International Journal for Uncertainty Quantification, 10(2):165–193 (2020)

ENHANCED ADAPTIVE SURROGATE MODELS WITH APPLICATIONS IN UNCERTAINTY QUANTIFICATION FOR NANOPLASMONICS

***Niklas Georg,*[1,2,3,*] *Dimitrios Loukrezis,*[2,3] *Ulrich Römer,*[1] *& Sebastian Schöps*[2,3]**

[1]*Institut für Dynamik und Schwingungen, Technische Universität Braunschweig, Braunschweig, Germany*

[2]*Centre for Computational Engineering, Technische Universität Darmstadt, Darmstadt, Germany*

[3]*Institut für Teilchenbeschleunigung und Elektromagnetische Felder (TEMF), Technische Universität Darmstadt, Darmstadt, Germany*

*Address all correspondence to: Niklas Georg, Institut für Dynamik und Schwingungen, Technische Universität Braunschweig, Schleinitzstraße 20, D-38106 Braunschweig, Germany, E-mail: n.georg@tu-braunschweig.de

Original Manuscript Submitted: 8/5/2019; Final Draft Received: 4/5/2020

We propose an efficient surrogate modeling technique for uncertainty quantification. The method is based on a well-known dimension-adaptive collocation scheme. We improve the scheme by enhancing sparse polynomial surrogates with conformal maps and adjoint error correction. The methodology is applied to Maxwell's source problem with random input data. This setting comprises many applications of current interest from computational nanoplasmonics, such as grating couplers or optical waveguides. Using a nontrivial benchmark model, we show the benefits and drawbacks of using enhanced surrogate models through various numerical studies. The proposed strategy allows us to conduct a thorough uncertainty analysis, taking into account a moderately large number of random parameters.

KEY WORDS: *adaptivity, adjoint error indicator, conformal maps, hierarchical interpolation, stochastic sparse grid collocation, Maxwell's source problem, plasmonics*

1. INTRODUCTION

The numerical solution of partial differential equations (PDEs) with random input data has been receiving considerable attention in the last decades in the context of uncertainty quantification (UQ). Numerical UQ methods are continuously improved to address large-scale problems with many input parameters, which still pose a computational challenge nowadays. The key property to reduce computational costs in high dimensions is a holomorphic dependency of the PDE solution on the input parameters. Such holomorphy results have been established for a variety of different problem classes [1,2] and allow the use of spectral stochastic methods [3–8] in combination with adaptive schemes, see, e.g., [9–13] for the case of stochastic collocation.

An alternative, sometimes complementary, approach for the numerical solution of parametric problems is model order reduction, see [14] and the references therein. Model order reduction based on moment-matching [15,16] can be used to derive a rational parametric approximation, which is appealing in the case of reduced parametric regularity. Rational Padé-type approximations have recently been employed for a stochastic Helmholtz problem [17]. Moreover, in [18], a Padé-Legendre method was introduced to cope with discontinuous response surfaces, where it was also noted that high-dimensional settings are still difficult to address. Another alternative are (multilevel) Monte Carlo methods, which can handle high-dimensional parameter spaces and more general settings with reduced smoothness.

2152–5080/20/$35.00 © 2020 by Begell House, Inc. www.begellhouse.com

A recent analysis of such an approach for a Helmholtz transmission problem, with point evaluations as quantitites of interest, was presented in [19]. The Helmholtz equation is recovered when two-dimensional versions of our model problem are considered. However, we quantify uncertainties in scattering parameters, which are more regular. The last class of methods that are mentioned here are perturbation methods [20,21]. A perturbation approach can lead to very efficient numerical methods, but is also not considered here, because the uncertainty in the input parameters of our models can be quite large.

The present study was motivated by the fact that the computational cost of constructing a sparse surrogate model can still be quite high for various applications. In particular, even if sufficient smoothness is present to allow for sparse approximation, achieving a reasonable error level in practical applications may require collocation grids with many points. Hence, in this work, we improve a state-of-the-art adaptive stochastic collocation method, based on dimension-adaptivity [10] and weighted Leja interpolation [11]. To this end, we combine conformal maps and adjoint error estimation and correction. Conformal maps have been put forth in [22,23] for the acceleration of interpolation and quadrature methods, but have not received much attention in the UQ context thus far. Adjoint error correction in turn was considered in [24,25] in the context of Clenshaw Curtis collocation and the stochastic Galerkin method. The combination of both methods in the context of uncertainty quantification has not been considered, to the best of our knowledge. The resulting collocation scheme is able to address a moderately high number of random model parameters. Moreover, weighted Leja nodes can handle almost arbitrary input probability distributions [11,26–30] and are ideally suited for adaptivity [9]. In order to efficiently steer the adaptivity, we derive an adjoint representation of the stochastic error. On the basis of this error formula, the convergence order is enhanced through extrapolation. Finally, conformal maps offer the potential to further enhance the numerical accuracy by suitably transforming the required region of holomorphy.

We consider Maxwell's source problem as a model class that is relevant for a wide variety of applications. This model problem is particularly important in computational nanoplasmonics. Plasmonic structures offer great potential for subwavelength optics and optoelectronics [31] and have been intensively studied from both a fundamental and an application point of view in recent years. With the aforementioned UQ methods, studying stochastic parameter variations within the numerical simulation of plasmonic structures comes into reach. This is highly relevant, as relatively large variabilities can be observed, see, e.g., [32]. Although, not considered in the present work, the inverse UQ problem is also of high relevance, due to the intrinsic difficulty in measuring material dispersion properties. Instead, we focus on the propagation of uncertainties from the model inputs to the outputs. In particular, the proposed framework allows one to compute moments, probability distributions, failure probabilities, and global sensitivities for physical quantities of interest (QoIs). In similar physical settings, UQ studies have been conducted in recent works [33–35], which however employ less advanced numerical methods. In [36], UQ for a silicon photonic device has been addressed, considering a low-dimensional correlated random input parameter vector. Additionally, in comparison to recent theoretical studies [2,37], we employ additional techniques for convergence acceleration and consider a more complex numerical example.

The rest of this paper is structured as follows. In Section 2, we describe the numerical method for enhanced surrogate modeling. In Section 3, we introduce Maxwell's source problem, its finite element discretization, and its parametrization. In Section 4, the developed method is used to conduct a UQ study for a nontrivial nanoplasmonics application, namely, an optical grating coupler. In Section 5, we give some concluding remarks.

2. ENHANCED SURROGATE MODELING

In this section, we consider the general parametric problem of finding

$$\mathbf{u}(\mathbf{y}) \in V \text{ s.t.}, \quad a_{\mathbf{y}}(\mathbf{u}(\mathbf{y}), \mathbf{v}) = l_{\mathbf{y}}(\mathbf{v}), \quad \forall \mathbf{v} \in V, \tag{1}$$

where V denotes a suitable Hilbert space and $\mathbf{y} \in \Xi \subset \mathbb{R}^N$ denotes the input parameter vector. Problem (1) may represent the model of Section 3.2 or other parametrized differential equations with a continuous sesquilinear form $a_{\mathbf{y}}(\cdot,\cdot)$ and a continuous (anti)linear form $l_{\mathbf{y}}(\cdot)$. Note that boldface letters are used to indicate matrices and vectors. Since the solutions governed by Maxwell's equations are typically vector-valued, this convention is also used for $\mathbf{u}$ and $\mathbf{v}$ in Eq. (1). We assume the map $\mathbf{u} : \Xi \to V$ to be well-defined and smooth, which is often the case for

parametrized differential equations, see, e.g., [38] for elliptic problems and [1] for other types of PDEs. We are interested in the model's response, which may be the solution $\mathbf{u}(\mathbf{y})$ itself or a bounded linear functional $J_{\mathbf{y}}(\mathbf{u}(\mathbf{y}))$, commonly referred to as the QoI. In this work, we focus on single-valued and complex QoIs, i.e., $J_{\mathbf{y}}(\mathbf{u}(\mathbf{y})) \in \mathbb{C}$. For brevity of notation and due to the well-posedness of the system, we shall replace $J_{\mathbf{y}}(\mathbf{u}(\mathbf{y}))$ with $J(\mathbf{y})$, where J can be understood as an abstract representation of the map from the input parameters to the QoI.

We now assume that the input parameters are given as independent random variables (RVs) Y_n, $n = 1, 2, \ldots, N$. We introduce the random vector $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_N)^\top$, defined on the probability space (Θ, Σ, P), where Θ denotes the sample space, Σ the sigma-algebra of events, and P the probability measure, its image set $\Xi = \Xi_1 \times \Xi_2 \cdots \times \Xi_N \subset \mathbb{R}^N$ and its probability density function (PDF) $\varrho(\mathbf{y}) = \prod_{n=1}^{N} \varrho_n(y_n)$, such that $\mathbf{Y} : \Theta \to \Xi$ and $\varrho : \Xi \to \mathbb{R}_+$. Then, the parameter vector represents a realization of the random vector, i.e., $\mathbf{y} = \mathbf{Y}(\theta) \in \Xi$, $\theta \in \Theta$. Assuming independence is necessary for the tensor-product constructions in the collocation method, however, dependence could also be taken into account through a suitable transformation, for instance, Rosenblatt or Nataf transformations [39,40]. In view of this transformation, we assume in this section that the image set Ξ is given as the hypercube $[-1, 1]^N$, for simplicity.

Now, the QoI is itself a RV and we are interested in quantifying uncertainty, e.g., by computing its moments, PDF, quantiles, etc. In the case where the QoI is smooth (ideally, analytic) with respect to the input RVs, spectral UQ methods [41,42] may be employed. Then, J is particularly well-suited to be approximated by polynomials such that

$$J(\mathbf{y}) \approx \widetilde{J}(\mathbf{y}) = \sum_{m=0}^{M} s_m \Psi_m(\mathbf{y}), \tag{2}$$

where $\Psi_m : \Xi \to \mathbb{R}$ are multivariate polynomials and $s_m \in \mathbb{C}$ the associated coefficients, and fast convergence can be expected. Once an approximation in the form of Eq. (2) is available, it can be used as an inexpensive substitute of the original computational model for sampling-based computations. Alternatively, some statistical information regarding the QoI can be derived directly from the coefficients. In the context of the present work, we will use approximations in the form of Eq. (2) based on sparse grid interpolation [6,8,9,11,43–47], which can be combined with a conformal mapping.

2.1 Univariate Interpolation and Conformal Maps

We first discuss univariate interpolation in some detail, since it is also the key building block for the tensor product constructions used in the multivariate case. In particular, we consider the univariate function $f : [-1, 1] \to \mathbb{C}$

$$f(y) := J(y, 0, \ldots, 0). \tag{3}$$

We assume that f is analytic on $[-1, 1]$ and can be analytically extended onto E_r, where E_r refers to an open Bernstein ellipse of size $r > 1$, i.e., an ellipse in the complex plane with foci at ± 1 and the semi-minor and semi-major axes summing up to r, as illustrated in Fig. 1. Then, cf. [23] (Theorem 8.2), the error of the univariate polynomial best approximation f_M^* of degree M can be estimated as follows:

$$\|f - f_M^*\|_\infty \leq \frac{C_{\mathrm{B}} r^{-M}}{r - 1}, \tag{4}$$

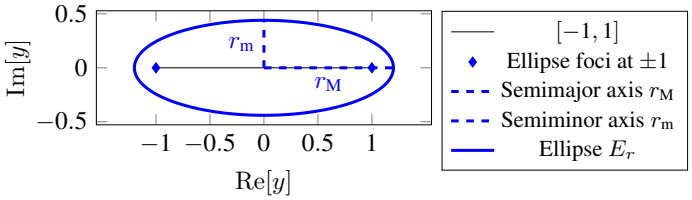


FIG. 1: Bernstein ellipse E_r of size $r = r_M + r_m$

where $\|\cdot\|_\infty$ denotes the supremum-norm on $[-1,1]$ and the constant $C_\mathrm{B} > 0$ depends on the uniform bound of the analytic continuation of f in E_r. We consider a polynomial interpolant

$$f_M(y) := \sum_{i=0}^{M} f(y^{(i)}) l_i(y), \tag{5}$$

where $\{l_i\}_{i=0}^M$ and $\{y^{(i)}\}_{i=0}^M$ denote univariate Lagrange polynomials and a set of distinct nodes, respectively. There holds

$$\|f - f_M\|_\infty \leq (1 + \Delta_M)\|f - f_M^*\|_\infty \leq (1 + \Delta_M)\frac{C_\mathrm{B} r^{-M}}{r-1}, \tag{6}$$

where

$$\Delta_M := \max_{y\in[-1,1]} \sum_{i=0}^{M} |l_i(y)|, \tag{7}$$

denotes the Lebesgue constant. If Δ_M grows sub-exponentially, the polynomial interpolation converges uniformly (for analytic functions). However, the convergence rate depends on the regularity of the analytic continuation of f in the complex plane. This is illustrated by considering the Runge function

$$f_\mathrm{R}(y;c) = \frac{1}{1+cy^2}, \quad c \in \mathbb{R}^+, \quad y \in [-1,1], \tag{8}$$

which is shown in Fig. 2(a), as a benchmark example. This function is analytic on $[-1,1]$, but the analytic continuation has a complex conjugate pole pair at $y = \pm i(1/\sqrt{c})$, limiting the size of the largest Bernstein ellipse, where the function f_R is analytic. Figure 2(b) demonstrates the effect on the convergence rate, where for increasing constants c, corresponding to a reduced size of the region of analyticity, a reduced convergence rate can be observed. The plot shows the convergence of the polynomial interpolant associated to unweighted Leja points in the empirical supremum-norm with a cross-validation sample of size 1×10^3.

Hale and Trefethen [48] have raised and discussed the question, whether polynomial methods are an optimal choice for functions analytic in an ϵ-neighborhood, in the context of numerical quadrature. Such a neighborhood is depicted in Fig. 3 together with the largest Bernstein ellipse contained in its interior. They have pointed out that, in

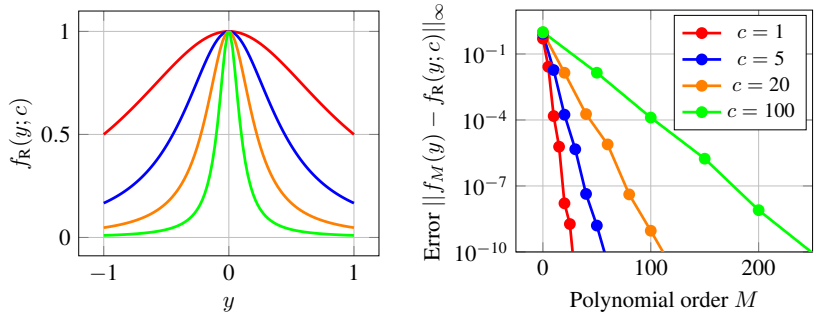


FIG. 2: (a) Runge function $f_\mathrm{R}(y;c)$ for different $c \in \mathbb{R}^+$ and $y \in [-1,1]$ and (b) geometric convergence rates of Leja interpolants

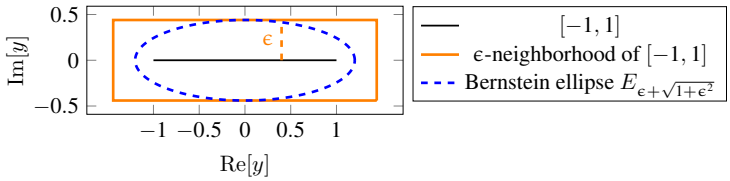


FIG. 3: ϵ-neighborhood and largest interior Bernstein ellipse

this case, superior methods to Gauss quadrature can be derived by conformally mapping the Bernstein ellipse E_r to a straighter region $\Omega_r = g(E_r)$, as illustrated in Fig. 4. As will be discussed in the following, this approach is also beneficial for (polynomial) interpolation. In accordance with [48], we focus in this work on conformal mappings $g : E_r \to \Omega_r$, which map the unit interval to itself, i.e., as follows:

$$g([-1,1]) = [-1,1], \tag{9}$$

and also fulfill

$$g(\pm 1) = \pm 1. \tag{10}$$

This ensures that the transplanted interpolation nodes

$$\{\hat{y}^{(i)}\}_{i=0}^{M} := \{g(y^{(i)})\}_{i=0}^{M}, \tag{11}$$

are still real numbers contained in the considered image set Ξ_n. There are various choices for g, see, e.g., [49]; however, in this work we focus on the sausage mapping proposed in [48]. It is defined by a dth-order Maclaurin expansion of the inverse sine function, which is then normalized such that (10) is fulfilled, as follows:

$$g_{\mathrm{S}}(y;d) = \left(\sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \frac{(2i)!}{4^i(2i+1)(i!)^2} \right)^{-1} \sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \frac{(2i)!}{4^i(2i+1)(i!)^2} y^{2i+1}. \tag{12}$$

An alternative mapping, due to Kosloff and Tal-Ezer [50], is given by

$$g_{\mathrm{KTE}}(y;\alpha) = \frac{\arcsin(\alpha y)}{\arcsin \alpha}, \quad \alpha \in (0,1). \tag{13}$$

It can be observed that the transplanted nodes are more evenly distributed, see Fig. 5.

We interpolate the transplanted knots $\{\hat{y}^{(i)}\}$ using mapped Lagrange polynomials $\hat{l}_i = l_i \circ g^{-1}$, shown in Fig. 6(a). Obviously, the mapped Lagrange polynomials also have the property

$$\hat{l}_j(\hat{y}^{(i)}) = l_j \circ g^{-1}(\hat{y}^{(i)}) = l_j(y^{(i)}) = \delta_{ij}, \tag{14}$$

where δ_{ij} denotes the Kronecker delta. Thus, the mapped interpolant $\hat{f}_M$ is defined by

$$\hat{f}_M(y) = \sum_{i=0}^{M} f(\hat{y}^{(i)})\hat{l}_i(y). \tag{15}$$

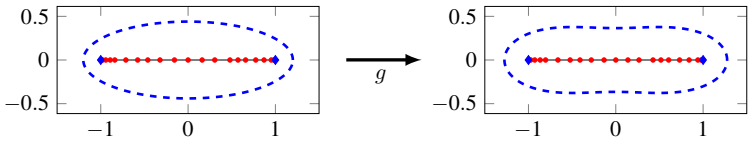


FIG. 4: Conformal map of a Bernstein ellipse

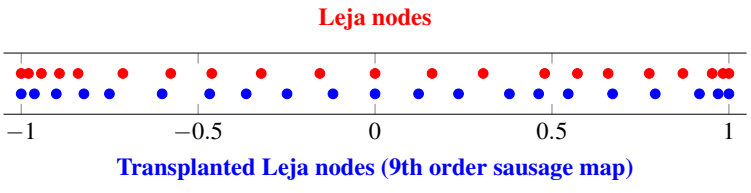
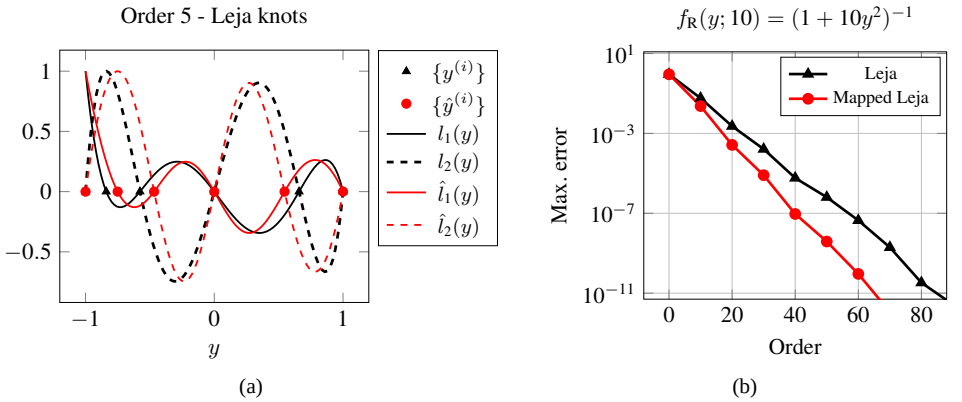


FIG. 5: Leja and transplanted Leja interpolation nodes



(a) (b)

FIG. 6: Mapped Leja interpolation polynomials and numerical orders for increasing polynomial degrees: (a) standard and mapped Lagrange polynomials and (b) convergence of (mapped) interpolation of $f_{\mathrm{R}}(y, 10)$

To derive an error bound for the transplanted interpolation, we first introduce the function $h := f \circ g$. We assume that h can be continued analytically to $E_{\hat{r}}$, where it is uniformly bounded. Let h_M be the Mth-order polynomial interpolant of h on the original nodes $\{y^{(i)}\}_{i=0}^{M}$. We observe that the mapped interpolant $\hat{f}_M$ is equivalent to $h_M \circ g^{-1}$ as follows:

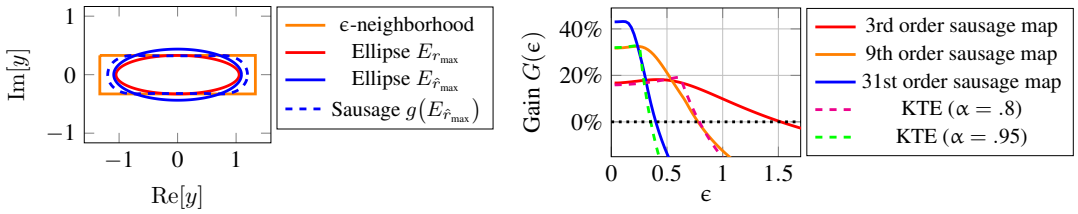
$$\hat{f}_M = \sum_{i=0}^{M} f\left(g\left(y^{(i)}\right)\right) l_i \circ g^{-1} = h_M \circ g^{-1}. \tag{16}$$

Because of Eq. (9), we obtain

$$\begin{aligned} ||f - \hat{f}_M||_\infty &= ||(f - \hat{f}_M) \circ g \circ g^{-1}||_\infty && (17)\\ &= ||(h - h_M) \circ g^{-1}||_\infty && (18)\\ &= ||h - h_M||_\infty && (19)\\ &\leq (1 + \Delta_M)||h - h_M^*||_\infty && (20)\\ &\leq (1 + \Delta_M)\frac{\hat{C}_{\mathrm{B}}\hat{r}^{-M}}{\hat{r} - 1}. && (21) \end{aligned}$$

The convergence rate is improved if $\hat{r} > r$, which we confirm numerically in Fig. 6(b), where the ninth-order sausage map $g_{\mathrm{S}}(y; 9)$ is employed. Since $g_{\mathrm{S}}(y; 9)^{-1}$ is not known analytically, we approximate the inverse mapping by a Chebyshev approximation of order 100 (up to machine precision).

If the region of analyticity is known, one can estimate the gain of employing a conformal mapping *a priori* (based on the convergence estimates), as illustrated in Fig. 7(a). Let the size of the largest Bernstein ellipse in the region of analyticity be $r_{\max}$ and the size of the largest Bernstein ellipse, which is fully mapped into this region, be



(a) (b)

FIG. 7: Convergence gain G by employing mapped approximations: (a) illustration of geometric gain estimation for $\epsilon = 0.3294$ and (b) gain in convergence for different mappings

$\hat{r}_{\max}$. The convergence of polynomial interpolation is then given by $\mathcal{O}\big[(1+\Delta_M)\exp\left(-\log(r_{\max})M\right)\big]$ according to Eq. (6); whereas the mapped interpolation converges as $\mathcal{O}\big[(1+\Delta_M)\exp\left(-\log(\hat{r}_{\max})M\right)\big]$, see Eq. (21). Assuming a sufficiently slowly growing Lebesgue constant Δ_M, we consider the relative improvement in the asymptotic rate of geometric convergence (see [51], Definition 6), given by

$$G = \frac{\log \hat{r}_{\max}}{\log r_{\max}} - 1, \tag{22}$$

which can be attributed to the use of conformal maps. We evaluate the gain G for functions that are analytic in ϵ-neighborhoods of $[-1,1]$, see Fig. 7(b), by numerically computing $\hat{r}_{\max}$ for different mappings. It should be noted that higher gains can be expected, if mappings would be employed, which are specifically tailored to the positions of the poles in the complex plane. However, usually the exact position of these poles is not known *a priori* and this approach is therefore not pursued any further. The interested reader is referred to [49]. In the remaining part of the paper, we work with the ninth-th order sausage mapping $g_{\mathrm{S}}(y;9)$ since a detailed comparison of different mappings is not in the scope of the present paper. The particular mapping is selected because it has already been established in [48,49], and Fig. 7(b) confirms a significant gain in convergence for a substantial range of ϵ-neighborhoods. Additionally, in contrast to the map (13), it does not introduce an artificial singularity.

2.2 Sparse Grid Interpolation

Approximations based on sparse grid interpolation are commonly referred to as sparse grid stochastic collocation methods [6,8,46]. Those methods are based on combinations of univariate interpolation rules, defined by an interpolation level $\ell_n \in \mathbb{N}_0$, a monotonically increasing level-to-nodes function $m_n : \mathbb{N}_0 \to \mathbb{N}$, where $m_n(\ell_n) =: m_{\ell_n}$ and $m_n(0) = 1$ and a grid of m_{ℓ_n} (mapped) interpolation nodes

$$Z_{\ell_n} = \left\{ \hat{y}_n^{(i_n)} \right\}_{i_n=0}^{m_{\ell_n}-1}. \tag{23}$$

Introducing the multi-index $\boldsymbol{\ell} = (\ell_1, \ell_2, \ldots, \ell_N) \in \mathbb{N}_0^N$, the tensor-product multivariate approximation is obtained as follows:

$$J(\mathbf{y}) \approx \widetilde{J}(\mathbf{y}) = \sum_{\mathbf{i}:\hat{\mathbf{y}}^{(\mathbf{i})} \in Z_{\boldsymbol{\ell}}} J\left(\hat{\mathbf{y}}^{(\mathbf{i})}\right) \hat{L}_{\boldsymbol{\ell},\mathbf{i}}(\mathbf{y}), \tag{24}$$

where $\hat{\mathbf{y}}^{(\mathbf{i})} = \left(\hat{y}_1^{(i_1)}, \hat{y}_2^{(i_2)}, \ldots, \hat{y}_N^{(i_N)}\right) \in Z_{\boldsymbol{\ell}}$ are multivariate interpolation nodes, uniquely identified by the multi-index $\mathbf{i} = (i_1, i_2, \ldots, i_N) \in \mathbb{N}_0^N$, and $Z_{\boldsymbol{\ell}} = Z_{\ell_1} \times Z_{\ell_2} \times \cdots \times Z_{\ell_N}$ is the tensor grid of interpolation nodes.

Moreover, $\hat{L}_{\boldsymbol{\ell},\mathbf{i}}$ are mapped multivariate Lagrange polynomials, obtained by the composition $\hat{L}_{\boldsymbol{\ell},\mathbf{i}} = L_{\boldsymbol{\ell},\mathbf{i}} \circ \mathbf{g}^{-1}$ with

$$L_{\boldsymbol{\ell},\mathbf{i}}(\mathbf{y}) = \prod_{n=1}^{N} l_{\ell_n,i_n}(y_n), \tag{25}$$

where

$$l_{\ell_n,i_n}(y_n) := \begin{cases} \displaystyle\prod_{k=0,k\neq i_n}^{m_{\ell_n}-1} \frac{y_n - y_n^{(k)}}{y_n^{(i_n)} - y_n^{(k)}}, & \ell_n \neq 0, \\ 1, & \ell_n = 0. \end{cases}$$

Obviously, for the trivial mapping, $\mathbf{g} : \mathbf{y} \mapsto \mathbf{y}$, it holds $\hat{L}_{\boldsymbol{\ell},\mathbf{i}} = L_{\boldsymbol{\ell},\mathbf{i}}$. More details on the multivariate coordinatewise conformal mapping $\mathbf{g}$ will be given in Section 2.2.1. It should be noted that (25) is used for the ease of exposition; in the actual implementation, the barycentric representation should be used [52]. Since $J(\mathbf{y})$ has to be evaluated for each $\mathbf{y}^{(\mathbf{i})} \in Z_{\boldsymbol{\ell}}$, the complexity of the tensor-product approach is $\mathcal{O}\big(m_k^N\big)$, where

$$m_k := \max_n m_{\ell_n}. \tag{26}$$

This complexity can be mitigated to $\mathcal{O}\Big(m_k(\log m_k)^{N-1}\Big)$ by employing Smolyak sparse grids [53], which typically result in an acceptable trade-off between approximation accuracy and complexity. We introduce the approximation level $k \in \mathbb{N}_0$ and define the multi-index set Λ_k, such that

$$\Lambda_k := \{\boldsymbol{\ell} : |\boldsymbol{\ell}| = \ell_1 + \ell_2 + \cdots + \ell_N \leq k\}. \tag{27}$$

Then, the sparse grid of multivariate interpolation nodes Z_{Λ_k} is constructed as follows:

$$Z_{\Lambda_k} = \bigcup_{k-N+1 \leq |\boldsymbol{\ell}| \leq k} Z_{\boldsymbol{\ell}}, \tag{28}$$

and the interpolation is given by

$$\mathcal{I}_{\Lambda_k}[J](\mathbf{y}) = \sum_{\mathbf{i}:\hat{\mathbf{y}}^{(\mathbf{i})} \in Z_{\Lambda_k}} J\Big(\hat{\mathbf{y}}^{(\mathbf{i})}\Big)\hat{L}_{\boldsymbol{\ell},\mathbf{i}}(\mathbf{y}). \tag{29}$$

2.2.1 Mapped Leja Nodes, Hierarchical Interpolation and Adaptivity

As shown in [43], Smolyak formulas are in general not interpolatory, unless based on nested sequences of univariate interpolation nodes, such that $Z_{\ell_n-1} \subset Z_{\ell_n}$. Moreover, to ensure accuracy and fast convergence of the approximation, the interpolation nodes should be chosen in agreement with the PDFs $\varrho_n(y_n)$. We opt for weighted Leja interpolation nodes, as in [11]. Omitting conformal mappings for the moment, we consider a univariate, continuous, and positive weight function. Here, this weight function is given by a univariate PDF $\varrho_n(y_n)$, $\varrho_n : \Xi_n \to \mathbb{R}_+$. A sequence of univariate Leja nodes $y_n^{(k)} \in \Xi_n$, $k = 0, 1, 2, \ldots$, can be constructed by solving the optimization problem

$$y_n^{(K)} = \underset{y_n \in \Xi_n}{\arg\max} \sqrt{\varrho_n(y_n)} \prod_{k=0}^{K-1} \Big|y_n - y_n^{(k)}\Big|, \tag{30}$$

where the starting node $y_n^{(0)}$ is arbitrarily chosen. For further details on the construction of weighted Leja nodes and an analysis of their properties, see [11]. We justify the choice of Leja nodes as follows. First of all, Leja nodes satisfy the nestedness requirement by construction. Second, they allow complete freedom in the choice of the level-to-nodes function $m_n(\ell_n)$. Finally, they can be tailored to any given PDF. In comparison, the commonly employed Clenshaw-Curtis nodes would restrict us to the rapidly growing level-to-nodes function $m_n(\ell_n) = 2^{\ell_n} + 1$. In the following, we employ the level-to-nodes function $m_n(\ell_n) = \ell_n + 1$, $\ell_n \in \mathbb{N}_0$, and denote with $y_n^{(\ell_n)}$ the single extra node corresponding to interpolation level ℓ_n, i.e., $y_n^{(\ell_n)} = Z_{\ell_n} \setminus Z_{\ell_n-1}$. We also introduce for each parameter a conformal map g_n, as discussed in Section 2.1. Then, the mapped univariate Leja nodes $\hat{y}_n^{(k)}$ are obtained as $\hat{y}_n^{(k)} = g_n(y_n^{(k)})$. Of course, for the trivial map $g_n : y_n \mapsto y_n$ we recover the original Leja nodes $\hat{y}_n^{(k)} = y_n^{(k)}$. The multivariate mapping is then obtained as follows:

$$\mathbf{g}(\mathbf{y}) = g_1(y_1) \cdots g_N(y_N). \tag{31}$$

We note that the multivariate mapping $\mathbf{g}$ is conformal in each coordinate y_n.

In the multivariate case, nested grids of multivariate interpolation nodes can be constructed by enforcing the use of downward-closed (also, monotone or lower) multi-index sets [9,10]. Such sets are known to preserve the telescopic properties of the series in Eq. (29) [10]. Moreover, sequences of nested, downward-closed multi-index sets result in polynomial approximations of increasing accuracy [9]. Given a multi-index set Λ, let us first define its forward and backward neighbor multi-index sets, Λ_+ and Λ_-, respectively, such that

$$\Lambda_+ := \{\boldsymbol{\ell} + \mathbf{e}_n, \forall \boldsymbol{\ell} \in \Lambda, \forall n = 1, \ldots, N\}, \tag{32a}$$

$$\Lambda_- := \{\boldsymbol{\ell} - \mathbf{e}_n, \forall \boldsymbol{\ell} \in \Lambda, \forall n = 1, \ldots, N : \ell_n > 0\}, \tag{32b}$$

where $\mathbf{e}_n$ is the nth unit vector. Then, Λ is said to be downward-closed if and only if

$$\Lambda_- \subset \Lambda. \tag{33}$$

Assuming now a multi-index $\boldsymbol{\ell} \notin \Lambda$ such that $\Lambda \cup \boldsymbol{\ell}$ is downward-closed, it holds that $Z_\Lambda \subset Z_{\Lambda\cup\boldsymbol{\ell}}$ and

$$\hat{\mathbf{y}}^{(\boldsymbol{\ell})} = Z_{\Lambda\cup\boldsymbol{\ell}} \setminus Z_\Lambda, \tag{34}$$

where

$$Z_\Lambda = \bigcup_{\boldsymbol{\ell}\in\Lambda} Z_{\boldsymbol{\ell}}. \tag{35}$$

Then, Eq. (29) can be naturally transformed into the hierarchical interpolation

$$\mathcal{I}_{\Lambda\cup\boldsymbol{\ell}}[J](\mathbf{y}) = \mathcal{I}_\Lambda[J](\mathbf{y}) + s_{\boldsymbol{\ell}}\, \hat{H}_{\boldsymbol{\ell}}(\mathbf{y}), \tag{36}$$

where the coefficients $s_{\boldsymbol{\ell}} \in \mathbb{C}$, known as hierarchical surpluses, are given by

$$s_{\boldsymbol{\ell}} = J\Big(\hat{\mathbf{y}}^{(\boldsymbol{\ell})}\Big) - \mathcal{I}_\Lambda[J]\Big(\hat{\mathbf{y}}^{(\boldsymbol{\ell})}\Big), \tag{37}$$

and $\hat{H}_{\boldsymbol{\ell}}$ are multivariate mapped hierarchical polynomials, defined as follows:

$$\hat{H}_{\boldsymbol{\ell}}(\mathbf{y}) = \prod_{n=1}^{N} \hat{h}_{\ell_n}(y_n), \tag{38}$$

where

$$\hat{h}_{\ell_n}(y_n) := \begin{cases} \displaystyle\prod_{k=0}^{\ell_n-1} \frac{g_n^{-1}(y_n) - y_n^{(k)}}{y_n^{(\ell_n)} - y_n^{(k)}}, & \ell_n \neq 0, \\ 1, & \ell_n = 0. \end{cases}$$

Again, by choosing g_n as the identity map, we recover standard hierarchical Lagrange polynomials.

The use of (mapped) hierarchical polynomials has the advantage that the basis functions do not change as new nodes are added. Moreover, the hierarchical surpluses $s_{\boldsymbol{\ell}}$ can be interpreted as error indicators, quantifying the contribution of the interpolation node $\hat{\mathbf{y}}^{(\boldsymbol{\ell})}$ to the already available approximation. This interpretation motivates the adaptive construction of the sparse grid approximation based on *a posteriori* error estimates. We consider a dimension-adaptive scheme, similar to the ones employed in [9–11,45,47], with minor modifications to address the case of complex QoIs. The scheme is presented in Algorithm 1. A detailed description follows.

Given a downward-closed multi-index set Λ, as well as the corresponding approximation $\mathcal{I}_\Lambda[J]$ and grid Z_Λ, we define the set of admissible neighbors Λ_+^{adm}, such that

$$\Lambda_+^{\mathrm{adm}} := \{\boldsymbol{\ell} \in \Lambda_+ \,:\, \boldsymbol{\ell} \notin \Lambda \;\text{ and }\; \{\boldsymbol{\ell}\}_- \subset \Lambda\}. \tag{39}$$

Algorithm 1: Dimension-adaptive interpolation

Data: QoI $J(\mathbf{y})$, conformal map $\mathbf{g}$, multi-index set Λ, budget B
Result: sparse grid $Z_{\Lambda\cup\Lambda_+^{\mathrm{adm}}}$, approximation $\mathcal{I}_{\Lambda\cup\Lambda_+^{\mathrm{adm}}}[J]$
repeat
  Compute the admissible set Λ_+^{adm}, as in Eq. (39).
  Compute the hierarchical surpluses $s_{\boldsymbol{\ell}}$, $\forall \boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$, as in Eq. (37).
  Find the multi-index $\boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$ with the maximum error indicator $|s_{\boldsymbol{\ell}}|$.
  Compute the approximation $\mathcal{I}_{\Lambda\cup\boldsymbol{\ell}}$, as in Eq. (36).
  Set $\Lambda = \Lambda \cup \boldsymbol{\ell}$.
until *simulation budget B is reached*;

Expanding Λ with admissible multi-indices $\boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$ guarantees that (33) is satisfied, and we thus construct a sequence of nested downward-closed sets [9]. In this work, the error indicator corresponding to each multi-index $\boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$ is chosen to be the modulus $|s_{\boldsymbol{\ell}}|$ of the corresponding complex hierarchical surplus; however, other choices are possible, e.g., $\max(|\mathrm{Re}\{s_{\boldsymbol{\ell}}\}|, |\mathrm{Im}\{s_{\boldsymbol{\ell}}\}|)$. We update Λ with the multi-index $\boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$ corresponding to the maximum error indicator $|s_{\boldsymbol{\ell}}|$. The grid of interpolation nodes Z_Λ and the approximation $\mathcal{I}_\Lambda$ are updated accordingly. This procedure is continued iteratively, until a budget of model evaluations B is reached. This criterion can be formulated as follows:

$$\#Z_{\Lambda \cup \Lambda_+^{\mathrm{adm}}} \geq B, \tag{40}$$

where $\#$ denotes the cardinality of a set. If an approximation is not readily available, then the algorithm is initiated with $\Lambda = \{(0, 0, \ldots, 0)\}$. After the termination of the algorithm, the approximation is constructed using the set $\Lambda \cup \Lambda_+^{\mathrm{adm}}$.

2.3 Adjoint Error Estimation and Adaptivity

We aim to improve Algorithm 1 by using an adjoint error indicator to steer adaptivity. Adjoint error estimation is well-established in the context of the finite element method (FEM); see [54] and the references therein. It has been considered in a stochastic/parametric context [25,55,56], as well as for Clenshaw-Curtis adaptivity [24,47]. Because of the exponential growth of Clenshaw-Curtis nodes, adjoint error estimation can result in a significant reduction of computational cost. In this work, we demonstrate that adjoint techniques can be beneficial for Leja adaptivity, too.

In this section, we rely on the fact that $J(\mathbf{y}) = J_{\mathbf{y}}(\mathbf{u}(\mathbf{y}))$, $J_{\mathbf{y}} : V \to \mathbb{C}$, is a linear functional with respect to $\mathbf{u}(\mathbf{y})$. However, generalizations to nonlinear functionals are also possible, as in [57] (Chapter 3.2). We rewrite the primal problem (1) as an operator equation: $\forall \mathbf{y} \in \Xi$, find $\mathbf{u}(\mathbf{y}) \in V$, such that

$$\langle L_{\mathbf{y}} \mathbf{u}(\mathbf{y}), \mathbf{v} \rangle = a_{\mathbf{y}}(\mathbf{u}(\mathbf{y}), \mathbf{v}) = l_{\mathbf{y}}(\mathbf{v}), \ \ \forall \mathbf{v} \in V, \tag{41}$$

where $L_{\mathbf{y}} : V \to V^*$ denotes the primal operator and V^* the dual space to V. The dual problem is given as, for all $\mathbf{y} \in \Xi$, find $\mathbf{z}(\mathbf{y}) \in V$, such that

$$\langle \mathbf{w}, L_{\mathbf{y}}^* \mathbf{z}(\mathbf{y}) \rangle = a_{\mathbf{y}}(\mathbf{w}, \mathbf{z}(\mathbf{y})) = J_{\mathbf{y}}(\mathbf{w}), \ \ \forall \mathbf{w} \in V, \tag{42}$$

where $L_{\mathbf{y}}^* : V \to V^*$ denotes the adjoint operator defined by

$$\langle L_{\mathbf{y}} \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, L_{\mathbf{y}}^* \mathbf{v} \rangle, \ \ \forall \mathbf{u}, \mathbf{v} \in V, \ \ \forall \mathbf{y} \in \Xi. \tag{43}$$

The so-called primal-dual equivalence

$$J_{\mathbf{y}}(\mathbf{u}(\mathbf{y})) = \langle \mathbf{u}(\mathbf{y}), L_{\mathbf{y}}^* \mathbf{z}(\mathbf{y}) \rangle = \langle L_{\mathbf{y}} \mathbf{u}(\mathbf{y}), \mathbf{z}(\mathbf{y}) \rangle = l_{\mathbf{y}}(\mathbf{z}(\mathbf{y})), \tag{44}$$

follows directly from these definitions. Given (mapped) polynomial approximations $\tilde{\mathbf{u}}, \tilde{\mathbf{z}}$ of the mappings $\mathbf{u}, \mathbf{z} : \Xi \to V$, we are interested in the error

$$\eta(\mathbf{y}) = J_{\mathbf{y}}\big(\mathbf{u}(\mathbf{y}) - \tilde{\mathbf{u}}(\mathbf{y})\big) = a_{\mathbf{y}}\big(\mathbf{u}(\mathbf{y}) - \tilde{\mathbf{u}}(\mathbf{y}), \mathbf{z}(\mathbf{y})\big) = l_{\mathbf{y}}\big(\mathbf{z}(\mathbf{y})\big) - a_{\mathbf{y}}\big(\tilde{\mathbf{u}}(\mathbf{y}), \mathbf{z}(\mathbf{y})\big). \tag{45}$$

Even if $\tilde{\mathbf{u}}, \tilde{\mathbf{z}}$ are replaced by their finite element counterparts, the error according to Eq. (45) is not readily computable, as it would require the computation of the adjoint $\mathbf{z}$ for all $\mathbf{y} \in \Xi$. Following [25,55], we propose to use the error indicator

$$\tilde{\eta}(\mathbf{y}) = a_{\mathbf{y}}\big(\mathbf{u}(\mathbf{y}) - \tilde{\mathbf{u}}(\mathbf{y}), \tilde{\mathbf{z}}(\mathbf{y})\big) = l_{\mathbf{y}}\big(\tilde{\mathbf{z}}(\mathbf{y})\big) - a_{\mathbf{y}}\big(\tilde{\mathbf{u}}(\mathbf{y}), \tilde{\mathbf{z}}(\mathbf{y})\big). \tag{46}$$

By exploiting the continuity of the sesquilinear form $a_{\mathbf{y}}(\cdot, \cdot)$, it can be shown that the error indicator (46) converges faster than the mapped polynomial approximations $\tilde{\mathbf{u}}, \tilde{\mathbf{z}}$

$$|\eta(\mathbf{y}) - \tilde{\eta}(\mathbf{y})| = |a_{\mathbf{y}}\big(\mathbf{u}(\mathbf{y}) - \tilde{\mathbf{u}}(\mathbf{y}), \mathbf{z}(\mathbf{y}) - \tilde{\mathbf{z}}(\mathbf{y})\big)| \leq C \|\mathbf{u}(\mathbf{y}) - \tilde{\mathbf{u}}(\mathbf{y})\|_V \|\mathbf{z}(\mathbf{y}) - \tilde{\mathbf{z}}(\mathbf{y})\|_V. \tag{47}$$

In particular, considering for the moment the univariate case $N = 1$, for simplicity, and assuming that $\mathbf{u}, \mathbf{z}$ can be extended analytically onto open Bernstein ellipses $E_{\hat{r}_u}, E_{\hat{r}_z}$, respectively, and that there exist uniform bounds on their extensions, then, we obtain

$$\|\eta - \tilde{\eta}\|_\infty \leq C_1(1 + \Delta_M)^2 \frac{C_2(\hat{r}_u \hat{r}_z)^{-M}}{(1 - \hat{r}_u)(1 - \hat{r}_z)}, \tag{48}$$

for M-point approximations of both $\mathbf{u}$ and $\mathbf{z}$. Hence, for $\hat{r}_u = \hat{r}_z$, $\tilde{\eta}$ exhibits twice the rate of geometric convergence.

We proceed by discussing the necessary adaptations to Algorithm 1, in order to incorporate the adjoint error indicator (46). Additionally to the (mapped) polynomial approximation (36) of the single-valued and complex QoI, one needs to create (mapped) polynomial approximations $\tilde{\mathbf{u}}(\mathbf{y}), \tilde{\mathbf{z}}(\mathbf{y})$ of the primal and dual solution. The approximations are constructed with the same multi-index set Λ as for the QoI, using the same mapped polynomials $\hat{H}_{\boldsymbol{\ell}}(\mathbf{y})$.

Following [24], we carry out the algorithmic modifications in the dimension-adaptive scheme. While Algorithm 1 uses the error indicators $|s_{\boldsymbol{\ell}}|, \forall \boldsymbol{\ell} \in \Lambda_+^{\text{adm}}$, by solving the respective linear problem, we suggest the use of the adjoint-based error indicators $|\tilde{s}_{\boldsymbol{\ell}}|$, where $\tilde{s}_{\boldsymbol{\ell}} = \tilde{\eta}\big(\hat{\mathbf{y}}^{(\boldsymbol{\ell})}\big)$. As before, we choose the multi-index with the maximum error indicator, solve the corresponding linear system, and update the approximations of the primal and the dual solution, as well as of the QoI. This scheme is summarized in Algorithm 2. After the termination of the algorithm, the approximation can be constructed with the set $\Lambda \cup \Lambda_+^{\text{adm}}$, such that the already computed adjoint-based error indicators are used as the hierarchical surpluses corresponding to the admissible neighbors, i.e., $s_{\boldsymbol{\ell}} = \tilde{s}_{\boldsymbol{\ell}}, \ \forall \boldsymbol{\ell} \in \Lambda_+^{\text{adm}}$. The error indicator (46) can be further employed in order to improve the (mapped) polynomial surrogate model of the QoI. In particular, one can replace the single-valued QoI $J(\mathbf{y})$ by

$$\tilde{J}(\mathbf{y}) = \mathcal{I}_\Lambda[J](\mathbf{y}) + \tilde{\eta}(\mathbf{y}), \tag{49}$$

such that the computed mapped polynomial approximation is corrected by the adjoint-error indicator, before continuing with further approximation refinements using Algorithm 1. We emphasize that no additional linear equation system has to be solved in order to evaluate (49).

3. MAXWELL'S SOURCE PROBLEM

In this section, we introduce the model problem, i.e., Maxwell's source problem with periodic boundary conditions. Such a model can be used, for instance, to describe the coupling into a plasmonic grating coupler, which will be considered in Section 4. We also introduce the finite element (FE) approximation and a parametric version of the model.

Algorithm 2: Adjoint error-based, dimension-adaptive interpolation

Data: $\mathbf{g}$, Λ, B and $a_{\mathbf{y}}$, $l_{\mathbf{y}}$, $J_{\mathbf{y}}$ as defined in Eqs. (41) and (42)
Result: sparse grid $Z_{\Lambda \cup \Lambda_+^{\text{adm}}}$, approximation $\mathcal{I}_{\Lambda \cup \Lambda_+^{\text{adm}}}[J]$
repeat
- Compute the admissible set Λ_+^{adm}, as in Eq. (39).
- Compute the error indicators $|\tilde{s}_{\boldsymbol{\ell}}|$, where $\tilde{s}_{\boldsymbol{\ell}} = \tilde{\eta}\big(\mathbf{y}^{(\boldsymbol{\ell})}\big), \forall \boldsymbol{\ell} \in \Lambda_+^{\text{adm}}$.
- Find the multi-index $\boldsymbol{\ell} \in \Lambda_+^{\text{adm}}$ with the maximum error indicator.
- Compute the hierarchical surpluses $s_{\boldsymbol{\ell}}$, $\mathbf{u}_{\boldsymbol{\ell}}$, $\mathbf{z}_{\boldsymbol{\ell}}$ as in Eq. (37), by solving the linear problems for primal and dual solution.
- Compute the approximation $\mathcal{I}_{\Lambda \cup \boldsymbol{\ell}}$, as in Eq. (36), and the corresponding approximations of primal and dual solution.
- Set $\Lambda = \Lambda \cup \boldsymbol{\ell}$.

until *stopping criterion fulfilled*;

3.1 Deterministic Problem

We start with the time-harmonic Maxwell's equations

$$\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H} \quad \text{in } D, \tag{50a}$$
$$\nabla \times \mathbf{H} = \mathbf{J}_s + j\omega\varepsilon\mathbf{E} \quad \text{in } D, \tag{50b}$$
$$\nabla \cdot (\varepsilon\mathbf{E}) = \rho \quad \text{in } D, \tag{50c}$$
$$\nabla \cdot (\mu\mathbf{H}) = 0 \quad \text{in } D, \tag{50d}$$

where $\mathbf{E}$ denotes the electric field phasor, $\mathbf{H}$ the magnetic field phasor, $\mathbf{J}_s$ the source current phasor, ρ the charge density phasor, ω the angular frequency, ε the dispersive complex-valued permittivity, μ the permeability, and D the computational domain to be specified. The permeability $\mu = \mu_0\mu_r$, where μ_r and μ_0 represent the relative and vacuum permeability, respectively, is assumed to be nondispersive. In absence of charges and source currents, i.e., $\rho = 0$ and $\mathbf{J}_s = 0$, the so-called curl-curl equation reads as follows:

$$\nabla \times \left(\mu_r^{-1}\nabla \times \mathbf{E}\right) - \omega^2\varepsilon\mu_0\mathbf{E} = 0 \quad \text{in } D, \tag{51}$$

to be endowed with appropriate boundary conditions.

Given an infinitely periodic structure and a periodic excitation, the computational domain D can be confined to a single unit cell of the periodic structure, based on Floquet's Theorem ([58], Chapter 13). The unit cell is illustrated in Fig. 8. Without loss of generality, we assume periodicity in the x and y directions; whereas, Γ_{z^+} and Γ_{z^-} denote the boundaries in the nonperiodic direction. At Γ_{z^+}, the structure is excited by an incident plane wave

$$\mathbf{E}^{\text{inc}} = \mathbf{E}_0 e^{-j\mathbf{k}^{\text{inc}}\cdot\mathbf{r}}, \quad \mathbf{k}^{\text{inc}} = \begin{bmatrix} k_x^{\text{inc}} \\ k_y^{\text{inc}} \\ k_z^{\text{inc}} \end{bmatrix} = -k_0 \begin{bmatrix} \sin\theta^{\text{inc}}\cos\phi^{\text{inc}} \\ \sin\theta^{\text{inc}}\sin\phi^{\text{inc}} \\ \cos\theta^{\text{inc}} \end{bmatrix}, \tag{52}$$

where $\theta^{\text{inc}},\ \phi^{\text{inc}}$ are the angles of incidence and $k_0 = \omega\sqrt{\mu_0\varepsilon_0}$ is the wavenumber in vacuum. It is worth noting that, due to the oblique angles, the periodicity of the excitation differs from the geometrical periodicity of the structure. According to Floquet's theorem, we need to enforce periodic phase-shift boundary conditions on $\Gamma_{x^+} \cup \Gamma_{x^-}$ and on $\Gamma_{y^+} \cup \Gamma_{y^-}$, i.e.,

$$\mathbf{E}|_{\Gamma_{x^+}} = \mathbf{E}|_{\Gamma_{x^-}} e^{j\psi_x}, \quad \psi_x = -k_x^{\text{inc}} d_x, \tag{53a}$$
$$\mathbf{E}|_{\Gamma_{y^+}} = \mathbf{E}|_{\Gamma_{y^-}} e^{j\psi_y}, \quad \psi_y = -k_y^{\text{inc}} d_y, \tag{53b}$$

where the phase-shifts $\psi_x,\ \psi_y$ depend only on the wavevector $\mathbf{k}^{\text{inc}}$ of the incident wave at Γ_{z^+} and on the dimensions $d_x,\ d_y$ of the unit cell.

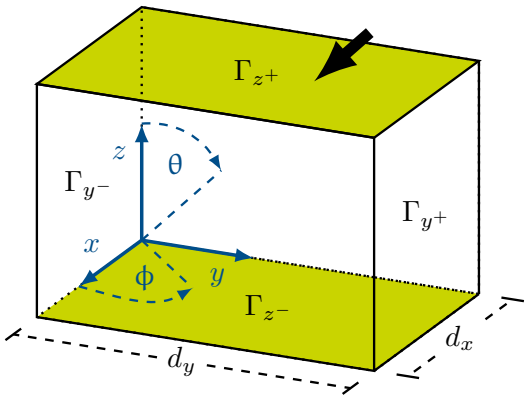


FIG. 8: Sketch of a unit cell representing the computational domain D. The black arrow indicates the incident wavevector $\mathbf{k}^{\text{inc}}$.

To truncate the structure in the nonperiodic direction at Γ_{z^+}, we employ a Floquet absorbing boundary condition [58] as derived in Appendix A. At Γ_{z^-}, a perfect electric conductor (PEC) boundary condition is applied to truncate the structure; however, different boundary conditions are also possible, e.g., again a Floquet absorbing boundary condition or perfectly matched layers (PMLs) [58]. In summary, we are concerned with the boundary value problem

$$\nabla \times \left(\mu_r^{-1}\nabla \times \mathbf{E}\right) - \omega^2 \varepsilon \mu_0 \mathbf{E} = 0 \quad \text{in } D, \tag{54a}$$

$$\mathbf{E}|_{\Gamma_{x^+}} e^{-j\psi_x} = \mathbf{E}|_{\Gamma_{x^-}} \quad \text{on } \Gamma_{x^+} \cup \Gamma_{x^-}, \tag{54b}$$

$$\mathbf{E}|_{\Gamma_{y^+}} e^{-j\psi_y} = \mathbf{E}|_{\Gamma_{y^-}} \quad \text{on } \Gamma_{y^+} \cup \Gamma_{y^-}, \tag{54c}$$

$$\mathbf{n} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{z^-}, \tag{54d}$$

$$\mathbf{n} \times \mathbf{H} + \boldsymbol{\mathcal{G}}(\mathbf{E}) = \boldsymbol{\mathcal{F}}^{\text{inc}} \quad \text{on } \Gamma_{z^+}, \tag{54e}$$

where $\boldsymbol{\mathcal{G}}(\mathbf{E})$ and $\boldsymbol{\mathcal{F}}^{\text{inc}}$ are derived and defined in Appendix A, see Eqs. (A.10)–(A.12).

3.1.1 Weak Formulation and Discretization

To simplify the notation, we introduce the traces

$$\mathbf{u}_{\text{T}} := (\mathbf{n}_\Gamma \times \mathbf{u}|_\Gamma) \times \mathbf{n}_\Gamma, \tag{55a}$$

$$\mathbf{u}_{\text{t}} := \mathbf{n}_\Gamma \times \mathbf{u}|_\Gamma, \tag{55b}$$

where $\Gamma := \partial D$ denotes the boundary of D and $\mathbf{n}_\Gamma$ refers to its outer unit normal. Note that the trace operators are denoted by subscripts, for brevity of notation.

By building the inner product of Eq. (54a) with tests function $\mathbf{E}' \in V$, where V is to be determined, and integration by parts, we obtain

$$\left(\mu_{\text{r}}^{-1}\nabla \times \mathbf{E}, \nabla \times \mathbf{E}'\right)_D - \omega^2 \mu_0 (\varepsilon \mathbf{E}, \mathbf{E}')_D - j\omega\mu_0 (\mathbf{H}_{\text{t}}, \mathbf{E}'_{\text{T}})_\Gamma = 0. \tag{56}$$

The boundary integral can be further simplified, i.e., the contributions on Γ_{x^+}, Γ_{x^-} and Γ_{y^+}, Γ_{y^-} cancel each other due to the periodic phase-shift boundary conditions (54b) and (54c) of trial and test functions. We further eliminate the portion of the integral on Γ_{z^-} by demanding that the test functions $\mathbf{E}'$ fulfill the PEC boundary condition (54d).

The appropriate function space V for a weak formulation is a subspace of $H(\text{curl}; D)$, i.e., the (complex) vector function space of square-integrable functions with square-integrable curl. For more details on function spaces in the context of Maxwell's source problem, the reader is referred to [59] (Chapter 3). To account for the boundary conditions in Eq. (54), the function space is chosen as

$$\begin{aligned} V := \Big\{ \mathbf{v} \in H(\text{curl}; D) :\ & \mathbf{v}_{\text{T}}|_{\Gamma_{z^-}} = 0 \wedge\ \mathbf{v}_{\text{T}}|_{\Gamma_{x^+}} = -\mathbf{v}_{\text{T}}|_{\Gamma_{x^-}} e^{j\psi_x} \\ & \wedge\, \mathbf{v}_{\text{T}}|_{\Gamma_{y^+}} = -\mathbf{v}_{\text{T}}|_{\Gamma_{y^-}} e^{j\psi_y} \wedge\ \mathbf{v}_{\text{T}}|_{\Gamma_{z^+}} \in \left(L^2(\Gamma_{z^+})\right)^3 \Big\}, \end{aligned} \tag{57}$$

where the condition $\mathbf{v}_{\text{T}}|_{\Gamma_{z^+}} \in \left(L^2(\Gamma_{z^+})\right)^3$ is required to obtain a well-defined boundary integral. Employing the Floquet absorbing boundary condition (54e) on Γ_{z^+} yields the weak formulation: find $\mathbf{E} \in V$ s.t.

$$\left(\mu_{\text{r}}^{-1}\nabla \times \mathbf{E}, \nabla \times \mathbf{E}'\right)_D - \omega^2 \mu_0 (\varepsilon \mathbf{E}, \mathbf{E}')_D + j\omega\mu_0 (\boldsymbol{\mathcal{G}}(\mathbf{E}), \mathbf{E}'_{\text{T}})_{\Gamma_{z^+}} = j\omega\mu_0 \left(\boldsymbol{\mathcal{F}}^{\text{inc}}, \mathbf{E}'_{\text{T}}\right)_{\Gamma_{z^+}} \quad \forall \mathbf{E}' \in V. \tag{58}$$

To ensure a curl-conforming discretization of Eq. (58), we approximate the electric field $\mathbf{E}$ numerically as follows:

$$\mathbf{E}_h(\mathbf{x}) = \sum_{j=1}^{N_h} c_j \mathbf{N}_j(\mathbf{x}), \tag{59}$$

where $\mathbf{N}_j$ denotes Nédélec basis functions of the first kind [59,60] and first- or second-order, defined on a tetrahedral mesh of the domain D. Further details on the discretization are given in Appendix B.

In practice, one is often interested in reflection and transmission coefficients, in addition to the field solution $\mathbf{E}$ itself. Therefore, we define the (complex-valued) scattering parameters as (affine-) linear functionals of $\mathbf{E}$

$$S_{\alpha,mn} := \left(\mathbf{E}_{\mathrm{T}} - \mathbf{E}_{\mathrm{T}}^{\mathrm{inc}}, \pi_{\mathrm{T}}[\mathbf{E}_{\alpha,mn}]\right)_{\Gamma_{z^+}}, \tag{60}$$

where $\alpha \in \{\mathrm{TE}, \mathrm{TM}\}, m \in \mathbb{Z}, n \in \mathbb{Z}$, and $\mathbf{E}_{\alpha,mn}$ are Floquet modes defined in Appendix A. The scattering parameters are considered as QoIs, in the context of the present work.

3.2 Parametrized Model

In this section, we specify the material distribution of the complex permittivity ε. In particular, we assume a linear material behavior for ε and μ inside D. Let the domain D be composed of M nonoverlapping subdomains D_m, i.e., $\overline{D} = \bigcup_{m=1}^{M} \overline{D_m}$. We further assume that the dispersive permittivity $\varepsilon(\mathbf{x}, \omega)$ is spatially piecewise constant on each subdomain D_m and depends smoothly on a given vector of N parameters $\mathbf{y} \in \Xi \subset \mathbb{R}^N$

$$\varepsilon(\mathbf{x}, \omega, \mathbf{y}) = \sum_{m=1}^{M} \varepsilon_m(\omega, \mathbf{y}) 1_m(\mathbf{x}, \mathbf{y}), \tag{61}$$

where

$$1_m(\mathbf{x}, \mathbf{y}) = \begin{cases} 1, & \mathbf{x} \in D_m(\mathbf{y}), \\ 0, & \mathbf{x} \notin D_m(\mathbf{y}). \end{cases}$$

On the one hand, the parameter vector $\mathbf{y}$ can be used to represent variations in the material parameters, e.g., different permittivities, refractive indices, or extinction coefficients, by changing the coefficients $\varepsilon_m(\omega, \mathbf{y})$. On the other hand, it also represents geometric variations of the structure inside the unit cell, since the subdomains $D_m(\mathbf{y})$ for each material depend on $\mathbf{y}$ as well.

The parametrized weak formulation reads: find $\mathbf{E}(\mathbf{y}) \in V$ s.t.

$$a_{\mathbf{y}}(\mathbf{E}(\mathbf{y}), \mathbf{E}') = l(\mathbf{E}') \quad \forall \mathbf{E}' \in V, \tag{62}$$

where

$$a_{\mathbf{y}}(\mathbf{E}, \mathbf{E}') := \left(\mu_{\mathrm{r}}^{-1} \nabla \times \mathbf{E}(\mathbf{y}), \nabla \times \mathbf{E}'\right)_D - \omega^2 \mu_0 (\varepsilon(\mathbf{y}) \mathbf{E}(\mathbf{y}), \mathbf{E}')_D + j\omega\mu_0 \left(\boldsymbol{\mathcal{G}}\left(\mathbf{E}(\mathbf{y})\right), \mathbf{E}'_{\mathrm{T}}\right)_{\Gamma_{z^+}}. \tag{63}$$

The parameter-dependent scattering parameters are given as follows:

$$S_{\alpha,mn}(\mathbf{y}) = \left(\mathbf{E}_{\mathrm{T}}(\mathbf{y}) - \mathbf{E}_{\mathrm{T}}^{\mathrm{inc}}, \pi_{\mathrm{T}}[\mathbf{E}_{\alpha,mn}]\right)_{\Gamma_{z^+}}, \tag{64}$$

where $\alpha \in \{\mathrm{TE}, \mathrm{TM}\}$.

Remark 1. Relating the model problem of scattering in periodic media to the UQ methodology of the previous section, the linear functional $J_{\mathbf{y}}(\cdot)$ is given by

$$S_{\alpha,mn}(\mathbf{y}) = \underbrace{(\mathbf{E}_{\mathrm{T}}(\mathbf{y}), \pi_{\mathrm{T}}[\mathbf{E}_{\alpha,mn}])_{\Gamma_{z^+}}}_{=J_{\mathbf{y}}(\mathbf{E})} - \left(\mathbf{E}_{\mathrm{T}}^{\mathrm{inc}}, \pi_{\mathrm{T}}[\mathbf{E}_{\alpha,mn}]\right)_{\Gamma_{z^+}},$$

where $\alpha \in \{\mathrm{TE}, \mathrm{TM}\}$. The strong formulation of the adjoint problem (42) reads

$$\nabla \times \left(\frac{1}{\mu_r^*} \nabla \times \mathbf{z}\right) - \omega^2 \mu_0 \varepsilon^* \mathbf{z} = 0 \quad \text{in } D, \tag{65a}$$

$$\mathbf{z}_{\mathrm{T}}|_{\Gamma_{x^+}} e^{-j\psi_x} = \mathbf{z}_{\mathrm{T}}|_{\Gamma_{x^-}} \quad \text{on } \Gamma_{x^+} \cup \Gamma_{x^-}, \tag{65b}$$

$$\mathbf{z}_{\mathrm{T}}|_{\Gamma_{y^+}} e^{-j\psi_y} = \mathbf{z}_{\mathrm{T}}|_{\Gamma_{y^-}} \quad \text{on } \Gamma_{y^+} \cup \Gamma_{y^-}, \tag{65c}$$

$$\mathbf{z}_{\mathrm{t}} = 0 \quad \text{on } \Gamma_{z^-}, \tag{65d}$$

$$\mathbf{e}_z \times \left(\frac{j}{\omega\mu_0} \nabla \times \mathbf{z}\right) + \overline{\boldsymbol{\mathcal{G}}} = \overline{\boldsymbol{\mathcal{F}}} \quad \text{on } \Gamma_{z^+}, \tag{65e}$$

where $\overline{\mathcal{G}}$ and $\overline{\mathcal{F}}$ are defined in Appendix A.

Introducing $\mathbf{A}_{\mathrm{dof}}$ as system matrix arising from the discretization of Eq. (58), the discrete primal problem is given by

$$\mathbf{A}_{\mathrm{dof}}\mathbf{c}_{\mathrm{dof}} = \mathbf{f}_{\mathrm{dof}}, \tag{66}$$

as derived in Appendix B. Discretization of the adjoint problem (42) yields the discrete matrix equation

$$\mathbf{A}_{\mathrm{dof}}^{\mathrm{H}}\mathbf{z}_{\mathrm{dof}} = \mathbf{J}_{\mathrm{dof}}, \tag{67}$$

and the discrete version of the error indicator (46) reads

$$\tilde{\eta}_h(\mathbf{y}) = \tilde{\mathbf{z}}_{\mathrm{dof}}^{\mathrm{H}}(\mathbf{y})\mathbf{f}_{\mathrm{dof}}(\mathbf{y}) - \tilde{\mathbf{z}}_{\mathrm{dof}}^{\mathrm{H}}(\mathbf{y})\mathbf{A}_{\mathrm{dof}}(\mathbf{y})\tilde{\mathbf{c}}_{\mathrm{dof}}(\mathbf{y}). \tag{68}$$

Note that the dual solution can be obtained with negligible cost in many cases, e.g., if the primal problem is solved with a sparse LU decomposition $\mathbf{A}_{\mathrm{dof}} = \mathbf{L}\mathbf{U}$, for the respective dual problem, we obtain

$$\mathbf{A}_{\mathrm{dof}}^{\mathrm{H}} = (\mathbf{L}\mathbf{U})^{\mathrm{H}} = \mathbf{U}^{\mathrm{H}}\mathbf{L}^{\mathrm{H}}. \tag{69}$$

Remark 2. To prove that Eq. (62) is well-posed and analytic with respect to the model parameters, which are the main working assumptions of the paper, requires special care due to the presence of the general boundary operator $\mathcal{G}$. Here we only refer to [59] (Chapter 4) and [61] (Section 5) for a numerical analysis of well-posedness in the deterministic setting with simpler boundary conditions. Also, recently, shape holomorphy for Maxwell's source problem was established in [62], considering bi-Lipschitz shape transformations and holomorphic material parameters, which are bounded away from zero. However, the analysis was, again, carried out with homogeneous Dirichlet boundary conditions.

4. APPLICATION

We apply the enhanced surrogate modeling presented in Section 2 to a nontrivial benchmark application from nanoplasmonics, namely, an optical grating coupler [32,63]. We report some details on modeling uncertainties in material and geometric input data. We also describe how parametric variations are realized numerically and finally quantify uncertainties in the computational model.

For general periodic structures, we must distinguish between two types of uncertainties. In this work, we focus on global uncertainties, i.e., we assume that all unit cells are identically affected, modeling a systematic offset in the fabrication process. We do not address local uncertainties leading to a violation of the periodicity and different unit cells. Readers interested in the latter case are referred to [34] for a relevant study.

4.1 Numerical Model

The considered grating [32] couples power from an incident transverse magnetic (TM) polarized plane wave, such that

$$\pi_{\mathrm{T}}\left[\mathbf{E}^{\mathrm{inc}}\right] = \pi_{\mathrm{T}}[\mathbf{E}_{\mathrm{TM},00}], \quad \text{at} \quad \Gamma_{z^+},$$

with propagation direction $\theta^{\mathrm{inc}} = 53°$, $\phi^{\mathrm{inc}} = 0°$, directly into a metal-insulator-metal (MIM) plasmon mode, which is illustrated in Fig. 9(a).

The structure's design, shown in Fig. 9(b), is assumed to be periodic in the x direction and infinitely extended in the y direction. The reflection coefficients (60) at the upper boundary Γ_{z^+} correspond to the coupling efficiency of the structure, such that larger reflection coefficients indicate a lower coupling efficiency. Therefore, the scattering parameter $S := S_{\mathrm{TM},00}$ is considered as the QoI in the following. Note that we focus on the fundamental reflection coefficient $S_{\mathrm{TM},00}$ because, for this particular model, all other scattering parameters have negligible amplitudes.

We model the material properties based on measurement data for noble metals provided by Johnson and Christy [64] and presented in Table 1. We focus on the frequency range of $f_{\mathrm{min}} = 400$ THz to $f_{\mathrm{max}} = 430$ THz (see Fig. 10).

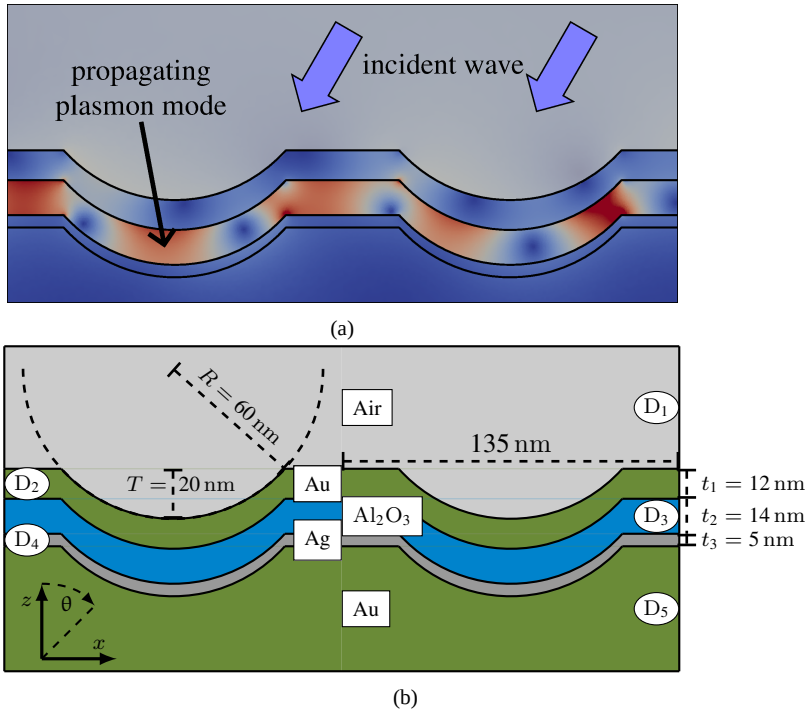


FIG. 9: An optical grating coupler [32] couples power from an incident plane wave in free space directly into a MIM plasmon mode, propagating in horizontal direction: (a) illustration of the magnitude of the electric field at a specific (arbitrary chosen) point in time and (b) shows the design of the considered grating coupler. Note that the structure is periodically extended in the horizontal direction; only two unit cells are shown.

TABLE 1: Material data, based on the results of [64]

Index i	**Energy (eV)**	**Frequency** f_i **(THz)**	**Refractive index** n_i^{Au}	**Extinction coefficient** κ_i^{Au}	**Refractive index** n_i^{Ag}	**Extinction coefficient** κ_i^{Ag}
0	1.64	396.55	0.14 ± 0.02	4.542 ± 0.015	0.03 ± 0.02	5.242 ± 0.015
1	1.76	425.57	0.13 ± 0.02	4.103 ± 0.010	0.04 ± 0.02	4.838 ± 0.010
2	1.88	454.58	0.14 ± 0.02	3.697 ± 0.007	0.05 ± 0.02	4.483 ± 0.007

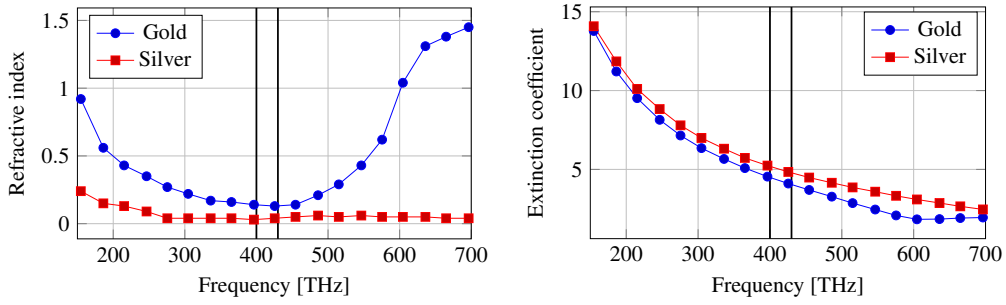


FIG. 10: Dispersive optical constants of gold and silver [64]. The black vertical lines define the considered frequency range.

The data are experimentally determined by reflectivity studies and therefore given in terms of the refractive indices n and the extinction coefficients κ for gold and silver, respectively. From those, one obtains the complex permittivity as in [65] (Chapter 1.1), i.e.,

$$\varepsilon = \left(n^2 - \kappa^2 - j(2n\kappa)\right)\varepsilon_0. \tag{70}$$

We interpolate the given material β_i^α data as follows:

$$\beta^\alpha(\omega) = \sum_{i=0}^{2} \beta_i^\alpha l_i(\omega), \tag{71}$$

where $\alpha \in \{\text{Au}, \text{Ag}\}$, $\beta \in \{n, \kappa\}$. Also,

$$l_i(\omega) = \prod_{j=0, j\neq i}^{2} \frac{\omega - \omega_j}{\omega_i - \omega_j}, \quad \omega_i = 2\pi f_i, \tag{72}$$

are second-order Lagrange polynomials and f_i, n_i^{Au}, κ_i^{Au}, n_i^{Ag}, κ_i^{Ag}, and $i = 0, 1, 2$ are given.

We proceed with the description of the deterministic numerical model, as well as its parametrization. The periodic mesh for the nominal design is created using GMSH [66]. Since, for this particular structure, only the fundamental Floquet modes propagate and all higher order modes are attenuated to a negligible amplitude at Γ_{z^+}, we can use the first-order Floquet boundary condition (A.12). We use FENICS [67] as the FE library to assemble the FE matrix $\mathbf{A}$ and right-hand side (rhs) $\mathbf{f}$ [see Eq. (B.1)], as well as the linear functional $\mathbf{J}_{\text{dof}}$ used for the numerical approximation of the scattering parameter

$$S_{\text{TM},00} = \left(\mathbf{E}_{\text{T}} - \mathbf{E}_{\text{T}}^{\text{inc}}, \pi_{\text{T}}[\mathbf{E}_{\text{TM},00}]\right)_{\Gamma_{z^+}} \approx \mathbf{J}_{\text{dof}}^{\text{H}} \mathbf{c}_{\text{dof}} - 1.$$

Since FENICS 2017.2.0 is not able to deal with complex numbers, we assemble the real and the imaginary parts of the matrix and the vectors separately. We then use NUMPY and SCIPY to impose the quasi-periodic boundary conditions (53a) and (53b) and solve the resulting linear system (B.1) with a sparse LU decomposition. Using second-order Nédélec elements of the first kind, we end up with 56,200 degrees of freedom (DoF)s and achieve an accuracy of $\approx 10^{-3}$ in the scattering parameter. The reference solutions for different frequency sample points are computed with a commercial software [68] employing an adaptively refined mesh of higher order curved elements. Since a sparse LU decomposition is used to solve the resulting linear system, the adjoint solution $\mathbf{z}_{\text{dof}}$ is obtained with negligible costs.

To incorporate changes in the geometry parameters without the need to remesh, a design element approach is applied [69]. We describe all (material) interfaces, illustrated in Fig. 11, using nonuniform rational B-splines (NURBS) [70]. Each NURBS curve

$$\mathbf{C}_i(\xi; \mathbf{y}) = \sum_{j=0}^{n} R_j(\xi)\mathbf{P}_{j,i}(\mathbf{y}), \quad \xi \in [0, 1], \tag{73}$$

is a superposition of rational basis functions $R_j(\xi)$ weighted by control points $\mathbf{P}_j$. We then define mappings

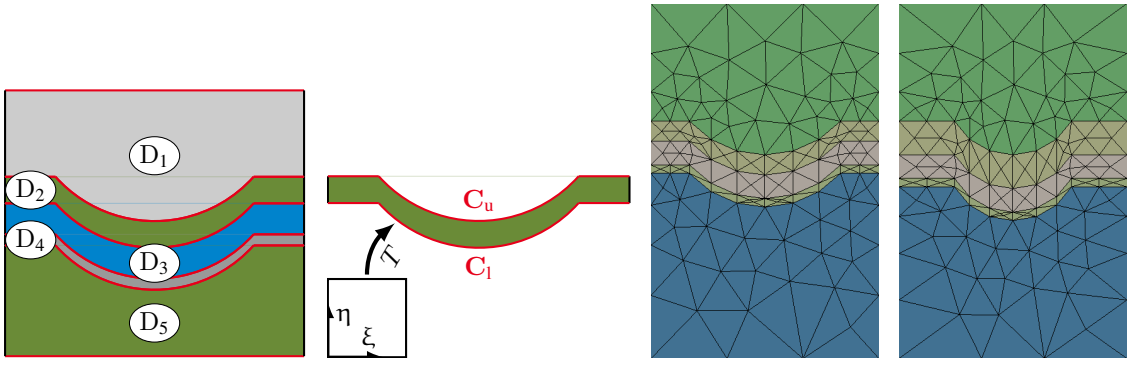


FIG. 11: (a) Design elements, (b) mapping from unit square, (c) initial mesh (coarse for illustration) for nominal design $\mathbf{y}^{\text{nominal}}$, and (d) deformed mesh for $R = 40$ nm and $t_1 = 20$ nm

$$\mathbf{T}_m(\xi,\eta;\mathbf{y}) = \eta\mathbf{C}_{m,\mathrm{u}}(\xi;\mathbf{y}) + (1-\eta)\mathbf{C}_{m,\mathrm{l}}(\xi;\mathbf{y}), \quad m = 1,\dots,M \tag{74}$$

from the unit square $[0 \le \xi \le 1] \times [0 \le \eta \le 1]$ to each design element $D_i(\mathbf{y})$ (see Fig. 11). Thereby, the subscripts u and l refer to the upper and lower NURBS curve of the design element, respectively. Given the initial mesh, for each mesh node the respective coordinates on the unit square are found by solving a nonlinear root-finding problem. We can then deform the mesh by moving the mesh nodes to the new coordinates obtained by evaluating the mapping (74) for different geometry parameters $\mathbf{y}$.

We consider a fixed frequency $\omega = 2\pi(414\,\mathrm{THz})$ and $N = 17$ random input parameters $\mathbf{Y}$. These are the five geometrical parameters presented in Table 2 and the 12 material parameters given in Table 1. As introduced in Section 3, both the uncertain geometry and the uncertain material coefficients are modeled by an uncertain complex permittivity $\varepsilon(\mathbf{x},\mathbf{y})$; see Eq. (61).

We assume that the RVs Y_n, $n = 1,2,\dots,N = 17$, are independent and distributed in the ranges defined by their nominal values and variations. The variations of the material parameters are chosen according to the error estimate provided by Johnson and Christy [64] "*based on the instrumental accuracy of the reflection and transmission measurements*." Since no further information on the distributions of those measurement uncertainties are specified, the given error estimate is assumed to correspond to a 2σ interval. For the geometrical parameters, only small variations in the range of $\pm$ 1.5 nm are considered (with a 2σ interval of $\pm$ 1 nm).

We opt for beta distributions, which have bounded support and can approximate normal distributions for suitable choices of their shape parameters (see Appendix B in [42]). The shape parameters are chosen based on the results of a series of Kolmogorov-Smirnov fitting tests [71].

4.1.1 Numerical Results

To illustrate the benefits of using the adjoint error indicator presented in Section 2.3, we consider here only the thickness of the upper gold layer t_1 and the thickness of the dielectric layer t_2 as input parameters. We have observed, numerically, that the QoI is particularly sensitive with respect to these parameters, with slow associated univariate convergence rates. Figure 12(a) shows the S-parameter with respect to small variations of these two parameters. We construct (mapped) adaptive Leja approximations using Algorithm 1 and the adjoint-based Algorithm 2. The accuracy of the surrogate models is measured using a cross-validation set of $N^{\mathrm{cv}} = 1 \times 10^3$ parameter realizations $S^{(i)} := S(\mathbf{y}^{(i)})$, $i = 1,\dots,N^{\mathrm{cv}}$, drawn according to the underlying PDF, which is used to compute a discrete approximation of the L^1_ϱ error

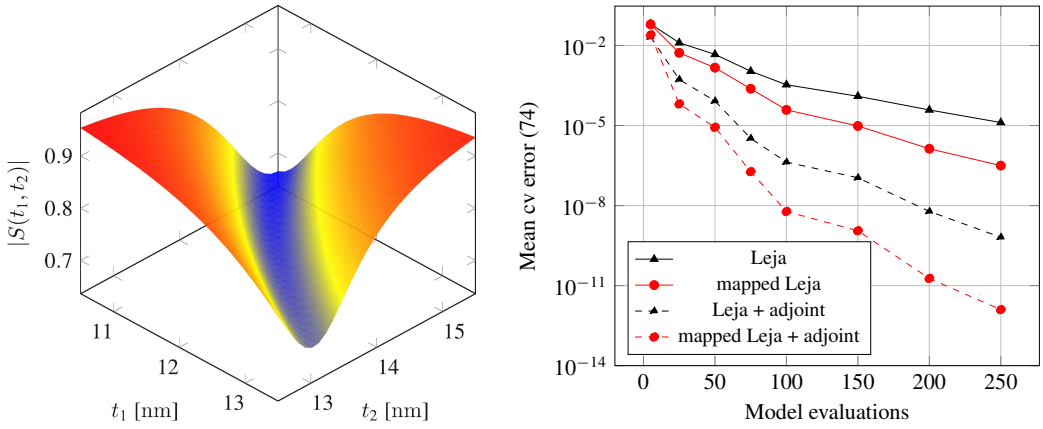
$$\mathbb{E}[|S - \tilde{S}|] \approx \frac{1}{N^{\mathrm{cv}}} \sum_{i=1}^{N^{\mathrm{cv}}} |S^{(i)} - \tilde{S}^{(i)}|. \tag{75}$$

The error (75) is computed for the mapped and adjoint-based approximation, as well as for nonmapped and/or nonadjoint-based variants, for increasing numbers of model evaluations. The corresponding results are shown in Fig. 12(b). The plot numerically confirms (47) and shows the doubled convergence order of the adjoint-error indicator. Additionally, it can already be observed that employing the conformal sausage map $g_{\mathrm{S}}(\cdot;9)$, defined in Eq. (12), yields to a significant improvement of both convergence orders in the considered setting.

Next, we consider all $N = 17$ input parameters and construct different polynomial approximations. As a reference, we compute two nonadaptive approximations, based on generalized polynomial chaos (gPC) [7] and on

TABLE 2: Uncertain geometrical parameters

Parameter	Nominal value (nm)	Variation (nm)
Grating radius, R	60	$\pm$ 1.5
Gold layer thickness, t_1	12	$\pm$ 1.5
Alumina layer thickness, t_2	14	$\pm$ 1.5
Silver layer thickness, t_3	5	$\pm$ 1.5
Grating depth, T	20	$\pm$ 1.5



(a) (b)

FIG. 12: Considering a two-dimensional parameter space and a fixed frequency of 414 THz. (a) Reflection coefficient $|S|$ and (b) improved convergence of (mapped) Leja approximations by employing an adjoint-error indicator

isotropic Smolyak sparse-grid interpolation [72]. These are compared to the proposed enhanced surrogate modeling, i.e., (mapped) Leja adaptive approximations, using both Algorithm 1 and the adjoint-based Algorithm 2 for the latter. Chaospy [73] is used for the gPC case, the Sparse-Grid-MATLAB-Kit [72] is employed for the Smolyak sparse-grid interpolation, while an in-house code was developed for both Leja adaptive algorithms [74]. We compare the resulting surrogate models with respect to accuracy and computational costs.

The computational costs refer to the number of model evaluations needed for the construction of the approximation. Although straightforward for the gPC, Smolyak sparse-grid, and the Leja adaptive Algorithm 1, the estimation of costs is more involved in the case of the adjoint-based Algorithm 2. First, in order to evaluate the duality-based error indicator (68) at a candidate point, it is sufficient to evaluate a residual of Eq. (62). Therefore, we distinguish between residual evaluations and solver calls, where in most cases the costs to evaluate the residuals are almost negligible compared to the solver costs, i.e., assembly and sparse LU decomposition of the system matrices $\mathbf{A}_{\text{dof}}(\mathbf{y})$. Second, the additional costs for computing the dual solution $\mathbf{z}$ by forward and backward substitution can also be neglected in most cases, since the primal problem is solved with a sparse LU decomposition.

As before, the accuracy of the surrogate models is measured using a cross-validation set of $N^{\text{cv}} = 1 \times 10^3$ parameter realizations $S^{(i)} := S(\mathbf{y}^{(i)})$, $i = 1, \ldots, N^{\text{cv}}$, drawn according to the underlying PDF. In addition to Eq. (75), we also consider the maximum error over all sample points

$$\max_{i=1,\ldots,N^{\text{cv}}} |S^{(i)} - \tilde{S}^{(i)}|. \tag{76}$$

All accuracy and cost results are presented in Table 3. First, a gPC approximation with a second-order total-degree polynomial basis, i.e., 171 Jacobi polynomials, is constructed. The polynomial coefficients are computed with a sparse second-order Gauss quadrature formula, resulting in 613 quadrature nodes, accordingly, model evaluations. Second, we employ interpolation on an isotropic Smolyak sparse-grid of level 2 based on Clenshaw Curtis nodes, which requires 613 model evaluations as well. Accordingly, we set a budget $B = 613$ for the classical, i.e., without adjoints or conformal maps, Leja adaptive Algorithm 1, such that its costs are identical to the gPC and the Smolyak sparse-grid interpolation. As can be seen in Table 3, the Leja adaptive approximation is about one order of magnitude more accurate than the gPC and also significantly better than the isotropic sparse-grid interpolation.

Next, we use again Algorithm 1 but employ conformal maps. In particular, we refer with iso-mapped to applying the conformal sausage map $g_{\text{S}}(\cdot; 9)$ for all parameters while aniso-mapped refers to the application of the conformal map only to the parameters t_1, t_2, T (parameters with a particularly slow univariate convergence rate). It can be observed that both approaches yield a similar improvement in terms of accuracy, without (relevant) extra computational cost.

TABLE 3: Accuracy and computational cost of different polynomial approximations for 17 input RVs. #LU refers to the dominating costs for the assembly and sparse LU decomposition of the system matrices. #FB and #Res denote the number of forward-backward substitutions and residual evaluations, respectively

	#LU	**#FB**	**#Res**	**Max. Error (75)**	**Mean Error (74)**
Total-degree gPC (without maps)	***613***	*613*	0	7.61×10^{-1}	1.92×10^{-1}
Level 2 Smolyak sparse-grid (without maps)	***613***	*613*	0	1.73×10^{-1}	4.45×10^{-2}
Ad. Leja (without adjoints/maps)	***613***	*613*	0	8.56×10^{-2}	5.53×10^{-3}
Ad. iso-mapped Leja (without adjoints)	***613***	*613*	0	3.59×10^{-2}	4.10×10^{-3}
Ad. aniso-mapped Leja (without adjoints)	***613***	*613*	0	3.85×10^{-2}	4.15×10^{-3}
Ad. Leja (with adjoints; without maps)	**558**	1116	*613*	8.46×10^{-2}	5.49×10^{-3}
Ad. iso-mapped Leja (with adjoints)	**563**	1126	*613*	3.57×10^{-2}	4.09×10^{-3}
Ad. aniso-mapped Leja (with adjoints)	**563**	1126	*613*	3.82×10^{-2}	4.15×10^{-3}
Ad. aniso-mapped Leja (with adjoints and error correction)	$\mathbf{3 \times 10^3}$	6×10^3	3×10^4	1.25×10^{-3}	1.20×10^{-4}

For the adjoint-based Algorithm 2, we then compute approximations using again 613 (mapped) polynomials, resulting in errors almost identical to the nonadjoint case. However, since the costs can be predominantly attributed to the $\approx$ 560 solver calls, the costs are reduced. Note that greater (relative) gains can be observed in different settings, e.g., when a smaller computational budget is used or less parameter anisotropy is present in the considered model. In particular, as a numerical test case, we increased material uncertainties by one-third and reduced geometric variations to the range of $\pm$ 0.25 nm. In that case, the respective computational cost was reduced by $>$ 50%.

The convergence of the mean error (75) with respect to function calls (corresponding to the number of LU decompositions) of the investigated spectral methods is additionally shown in Fig. 13. Isotropic gPC does not seem to show a proper convergence. However, it should be noted that we were only able to compute approximations up to order 3 due to the larger number of parameters and hence, the error decay is probably pre-asymptotic. Additional convergence results supporting this hypothesis are reported in Appendix C. It can be concluded that the nonadaptive gPC reference solution achieves a very poor accuracy with the given computational budget. All considered dimension-adaptive schemes greatly outperform the isotropic approaches. In accordance with the results in Table 3, Fig. 13 illustrates that the application of the conformal map leads to improvements compared to the classical Leja algorithm; whereas in this setting, there is a negligible difference between the iso-mapped and aniso-mapped approaches. It

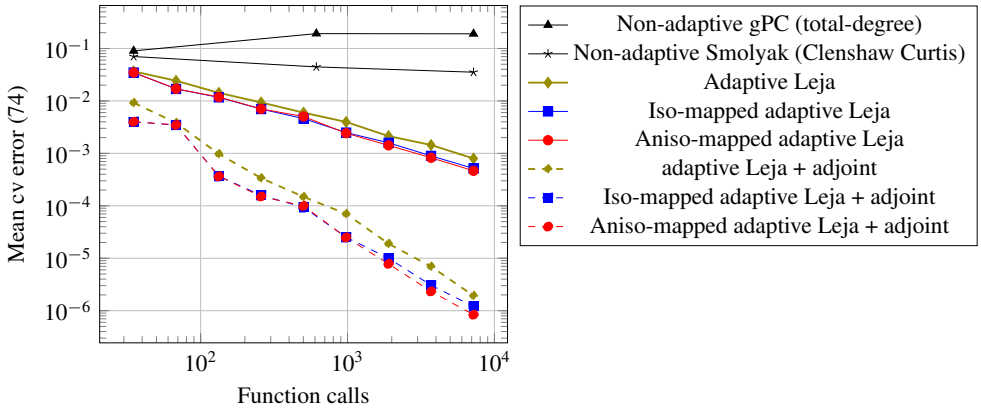


FIG. 13: Convergence study for the single frequency setting. The adaptive schemes clearly outperform the isotropic approach. Note that the dashed lines correspond to the error of improved surrogate models, which require the evaluation of FE residuals at each cross-validation point.

should be noted that, in contrast to Table 3, the dashed lines in Fig. 13 correspond to the error of the respective adjoint-based error indicator (68). Therefore, they do not correspond to (mapped) polynomial surrogate models but require the evaluation of a residual of (62) at each cross-validation sample point $\mathbf{y}^{(i)}$, $i = 1, \dots, N^{\text{cv}}$.

Finally, as shown in the last row of Table 3, we compute a very accurate surrogate model, by using the adjoint-based Algorithm 2 with a computational budget of 3×10^3 LU decomposition. The adjoint-based approximation is then refined by employing (49) until 3×10^4 polynomials are used, further reducing the error by more than one-order of magnitude. It shall be highlighted that the adjoint-based approach results in tremendous computational savings compared to the classical Leja Algorithm 1 since 27×10^3 full model evaluations could be avoided in this particular setting.

As often pointed out in the literature (see, e.g., [24]), it is inefficient to reduce the stochastic error below the discretization error. Therefore, the stochastic approximation is not further refined and the most accurate surrogate model (Table 3, last row) is in Section 4.1.2 used to compute statistical measures of the absolute value of the scattering parameter $|S|$.

4.1.2 Postprocessing the Surrogate Model

Since the (mapped) polynomial surrogate model $\tilde{S}(\mathbf{y})$ can be evaluated inexpensively, we employ a Monte Carlo-based approach by evaluating the surrogate model on a large number of N^{MC} parameter sample points, drawn from the joint PDF $\varrho(\mathbf{y})$. We then use the sample evaluations to estimate statistical moments of $|S|$, its PDF, failure probabilities based on specific design criteria, and its sensitivity with respect to the input parameters.

The expected value $\mathbb{E}[|S|]$ and the variance $\mathbb{V}[|S|]$ are estimated as follows:

$$\mathbb{E}[|S|] = \int_{\Xi} |S(\mathbf{y})| \varrho(\mathbf{y}) \, \mathrm{d}\mathbf{y} \approx \frac{1}{N^{\text{MC}}} \sum_{i=1}^{N^{\text{MC}}} |\tilde{S}^{(i)}| =: E^{\text{MC}}[|\tilde{S}|], \tag{77a}$$

$$\mathbb{V}[|S|] = \int_{\Xi} (|S(\mathbf{y})| - \mathbb{E}[|S|])^2 \varrho(\mathbf{y}) \, \mathrm{d}\mathbf{y} \approx \frac{1}{N^{\text{MC}} - 1} \sum_{i=1}^{N^{\text{MC}}} \Big(|\tilde{S}^{(i)}| - E^{\text{MC}}[|\tilde{S}|] \Big)^2. \tag{77b}$$

We estimate the failure probability $\mathcal{F} = P(|S| \geq 1 - \alpha)$ as follows:

$$\mathcal{F} = P(|S| \geq 1 - \alpha) = \int_{\Xi} \mathcal{I}_{\mathcal{F}}(S(\mathbf{y})) \varrho(\mathbf{y}) \; \mathrm{d}\mathbf{y} \approx \frac{1}{N^{\text{MC}}} \sum_{i=1}^{N^{\text{MC}}} \mathcal{I}_{\mathcal{F}} \Big(\tilde{S}^{(i)} \Big), \tag{78}$$

where ϱ_S denotes the PDF of $|S|$ and $\mathcal{I}_{\mathcal{F}}$ denotes the indicator function

$$\mathcal{I}_{\mathcal{F}}(S) = \begin{cases} 1, & |S| \in [1 - \alpha, 1], \\ 0, & |S| \in [0, 1 - \alpha). \end{cases} \tag{79}$$

Monte Carlo sampling in combination with surrogate modeling is used for simplicity here. However, it should be noted that equality in (78) for $N^{\text{cv}} \to \infty$ cannot be guaranteed, in general; see [75] for counterexamples and possible extensions.

The PDF ϱ_S of $|S|$ is estimated by employing a kernel density estimator

$$\varrho_S \approx \tilde{\varrho}_T := \frac{1}{h N^{\text{MC}}} \sum_{i=1}^{N^{\text{MC}}} K \left(\frac{T - |\tilde{S}^{(i)}|}{h} \right), \tag{80}$$

with $N^{\text{MC}} = 10^7$ samples, bandwidth $h = 10^{-3}$, and the Epanechnikov kernel [76]

$$K(T) := \begin{cases} \frac{3}{4}(1 - T^2), & T \in [-1, 1], \\ 0, & \text{else}. \end{cases} \tag{81}$$

The estimated expected values, standard deviations $\sqrt{\mathbb{V}}$, and failure probabilities for an increasing sample size N^{MC} and $\alpha = 0.1$ are given in Table 4. The estimated PDF $\tilde{\varrho}_S$ is shown in Fig. 14(a).

Sensitivity analysis is based on an analysis of variances (ANOVA) [77]. The related metrics are commonly known as Sobol indices, where we will focus on the so-called main-effect (first-order) and total-effect (total order) indices, defined in [78]. In the context of the present work, estimations of the Sobol indices for the magnitude of the scattering parameter shall be based on sampling of the (mapped) polynomial approximation $\tilde{S} : \Xi \to \mathbb{C}$. We use Saltelli's algorithm [79] with $N^{\mathrm{sens}} = 10^5$ sample points, resulting in $2(17+1)10^5 = 3.6 \times 10^6$ surrogate model evaluations. The main-effect and total-effect Sobol indices for each parameter are given in Fig. 14(b). The thickness of the dielectric layer t_2, the thickness of the upper gold layer t_1, the grating depth T, and the refractive index of the upper gold layer n_1^{Au} are identified as the most sensitive parameters. Moreover, since the sum of all main-effect sensitivity indices is $\sim 33\%$, the remaining 67% indicates higher order interactions, and thus strong coupling among the input parameters.

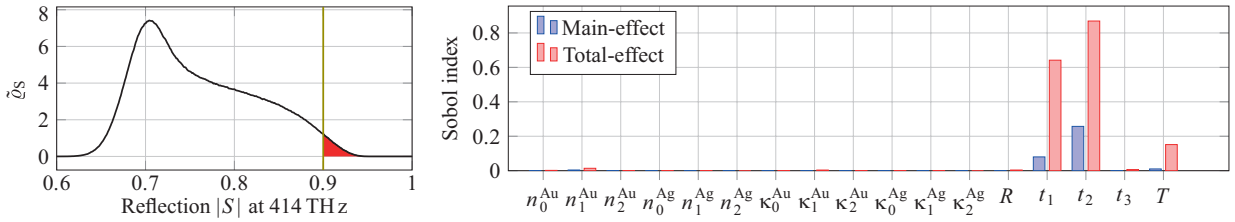
It is found that the considered model is highly sensitive to small geometrical variations. In particular, while geometrical variations in a range of only $\pm$ 1.5 nm are considered, their impact is significantly higher than the one attributed to material uncertainty, which was modeled based on the measurement error provided by [64].

5. CONCLUSION

In this work, we presented an efficient method to quantify uncertainties in the Maxwell source problem, assuming a moderately large number of input RVs. Dimension adaptivity in combination with adjoint error correction and conformal maps is confirmed to be a promising technique to delay the curse of dimensionality. For the considered FE model from nanoplasmonics, the comparison of the proposed adaptive algorithm with total degree gPC and isotropic Smolyak sparse grids shows significant gains in both accuracy and computational costs. In particular, with the adaptive scheme we were able to consider up to 17 parameters and achieve an accuracy of $\approx 10^{-3}$ with a few thousand numerical solutions of the deterministic model. To consider wider frequency ranges with possible poles in combination with large geometric uncertainties, a combination of polynomial and rational approximations is a topic of future research.

TABLE 4: Expectation, standard deviation, and failure probability, i.e., $\mathcal{F} = P(|S| \geq 1 - \alpha)$ for $\alpha = 0.1$; surrogate-based Monte Carlo estimation using N^{MC} sample points

N^{MC}	$\mathbb{E}$	$\sqrt{\mathbb{V}}$	$\mathcal{F}$ (%)
10^3	0.7595	0.0661	2.20
10^4	0.7605	0.0658	1.85
10^5	0.7606	0.0660	2.04
10^6	0.7607	0.0660	2.07
10^7	0.7607	0.0660	2.06



(a) (b)

FIG. 14: PDF, expectation, standard deviation, failure probability, and Sobol indices for 17 beta-distributed input parameters: (a) estimated PDF of scattering parameter, failure probability is highlighted and (b) Sobol indices computed with 10^5 sample points

For the considered optical grating coupler, according to Sobol-sensitivity measures, geometrical parameters have been found to be the dominant source of input uncertainty. Although the modeling of their probability distributions could not be based on measurement data yet, this conclusion is substantiated by the very conservative choice of geometrical standard deviations.

ACKNOWLEDGMENTS

The authors would like to thank R. Schuhmann for valuable discussions on the the topic of UQ in plasmonics and L. Scarabosio for helpful discussions on sparse approximation. U. Römer and N. Georg acknowledge the funding by the Deutsche Forschungsgemeinschaft [(DFG) German Research Foundation] Grant No. RO4937/1-1. The work of N. Georg, D. Loukrezis, and S. Schöps is also supported by the Excellence Initiative of the German Federal and State Governments and the Graduate School of Computational Engineering at Technische Universität Darmstadt. The work of D. Loukrezis is further supported by BMBF via the Research Contract No. 05K19RDB.

REFERENCES

1. Chkifa, A., Cohen, A., and Schwab, C., Breaking the Curse of Dimensionality in Sparse Polynomial Approximation of Parametric PDEs, *J. Math. Pures Appl.*, **103**(2):400–428, 2015.
2. Scarabosio, L., Shape Uncertainty Quantification for Scattering Transmission Problems, PhD, ETH Zurich, 2016.
3. Xiu, D., Fast Numerical Methods for Stochastic Computations: A Review, *Commun. Comput. Phys.*, **5**(2-4):242–272, 2009.
4. Ghanem, R.G. and Spanos, P.D., *Stochastic Finite Elements: A Spectral Approach*, New York: Springer, 1991.
5. Babuška, I., Tempone, R., and Zouraris, G.E., Galerkin Finite Element Approximations of Stochastic Elliptic Partial Differential Equations, *SIAM J. Numer. Anal.*, **42**(2):800–825, 2004.
6. Babuška, I., Nobile, F., and Tempone, R., A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data, *SIAM Rev.*, **2**:317–355, 2010.
7. Xiu, D. and Karniadakis, G.E., The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations, *SIAM J. Sci. Comput.*, **24**(2):619–644, 2002.
8. Xiu, D. and Hesthaven, J.S., High-Order Collocation Methods for Differential Equations with Random Inputs, *SIAM J. Sci. Comput.*, **27**(3):1118–1139, 2005.
9. Chkifa, A., Cohen, A., and Schwab, C., High-Dimensional Adaptive Sparse Polynomial Interpolation and Applications to Parametric PDEs, *Found. Comput. Math.*, **14**(4):601–633, 2014.
10. Gerstner, T. and Griebel, M., Dimension-Adaptive Tensor-Product Quadrature, *Computing*, **71**(1):65–87, 2003.
11. Narayan, A. and Jakeman, J.D., Adaptive Leja Sparse Grid Constructions for Stochastic Collocation and High-Dimensional Approximation, *SIAM J. Sci. Comput.*, **36**(6):A2952–A2983, 2014.
12. Nobile, F., Tempone, R., and Webster, C.G., An Anisotropic Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data, *SIAM J. Numer. Anal.*, **46**(5):2411–2442, 2008.
13. Ernst, O.G., Sprungk, B., and Tamellini, L., Convergence of Sparse Collocation for Functions of Countably Many Gaussian Random Variables (with Application to Elliptic PDEs), *SIAM J. Numer. Anal.*, **56**(2):877–905, 2018.
14. Benner, P. and Schneider, J., Uncertainty Quantification for Maxwell's Equations Using Stochastic Collocation and Model Order Reduction, *Int. J. Uncertain. Quantif.*, **5**(3):195–208, 2015.
15. Benner, P., Gugercin, S., and Willcox, K., A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems, *SIAM Rev.*, **57**(4):483–531, 2015.
16. Bodendiek, A. and Bollhöfer, M., Adaptive-Order Rational Arnoldi-Type Methods in Computational Electromagnetism, *BIT Numer. Math.*, **54**(2):357–380, 2014.
17. Bonizzoni, F., Nobile, F., Perugia, I., and Pradovera, D., Least-Squares Padé Approximation of Parametric and Stochastic Helmholtz Maps, *Numer. Anal.*, arXiv:1805.05031, 2018.
18. Chantrasmi, T., Doostan, A., and Iaccarino, G., Padé–Legendre Approximants for Uncertainty Analysis with Discontinuous Response Surfaces, *J. Comput. Phys.*, **228**(19):7159–7180, 2009.

19. Scarabosio, L., Multilevel Monte Carlo on a High-Dimensional Parameter Space for Transmission Problems with Geometric Uncertainties, *Int. J. Uncertain. Quantif.*, **9**(6):515–541, 2019.
20. Silva-Oelker, G., Aylwin, R., Jerez-Hanckes, C., and Fay, P., Quantifying the Impact of Random Surface Perturbations on Reflective Gratings, *IEEE Trans. Antennas Propag.*, **66**(2):838–847, 2017.
21. Dölz, J., A Higher Order Perturbation Approach for Electromagnetic Scattering Problems on Random Domains, *Numer. Anal.*, arXiv:1907.05501, 2019.
22. Jantsch, P. and Webster, C., Sparse Grid Quadrature Rules based on Conformal Mappings, in *Sparse Grids and Applications-Miami 2016*, New York: Springer, pp. 117–134, 2018.
23. Trefethen, L.N., *Approximation Theory and Approximation Practice*, Vol. 128, Philadelphia: SIAM, 2013.
24. Jakeman, J.D. and Wildey, T., Enhancing Adaptive Sparse Grid Approximations and Improving Refinement Strategies Using Adjoint-Based a Posteriori Error Estimates, *J. Comput. Phys.*, **280**:54–71, 2015.
25. Butler, T., Dawson, C., and Wildey, T., Propagation of Uncertainties Using Improved Surrogate Models, *SIAM/ASA J. Uncertain.*, **1**(1):164–191, 2013.
26. Loukrezis, D., Römer, U., and De Gersem, H., Assessing the Performance of Leja and Clenshaw-Curtis Collocation for Computational Electromagnetics with Random Input Data, *Int. J. Uncertain. Quantif.*, **9**(1):33–57, 2019.
27. Loukrezis, D. and De Gersem, H., Approximation and Uncertainty Quantification of Stochastic Systems with Arbitrary Input Distributions Using Weighted Leja Interpolation, *Numer. Anal.*, arXiv:1904.07709, 2020.
28. Farcas, I.G., Latz, J., Ullmann, E., Neckel, T., and Bungartz, H.J., Multilevel Adaptive Sparse Leja Approximations for Bayesian Inverse Problems, *Stat. Comput.*, arXiv:1904.12204, 2019.
29. van den Bos, L., Sanderse, B., Bierbooms, W., and van Bussel, G., Bayesian Model Calibration with Interpolating Polynomials based on Adaptively Weighted Leja Nodes, *Commun. Comput. Phys.*, **27**(1):33–69, 2019.
30. Loukrezis, D. and De Gersem, H., Adaptive Sparse Polynomial Chaos Expansions via Leja Interpolation, *Numer. Anal.*, arXiv:1911.08312, 2019.
31. Genet, C. and Ebbesen, T.W., Light in Tiny Holes, *Nature*, **445**:39–46, 2007.
32. Preiner, M.J., Shimizu, K.T., White, J.S., and Melosh, N.A., Efficient Optical Coupling into Metal-Insulator-Metal Plasmon Modes with Subwavelength Diffraction Gratings, *Appl. Phys. Lett.*, **92**(11):113109, 2008.
33. Pitelet, A., Schmitt, N., Loukrezis, D., Scheid, C., Gersem, H.D., Ciracì, C., Centeno, E., and Moreau, A., Influence of Spatial Dispersion on Surface Plasmons, Nanoparticles, and Grating Couplers, *J. Opt. Soc. Am. B*, **36**(11):2989–2999, 2019.
34. Schmitt, N., Georg, N., Brière, G., Loukrezis, D., Héron, S., Lanteri, S., Klitis, C., Sorel, M., Römer, U., Gersem, H.D., Vézian, S., and Genevet, P., Optimization and Uncertainty Quantification of Gradient Index Metasurfaces, *Opt. Mater. Express*, **9**(2):892–910, 2019.
35. Loukrezis, D., Galetzka, A., and De Gersem, H., Robust Adaptive Least Squares Polynomial Chaos Expansions in High-Frequency Applications, *Comput. Eng. Finance Sci.*, arXiv:1912.07725, 2019.
36. Weng, T.W., Zhang, Z., Su, Z., Marzouk, Y., Melloni, A., and Daniel, L., Uncertainty Quantification of Silicon Photonic Devices with Correlated and Non-Gaussian Random Parameters, *Opt. Express*, **23**(4):4242–4254, 2015.
37. Hiptmair, R., Scarabosio, L., Schillings, C., and Schwab, C., Large Deformation Shape Uncertainty Quantification in Acoustic Scattering, *Adv. Comput. Math.*, **44**:1475–1518, 2018.
38. Babuška, I., Nobile, F., and Tempone, R., A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data, *SIAM J. Numer. Anal.*, **45**(3):1005–1034, 2007.
39. Jankoski, R., Römer, U., and Schöps, S., Stochastic Modeling of Magnetic Hysteretic Properties by Using Multivariate Random Fields, *Int. J. Uncertain. Quantif.*, **9**(1):85–102, 2019.
40. Lebrun, R. and Dutfoy, A., Do Rosenblatt and Nataf Isoprobabilistic Transformations Really Differ, *Probabilist. Eng. Mech.*, **24**(4):577–584, 2009.
41. Le Maitre, O.P. and Knio, O.M., *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Heidelberg, Germany: Springer, 2010.
42. Xiu, D., *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton, NJ: Princeton University Press, 2010.
43. Barthelmann, V., Novak, E., and Ritter, K., High Dimensional Polynomial Interpolation on Sparse Grids, *Adv. Comput. Math.*, **12**(4):273–288, 2000.

44. Bungartz, H.J. and Griebel, M., Sparse Grids, *Acta Numer.*, **13**:147–269, 2004.
45. Klimke, A. and Wohlmuth, B.I., Algorithm 847: Spinterp: Piecewise Multilinear Hierarchical Sparse Grid Interpolation in MATLAB, *ACM Trans. Math. Softw.*, **31**(4):561–579, 2005.
46. Nobile, F., Tempone, R., and Webster, C.G., A Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data, *SIAM J. Numer. Anal.*, **46**(5):2309–2345, 2008.
47. Schieche, B., *Unsteady Adaptive Stochastic Collocation on Sparse Grids*, PhD, TU Darmstadt, 2012.
48. Hale, N. and Trefethen, L.N., New Quadrature Formulas from Conformal Maps, *SIAM J. Numer. Anal.*, **46**(2):930–948, 2008.
49. Hale, N., On the Use of Conformal Maps to Speed up Numerical Computations, PhD, Oxford University, 2009.
50. Kosloff, D. and Tal-Ezer, H., A Modified Chebyshev Pseudospectral Method with an $\mathcal{O}(N-1)$ Time Step Restriction, *J. Comput. Phys.*, **104**(2):457–469, 1993.
51. Boyd, J.P., *Chebyshev and Fourier Spectral Methods*, 2nd ed., Mineola, NY: Dover Publications, 2001.
52. Berrut, J. and Trefethen, L., Barycentric Lagrange Interpolation, *SIAM Rev.*, **46**(3):501–517, 2004.
53. Smolyak, S.A., Quadrature and Interpolation Formulas for Tensor Products of Certain Classes of Functions, *Dokl. Acad. Nauk SSSR*, **4**:240–243, 1963.
54. Becker, R. and Rannacher, R., An Optimal Control Approach to a Posteriori Error Estimation in Finite Element Methods, *Acta Numer.*, **10**:1–102, 2001.
55. Butler, T., Constantine, P., and Wildey, T., A Posteriori Error Analysis of Parameterized Linear Systems Using Spectral Methods, *SIAM J. Matrix Anal. Appl.*, **33**(1):195–209, 2012.
56. Römer, U. and Schöps, S., Adjoint Error Estimation for a Pseudo-Spectral Approach to Stochastic Field-Circuit Coupled Problems, *Proc. Appl. Math. Mech.*, **15**:711–714, 2015.
57. Teckentrup, A.L., Scheichl, R., Giles, M.B., and Ullmann, E., Further Analysis of Multilevel Monte Carlo Methods for Elliptic PDEs with Random Coefficients, *Numer. Math.*, **125**(3):569–600, 2013.
58. Jin, J.M., *The Finite Element Method in Electromagnetics*, Hoboken, NJ: Wiley, 2015.
59. Monk, P., *Finite Element Methods for Maxwell's Equations*, Oxford: Oxford University Press, 2003.
60. Nedelec, J.C., Mixed Finite Elements in R^3, *Numer. Math.*, **35**(3):315–341, 1980.
61. Hiptmair, R., Finite Elements in Computational Electromagnetism, *Acta Numer.*, **11**:237–339, 2002.
62. Aylwin, R., Jerez-Hanckes, C., Schwab, C., and Zech, J., Domain Uncertainty Quantification in Computational Electromagnetics, *SIAM/ASA J. Uncertain. Quantif.*, **8**(1):301–341, 2020.
63. Dassault Systemes, Optical Applications with CST Microwave Studio, Accessed March 12, 2018, from https://www.cst.com/content/events/downloads/euc2012/talk_5-3-1_cst_euc_2012.pdf
64. Johnson, P.B. and Christy, R.W., Optical Constants of the Noble Metals, *Phys. Rev. B*, **6**(12):4370–4379, 1972.
65. Maier, S.A., *Plasmonics: Fundamentals and Applications*, New York: Springer, 2007.
66. Geuzaine, C. and Remacle, J.F., GMSH: A 3-D Finite Element Mesh Generator with Built-In Pre-And Post-Processing Facilities, *Int. J. Numer. Meth. Eng.*, **79**(11):1309–1331, 2009.
67. Alnæs, M., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M.E., and Wells, G.N., The FEniCS Project Version 1.5, *Arch. Numer. Soft.*, **3**(100):9–23, 2015.
68. Dassault Systèmes, CST Studio Suite, Darmstadt, Germany, 2018.
69. Braibant, V. and Fleury, C., Shape Optimal Design Using B-Splines, *Comput. Methods Appl. Mech. Eng.*, **44**(3):247–267, 1984.
70. Piegl, L. and Tiller, W., *The NURBS Book*, 2nd ed., New York: Springer, 1997.
71. Lopes, R.H., Kolmogorov-Smirnov Test, *International Encyclopedia of Statistical Science*, New York: Springer, pp. 718–720, 2011.
72. Bäck, J., Nobile, F., Tamellini, L., and Tempone, R., *Stochastic Spectral Galerkin and Collocation Methods for PDEs with Random Coefficients: A Numerical Comparison, Spectral and High Order Methods for Partial Differential Equations*, J. Hesthaven and E. Ronquist, Eds., Vol. 76, Springer, pp. 43–62, 2011.
73. Feinberg, J. and Langtangen, H.P., Chaospy: An Open Source Tool for Designing Methods of Uncertainty Quantification, *J. Comput. Sci.*, **11**:46–57, 2015.

74. Loukrezis, D., Dimension-Adaptive Leja Interpolation (DALI), GitHub Repository, from https://github.com/dlouk/DALI3, 2019.
75. Li, J. and Xiu, D., Evaluation of Failure Probability via Surrogate Models, *J. Comput. Phys.*, **229**(23):8966–8980, 2010.
76. Epanechnikov, V.A., Non-Parametric Estimation of a Multivariate Probability Density, *Theor. Probab. Appl.*, **14**(1):153–158, 1969.
77. Sobol, I.M., Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates, *Math. Comput. Simul.*, **55**(1):271–280, 2001.
78. Homma, T. and Saltelli, A., Importance Measures in Global Sensitivity Analysis of Nonlinear Models, *Reliab. Eng. Syst. Safety*, **52**(1):1–17, 1996.
79. Saltelli, A., Making Best Use of Model Evaluations to Compute Sensitivity Indices, *Comput. Phys. Commun.*, **145**(2):280–297, 2002.
80. Bhattacharyya, A.K., *Phased Array Antennas: Floquet Analysis, Synthesis, BFNs and Active Array Systems*, Vol. 179, Hoboken, NJ: Wiley, 2006.
81. Zhu, Y. and Cangellaris, A.C., *Multigrid Finite Element Methods for Electromagnetic Field Modeling*, Vol. 28, Hoboken: Wiley, 2006.

APPENDIX A. FLOQUET BOUNDARY CONDITION

To truncate the structure in the nonperiodic direction at Γ_{z^+}, a Floquet absorbing boundary condition can be derived by splitting the electric field in the unbounded region $z \geq z^+$, where we assume vacuum permittivity ε_0 and vacuum permeability μ_0, as follows:

$$\mathbf{E} = \mathbf{E}^{\text{inc}} + \mathbf{E}^{\text{sc}}, \tag{A.1}$$

where $\mathbf{E}^{\text{inc}}$ and $\mathbf{E}^{\text{sc}}$ represent the known incident field and the unknown scattered field, respectively. As derived in [80] (Chapter 3) and [81] (Chapter 12.2.1), the scattered field $\mathbf{E}^{\text{sc}}$ can be represented as an infinite series of Floquet modes

$$\mathbf{E}^{\text{sc}} = \sum_{\substack{m,n\in\mathbb{Z} \\ \alpha\in\{\text{TE, TM}\}}} c_{\alpha,mn}\mathbf{E}_{\alpha,mn}e^{-j\kappa_{mn}(z-z^+)}, \tag{A.2}$$

where

$$\mathbf{E}_{\text{TE},mn} := \frac{e^{-j(k_{xm}x+k_{yn}y)}(k_{yn}\mathbf{e}_x - k_{xm}\mathbf{e}_y)}{\sqrt{d_x d_y}\sqrt{k_{xm}^2+k_{yn}^2}},$$

$$\mathbf{E}_{\text{TM},mn} := \frac{e^{-j(k_{xm}x+k_{yn}y)}\big(k_{xm}\mathbf{e}_x + k_{yn}\mathbf{e}_y - ((k_{xm}^2+k_{yn}^2)/\kappa_{mn})\mathbf{e}_z\big)}{\sqrt{d_x d_y}\sqrt{k_{xm}^2+k_{yn}^2}},$$

with

$$k_{xm} := k_x^{\text{inc}} + \frac{2\pi m}{d_x}, \quad k_{yn} := k_y^{\text{inc}} + \frac{2\pi n}{d_y}, \quad \kappa_{mn} := \sqrt{k_0^2 - k_{xm}^2 - k_{yn}^2}. \tag{A.3}$$

Thereby, we distinguish between transverse electric (TE) modes $\mathbf{E}_{\text{TE},mn}$ and TM modes $\mathbf{E}_{\text{TM},mn}$, fulfilling $\mathbf{E} \perp \mathbf{e}_z$ and $\mathbf{H} \perp \mathbf{e}_z$, respectively. There exists only a finite number of propagating modes, i.e., $\kappa_{mn} \in \mathbb{R}$, depending on the wavenumber k_0, the angles of incidence θ^{inc}, ϕ^{inc}, and the dimensions d_x, d_y of the unit cell.

We introduce the operators $\pi_{\text{t}}[\mathbf{u}] := \mathbf{e}_z \times \mathbf{u}$ and $\pi_{\text{T}}[\mathbf{u}] := (\mathbf{e}_z \times \mathbf{u}) \times \mathbf{e}_z$ such that

$$\pi_{\text{t}}\Big[\mathbf{H}_{\alpha,mn}e^{-j\kappa_{mn}(z-z^+)}\Big] = \pi_{\text{t}}\bigg[\frac{j}{\omega\mu}\nabla\times\Big(\mathbf{E}_{\alpha,mn}e^{-j\kappa_{mn}(z-z^+)}\Big)\bigg] = -Y_{\alpha,mn}\pi_{\text{T}}\Big[\mathbf{E}_{\alpha,mn}e^{-j\kappa_{mn}(z-z^+)}\Big], \tag{A.4}$$

with

$$Y_{\alpha,mn} := \begin{cases} \dfrac{\kappa_{mn}}{\omega\mu} & \text{for } \alpha = \text{TE}, \\ \dfrac{\omega\epsilon}{\kappa_{mn}} & \text{for } \alpha = \text{TM}. \end{cases}$$

The incident plane wave $\mathbf{E}^{\text{inc}}$ corresponds to the lowest order Floquet modes $\mathbf{E}_{\alpha,00}$ with modal admittance Y^{inc}

$$\pi_{\text{t}}\big[\mathbf{H}^{\text{inc}}\big] = Y^{\text{inc}}\pi_{\text{T}}\big[\mathbf{E}^{\text{inc}}\big], \quad Y^{\text{inc}} := \begin{cases} \dfrac{\sqrt{\epsilon}\cos(\theta^{\text{inc}})}{\sqrt{\mu}} & \text{for } \alpha = \text{TE}, \\ \dfrac{\sqrt{\epsilon}}{\sqrt{\mu}\cos(\theta^{\text{inc}})} & \text{for } \alpha = \text{TM}. \end{cases} \tag{A.5}$$

By taking the cross product of the curl of (A.1) with $\mathbf{e}_z$, the magnetic field above the structure is expressed as follows:

$$\pi_{\text{t}}[\mathbf{H}] + \sum_{\substack{m,n\in\mathbb{Z} \\ \alpha\in\{\text{TE, TM}\}}} \tilde{c}_{\alpha,mn} Y_{\alpha,mn}\pi_{\text{T}}[\mathbf{E}_{\alpha,mn}e^{-j\kappa_{mn}(z-z^+)}] = 2Y^{\text{inc}}\pi_{\text{T}}\big[\mathbf{E}^{\text{inc}}\big]. \tag{A.6}$$

For any $\mathbf{u}, \mathbf{v} \in \big(L^2(\Gamma_{z^+})\big)^3$, the space of square-integrable complex vector functions on Γ_{z^+}, we introduce the inner product

$$(\mathbf{u}, \mathbf{v})_{\Gamma_{z^+}} := \int_{\Gamma_{z^+}} \mathbf{u}\cdot\mathbf{v}^* \,\mathrm{d}\mathbf{x}, \tag{A.7}$$

where the superscript * denotes complex conjugation. Because of the orthogonality of the modal basis, i.e.,

$$(\pi_{\text{T}}[\mathbf{E}_{\text{TE},mn}], \pi_{\text{T}}[\mathbf{E}_{\text{TE},ij}])_{\Gamma_{z^+}} = \delta_{mi}\delta_{nj}, \tag{A.8a}$$

$$(\pi_{\text{T}}[\mathbf{E}_{\text{TM},mn}], \pi_{\text{T}}[\mathbf{E}_{\text{TM},ij}])_{\Gamma_{z^+}} = \delta_{mi}\delta_{nj}, \tag{A.8b}$$

$$(\pi_{\text{T}}[\mathbf{E}_{\text{TE},mn}], \pi_{\text{T}}[\mathbf{E}_{\text{TM},ij}])_{\Gamma_{z^+}} = 0, \tag{A.8c}$$

where δ denotes the Kronecker delta, the unknown coefficients $\tilde{c}_{\alpha,mn} \in \mathbb{C}$ of the modal expansion (A.6) can be obtained as follows:

$$\tilde{c}_{\alpha,mn} = (\pi_{\text{T}}[\mathbf{E}], \pi_{\text{T}}[\mathbf{E}_{\alpha,mn}])_{\Gamma_{z^+}} = \big(\pi_{\text{T}}\big[\mathbf{E}^{\text{inc}}\big], \pi_{\text{T}}[\mathbf{E}_{\alpha,mn}]\big)_{\Gamma_{z^+}} + \underbrace{(\pi_{\text{T}}[\mathbf{E}^{\text{sc}}], \pi_{\text{T}}[\mathbf{E}_{\alpha,mn}])_{\Gamma_{z^+}}}_{=c_{\alpha,mn}}. \tag{A.9}$$

Equation (A.6) represents the boundary condition to be imposed on Γ_{z^+}. In practice, the infinite sum of Floquet modes is truncated to $-m_{\max} \le m \le m_{\max}$, $-n_{\max} \le n \le n_{\max}$.

In that case, we obtain a boundary condition in the form of Eq. (54e) with

$$\boldsymbol{\mathcal{F}}^{\text{inc}} = 2Y^{\text{inc}}\pi_{\text{T}}\big[\mathbf{E}^{\text{inc}}\big], \tag{A.10}$$

$$\boldsymbol{\mathcal{G}}(\mathbf{E}) = \sum_{\substack{|m|\le m_{\max} \\ |n|\le n_{\max} \\ \alpha\in\{\text{TE, TM}\}}} \tilde{c}_{\alpha,mn} Y_{\alpha,mn}\pi_{\text{T}}[\mathbf{E}_{\alpha,mn}]. \tag{A.11}$$

Further simplifications are possible if the dimensions of the unit cell are small enough, such that only the fundamental modes $\mathbf{E}_{\alpha,00}$ propagate, and the boundary Γ_{z^+} is placed sufficiently far away from the structure, such that all higher order modes are attenuated to a negligible amplitude. In this case, the fundamental mode is of particular interest and we may omit all evanescent higher order modes in Eq. (A.11). In particular, we can employ the first-order absorbing boundary condition [58] (Chapter 13.4.1), i.e., Eq. (54e) with

$$\boldsymbol{\mathcal{G}}(\mathbf{E}) = -\frac{\mathbf{k}_{\text{t}}^{\text{inc}}}{\omega\mu k_z^{\text{inc}}}\big(\mathbf{k}_{\text{t}}^{\text{inc}}\cdot\pi_{\text{T}}[\mathbf{E}]\big) - \frac{k_z^{\text{inc}}}{\omega\mu}\pi_{\text{T}}[\mathbf{E}], \tag{A.12}$$

where $\mathbf{k}_{\text{t}}^{\text{inc}} := \pi_{\text{T}}[\mathbf{k}^{\text{inc}}]$.

The corresponding terms in the boundary conditions of the dual problem (65) are given as follows:

$$\overline{\boldsymbol{\mathcal{F}}} = -\frac{j}{\omega\mu_0}\pi_{\mathrm{T}}[\mathbf{E}_{\alpha,mn}], \tag{A.13}$$

and either

$$\overline{\boldsymbol{\mathcal{G}}} = -\sum_{\alpha,m,n} \tilde{d}^*_{\alpha,mn} Y^*_{\alpha,mn} \pi_{\mathrm{T}}[\mathbf{E}_{\alpha,mn}], \tag{A.14}$$

where $\tilde{d}_{\alpha,mn} = (\pi_{\mathrm{T}}[\mathbf{E}_{\alpha,mn}], \mathbf{z}_{\mathrm{T}})_{\Gamma_{z^+}}$, if Eq. (A.11) is used for the primal problem, or

$$\overline{\boldsymbol{\mathcal{G}}} = \frac{\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}}{\omega\mu_0 k_z^{\mathrm{inc}}}\left(\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}} \cdot \mathbf{z}_{\mathrm{T}}\right) + \frac{k_z^{\mathrm{inc}}}{\omega\mu_0}\mathbf{z}_{\mathrm{T}}, \tag{A.15}$$

if lowest order Floquet boundary conditions (A.12) are employed in Eq. (54).

APPENDIX B. DETAILS ON FE DISCRETIZATION

The mesh is assumed to be periodic, i.e., the surface meshes on Γ_{x^+} and Γ_{x^-}, as well as on Γ_{y^+} and Γ_{y^-}, are respectively identical. Without loss of generality, we further assume the vector of coefficients $\mathbf{c} \in \mathbb{C}^{N_h}$ to be ordered such that the boundary conditions imposed in Eq. (57) can be expressed as follows:

$$\mathbf{c} = \begin{bmatrix} \mathbf{c}_{\text{inner}} \\ \mathbf{c}_{\Gamma_{z^+}} \\ \mathbf{c}_{\Gamma_{z^-}} \\ \mathbf{c}_{\Gamma_{x^+}} \\ \mathbf{c}_{\Gamma_{x^-}} \\ \mathbf{c}_{\Gamma_{y^+}} \\ \mathbf{c}_{\Gamma_{y^-}} \\ \mathbf{c}_{\Gamma_{x^+}\cap\Gamma_{y^+}} \\ \mathbf{c}_{\Gamma_{x^-}\cap\Gamma_{y^+}} \\ \mathbf{c}_{\Gamma_{x^+}\cap\Gamma_{y^-}} \\ \mathbf{c}_{\Gamma_{x^-}\cap\Gamma_{y^-}} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & \mathbf{I}e^{-j\psi_x} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & \mathbf{I}e^{-j\psi_y} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I} \\ 0 & 0 & 0 & 0 & \mathbf{I}e^{-j\psi_x} \\ 0 & 0 & 0 & 0 & \mathbf{I}e^{-j\psi_y} \\ 0 & 0 & 0 & 0 & \mathbf{I}e^{-j(\psi_x+\psi_y)} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{\text{inner}} \\ \mathbf{c}_{\Gamma_{z^+}} \\ \mathbf{c}_{\Gamma_{x^+}} \\ \mathbf{c}_{\Gamma_{y^+}} \\ \mathbf{c}_{\Gamma_{x^+}\cap\Gamma_{y^+}} \end{bmatrix} = \mathbf{P}\mathbf{c}_{\text{dof}},$$

where we have introduced the reduced vector $\mathbf{c}_{\text{dof}} =\in \mathbb{C}^{N_{\text{DoF}}}$ of $N_{\text{DoF}} < N_h$ DoFs and $\mathbf{I}$ denotes an identity matrix of appropriate size [58] (Chapter 13.1.2).

Let $\mathbf{A} \in \mathbb{C}^{N_h \times N_h}$ and $\mathbf{f} \in \mathbb{C}^{N_h}$ be the system matrix and right-hand side vector, which are obtained by using Eq. (59) in Eq. (58), as well as Nédélec test functions. In the case of using the higher-order Floquet port boundary condition, i.e., Eq. (54e) with Eq. (A.10), the boundary integrals lead to dense sub-blocks in the matrix $\mathbf{A}$, whereas Eq. (A.12) preserves the sparsity of the FE matrix. The quasi-periodic and PEC boundary conditions (57) on ansatz and test functions can be imposed conveniently using the matrix $\mathbf{P} \in \mathbb{C}^{N_h \times N_{\text{DoF}}}$, leading to the reduced system

$$\mathbf{A}_{\text{dof}}\mathbf{c}_{\text{dof}} = \mathbf{P}^{\mathrm{H}}\mathbf{A}\mathbf{P}\mathbf{c}_{\text{dof}} = \mathbf{P}^{\mathrm{H}}\mathbf{f} = \mathbf{f}_{\text{dof}}, \tag{B.1}$$

where $\mathbf{P}^{\mathrm{H}}$ denotes the Hermitian transpose of $\mathbf{P}$. Functions spanned by the reduced DoF form a proper subspace of Eq. (57).

APPENDIX C. CONVERGENCE STUDY OF GPC AND SPARSE-GRID PROJECTION

Since the computational cost for a proper convergence study in the 17-dimensional setting is too high, we are restricted again to the two most sensitive parameters, i.e., t_1, t_2, and repeat the gPC convergence study with sparse Gaussian quadrature. Again, we use a random cross-validation sample of size $N^{\text{MC}} = 1 \times 10^3$ to compute the error (75). Results are presented in Fig. C1. It can be observed that, similar to Fig. 13, the error slightly increases up to order 3 before a convergent behavior can indeed be observed.

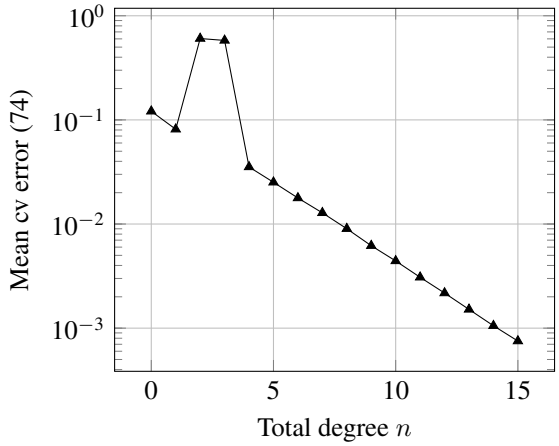


FIG. C1: Convergence study of a gPC approximation using pseudo-spectral projection and sparse Gauss quadrature