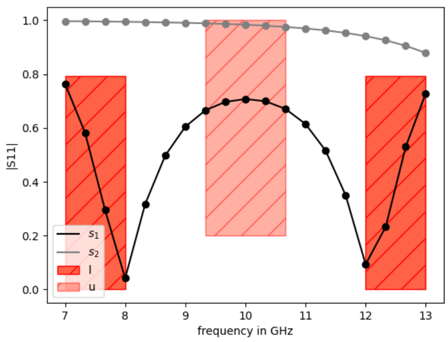

Fig. 3. Discrete response resonating twice in black, not resonating in grey

B. Necessary condition of realisability

In a simultaneous work carried out by Koziel [12], a multi-band antenna is optimised to match resonant frequencies with target locations $(f_r)_{r\in[1,R]}$. The objective function takes the form $\max_{r\in[1,R]} |S_{11}|(\boldsymbol{x}, f_r)$ and is minimised with a trust region algorithm. The authors observed that convergence of the local algorithm is improved by penalising the objective function by a necessary condition on resonant frequencies.

Our contribution follows a similar approach for wideband applications where the target resonant frequencies are specified through the mask. Instead of extracting resonant frequencies, we define the infraction of the convolved response by the convolved frame as:

$$\boldsymbol{\mathcal{H}}(\boldsymbol{x}) = \max\left(\boldsymbol{0}, \boldsymbol{\mathcal{L}} - \boldsymbol{\mathcal{S}}(\boldsymbol{x}), \boldsymbol{\mathcal{S}}(\boldsymbol{x}) - \boldsymbol{\mathcal{U}}\right) \tag{7}$$

where $\boldsymbol{\mathcal{L}} = \boldsymbol{l} * \boldsymbol{l}$, $\boldsymbol{\mathcal{U}} = \boldsymbol{u} * \boldsymbol{u}$ and $\boldsymbol{\mathcal{S}}(\boldsymbol{x}) = (\boldsymbol{s} * \boldsymbol{s})(\boldsymbol{x})$ with $*$ denoting the convolution operator over frequency. We illustrate in Figure 4 the effect of convolutions on the bounds and on the two responses already presented in Figure 3, $\boldsymbol{\mathcal{S}_1} = \boldsymbol{\mathcal{S}}(\boldsymbol{x_1})$ and $\boldsymbol{\mathcal{S}_2} = \boldsymbol{\mathcal{S}}(\boldsymbol{x_2})$. Clearly, we have $||\boldsymbol{\mathcal{H}}(\boldsymbol{x}_1)|| \leq ||\boldsymbol{\mathcal{H}}(\boldsymbol{x}_2)||$ while $||\boldsymbol{h}(\boldsymbol{x}_1)|| \geq ||\boldsymbol{h}(\boldsymbol{x}_2)||$. Moreover, the piecewise constant bounds $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{U}}$ are made continuous by the autoconvolution.

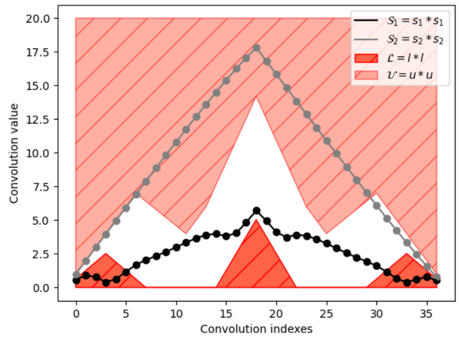

Fig. 4. Convolved response resonating twice in black, not resonating in grey

From the properties of convolutions, if $\boldsymbol{0} \leq \boldsymbol{l} \leq \boldsymbol{s} \leq \boldsymbol{u} \leq \boldsymbol{1}$ then $\boldsymbol{\mathcal{L}} \leq \boldsymbol{\mathcal{S}}(\boldsymbol{x}) \leq \boldsymbol{\mathcal{U}}$. Therefore we have the following necessary condition of realisability:

$$||\boldsymbol{\mathcal{H}}(\boldsymbol{x})|| = 0 \impliedby \boldsymbol{l} \leq \boldsymbol{s}(\boldsymbol{x}) \leq \boldsymbol{u} \tag{8}$$

The condition $||\boldsymbol{\mathcal{H}}(\boldsymbol{x})|| = 0$ defines a large region around the global minimum. Problem (1) is then equivalently written as:

$$\begin{gathered} \min_{\boldsymbol{x}\in\mathcal{X}} \; ||\boldsymbol{h}(\boldsymbol{x})|| \\ \text{such that } ||\boldsymbol{\mathcal{H}}(\boldsymbol{x})|| = 0 \end{gathered} \tag{9}$$

In next section, we solve problem (7) by penalising the objective $||\boldsymbol{h}(\boldsymbol{x})||$ with $||\boldsymbol{\mathcal{H}}(\boldsymbol{x})||$ and expect to improve the optimisation process in two ways. First, designs with non-resonating response become the most penalised, which limits their exploration to favor resonating responses. Second, the optimised objective function is more regular and should therefore be better captured by the gaussian process.

IV. Numerical Results

In this section, the algorithm EGO is applied to solve problem (1) formulated as an unconstrained optimisation problem. The optimisation process starts by evaluating the response at $n = 4$ initial points. We generate 100 designs of experiments with randomised Sobol sequences [8]. The optimisation stops when the objective function value is smaller than $\varepsilon = 10^{-6}$ or when the maximum number of evaluations reaches a budget of 200.

We start by minimising $||\boldsymbol{h}(\boldsymbol{x})||_\infty$, $||\boldsymbol{h}(\boldsymbol{x})||_1$ and $||\boldsymbol{h}(\boldsymbol{x})||_2$ to assess the impact of the objective function. Since $||\boldsymbol{h}(\boldsymbol{x})||_1$ is the best performing necessary and sufficient condition, we penalise it with $||\boldsymbol{\mathcal{H}}(\boldsymbol{x})||_1$ to evaluate the benefits of the regularisation. The mean minimum value of the objective obtained at each iteration is presented in Figure 5 for the five approaches.

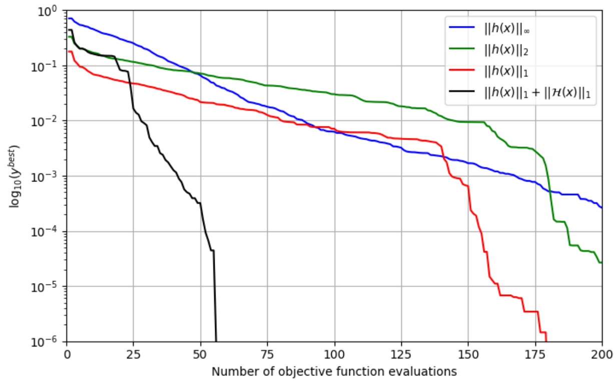

Fig. 5. Mean minimum value of the objective function through the iterations.

With $||\boldsymbol{h}(\boldsymbol{x})||_\infty$ and $||\boldsymbol{h}(\boldsymbol{x})||_2$ some runs of EGO exhausted the budget. On the other hand, all the runs converged with $||\boldsymbol{h}(\boldsymbol{x})||_1$ and $||\boldsymbol{h}(\boldsymbol{x})||_1 + ||\boldsymbol{\mathcal{H}}(\boldsymbol{x})||_1$. We observe a significant improvement in the average convergence speed when penalising $||\boldsymbol{h}(\boldsymbol{x})||_1$ with $||\boldsymbol{\mathcal{H}}(\boldsymbol{x})||_1$. The percentage of converged runs after a given budget of evaluation is presented in Figure 6. The worst run strongly impacts the mean of the objective function.

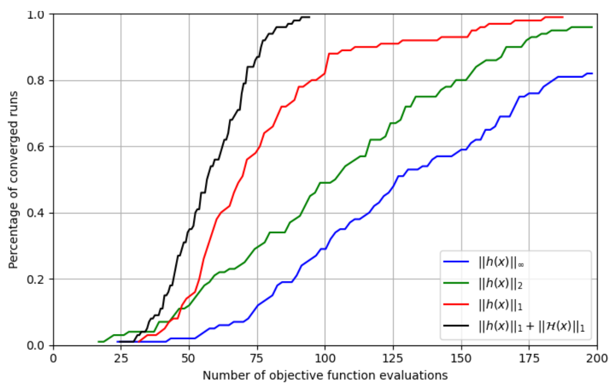

Fig. 6. Cumulative distribution of number of evalations before convergence.

We also optimised $||\boldsymbol{h}(\boldsymbol{x})||_1$ and $||\boldsymbol{h}(\boldsymbol{x})||_1+||\boldsymbol{\mathcal{H}}(\boldsymbol{x})||_1$ using Particle Swarm Optimisation [13] (PSO) algorithm. With this evolutionary strategy, an initial positions of a population is updated over the iterations. We heuristically chose an initial population size of $n = 20$ and keep unchanged the other hyperparameters. The median PSO run used 12 iterations to find the global optimum for $||\boldsymbol{h}(\boldsymbol{x})||_1$ and 11 iterations for $||\boldsymbol{h}(\boldsymbol{x})||_1 + ||\boldsymbol{\mathcal{H}}(\boldsymbol{x})||_1$. Therefore the proposed approach uses 220 evaluations instead of 240 for the median run.

Bayesian Optimisation is a more sample efficient method, with median runs converging in 56 and 67 evaluations for $||\boldsymbol{h}(\boldsymbol{x})||_1$ and $||\boldsymbol{h}(\boldsymbol{x})||_1+||\boldsymbol{\mathcal{H}}(\boldsymbol{x})||_1$. In practice, PSO is easily parallelisable, while standard Bayesian Optimisation is not. To parallelise bayesian optimisation, a non-myopic acquisition function such as $\alpha_{qEI}(\boldsymbol{x})$ [14] is maximised to obtain a batch of q candidate points.

Conclusion

In this work, we applied unconstrained bayesian optimisation to optimise a periodic multi-layer FSS over two parameters. We proposed a necessary condition of realisability and used it to penalise the objective function. The number of evaluations before convergence is decreased using the penalisation.

In an upcoming work, we will propose and compare new objective functions for wideband filtering and optimise three FSS of growing order. The optimisation will be conducted on HFSS numerical solver to evaluate the benefits of the approach in terms of computational time.

Acknowledgements

This work was financed by Thales DMS France.

References

[1] B. A Munk, *Frequency selective surfaces: theory and design*, John Wiley and Sons, DOI: 10.1002/0471723770, 2000.

[2] D.R. Jones, *Efficient Global Optimization of Expensive Black-Box Functions*, Journal of Global Optimization, DOI: 10.1023/A:1008306431147 1998.

[3] J. Sacks, W. Welch *Design and analysis of computer experiments*, Statistical Science, ISBN:978-1-4757-3799-8, 1989.

[4] C. Rasmussen, C. Williams, *Gaussian processes for machine learning*, MIT Press, ISBN: 978-3-540-23122-6, 2006.

[5] M. Al-Joumayly, N. Behdad, *A New Technique for Design of Low-Profile, Second-Order, Bandpass Frequency Selective Surfaces*, IEEE Transactions on Antennas and Propagation, DOI:10.1109/TAP.2008.2011382, 2009.

[6] A. Guarriello, G. Courtin, R. Loison, R. Gillard, *A general equivalent circuit model for Phoenix cells*, IEEE Transactions on Antennas and Propagation, DOI: 10.1109/TAP.2021.3076676, 2021.

[7] A. V. Stankovsky, Y. A. Litinskaya, A. D. Nemshon, S. V. Polenga, Y. P. Salomatov, *Synthesis of Spatial Band-Pass Filter Based on Frequency-Selective Surfaces* presented in 2021 at USBEREIT.

[8] Pierre l'Ecuyer *Randomized Quasi-Monte Carlo: An Introduction for Practitioners*, International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, DOI: 10.1007/978-3-319-91436-7_2, 2016

[9] J. Nocedal, S. Wright, *Numerical Optimization*, Springer series in operation research and financial engineering, ISBN: 978-0-387-40065-5, 2006.

[10] J. P. Jacobs, S. Koziel, L. Leifsson, *Bayesian Support Vector Regression Modeling of Microwave Structures for Design Applications*, Springer, DOI:10.1007/978-1-4614-7551-4_6, 2013

[11] J. Bandler, R. Biernacki, *Electromagnetic Optimization Exploiting Aggressive Space Mapping*, IEEE Transactions on Microwave Theory and Techniques, DOI:10.1109/22.475649, 1996.

[12] S. Koziel, A. Pietrenko-Dabrowska, *Reliable Parameter Tuning of Multi-Band Antennas Using Penalty-Based Regularization*, International Applied Computational Electromagnetics Society Symposium, DOI: 10.1109/ACES53325.2021.00055, 2021.

[13] S. Genovesi, *Particle Swarm Optimization for the Design of Frequency Selective Surfaces*, IEEE Antennas and Wireless Propagation Letters, DOI: 10.1108/COMPEL-10-2015-0362, 2006.

[14] S. Marmin, C. Chevalier, D. Ginsbourger, *Differentiating the multipoint Expected Improvement for optimal batch design*, hal-01133220v3, 2015.