Opt Quant Electron (2016) 48:250
DOI 10.1007/s11082-016-0530-1

Reduced basis method for the electromagnetic scattering problem: a case study for FinFETs

Martin Hammerschmidt[1] · Sven Herrmann[1] · Sven Burger[1,2] · Jan Pomplun[2] · Frank Schmidt[1,2]

Received: 7 October 2015 / Accepted: 18 March 2016 / Published online: 25 March 2016

Abstract Optical 3D simulations in many-query and real-time contexts require new solution strategies. We study an adaptive, error controlled reduced basis method for solving parameterized time-harmonic optical scattering problems. Application fields are, among others, design and optimization problems of nano-optical devices as well as inverse problems for parameter reconstructions occurring e. g. in optical metrology. The reduced basis method presented here relies on a finite element modeling of the scattering problem with parameterization of materials, geometries and sources.

Keywords Finite element method · Reduced basis method · Electromagnetic field · Model reduction · Optical critical dimension metrology

1 Introduction

Electromagnetic field solvers must be efficient to be of use for optimization tasks of 3D structures where computation times are of importance as typically a large number of computations with varying parameters have to be performed until the optimal structure is found. The same holds true for inverse problems where measured data is given and structural details are sought after. Hence there is a demand for highly accurate, error

This article is part of the Topical Collection on Numerical Simulation of Optoelectronic Devices, NUSOD'15.

Guest edited by Julien Javaloyes, Weida Hu, Slawek Sujecki and Yuh-Renn Wu.

✉ Martin Hammerschmidt
hammerschmidt@zib.de

[1] Zuse Institute Berlin, Takustrasse 7, 14195 Berlin, Germany

[2] JCMwave GmbH, Bolivarallee 22, 14050 Berlin, Germany

controlled results at low computation times. Especially for real time applications, like online process control in semiconductor manufacturing, new solution strategies for a repeated solution of 3D problems are required.

The reduced basis method (RBM) is such a strategy. Here we consider application of the RBM to time-harmonic Maxwell's equations based on a finite element (FEM) discretization. The setup discussed here is light scattering off a fin field-effect transistor (FinFET) (analyzed previously without the RBM (Burger et al. 2015)). FinFETs are realized on wafers and measured after manufacturing by optical methods to control the actual geometrical shape. The measurement consists of an illumination of an array of periodically placed FinFETs and a determination of the reflectance in dependence of the wavelength and polarization.

2 Reduced basis method

The reduced basis method (Prudhomme et al. 2002; Rozza et al. 2008) allows for the construction of error controlled approximations to the relationship of an input, a specific configuration of parameters $\mu \in \mathbb{R}^d$, to an output quantity $s(\mu) = s(\boldsymbol{E}(\mu))$ derived from a solution $\boldsymbol{E}$ of the parameter dependent Maxwell's equations. In the following we describe the method only briefly and refer to Pomplun (2010) and Pomplun et al. (2011) for details.

In this work we consider electromagnetic scattering problems where a parameter $\mu \in \mathfrak{D} \subset \mathbb{R}^d$ may describe geometry and/or properties of the illumination or materials. The weak formulation for an $\mathcal{N}$ dimensional FEM space V_h reads:

For $\mu \in \mathfrak{D}$ find $\boldsymbol{E} \in V_h$ such that:

$$a(\varphi, \boldsymbol{E}; \mu) = f(\varphi; \mu) \quad \forall\, \varphi \in V_h. \tag{1}$$

The natural space for solutions of the electromagnetic scattering problem 1 is $H(\mathbf{curl}, \Omega) = \{\boldsymbol{u} \in (L^2(\Omega))^3 \,|\, \nabla \times \boldsymbol{u} \in (L^2(\Omega))^3\}$ with the corresponding norm $\|u\|_{H(\mathbf{curl},\Omega)} = (\|\boldsymbol{u}\|^2_{(L^2(\Omega))^3} + \|\nabla \times \boldsymbol{u}\|^2_{(L^2(\Omega))^3})^{1/2}$.

Essential for the RBM is the splitting of the solution process into two phases: a computationally expensive *offline phase* executed once and an inexpensive, fast *online phase*.

During the *offline* construction of the reduced basis the underlying, high-dimensional FEM model (commonly referred to as the *truth approximation*) is solved multiple times for certain parameter configurations to compute so-called snapshots. The selection of snapshots spanning the reduced basis can be carried out self-adaptively through a Greedy strategy by selecting the worst resolved parameter configuration contained within a finite training set $\mathfrak{D}_{train} \subset \mathfrak{D}$ of the parameter domain. An efficient and cheap a posteriori error estimator for the reduced basis approximation error is crucial for this selection (Pomplun and Schmidt 2010).

Subsequently, in the *online* phase, solutions are sought in the linear space $X_N \subset V_h$ spanned by these snapshots via a Galerkin projection. This *reduced problem* reads:

For $\mu \in \mathfrak{D}$ find $\boldsymbol{E}_N \in X_N$ such that:

$$a(\varphi, \boldsymbol{E}_N; \mu) = f(\varphi; \mu) \quad \forall\, \varphi \in X_N. \tag{2}$$

As the number of snapshots N is typically in the order of 10–100, the reduced problem stated on the reduced basis requires only a low dimensional linear system to be solved. This operation can be done in real time and for many parameters if the assembly of the

reduced system in the *online* phase remains independent of the FEM dimension $\mathcal{N}$ and only depends on the reduced basis dimension N.

Essential in this offline–online splitting is the parameterization of the Maxwell operator $a(\cdot,\cdot;\mu)$. It must thus be parameterized in a way that permits an effective offline–online decomposition. This is readily achieved for example for changes in the (scalar) permittivities of isotropic materials where the operator can be decomposed affinely, i.e. $a(v,u;\mu) = a_1(v,u;\mu) + \varepsilon a_2(v,u;\mu)$. However, for geometrical variations or parameterizations of the illumination the operator is generally not affine in μ and has to be approximated by the Empirical Interpolation Method (EIM) (Barrault et al. 2004) to guarantee an affine decomposition.

In addition to the Maxwell operator also the output functional $s(\mu) = s(\boldsymbol{E}(\mu))$ must admit an offline–online decomposition for the *online* phase to be independent of $\mathcal{N}$. This is trivial for linear outputs like the Fourier Transform of the electromagnetic field. We call the reduced basis together with the reduced output functional the *reduced model*.

3 Results

3.1 Optical model

The investigated structure is a periodic array of FinFETs (at the 22 nm technology node). These complex 3D architectures have become a crucial driver for down-scaling of structures on integrated circuits (Bunday et al. 2013).

The device geometry of the unit cell is depicted as a FEM mesh in Fig. 1 (left) with designations and materials of the components. The device dimensions and optical parameters are taken from Bunday et al. (2013) and Palik (1998). The silicon fin sits on a 200 nm thick silicon oxide substrate (only partially shown) and the gate (silicon) wraps around the fin orthogonal to its axis in the center of the unit cell. A SiN layer (dark red) of

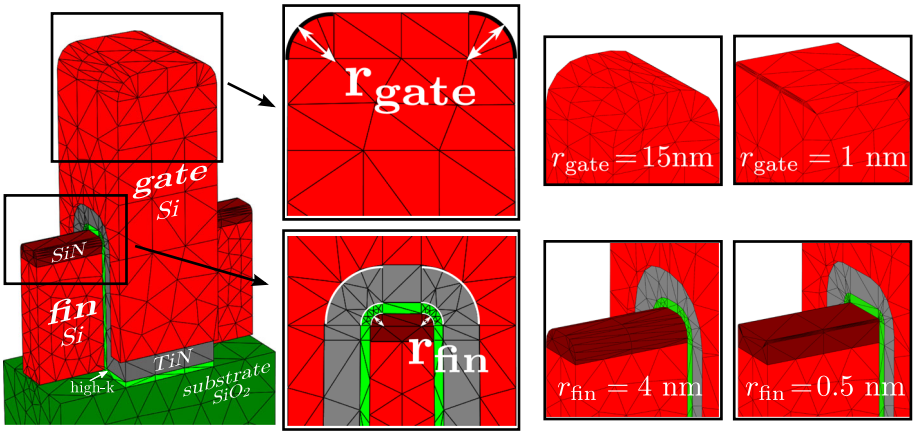

Fig. 1 *Left* FEM-mesh side view of the FinFET geometry. The used designation (*gate* and *fin*) along with the material are inscribed on each particular component of the geometry. The surrounding air-domain is not shown. *Center* detailed view of FEM-mesh regions with geometrical meaning of radii r_{gate} (*top*) and r_{fin} (*bottom*). r_{fin} also affects the roundings of adjacent layers. *Right* detailed view of displaced FEM-mesh to the most extreme values of $r_{gate} = \{15.0, 1.0\,\text{nm}\}$ and $r_{fin} = \{4.0, 0.5\,\text{nm}\}$ respectively. (Color figure online)

5 nm, a high-k layer (green) of 2 nm and a TiN layer (gray) of 7 nm complete the structure. The pitches of fin and gate are 44 and 88 nm respectively. The fin is 12.7 nm wide and 40 nm high with an undercut of 2.1 nm and a side-wall angle of 89.5 deg. The gate is 40 nm wide and 95 nm high with an undercut of 2 nm. Additionally a rounding of the fin and gate top edges is used. The FinFET is discretized with a parameterized tetrahedral mesh which accurately resolves the undercuts and corner roundings. The transparent boundary conditions at the top and bottom are realized with Perfectly Matched layers (Zschiedrich 2009).

The structure is illuminated by a plane wave at 390 nm and tilted 30 deg along the gate direction. Using 4th order ansatz functions and 190,089 degrees of freedom the FEM problem is solved in ≈270 s of CPU time.

3.2 Reduced basis construction

We build a reduced basis for the corner rounding radii r_{gate} and r_{fin} in the parameter domain $\mathfrak{D} = [1\,\text{nm}, 15\,\text{nm}] \times [.5\,\text{nm}, 4\,\text{nm}]$. The definition of radii is depicted in Fig. 1 (center). A detailed view of the FEM-mesh for the extreme values is depicted on the right in Fig. 1. The reduced model is constructed for the Fourier transform of the electromagnetic field which is used to compute the reflectance of the structure.

The training set $\mathfrak{D}_{train}$ used for the Greedy searches is adaptively refined from a fine initial mesh of 11 by 11 equidistant values depicted as crosses in Fig. 2b. The EIM approximation uses 13 snapshots for the system matrix with an estimated relative approximation error of 5.3×10^{-7} in the L_∞ norm over the training set. A single snapshot is sufficient to approximate the right-hand side.

In the offline phase 38 truth solutions are computed to construct the reduced basis. Figure 2b depicts the snapshot order and location in the parameter domain $\mathfrak{D}$ (red numbered circles). The first snapshot is chosen in the center and the subsequent snapshots cover the corner of $\mathfrak{D}$ first. Generally the snapshot locations follow the boundaries of $\mathfrak{D}$ with a slight clustering at the corners.

The Greedy algorithm choses in each iteration the worst resolved parameter in the training set as the new snapshot location. Thus the error estimate generally decreases with increasing reduced basis dimension N. In Fig. 2 the residual error estimate (solid black line) with increasing N is shown in a semi-logarithmic plot. The initial estimate is used for

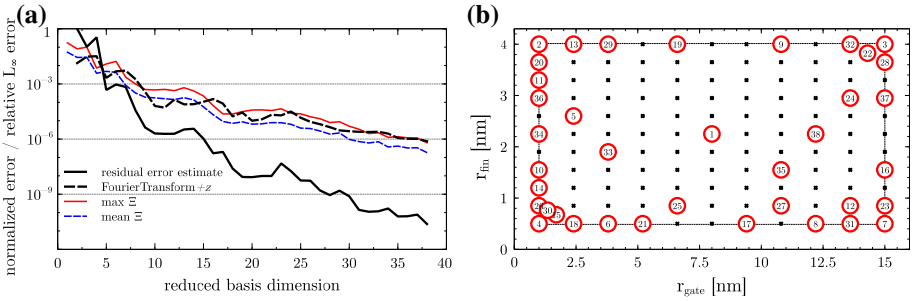

Fig. 2 **a** Error estimates with increasing reduced basis dimension in a semi-logarithmic plot. The residual error estimate (*black solid line*) is normalized with the estimate for N = 1. The error estimate for the Fourier transform (*black dashed line*) estimates the relative error in the L_∞ norm. **b** Snapshot positions in parameter space are marked as *red circles* and the *number* indicates the sequence of snapshot selection. The locations contained in the initial training set $\mathfrak{D}_{train}$ are marked by *crosses*. (Color figure online)

normalization. The normalized estimate converges exponentially with N up to an estimated error of less than 2.5×10^{-11}. The adaptive refinement of the training set leads to the non-monotonic convergence as new values are included.

The estimated error in the output quantity (dashed black line, Fourier transform in the z direction) shows a similar decrease although with a different rate due to different norms employed (relative L_∞ norm and H(**curl**, Ω) norm). The error estimate convergences up to an estimated error of less than 1×10^{-6}.

3.3 Online evaluation of the reduced basis

In the online phase we evaluated the reduced system on 21 by 21 equidistant points evenly spread over the whole parameter space $\mathfrak{D}$. Each evaluation takes ≈200 ms. The contour plots in Fig. 3 (center column) show the computed reflectance of P and S polarization respectively. The plots next to the contour plots show the horizontal and vertical (left and right) cross-sections through the center of the parameter space supplemented by reference solutions (blue dots) and in the outermost plots the linear trend of the reference data is removed to highlight deviations. We observe a much stronger dependence on the corner rounding radius of the gate than of the fin in both polarizations. In both polarizations the deviation compared to the reference is on the order of 1×10^{-5}.

3.4 Error analysis

In the following we investigate these different errors by comparing the reduced basis approximations to reference solutions for a randomly selected set Ξ of 100 parameter values. For $\mu \in \Xi$ we solve the truth approximation including the EIM approximation and compare the results.

In Fig. 2 the relative errors in the Fourier transform are shown in addition to the error estimate on the left. The mean relative error over Ξ is depicted by the blue dashed linear whereas the maximum is shown in red. Both lines follow the error estimate very closely.

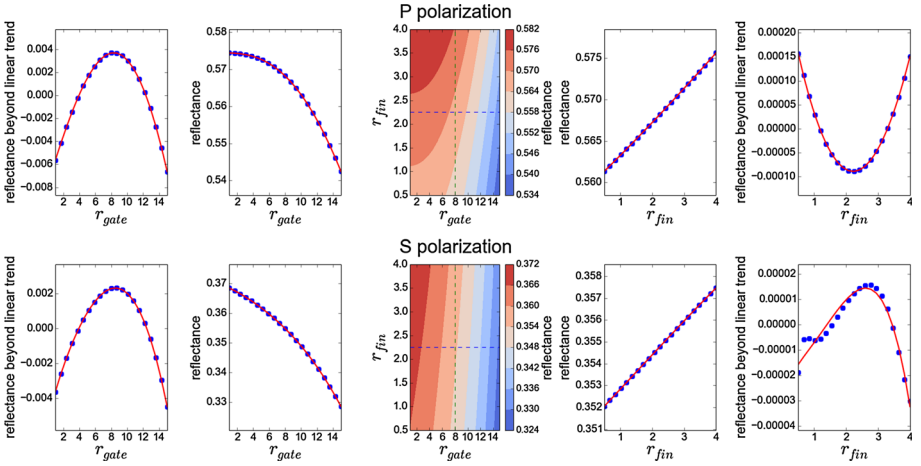

Fig. 3 Reflectance for the P (*top row*) and S polarization (*bottom row*) over radii as contour plots (*central column*), as cross-sections through the parameter domain (*left and right from center columns*) and with removed linear trend (*outermost columns*). (Color figure online)

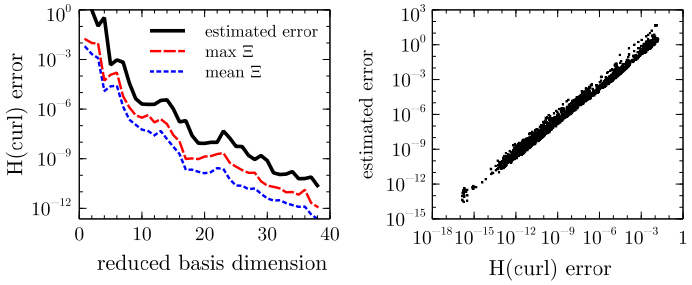

Fig. 4 Maximum and mean error of $\Xi \subset \mathfrak{D}^N_{train}$ in H(**curl**, Ω) norm over the reduced basis dimension (*left*) and estimated reduced basis error over the error in H(**curl**, Ω) norm (*right*). The maximal and mean error decrease with increased reduced basis dimension. The estimated and actual error are highly correlated. (Color figure online)

The errors are reduced from about 1×10^{-1} to 1×10^{-6} with increasing reduced basis dimension. We conclude that the error estimator is very efficient for the given problem.

In Fig. 4 the max and mean of of the approximation error over Ξ in the H(**curl**, Ω) norm are shown as well as a scatter plot relating the estimated error and the true error in the H(**curl**, Ω) norm. The maximum as well as the mean error exhibit identical convergence trends to the estimated error. With increasing N the error decreases exponentially from 1×10^{-2} to 1.2×10^{-12} at most with the mean being an order of magnitude smaller. The scatter plot demonstrates the good performance of the error estimator. The error and the estimate are highly correlated over more than 9 orders of magnitude.

4 Conclusions

The reduced basis method is well suited for many-query and real-time simulation tasks. We have demonstrated its application to optical scattering simulations of a parameterized FinFET where it allows for the computation of online solutions with a three order of magnitude speed up. The errors in both the field solution and the output are controlled satisfactorily. As a result we have an efficient, error controlled, real-time capable procedure for situations where the application of the direct problem would not be realistic.

Acknowledgments The results were obtained at the Berlin Joint Lab for Optical Simulations for Energy Research (BerOSE) of Helmholtz-Zentrum Berlin für Materialien und Energie, Zuse Institute Berlin and Freie Universität Berlin. This research was carried out in the framework of MATHEON supported by Einstein Foundation Berlin through ECMath within subprojects SE6 and OT5.

References

Barrault, M., Maday, Y., Nguyen, N.C., Patera, A.T.: An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations. C. R. Math. **339**, 667–672 (2004)

Bunday, B., Germer, T.A., Vartanian, V., Cordes, A., Cepler, A., Settens, C.: Gaps analysis for CD metrology beyond the 22nm node. In: Proceedings of SPIE, vol. 8681 (2013)

Burger, S., Zschiedrich, L., Pomplun, J., Herrmann, S., Schmidt, F.: Hp-finite element method for simulating light scattering from complex 3D structures. In: Proceedings of SPIE, vol. 9424 (2015)

Palik, E.D.: Handbook of Optical Constants of Solids. Bd. 3. Academic Press, Cambridge (1998)

Pomplun, J.: Reduced Basis Method for Electromagnetic Scattering Problems. Ph.D. thesis, Free University Berlin (2010)

Pomplun, J., Schmidt, F.: Accelerated a posteriori error estimation for the reduced basis method with application to 3D electromagnetic scattering problems. SIAM J. Sci. Comput. **32**, 498–520 (2010)

Pomplun, J., Burger, S., Zschiedrich, L., Schmidt, F.: Reduced basis method for real-time inverse scatterometry. In: Proceedings of SPIE, vol. 8083 (2011)

Prudhomme, C., Rovas, D.V., Veroy, K., Machiels, L., Maday, Y., Patera, A.T., Turinici, G.: Reliable real-time solution of parametrized partial differential equations: reduced-basis output bound methods. J. Fluids Eng. **124**, 70–80 (2002)

Rozza, G., Huynh, D.B.P., Patera, A.T.: Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. Arch. Comput. Methods Eng. **15**, 229–275 (2008)

Zschiedrich, L.: Transparent Boundary Conditions for Maxwells Equations: Numerical Concepts Beyond the PML Method. Ph.D. thesis, FU Berlin (2009)