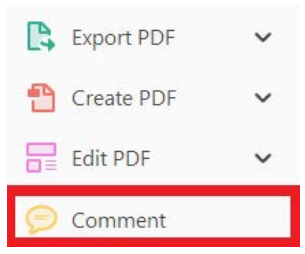
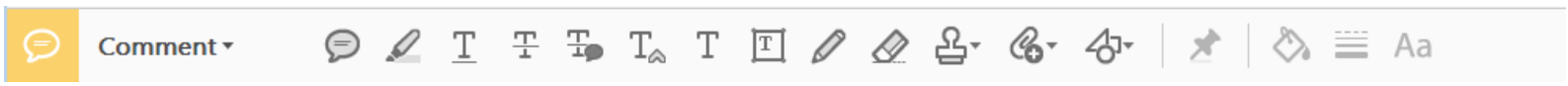
USING e-ANNOTATION TOOLS FOR ELECTRONIC PROOF CORRECTION

Required software to e-Annotate PDFs: Adobe Acrobat Professional or Adobe Reader (version 11 or above). (Note that this document uses screenshots from Adobe Reader DC.)
The latest version of Acrobat Reader can be downloaded for free at: http://get.adobe.com/reader/

Once you have Acrobat Reader open on your computer, click on the Comment tab (right-hand panel or under the Tools menu).

This will open up a ribbon panel at the top of the document. Using a tool will place a comment in the right-hand panel. The tools you will use for annotating your proof are shown below:



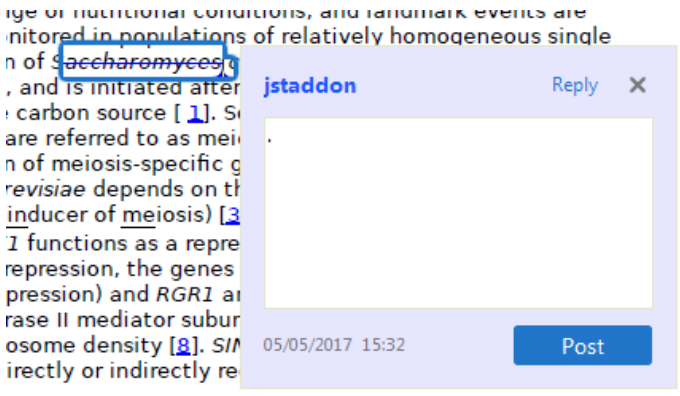


1. Replace (Ins) Tool – for replacing text.

Strikes a line through text and opens up a text box where replacement text can be entered.

How to use it:

- Highlight a word or sentence.
- Click on .
- Type the replacement text into the blue box that appears.

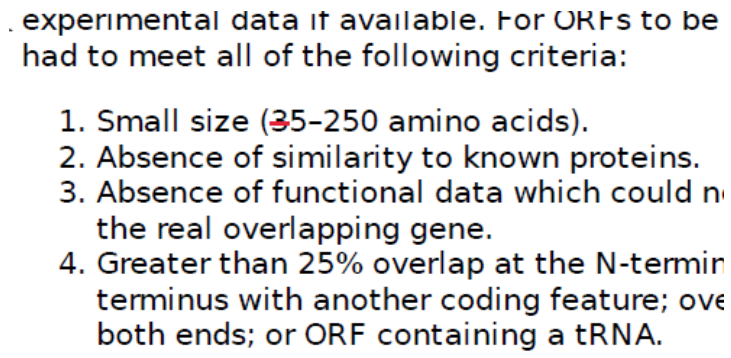


2. Strikethrough (Del) Tool – for deleting text.

Strikes a red line through text that is to be deleted.

How to use it:

- Highlight a word or sentence.
- Click on .
- The text will be struck out in red.


experimental data if available. For ORFs to be had to meet all of the following criteria:

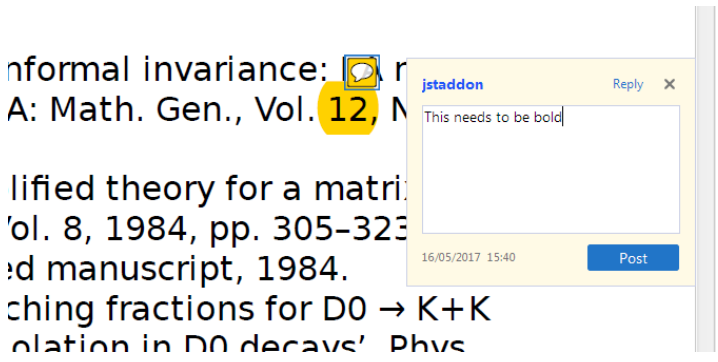
1. Small size (~~3~~5–250 amino acids).
2. Absence of similarity to known proteins.
3. Absence of functional data which could n the real overlapping gene.
4. Greater than 25% overlap at the N-termir terminus with another coding feature; ove both ends; or ORF containing a tRNA.

3. Commenting Tool – for highlighting a section to be changed to bold or italic or for general comments.

Use these 2 tools to highlight the text where a comment is then made.

How to use it:

- Click on .
- Click and drag over the text you need to highlight for the comment you will add.
- Click on .
- Click close to the text you just highlighted.
- Type any instructions regarding the text to be altered into the box that appears.

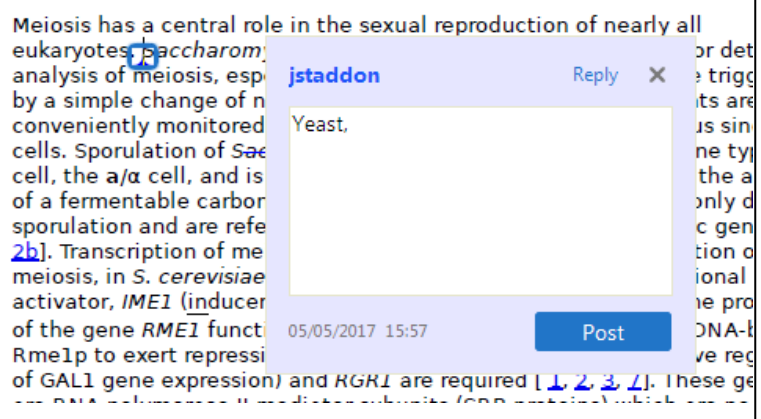


4. Insert Tool – for inserting missing text at specific points in the text.

Marks an insertion point in the text and opens up a text box where comments can be entered.

How to use it:

- Click on .
- Click at the point in the proof where the comment should be inserted.
- Type the comment into the box that appears.



WILEY

USING e-ANNOTATION TOOLS FOR ELECTRONIC PROOF CORRECTION

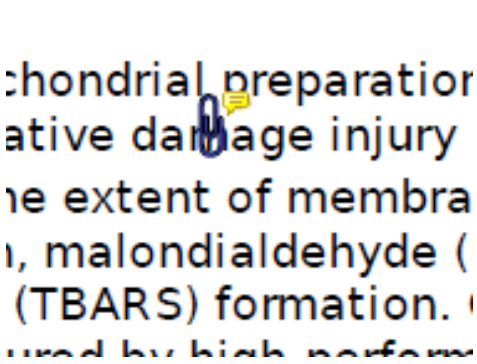
5. Attach File Tool – for inserting large amounts of text or replacement figures.

Inserts an icon linking to the attached file in the appropriate place in the text.

How to use it:

- Click on  .
- Click on the proof to where you'd like the attached file to be linked.
- Select the file to be attached from your computer or network.
- Select the colour and type of icon that will appear in the proof. Click OK.

  The attachment appears in the right-hand panel.

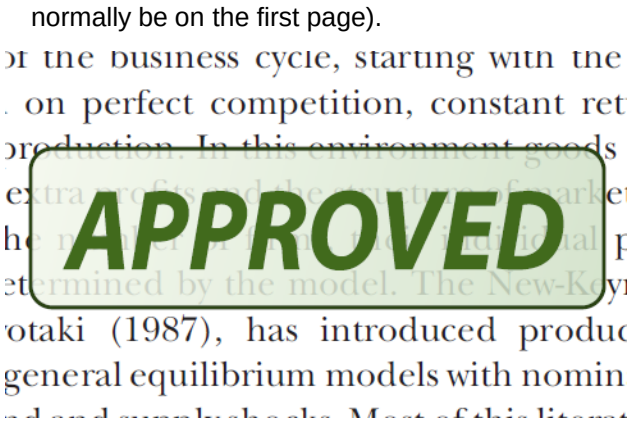


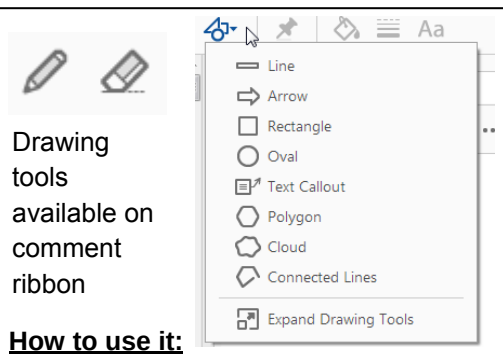
6. Add stamp Tool – for approving a proof if no corrections are required.

Inserts a selected stamp onto an appropriate place in the proof.

How to use it:

- Click on  .
- Select the stamp you want to use. (The Approved stamp is usually available directly in the menu that appears. Others are shown under *Dynamic*, *Sign Here*, *Standard Business*).
- Fill in any details and then click on the proof where you'd like the stamp to appear. (Where a proof is to be approved as it is, this would normally be on the first page).





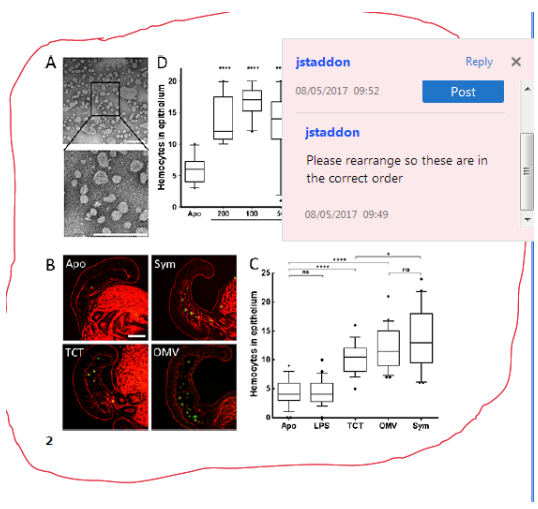
Drawing tools available on comment ribbon

7. Drawing Markups Tools – for drawing shapes, lines, and freeform annotations on proofs and commenting on these marks.

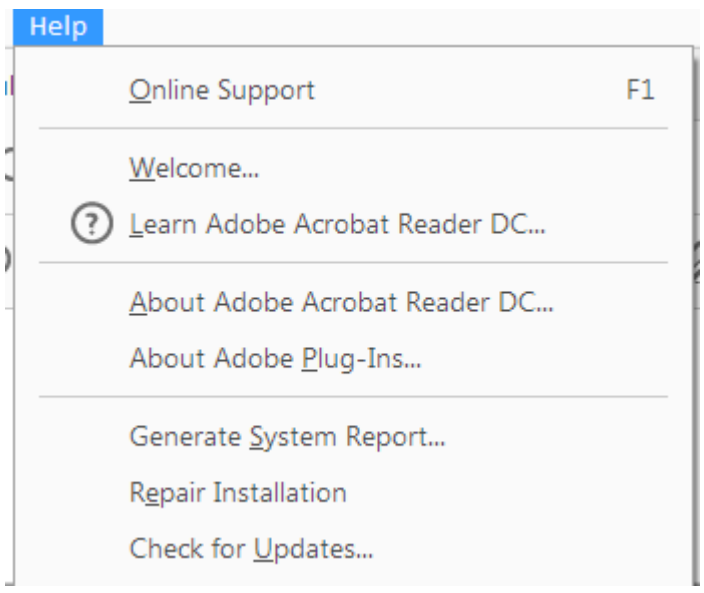
Allows shapes, lines, and freeform annotations to be drawn on proofs and for comments to be made on these marks.

How to use it:

- Click on one of the shapes in the Drawing Markups section.
- Click on the proof at the relevant point and draw the selected shape with the cursor.
- To add a comment to the drawn shape, right-click on shape and select *Open Pop-up Note.*
- Type any text in the red box that appears.



For further information on how to annotate proofs, click on the Help menu to reveal a list of further options:



WILEY

Author Query Form

Journal: International Journal of Numerical Modelling: Electronic Networks, Devices and Fields

Article: jnm_2584

Dear Author,

During the copyediting of your paper, the following queries arose. Please respond to these by annotating your proofs with the necessary changes/additions.

- If you intend to annotate your proof electronically, please refer to the E-annotation guidelines.
- If you intend to annotate your proof by means of hard-copy mark-up, please use the standard proofing marks. If manually writing corrections on your proof and returning it by fax, do not write too close to the edge of the paper. Please remember that illegible mark-ups may delay publication.

Whether you opt for hard-copy or electronic annotation of your proofs, we recommend that you provide additional clarification of answers to queries by entering your answers on the query sheet, in addition to the text mark-up.

Query No.	Query	Remark
Q1	AUTHOR: Please verify that the linked ORCID identifiers are correct for each author.	
Q2	AUTHOR: Please confirm that forenames/given names (blue) and surnames/family names (vermilion) have been identified correctly.	
Q3	AUTHOR: The journal style allows up to 6 key words only. Please identify which key words to retain.	
Q4	AUTHOR: Please verify whether “defined under (1)” should be changed to “defined under Equation 1”.	
Q5	AUTHOR: Please define GND.	
Q6	AUTHOR: Please provide page number for reference 22.	
Q7	AUTHOR: "[34]" has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List.	

Please confirm that the funding sponsor list below was correctly extracted from your article: that it includes all funders and that the text has been matched to the correct FundRef Registry organization names. If a name was not found in the FundRef registry, it may not be the canonical name form, it may be a program name rather than an organization name, or it may be an organization not yet included in FundRef Registry. If you know of another name form or a parent organization name for a “not found” item on this list below, please share that information.

FundRef Name	**FundRef Organization Name**
Narodowe Centrum Nauki	Narodowe Centrum Nauki
Icelandic Centre for Research	Icelandic Centre for Research

SPi | Journal Code: J N M | Article ID: 2 5 8 4 | Dispatch: 20.02.1 | CE: Palang, Olivia | No. of Pages: 11 | ME:

Received: 9 October 2018 | Revised: 2 January 2019 | Accepted: 10 February 2019

DOI: 10.1002/jnm.2584

RESEARCH ARTICLE

WILEY

Enhanced uniform data sampling for constrained data-driven modeling of antenna input characteristics

Slawomir Koziel[1,2] | Ari T. Sigurðsson[1] | Anna Pietrenko-Dabrowska[2] | Stanislaw Szczepanski[2]

[1] Engineering Optimization & Modeling Center, Reykjavik University, Reykjavik, Iceland

[2] Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk, Poland

Correspondence
Slawomir Koziel, Engineering Optimization & Modeling Center, Reykjavik University, 103 Reykjavik, Iceland.
Email: koziel@ru.is

Funding information
Narodowe Centrum Nauki, Grant/Award Numbers: 2016/23/B/ST7/03733, 2011/03/B/ST7/03547 and 2015/17/B/ST6/01857; Icelandic Centre for Research, Grant/Award Number: 174114051

Abstract

Data-driven surrogates are the most popular replacement models utilized in many fields of engineering and science, including design of microwave and antenna structures. The primary practical issue is a curse of dimensionality, which limits the number of independent parameters that can be accounted for in the modeling process. Recently, a performance-driven modeling technique has been proposed where the constrained domain of the model is spanned by a set of reference designs optimized with respect to selected figures of interest. This approach allows for significant improvement of prediction power of the surrogates without the necessity of reducing the parameter ranges. Yet uniform allocation of the training data samples in the constrained domain remains a problem. Here, a novel design of experiments technique ensuring better sample uniformity is proposed. Our approach involves uniform sampling on the domain-spanning manifold and linear transformation of the remaining sample vector components onto orthogonal directions with respect to the manifold. Two antenna examples are provided to demonstrate the advantages of the technique, including application case studies (antenna optimization).

KEYWORDS

antenna design, constrained modeling, data-driven modeling, design of experiments, simulation-based design, surrogate modeling, uniform sampling

1 | INTRODUCTION

The most versatile and ubiquitous antenna design tools nowadays are full-wave electromagnetic (EM) simulators. EM analysis permits reliable performance evaluation when executed at sufficient discretization level of the structure. EM-driven design closure (primarily, adjustment of geometry parameters) is mandatory yet challenging stage of the design process. The primary problem is a high cost of simulation, which may be acceptable for simple designs but not so much for complex structures described by a large number of parameters. In particular, numerous evaluations required by, eg, conventional optimization algorithms,[1-5] may be impractical. The problem is even more pronounced for tasks involving massive simulations such as statistical analysis[6] or tolerance-aware design.[7,8] The most common work-around is an interactive design based on parameter sweeping; however, this approach has serious limitations: it fails to yield optimum designs, cannot handle design constraints, or cannot account for parameter interactions, to name just a few.

The difficulties outlined in the previous paragraph can be—to a certain extent—overcome by fast replacement models (surrogates). Surrogates can be roughly divided into approximation (or data-driven) ones[9] and physics-based ones.[10-13] The first group is far more popular because of a conceptual simplicity and wide accessibility of relevant computer codes. Data-driven models are obtained by approximating sampled simulation data. The most popular techniques include polynomial regression,[14] kriging,[15] neural networks,[16] Gaussian process regression,[17] or polynomial chaos expansion.[18] In the context of antenna modeling, where typical responses are highly nonlinear and tend to change rapidly as both a function of geometry parameters and the frequency, conventional approximation surrogates suffer from two serious (and related) limitations. These are (i) a low dimensionality of the design space and (ii) relatively narrow ranges of parameters in which the accurate model can be constructed using a reasonable number of training data points. The second issue is more important in practical applications because, for the model to be useful as an actual design aid, it has to be valid for wide variations of geometry parameters (eg, when redesigning a multi-band antenna for various operating frequencies[19]). Popular means for mitigating curse of dimensionality, eg, high-dimensional model representation (HDMR),[20] or principal component analysis (PCA),[21] are of little use for antenna modeling because the parameter redundancy and correlation are quite limited in typical designs.

The second group of models, physics-based surrogates, has certain advantages over the data-driven ones at the expense of reduced generality. A physics-based surrogate normally involves an auxiliary low-fidelity model (eg, equivalent network in case of microwave structures), which is appropriately corrected using sparsely sampled high-fidelity data. A few of the most popular approaches include space mapping[10] and various response correction methods.[22,23] It should be mentioned that because of a general lack of fast (and sufficiently reliable) low-fidelity representations of antenna structures, physics-based antenna modeling is of limited use.

In Koziel and Bekasiewicz,[24] a constrained modeling technique has been proposed (and further enhanced by Koziel[25]) that permits reduction of the computational cost of surrogate construction by appropriate definition of the model domain. The domain is defined using a set of reference designs, which are optimized for selected figures of interest (eg, operating frequencies) and/or material parameters (eg, substrate permittivity or thickness). The reference designs are subsequently triangulated, and the resulting manifold is extended in orthogonal directions to yield a set of dimensionality equal to that of the original parameter space. Although the domain obtained this way is "thin," it advantageously covers wide parameter ranges. At the same time, its volume is considerably (by order of magnitudes) smaller than the volume of the original space, which allows for a significant reduction of the number of training data points when constructing the model. A practical issue concerning the method of Koziel and Bekasiewicz[24] is a design of experiments (DoE). In particular, it is not possible to use established uniform sampling techniques such as Latin hypercube sampling (LHS),[9] or other space-filling DoE (eg, that of Santner et al[26] and Steponavice et al[27]) because geometry of the constrained domain is complex and not easily mapped onto a unit hypercube. The DoE approach assumed by Koziel[25] was quite rudimentary: iterative generation of random samples within the hypercube containing the model domain and accepting those that belong to the latter. Cleary, uniformity of the sample set obtained this way is poor and negatively affects the model predictive power.

The purpose of this paper is to propose a novel software-based DoE procedure that permits better uniformity of sample allocation in the constrained domain. This is a two-stage scheme that involves transformation of the uniform (LHS-based) data set generated in a unit hypercube into the surrogate domain. The first stage is a direct mapping of the first few components of the sample vector onto simplexes constituting the domain-defining manifold (cf Koziel[25]), whereas the second stage consists of a mapping of the remaining components into orthogonal directions using appropriately defined linear transformations. Our methodology is demonstrated using two antenna examples and shown to outperform the previously utilized DoE with respect to sample uniformity. This directly translates into an improved predictive power of the surrogate model. A comparison with conventional kriging modeling as well as application cases studies (antenna optimization) ~~is~~ also provided.

2 | PERFORMANCE-DRIVEN ANTENNA MODELING

Conventional data-driven modeling normally assumes that a model domain is a hypercube. In case of antennas, a vast majority of designs within a hypercube delimited by given lower or upper parameter bounds are poor. Yet standard uniform sampling techniques (eg, LHS,[9] orthogonal arrays,[28] and Hammersley sampling[29]) have been developed to handle such domains. "Good" designs (from the point of view of the considered operating conditions and/or material parameters) are typically allocated along certain manifolds being a result of particular parameter correlations and interactions.

Consequently, constructing the model in the entire space is a waste of resources. For practical antenna structures described by more than a few parameters, it may be simply infeasible because of an excessive computational expense related to training data acquisition.

The constrained modeling technique[25] has been developed to alleviate the difficulties highlighted in the previous paragraph. We denote by F_k, $k = 1, ..., N$, the performance figures or operating conditions that are of interest for a given antenna structure (eg, operating frequencies and substrate permittivity the antenna is to be implemented on). Constrained modeling restricts the surrogate domain to the most promising regions of the original design space, conventionally determined by the lower and upper bounds on antenna parameters.[25] For the sake of defining the constrained domain, a set of reference designs $\boldsymbol{x}^{(j)}$, $j = 1, ..., p$, is prepared by optimizing the antenna for selected values $\boldsymbol{F}^{(j)} = [F_1{}^{(j)}, ..., F_N{}^{(j)}]$. The vectors $\boldsymbol{F}^{(j)}$ should cover the ranges of the figures of interest that the surrogate is supposed to be valid for. The method of Koziel[25] permits arbitrary allocation of the set $\{\boldsymbol{F}^{(j)}\}_{j=1,...,p}$. The reference designs are subsequently structured by means of Delaunay triangulation.[30] It is used to create simplexes $S^{(k)} = \{\boldsymbol{x}^{(k.1)}, ..., \boldsymbol{x}^{(k.N+1)}\}$, $k = 1, ..., N_S$, with $\boldsymbol{x}^{(k.j)} \in \{\boldsymbol{x}^{(1)}, ..., \boldsymbol{x}^{(N)}\}$, $j = 1, ..., N + 1$, being vertices. The triangulation process is illustrated in Figure 1 for an exemplary case involving two figures of interest. **F1**

Having the reference designs, a manifold M is defined as a union of the convex hulls $h(S^{(k)})$ of the simplexes $S^{(k)}$

$$M = \bigcup_k \left\{ \mathbf{y} = \sum_{j=1}^{N+1} \alpha_j \mathbf{x}^{(k.j)} : 0 \le \alpha_j \le 1, \ \sum_{j=1}^{N+1} \alpha_j = 1 \right\}. \tag{1}$$

The surrogate model domain is defined by extending M into orthogonal directions (with respect to the spanning vectors of the simplexes $S^{(k)}$). The reason for this is the following. Although, by definition, M contains all of the reference designs (ie, optimum designs for all vectors $\boldsymbol{F}^{(j)}$), certain "thickness" is to be provided to ensure that the domain will contain the optimum designs corresponding to all combinations of figures of interest within the ranges of interest. The extension is arranged as described below. For a point $\boldsymbol{z}$, we consider the projection $P_k(\boldsymbol{z})$ onto the hyper-plane H_k containing the convex hull $h(S^{(k)})$. We use the following notation: $\boldsymbol{x}^{(0)} = \boldsymbol{x}^{(k.1)}$ will be referred to as the simplex "anchor," whereas $\boldsymbol{v}^{(j)} = \boldsymbol{x}^{(k.j+1)} - \boldsymbol{x}^{(0)}$, $j = 1, ..., N$, are the simplex spanning vectors. The expansion coefficients $\alpha^{(j)}$ defining the $P_k(\boldsymbol{z})$ are found by solving[25]

$$\arg \min_{\left[\alpha^{(1)},...,\alpha^{(N)}\right]} \left\| \mathbf{z} - \left[\mathbf{x}^{(0)} + \sum_{j=1}^{N} \alpha^{(j)} \mathbf{v}^{(j)} \right] \right\|^2. \tag{2}$$

Problem 2 can be solved analytically as

$$\left[\alpha^{(1)}, ..., \alpha^{(N)}\right]^T = \left(\mathbf{V}^T\mathbf{V}\right)^{-1}\mathbf{V}^T\left(\mathbf{z} - \mathbf{x}^{(0)}\right). \tag{3}$$

Note that $P_k(\boldsymbol{z}) \in h(S^{(k)})$ if and only if $\alpha^{(j)} \ge 0$ for $j = 1, ..., N$, and $\alpha^{(1)} + \cdots + \alpha^{(N)} \le 1$; ie, the projection is a convex combination of the spanning vectors. Let $\boldsymbol{x}_{\max} = \max\{\boldsymbol{x}^{(k)}, k = 1, ..., p\}$, $\boldsymbol{x}_{\min} = \min\{\boldsymbol{x}^{(k)}, k = 1, ..., p\}$, and the vector $\boldsymbol{dx} = \boldsymbol{x}_{\max} - \boldsymbol{x}_{\min}$ determines the range of variation of the antenna parameters within the manifold M.

Based on the above considerations, the surrogate model domain X_S can now be defined as an orthogonal extension of M with the "thickness" determined by a user-selected parameter $d_{\max}$. A point $\boldsymbol{y} \in X_S$ if and only if

1. set $K(\boldsymbol{y}) = \{k \in \{1, ..., N_S\}: P_k(\boldsymbol{y}) \in S^{(k)}\} \neq \varnothing$; and
2. $\min\{||\boldsymbol{y} - P_k(\boldsymbol{y})||/||\boldsymbol{dx}||: k \in K(\boldsymbol{y})\} \le d_{\max}$.

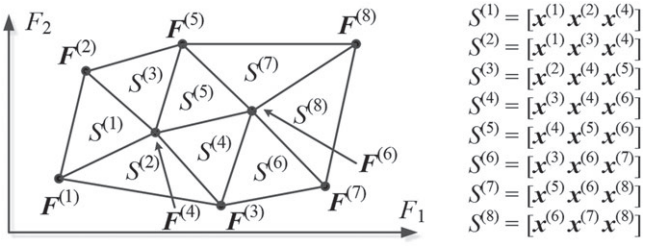


FIGURE 1 Conceptual illustration of reference designs and their triangulation in two-dimensional space of figures of interest

The first condition determines whether the point $\boldsymbol{y}$ is sufficiently close to M in the "tangential" sense, whereas the second condition assesses the "orthogonal" distance between $\boldsymbol{y}$ and the manifold.

There are two critical advantages of restricting the surrogate model domain to X_S with respect to the original space. First, the volume of X_S is considerably smaller than the volume of the interval $[\boldsymbol{x}_{\min}, \boldsymbol{x}_{\max}]$ (which is already a proper subset of a conventional domain). Consequently, the second significant advantage is secured, namely, the number of training data samples necessary to construct the surrogate can be greatly reduced. For highly dimensional spaces, the volume ratio can be many orders of magnitude. The flowchart of the modeling process is shown in Figure 2. F2

At the same time, the region of validity of the surrogate spans over the same ranges of parameters as in the original domain $[\boldsymbol{x}_{\min}, \boldsymbol{x}_{\max}]$. This means that the surrogate will cover the same ranges of figures of interest as the model defined in the original space. The difference is that X_S leaves out the designs that are poor with respect to these figures of interest. The surrogate itself is constructed using kriging interpolation.[9]

3 | UNIFORM SAMPLING IN CONSTRAINED DOMAIN

Given a relatively complex geometry of the constrained domain X_S defined in the previous section, the allocation of training data samples becomes a challenge. A rudimentary approach utilized by Koziel[25] was pure random sampling within $[\boldsymbol{x}_{\min}, \boldsymbol{x}_{\max}]$ with accepting samples that are in X_S. The process was repeated until a required number of samples have been found. Clearly, a uniformity of the data set obtained this way was poor. In Koziel et al,[31] a considerably better, two-stage approach has been proposed. In the first stage, the required number of points is assigned to the simplexes $S^{(k)}$ proportionally to their volumes. Then, in the second stage, the points are allocated on the simplexes using LHS (mapped from unit hypercubes of appropriate dimensions) and relocated in directions orthogonal to the respective simplexes. Both the relocation ranges and specific directions are determined randomly. This scheme improves sample set uniformity considerably leading to improvement of surrogate model predictive power.[31]

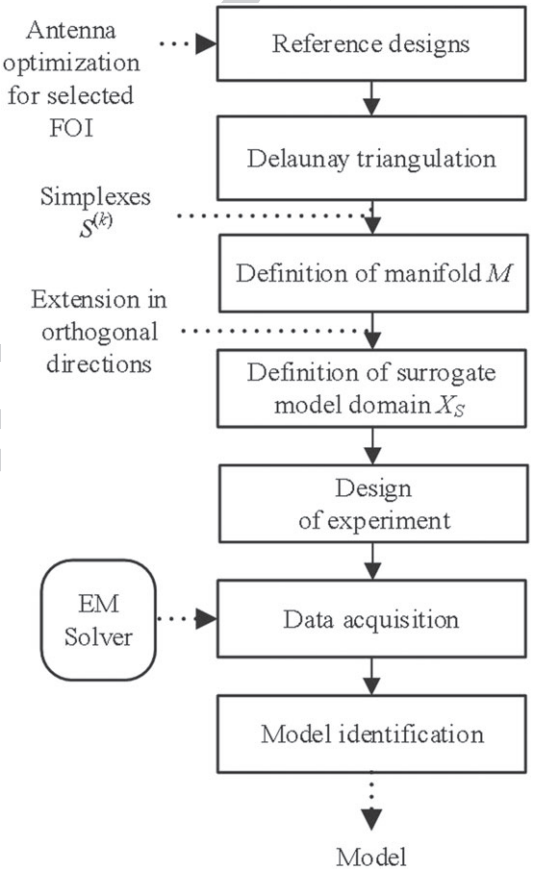


FIGURE 2 Flowchart of the constrained modeling procedure of Section 2. EM indicates electromagnetic; FOI, figure of interest

In this work, an alternative sampling scheme is proposed, supposedly leading to a further improvement of the data set uniformity over the method of Koziel et al.[31] The principal concept is to establish the mapping between LHS-sampled unit hypercube and orthogonally expanded simplex $S^{(k)}$. Detailed explanation of the procedure has been given in the remaining part of this section.

The main challenge regarding the mapping between the n-dimensional hypercube and the orthogonally extended N-simplex is that n-N coordinates of the hypercube-sampled points have to be transformed into all directions that are orthogonal to the simplex. We use the following notation. Let $\boldsymbol{e}^{(j)}$, $j = 1, ..., n$, denote the standard basis in R^n, and let $\boldsymbol{e}^{(j)\perp}$ denote the components of $\boldsymbol{e}^{(j)}$ that are orthogonal to the simplex $S^{(k)}$; ie, $\boldsymbol{e}^{(j)\perp} = \boldsymbol{e}^{(j)} - P_k(\boldsymbol{e}^{(j)})$ (as before, P_k denotes orthogonal projection onto $S^{(k)}$). Let $\boldsymbol{w}^{(j)}$, $j = 1, ..., n - N$, be the subset of the n-N (norm-wise) largest vectors $\boldsymbol{e}^{(j)\perp}$. The remaining N vectors $\boldsymbol{e}^{(j)\perp}$ will be referred to as dependent ones and denoted as $\boldsymbol{w}_d{}^{(k)}$, $k = 1, ..., N$. We consider a matrix $\boldsymbol{B}$ defined as a least-square solution of the problem

$$\mathbf{WB} = \left[\mathbf{w}^{(1)} \ \cdots \ \mathbf{w}^{(N)}\right]\mathbf{B} = \left[\mathbf{w}_{d^{(1)}} \ \cdots \ \mathbf{w}_{d^{(N)}}\right] = \mathbf{W}_d\mathbf{B}. \tag{4}$$

Problem 4 can be solved analytically as

$$\boldsymbol{B} = \left(\boldsymbol{W}^T\boldsymbol{W}\right)^{-1}\boldsymbol{W}\boldsymbol{W}_d. \tag{5}$$

Matrix $\boldsymbol{B}$ contains expansion coefficients of the dependent vectors with respect to the independent basis. Because for practical problems, we have $N < n/2$, wherein the representation is unique (the matrix $\boldsymbol{W}^T\boldsymbol{W}$ is nonsingular). The reason behind selecting the independent and dependent vectors is that for n-component sample vectors, the first N components will be reserved for allocating the samples on the simplexes $S^{(k)}$ and subsequent n-N orthogonal directions are necessary to account for the remaining n-N components. Although there are many different ways of selecting the independent vectors, the most natural choice (adopted here) is to have them as the largest orthogonal directions. The remaining ones are accounted for through their representation in the independent basis.

Let $\boldsymbol{d}$ be the $n \times 1$ perturbation vector, which is a user-defined parameter, in practice, $\boldsymbol{d} = \boldsymbol{dx}{\cdot}d_{\max}$. For $\boldsymbol{d}$, we repeat the above procedure of selecting the independent components $\boldsymbol{d}_{\text{ind}} = [d^{(1)}, ..., d^{(n-N)}]^T$ (associated with the independent vectors) and the dependent ones $\boldsymbol{d}_{\text{dep}} = [d_d{}^{(1)}, ..., d_d{}^{(N)}]^T$ (associated with the dependent vectors). Having $\boldsymbol{d}_{\text{ind}}$ and $\boldsymbol{d}_{\text{dep}}$, the overall perturbation vector $\boldsymbol{D}$ is calculated as

$$\boldsymbol{D} = [D_1, ..., D_{n-N}]^T = \boldsymbol{d}_{\text{ind}} + |\boldsymbol{B}|\boldsymbol{d}_{\text{dep}}. \tag{6}$$

In Equation 6, $|\boldsymbol{B}|$ denotes a matrix of the absolute values of $\boldsymbol{B}$. It should be noted that the dependent vectors will contribute to the sample perturbation by increasing the perturbations towards independent directions proportionally to the expansion coefficients of the latter.

At this point, we are ready to formulate the proposed uniform sampling procedure, which works as follows:

1. Calculate the volumes V_k of simplexes $S^{(k)}$.
2. Set $K_k = \lceil KV_k/\sum_j V_j \rceil$ (here, $\lceil\ \rceil$ is a ceiling function).
3. For each $k = 1, ..., N_S$,
   - allocate mK_k LHS[6] samples in a unit hypercube $[0, 1]^n$ (here, m is the volume ratio between the unit hypercube and the unit simplex, both of the dimension N);
   - choose samples $\boldsymbol{y}$ for which the vectors that consist of the first N components of $\boldsymbol{y}$ are allocated in the unit N-simplex;
   - map the samples selected in the previous step onto a convex hull of simplex $S^{(k)}$ as follows: $\boldsymbol{x} \rightarrow \boldsymbol{x}^{(0)} + \sum_{j=1,...,N} x_j\boldsymbol{v}^{(j)}$, where $\boldsymbol{x}^{(0)}$ and $\boldsymbol{v}^{(j)}$ are defined under (1), and $\boldsymbol{x} = [x_1, ..., x_n]^T$ (note that the first N components of $\boldsymbol{x}$ are used); and Q4
   - perturb the mapped samples by adding vectors $\boldsymbol{x}_d = \Sigma_{j=1,...,n-N}\, x_{N+j}D_j\boldsymbol{w}^{(j)}$.

The sampling procedure described above allows for mapping the LHS-allocated samples from a unit hypercube (of a dimension corresponding to that of the antenna parameter space) onto individual segments of the constrained domain. Figure 3 shows the flowchart of the sampling procedure. F3

Visualization of the three DoE methods (random sampling of Koziel,[25] improved sampling of Koziel et al,[31] and the procedure proposed here) is provided in Figure 4. For the sake of clarity, only one domain segment is shown. It can be F4

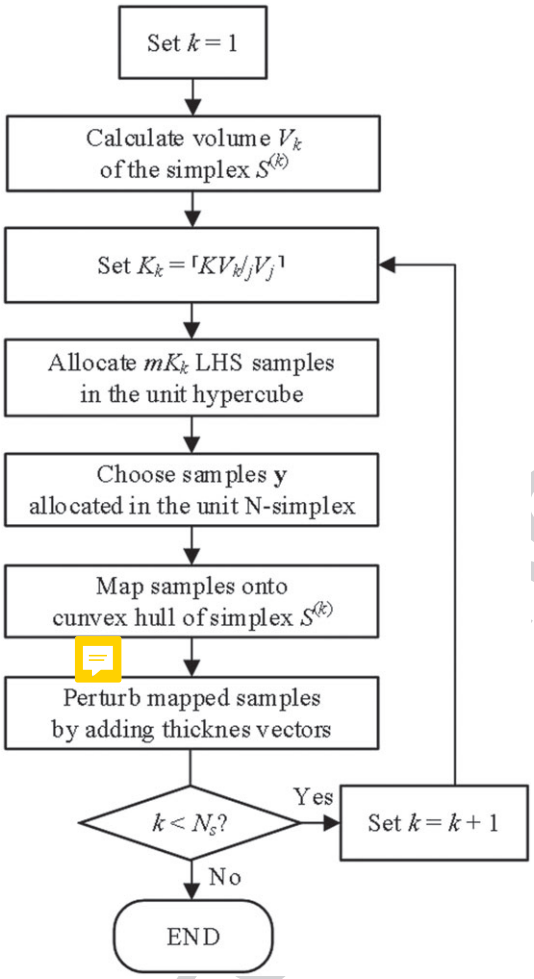


FIGURE 3 Flowchart of the proposed uniform sampling procedure. LHS indicates Latin hypercube sampling

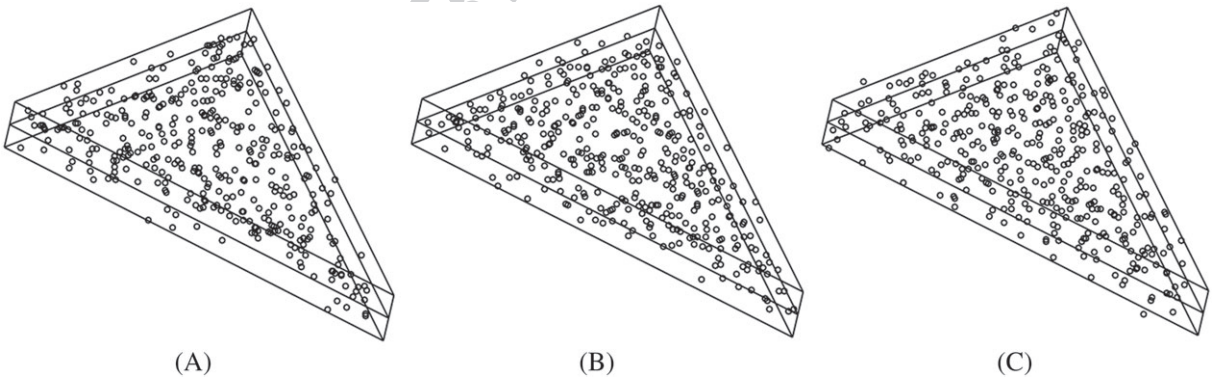


FIGURE 4 Example distribution of 500 samples for one simplex in three-dimensional parameter space: A, random sampling[25]; B, improved sampling[31]; C, uniform sampling proposed in this work

observed that sample uniformity is considerably improved for the method of Koziel et al[31] as compared with that for the random sampling.[25] The scheme proposed here provides further improvement. As demonstrated in Section 4, this will translate into a better accuracy of the surrogate model. Obviously, an additional improvement is expected to be mild given that relatively uniform sample allocation already provided by the method of Koziel et al.[31]

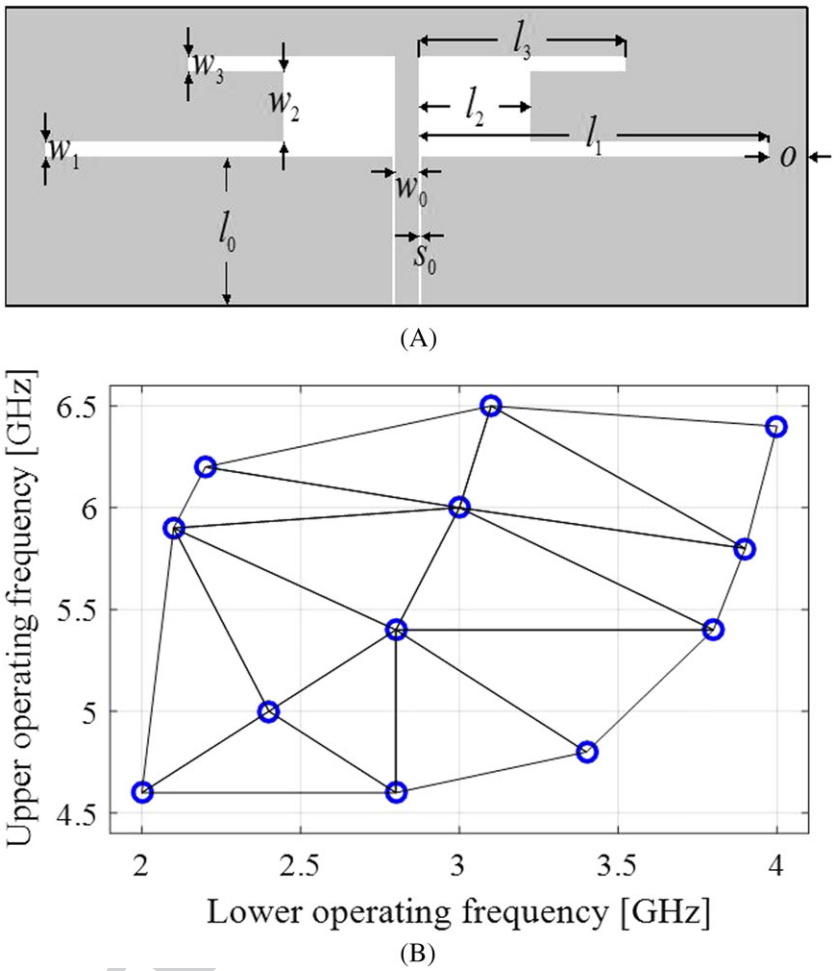


FIGURE 5 Dual-band uniplanar dipole antenna: A, geometry[32]; B, allocation of reference designs

TABLE 1 Surrogate model accuracy for antenna of Figure 5

Number of Training Samples	Average RMS Error			
	Unconstrained Surrogate, %	Constrained Surrogate		
		Random Sampling[25], %	Improved Sampling[31], %	Uniform Sampling (This Work), %
100	15.6	7.7	4.8	4.3
200	11.7	4.6	3.7	3.4
500	7.8	3.2	2.5	2.2

Abbreviation: RMS, root mean square.

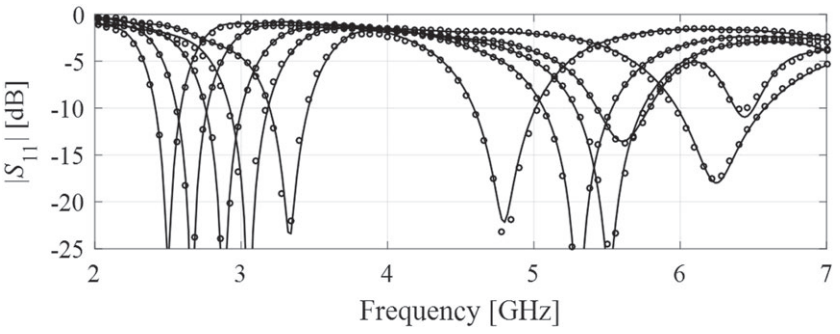


FIGURE 6 Responses of the antenna of Figure 5 at the selected test designs for $N = 500$: high-fidelity electromagnetic model (—) and surrogate with uniform sampling (o)

4 | CASE STUDIES AND RESULTS

For the sake of validation and benchmarking, the constrained surrogate modeling technique is applied here for two antenna structures, using the three DoE schemes of Koziel[25,31] and the one proposed here.

The first example is a uniplanar dual-band dipole antenna as shown in Figure 5A.[32] The structure is implemented on Taconic RF-35 substrate ($\varepsilon_r = 3.5$, $\tan\delta = 0.0018$, $h = 0.762$ mm). The design variables are $\boldsymbol{x} = [l_1\ l_2\ l_3\ w_1\ w_2\ w_3]^T$; other

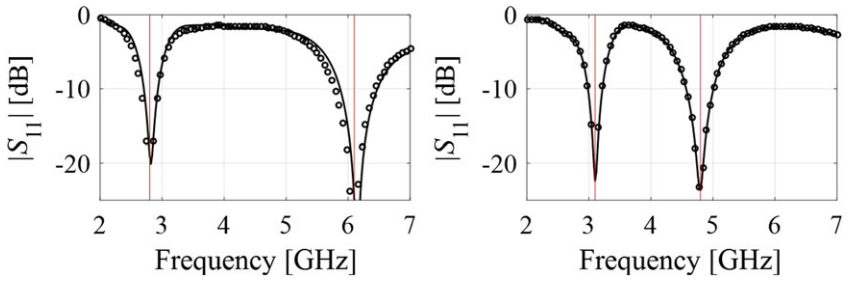
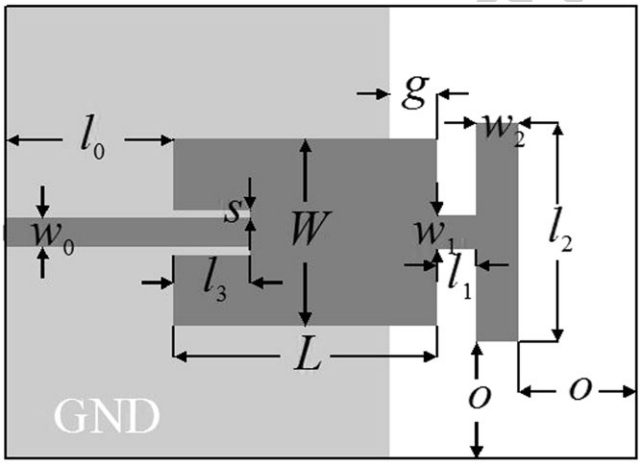
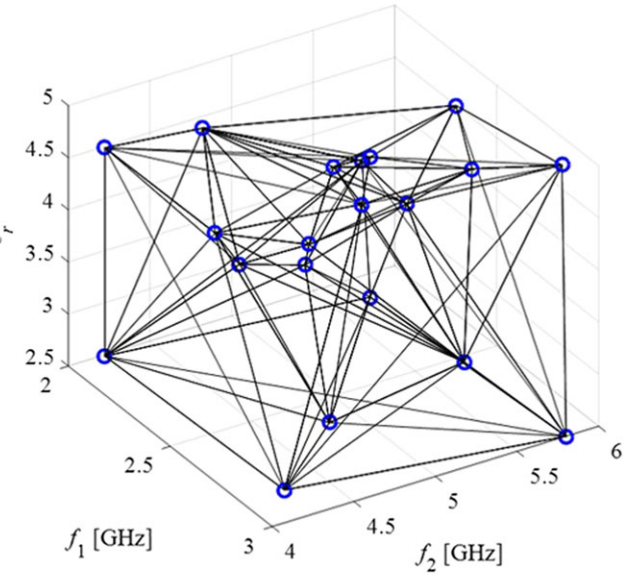


FIGURE 7 Surrogate (o) and electromagnetic model (—) responses at the two verification designs corresponding to $f_1 = 2.8$ GHz and $f_2 = 6.1$ GHz (left) and $f_1 = 3.1$ GHz and $f_2 = 4.8$ GHz (right). Vertical lines indicate the required operating frequencies



(A)



(B)

FIGURE 8 Geometry of the dual-band patch antenna[33]: A, geometry[33]; B, allocation of reference designs

parameters are fixed: $l_0 = 30$, $w_0 = 3$, $s_0 = 0.15$, and $o = 5$ (dimensions in mm). The computational model $\boldsymbol{R}$ (approximately 100 000 cells; simulation time 60 s) is implemented in CST. The goal of the modeling process is to construct the surrogate for the following ranges of operating frequencies 2.0 GHz $\leq f_1 \leq$ 4.0 GHz (lower band) and 4.5 GHz $\leq f_2 \leq$ 6.5 GHz (upper band). There are 12 reference designs selected and allocated as shown in Figure 5B. For the sake of computational efficiency, the designs have been obtained using a feature-based optimization algorithm.[2]

The numerical results have been gathered in Table 1, which shows the comparison of the (conventional) surrogate obtained in an unconstrained domain, as well as the constrained domain using random sampling,[25] improved sampling,[31] and uniform sampling proposed in this work. In order to investigate scalability of the models, the benchmarking has been executed for various numbers of training data samples (100, 200, and 500). As expected, training data allocation has a profound effect on the modeling error, especially when comparing the random sampling with the improved DoE of Koziel et al.[31] T1

The sampling scheme introduced here leads to further improvements of the predictive power, which is not as significant but noticeable. The surrogate and EM-simulated high-fidelity model response at the selected test designs are shown in Figure 6. F6

In order to demonstrate usefulness of the model for design purposes, the antenna has been optimized for the selected operating frequencies, $f_1 = 2.8$ GHz and $f_2 = 6.1$ GHz (Case I) and $f_1 = 3.1$ GHz and $f_2 = 4.8$ GHz (Case II). The results shown in Figure 7 indicate not only good quality responses but also excellent agreement between the surrogate and EM simulations of the antenna. F7

The second verification example is a dual-band planar antenna.[33] The antenna geometry is shown in Figure 8A. The structure is implemented on a 0.762-mm-thick substrate. The design variables are $\boldsymbol{x} = [L\ l_1\ l_2\ l_3\ W\ w_1\ w_2\ g]^T$. The parameters $o = 7$, $l_0 = 10$, and $s = 0.5$ are fixed; the feed line width w_0 is adjusted for a given substrate permittivity to ensure 50-Ω impedance. The EM model is implemented in CST (approximately 400 000 mesh cells, simulation time 3 min). For this example, we aim at constructing the surrogate for the following ranges of operating frequencies: 2.0 GHz $\leq f_1 \leq$ 3.0 GHz (lower band) and 4.0 GHz $\leq f_2 \leq$ 6.0 GHz (upper band). Furthermore, the model should be valid for a range of substrate permittivities: $2.5 \leq \varepsilon_r \leq 5.0$. Here, 20 reference designs have been selected and allocated as indicated (along with their triangulation) in Figure 8B. F8

Table 2 shows the model predictive power as well as results of benchmarking. It can be observed that the proposed uniform sampling leads to further improvement of the modeling accuracy (both over the random sampling and the improved scheme of Koziel et al[31]). Visualization of antenna characteristics at the selected designs is shown in Figure 9. T2 F9

TABLE 2 Surrogate model accuracy for antenna of Figure 8

Number of Training Samples	Average RMS Error			
	Unconstrained Surrogate, %	Constrained Surrogate		
		Random Sampling[25], %	Improved Sampling[31], %	Uniform Sampling (This Work), %
100	50.5	19.1	16.7	14.3
200	49.1	16.5	14.8	13.1
500	48.5	13.8	13.1	12.3

Abbreviation: RMS, root mean square.

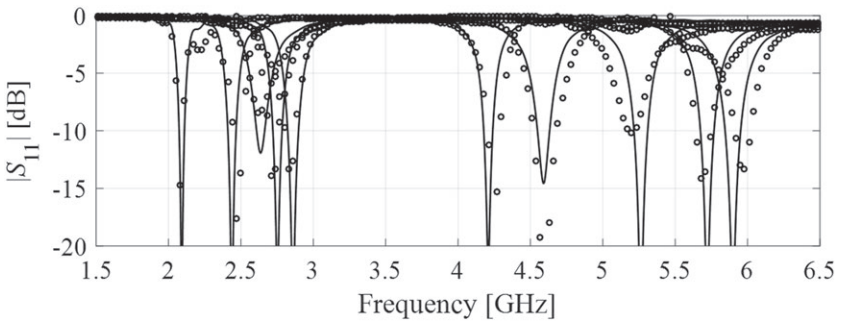


FIGURE 9 Responses of the antenna of Figure 8 at the selected test designs for $N = 500$: high-fidelity electromagnetic model (—) and surrogate with uniform sampling (o)

5 | CONCLUSION

A novel DoE approach for uniform data sampling in constrained domains has been proposed. Our methodology involves a mapping between a unit hypercube and individual domain segments, the latter being simplexes expanded in orthogonal directions. The scheme allows better sample uniformity than provided by a rudimentary random sampling as well as a recently reported enhanced procedure. As a result, improved predictive power of the surrogate model is observed as demonstrated using two examples of dual-band antenna structures. The proposed technique has been rigorously formulated, comprehensively validated, and benchmarked.

ACKNOWLEDGEMENTS

The authors would like to thank Dassault Systemes, France, for making CST Microwave Studio available. This work is partially supported by the Icelandic Centre for Research (RANNIS) grant 174114051 and by the National Science Centre of Poland (Narodowe Centrum Nauki) grants 2015/17/B/ST6/01857, 2011/03/B/ST7/03547, and 2016/23/B/ST7/03733.

ORCID

Slawomir Koziel https://orcid.org/0000-0002-9063-2647
Anna Pietrenko-Dabrowska https://orcid.org/0000-0003-2319-6782

REFERENCES

1. Nocedal J, Wright S. *Numerical Optimization*. 2nd ed. New York: Springer; 2006.
2. Fernandez Pantoja M, Rubio Bretones A, Gomez Martin R. Benchmark antenna problems for evolutionary optimization algorithms. *IEEE Trans Antennas Propag*. 2007;55(4):1111-1121.
3. Lalbakhsh A, Afzal MU, Esselle KP. Multiobjective particle swarm optimization to design a time-delay equalizer metasurface for an electromagnetic band-gap resonator antenna. *IEEE Antennad Wirel Propag Lett*. 2017;16:915-915.
4. Darvish A, Ebrahimzadeh A. Improved fruit-fly optimization algorithm and its applications in antenna array synthesis. *IEEE Trans Antennas Propag*. 2018;66(4):1756-1766.
5. Wang J, Yang XS, Wang BZ. Efficient gradient-based optimization of pixel antenna with large-scale connections. *IET Microwaves Antennas Propag*. 2018;12(3):385-389.
6. Ochoa JS, Cangellaris AC. Random-space dimensionality reduction for expedient yield estimation of passive microwave structures. *IEEE Trans Microwave Theory Tech*. 2013;61(12):4313-4321.
7. Kouassi A, Nguyen-Trong N, Kaufmann T, Lalléchère S, Bonnet P, Fumeaux C. Reliability-aware optimization of a wideband antenna. *IEEE Trans Antennas Propag*. 2016;64(2):450-460.
8. Abdel-Malek HL, Hassan ASO, Soliman EA, Dakroury SA. The ellipsoidal technique for design centering of microwave circuits exploiting space-mapping interpolating surrogates. *IEEE Trans Microwave Theory Tech*. 2006;54(10):3731-3738.
9. Simpson TW, Pelplinski JD, Koch PN, Allen JK. Metamodels for computer-based engineering design: survey and recommendations. *Eng Comput*. 2001;17(2):129-150.
10. Bandler JW, Georgieva N, Ismail MA, Rayas-Sánchez JE, Zhang QJ. A generalized space mapping tableau approach to device modeling. *IEEE Trans Microwave Theory Tech*. 2001;49(1):67-79.
11. Baratta IA, de Andrade CB, de Assis RR, Silva EJ. Infinitesimal dipole model using space mapping optimization for antenna placement. *IEEE Antennas Wirel Propag Lett*. 2018;17(1):17-20.
12. Xu J, Li M, Chen R. Space mapping optimisation of 2D array elements arrangement to reduce the radar cross-scattering. *IET Microwaves Antennas Propag*. 2017;11(11):1578-1582.
13. Su Y, Lin J, Fan Z, Chen R. Shaping optimization of double reflector antenna based on manifold mapping, *Int Applied Computational Electromagnetic Society Symp. (ACES)*, pp. 1-2, 2017.
14. Chavez-Hurtado JL, Rayas-Sanchez JE. Polynomial-based surrogate modeling of RF and microwave circuits in frequency domain exploiting the multinomial theorem. *IEEE Trans Microwave Theory Tech*. 2016;64(12):4371-4381.
15. Couckuyt I. Forward and inverse surrogate modeling of computationally expensive problems, Ph.D. Thesis, Ghent University, 2013.
16. Kabir H, Wang Y, Yu M, Zhang QJ. Neural network inverse modeling and applications to microwave filter design. *IEEE Trans Microwave Theory Tech*. 2008;56(4):867-879.
17. Jacobs JP, Koziel S. Reduced-cost microwave filter modeling using a two-stage Gaussian process regression approach. *Int J RF Microwave Comput Aided Eng*. 2014;25(5):453-462.

18. Zhang J, Zhang C, Feng F, Zhang W, Ma J, Zhang QJ. Polynomial chaos-based approach to yield-driven EM optimization. *IEEE Trans Microwave Theory Tech*. 2018;66(7):3186-3199.
19. Koziel S, Bekasiewicz A. Rapid dimension scaling of triple-band antennas by means of inverse surrogate modeling, *IEEE Antennas Prop Symp*, 2017.
20. Taskin G, Kaya H, Bruzzone L. Feature selection based on high dimensional model representation for hyperspectral images. *IEEE Trans Image Process*. 2017;26(6):2918-2928.
21. Ma S, Aybat NS. Efficient optimization algorithms for robust principal component analysis and its variants. *Proc IEEE*. 2018;106(8):1411-1426.
22. Koziel S, Bekasiewicz A. Computationally feasible narrow-band antenna modeling using response features. *Int J RF Microwave Comput Aided Eng*. 2017;27(4). Q6
23. Echeverria D, Lahaye D, Encica L, Lomonova EA, Hemker PW, Vandenput AJA. Manifold-mapping optimization applied to linear actuator design. *IEEE Trans Magn*. 2006;42(4):1183-1186.
24. Koziel S, Bekasiewicz A. On reduced-cost design-oriented constrained surrogate modeling of antenna structures. *IEEE Antennas Wirel Propag Lett*. 2017;16:1618-1621.
25. Koziel S. Low-cost data-driven surrogate modeling of antenna structures by constrained sampling. *IEEE Antennas Wirel Propag Lett*. 2017;16:461-464.
26. Santner TJ, Williams BJ, Notz WI. Space-filling designs for computer experiments. In: *The Design and Analysis of Computer Experiments*. *Springer Series in Statistics*. New York: Springer; 2003:121-161.
27. Steponavice I, Shirazi-Manesh M, Hyndman RJ, Smith-Miles K, Villanova L. On sampling methods for costly multi-objective black-box optimization. In: *Advances in Stochastic and Deterministic Global Optimization*. New York: Springer; 2016:273-296.
28. Leary S, Bhaskar A, Keane A. Optimal orthogonal-array-based Latin hypercubes. *J Appl Stat*. 2003;30(5):585-598.
29. Giunta AA, Wojtkiewicz SF, Eldred MS. Overview of modern design of experiments methods for computational simulations, *American Institute of Aeronautics and Astronautics*, Paper AIAA 2003-0649, 2003.
30. Borouchaki H, George PL, Lo SH. Optimal Delaunay point insertion. *Int J Numer Methods Eng*. 1996;39(20):3407-3437.
31. Koziel S, Sigurdsson AT, Szczepanski S. Uniform sampling in constrained domains for low-cost surrogate modeling of antenna input characteristics. *IEEE Antennas Wirel Propag Lett*. 2018;17(1):164-167.
32. Chen Y-C, Chen S-Y, Hsu P. Dual-band slot dipole antenna fed by a coplanar waveguide, *IEEE Int Symp Ant Prop*., pp. 3589-3592, 2006.
33. Koziel S, Bekasiewicz A, Leifsson L. Rapid EM-driven antenna dimension scaling through inverse modeling. *IEEE Antennas Wirel Propag Lett*. 2016;15:714-717.
34. Koziel S. Fast simulation-driven antenna design using response-feature surrogates. *Int J RF Microwave Comput Aided Eng*. 2015;25(5):394-402. Q7

How to cite this article: Koziel S, Sigurðsson AT, Pietrenko-Dabrowska A, Szczepanski S. Enhanced uniform data sampling for constrained data-driven modeling of antenna input characteristics. *Int J Numer Model*. 2019; e2584. https://doi.org/10.1002/jnm.2584