arXiv:1011.2654v1 [physics.optics] 11 Nov 2010

Reduced basis method for computational lithography

Jan Pomplun [ab], Lin Zschiedrich [ab], Sven Burger [ab], Frank Schmidt [ab]

[a] Zuse Institute Berlin, Takustraße 7, D – 14 195 Berlin, Germany

[b] JCMwave GmbH, Haarer Straße 14a, D – 85 640 Putzbrunn, Germany

This paper has been published in Proc. SPIE Vol. **7488** (2009) 74882B, (*Photomask Technology 2009, Larry S. Zurbrick; M. Warren Montgomery, Editors*) and is made available as an electronic preprint with permission of SPIE. Copyright 2009 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

ABSTRACT

A bottleneck for computational lithography and optical metrology are long computational times for near field simulations. For design, optimization, and inverse scatterometry usually the same basic layout has to be simulated multiple times for different values of geometrical parameters.

The reduced basis method allows to split up the solution process of a parameterized model into an expensive offline and a cheap online part. After constructing the reduced basis offline, the reduced model can be solved online very fast in the order of seconds or below. Error estimators assure the reliability of the reduced basis solution and are used for self adaptive construction of the reduced system.

We explain the idea of reduced basis and use the finite element solver JCMsuite constructing the reduced basis system. We present a 3D optimization application from optical proximity correction (OPC).

Keywords: inspection, computational lithography, reduced basis, model order reduction, optical proximity correction, scatterometry

1. INTRODUCTION

The importance of numerical simulations for design, optimization and metrology of photomasks has grown rapidly over the last years due to the ongoing miniaturization of integrated curcuits.

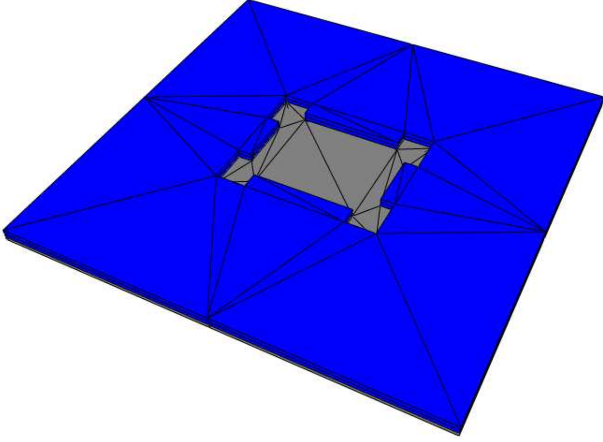
A bottleneck for these many-query and real-time applications are long computational times for rigorous simulations of the near field in the photomask. Usually the same basic layout has to be simulated multiple times for different values of geometrical parameters, e.g. line width, absorber edge angle, etc.

The reduced basis method[1–3] can be applied to this task. The solution process is decomposed into an expensive offline and a cheap online step. In the offline step the reduced basis is built self-adaptively by solving the underlying model rigorously several times. The full model is then projected onto the reduced basis. In the online step the assembled reduced system can be solved in the order of seconds independent on the size of the original problem. Furthermore methods from the well established field of a posteriori error estimation of finite element methods[4] can be applied to the reduced basis method to assure the reliability of the computed output and also for construction of the reduced basis.[5, 6] These are advantages in comparison to interpolation or table based methods for output data.

In the following we first describe the basics of the reduced basis method and then apply it to an OPC optimization problem of a contact hole array depicted in Fig. 1.

Corresponding author: J. Pomplun
URL: http://www.zib.de/nano-optics/
Email: pomplun@zib.de

(a)



(b)

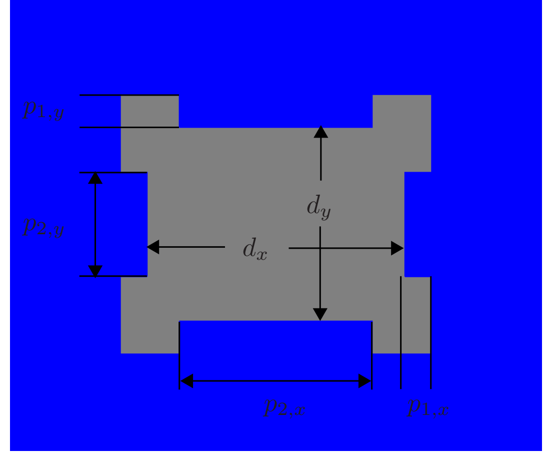


Figure 1. (a) 3D mask model used for finite element computation. (b) Definition of OPC parameters used for optimization.

2. INPUT-OUTPUT RELATIONSHIP

Before explaining the reduced basis method we first derive the mathematical setup for the scattering problem. Suppose we have a photomask with a parameterized geometry as depicted in Fig. 1. In the following the free geometrical parameters are denoted by a p-tuple ν in a bounded parameter space $D \subset \mathbb{R}^p$. Fixing the input parameters ν to the mask model, the scattering problem can be solved numerically with a Maxwell solver relying, e.g. on the finite element method (FEM), and outputs of interest, e.g. diffraction orders, can be computed. Hence the problem can be formulated as an input output relationship. In this example the input is a set of geometrical parameters, the output are the diffraction orders. For notational convenience we consider a single output of interest in the following. Mathematically the input-output relationship is given as follows:
Given geometrical parameters $\nu \in D$ determine the output of interest $s^{\mathcal{N}}(\nu)$:

$$s^{\mathcal{N}}(\nu) = l^{\mathrm{o}}\left(u^{\mathcal{N}}(\nu)\right), \tag{1}$$

where $u^{\mathcal{N}}(\nu)$ is the solution to following problem:
Find $u^{\mathcal{N}} \in X^{\mathcal{N}}$ such that:

$$a\left(u^{\mathcal{N}}, v; \nu\right) = f\left(v\right), \quad \forall v \in X^{\mathcal{N}}. \tag{2}$$

The electromagnetic scattering problem (2) is stated in so called "weak form"[7]. Roughly spoken the sesquilinear form $a\left(\cdot,\cdot;\nu\right)$ presents Maxwell's equations and the linear functional $f\left(\cdot\right)$ the incoming field, e.g. a plane wave. The finite element space $X^{\mathcal{N}} \subset \mathrm{H}\left(\mathbf{curl}\,,\Omega\right)$ is the space in which the electric field solution $u^{\mathcal{N}}(\nu)$ is determined, where Ω is the domain of interest. Since Maxwell's equations and therewith the sesquilinear form are parameter dependent, the solution $u^{\mathcal{N}}$ is also parameter dependent. The output of interest, e.g. a diffraction mode in a periodic setting, is described via a linear output functional applied to the electric field $u^{\mathcal{N}}$:

$$\begin{aligned} l^o : X^{\mathcal{N}} &\to \mathbb{C}, \\ u^{\mathcal{N}}(\nu) &\mapsto l^{\mathrm{o}}\left(u^{\mathcal{N}}(\nu)\right) = s^{\mathcal{N}}(\nu). \end{aligned} \tag{3}$$

Hence $l^o \in \left(X^{\mathcal{N}}\right)'$, which is the dual space of $X^{\mathcal{N}}$.[8] The discretized version of Maxwell's equations (2) will be referred to as "truth approximation". Usually the truth approximation has to be solved several times for

different parameter values ν in design and optimization applications or inverse scatterometry. Typically $\mathcal{N}$ is very large such that already a single solution process takes very long which prohibits many-query or real-time application.

3. REDUCED BASIS METHOD

The purpose of the reduced basis method is to construct an approximative input-output relationship which can be solved very fast. Furthermore it is important to construct error estimators which control and quantify the quality and reliability of the reduced basis solution. Finally we want that computation of the reduced basis solution and error estimation becomes independent on the number of finite element degrees of freedom $\mathcal{N}$.

The idea of the reduced basis method is simple.[9] Let us denote the manifold of all possible solutions of the truth approximation (2) by:

$$\mathcal{M}^{\mathcal{N}} = \left\{ u^{\mathcal{N}}(\nu) \text{ is a solution to (2)} \mid \nu \in D \right\}. \tag{4}$$

Now suppose $\mathcal{M}^{\mathcal{N}}$ can be approximated by a low dimensional space $X^N \approx \mathcal{M}^{\mathcal{N}}$, with

$$\dim X^N = N \ll \mathcal{N}. \tag{5}$$

Then it is reasonable to assume that the following reduced basis approximation to the truth approximation will give results of good quality:
Given geometrical parameters $\nu \in D$ determine the output of interest $s^N(\nu)$:

$$s^N(\nu) = l^{\mathrm{o}}\left(u^N(\nu)\right), \tag{6}$$

where $u^N(\nu)$ is the solution to following problem:
Find $u^N \in X^N$ such that:

$$a\left(u^N, v; \nu\right) = f(v), \quad \forall v \in X^N. \tag{7}$$

The space X^N is the reduced basis space and is a subspace of the finite element space $X^{\mathcal{N}}$.

3.1. Reduced basis space

An important question which arises is how to construct the reduced basis space X^N such that it gives a good approximation to the space of all possible solutions $\mathcal{M}^{\mathcal{N}}$. To answer this question we first define a sequence of hierarchical subsets of the parameter domain D. Let $\nu_i \in D$, $i = 1, \ldots, N_{\max}$. Then we define:

$$S_i = \{\nu_1, \ldots, \nu_i\}, \quad i = 1, \ldots, N_{\max}. \tag{8}$$

These sets have the following property:

$$S_1 \subset S_2 \subset \cdots \subset S_N \subset \cdots \subset S_{N_{\max}}. \tag{9}$$

The Lagrange reduced basis space $W_N^{\mathcal{N}}$ of dimension N is then defined by:

$$W_N^{\mathcal{N}} = \operatorname{span}\left\{ u^{\mathcal{N}}(\nu) \text{ is a solution to (2)} \mid \nu \in S_N \right\}, \tag{10}$$

hence it is spanned by solutions to the truth approximation for fixed parameter values. These solutions are called snapshot solutions. It is also possible to include first and higher derivatives of the field $u^{\mathcal{N}}$ with respect to parameters into the reduced basis space, which leads to so called Taylor and Hermite spaces. Here we notice the expensive “offline” costs of the reduced basis method. For construction of a Lagrange reduced basis space of dimension N the truth approximation, i.e. the full problem, has to be solved N times. However the reduced basis space has only to be assembled once and in principle a parallelization of this process is possible since computation of different snapshots is independent of each other.

3.2. Reduced basis system

In the following we construct the explicit form of the reduced basis system (7) which has to be solved in the online phase. Suppose we have given a basis

$$B_N^{\mathcal{N}} = \left\{ \zeta_q^{\mathcal{N}} \,|\, q = 1, \ldots, N \right\} \tag{11}$$

of X^N. Then we can expand the reduced basis solution into this basis:

$$u^N(\nu) = \sum_{q=1}^{N} \alpha_q(\nu) \zeta_q^{\mathcal{N}}. \tag{12}$$

Online the parameter dependent coefficients $\alpha_q(\nu)$ have to be computed. Therefore we insert ansatz (12) into the reduced basis system (7) and use the fact that if (7) holds for all $v \in B_N^{\mathcal{N}}$ then it holds for all $v \in X^N$:

$$\sum_{q=1}^{N} \alpha_q(\nu) a\left(\zeta_q^{\mathcal{N}}, \zeta_n^{\mathcal{N}}; \nu\right) = f\left(\zeta_n^{\mathcal{N}}\right), \quad n = 1, \ldots, N, \tag{13}$$

which gives a linear system of equations for the coefficients $\alpha_q(\nu)$. System (13) has to be assembled and solved online. On the left hand side we have a parameter dependent matrix:

$$A^N(\nu) = \left(a\left(\zeta_q^{\mathcal{N}}, \zeta_n^{\mathcal{N}}; \nu\right)\right)_{q,n=1,\ldots,N}, \tag{14}$$

and on the right hand side a parameter independent vector:

$$f^N = \left(f\left(\zeta_n^{\mathcal{N}}\right)\right)_{n=1,\ldots,N}. \tag{15}$$

The vector f^N can be assembled offline. However computation of each matrix entry:

$$a\left(\zeta_q^{\mathcal{N}}, \zeta_n^{\mathcal{N}}; \nu\right)$$

depends on the dimension of $\zeta_q^{\mathcal{N}}$, which is the number of finite element degrees of freedom $\mathcal{N}$. Since this assembling step has to be performed online we want to avoid any $\mathcal{N}$ dependence. In the following we explain how we can perform an online-offline decomposition of the assembling step.

3.3. Online-offline decomposition

For online-offline decomposition of the assembling step (14) we need an affine decomposition of the system sesquilinear form $a(\cdot, \cdot; \nu)$, which is defined as follows:

$$a(v, u; \nu) = \sum_{m=1}^{Q} \Theta_m(\nu) a_m(v, u), \tag{16}$$

where $\Theta_m(\nu)$ are parameter dependent functions and $a_m(\cdot, \cdot)$ parameter independent sesquilinear forms. If we can construct such a decomposition the following parameter independent matrices can be assembled offline:

$$A_m^N = \left(a_m(\zeta_q^{\mathcal{N}}, \zeta_n^{\mathcal{N}})\right)_{q,n=1,\ldots,N}, \quad m = 1, \ldots, Q.$$

The parameter dependent system matrix (14) is then assembled online according to:

$$A^N(\nu) = \sum_{m=1}^{Q} \Theta_m(\nu) A_m^N.$$

The costs are $O(N^2Q)$, where N is the reduced basis dimension and Q the number of terms in the affine decomposition (16). The solution of the reduced basis system has costs $O(N^3)$. For computation of the output of interest we get the following decomposition:

$$
\begin{aligned}
s^N(\nu) =& l^{\mathrm{o}}\left(u^N(\nu)\right) \\
=& l^{\mathrm{o}}\left(\sum_{q=1}^{N} \alpha_q(\nu)\zeta_q^{\mathcal{N}}\right) \\
=& \sum_{q=1}^{N} \alpha_q(\nu) l^{\mathrm{o}}\left(\zeta_q^{\mathcal{N}}\right),
\end{aligned} \tag{17}
$$

where the quantities $l^{\mathrm{o}}\left(\zeta_q^{\mathcal{N}}\right)$ can also be computed offline. For the output of interest we have costs $O(N)$. Hence with an affine decomposition (16) the costs of solving the reduced basis system are independent on $\mathcal{N}$ as desired.

In[5] we show how an affine decomposition can be constructed for electromagnetic scattering problems.

3.4. Error estimation

Since the reduced basis method is an approximation technique it is important to be able to quantify the reliability of the output. Let us define the error of the reduced basis solution by:

$$
e = u^{\mathcal{N}} - u^N. \tag{18}
$$

Then it can be shown[2, 10] that this error is bounded by:

$$
||e(\nu)||_{X^{\mathcal{N}}} \leq \Delta(\nu),
$$

with:

$$
\Delta(\nu) = \frac{1}{\beta(\nu)} \left|\left|r^{\mathrm{pr}}\left(\cdot; u^N; \nu\right)\right|\right|_{(X^{\mathcal{N}})'}. \tag{19}
$$

Here $r^{\mathrm{pr}}\left(\cdot; u^N; \nu\right)$ is the so called primal residuum of the reduced basis solution u^N. The parameter dependent constant $\beta(\nu)$ is the so called inf-sup constant of the sesquilinear form $a\left(\cdot,\cdot;\nu\right)$. It is important to note that all quantities of estimate (19) can be computed online[2, 11] without knowing the true solution. An appropriate online-offline decomposition furthermore makes computation of the bound independent on the number of finite element degrees of freedom such that an error bound can be computed for each reduced basis output online. Also this error estimator can be used in the offline construction of the reduced basis to explore the parameter space and find regions, where the reduced basis approximation has to be improved further, which is explained in the following section. This is a main advantage in comparison to interpolation methods. In[12] we will explain the subject of error estimation in more detail for electromagnetic scattering problems.

4. GREEDY CONSTRUCTION OF REDUCED BASIS SPACES

Before we apply the reduced basis method we comment on the construction of the Lagrange reduced basis space $X^{\mathcal{N}} = W_N^{\mathcal{N}}$ (10). How to choose the snapshot parameters ν_i in (8)? We use a greedy algorithm[2] whose idea is given in the following.

First we define a training set $\Xi_{\mathrm{train}} \in D$ of possible snapshot candidates. From this space we want to chose a number N of snapshot parameters for construction of the reduced basis. The first parameter ν_1 is chosen randomly. Then we construct a one dimensional reduced basis approximation corresponding to the snapshot $u^N(\nu_1)$. Now we evaluate the error estimator for this one dimensional reduced basis approximation on all candidate snapshots in the training sample Ξ_{train} and include the parameter value with the maximum error into the reduced basis because this is supposed to add a maximum of "new information" into the reduced basis. Then we have a two dimensional reduced basis and the process is continued iteratively. The process can be stopped e.g. if a certain maximum dimension is reached or if the error estimator gives sufficiently small bounds over the training set.

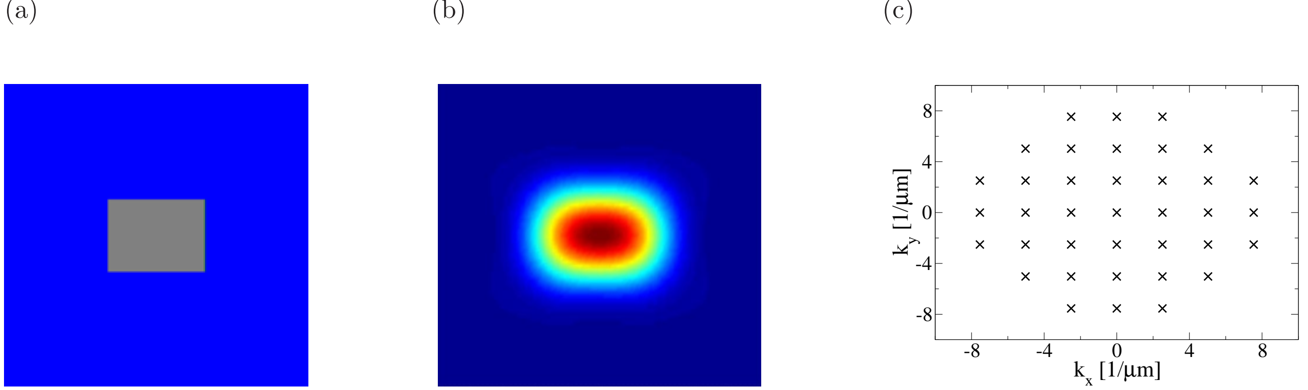


Figure 2. (a) Top view of mask layout for imaging of contact hole. Absorber material is shown in blue. (b) Far field image of contact hole shows corner rounding. (c) Position of incoming k-vectors used for conventional illumination.

5. NUMERICAL EXAMPLES

Figure 2 shows the aerial image of a rectangular contact hole. In the far field the sharp corners are washed out which is refered to as corner rounding. Correction of this rounding can be achieved with OPC methods like the introduction of serifs to the corners of the contact holes as depicted in Fig. 1. Our numerical example will be the optimization of this OPC structure to obtain a structure on the wafer which is closest to the desired rectangular shape.

The geometry of the contact hole is depicted in Fig. 1. The size of the computational domain is $2.5\,\mu\text{m}\times 2.5\,\mu\text{m}$ with a total height of 90 nm including silica substrate and air. The height of the Chromium absorber is 50 nm with a refractive index of:

$$n_{\mathrm{Cr}} = 0.84 - 1.65i$$

for a wavelength of $\lambda = 193\,\text{nm}$.[13] In x- and y-direction we apply periodic boundary conditions. The shape of serifs which we want to optimize is described by 4 input parameters $p_{1,x}$, $p_{2,x}$, $p_{1,y}$, and $p_{2,y}$. The dimensions of the contact hole itself are fixed at:

$$\begin{aligned} d_x &= 800\,\text{nm}, \\ d_y &= 600\,\text{nm}. \end{aligned}$$

As incoming light we use conventional illumination[13] which is modelled by a set of incoming plane waves whose incoming angles lie within a cone up to a certain maximum angle. The source is visualized by a set of points in the k_x-k_y-plane, as shown in Fig. 2(a). For each of these incoming directions we simulate two orthogonal polarization states to mimic unpolarized light. This gives $P = 74$ sources in total. For each of the sources we compute the near field and corresponding far field coefficients separately in order to determine a partially coherent intensity distribution. The far field of the mask passes an optical system with 4 to 1 reduction.

The finite element discretization gives a system with $\mathcal{N} = 474720$ unknowns. The for solution of the problem is about $9,300s$ (single CPU time). Hence optimization of this structure is extremely time using the truth approximation.

Therefore we construct a reduced basis approximation of above problem. For the parameter domain D we choose

$$\begin{aligned} p_{1x} &\in [145\,\text{nm}; 315\,\text{nm}], \\ p_{2x} &\in [100\,\text{nm}; 200\,\text{nm}], \\ p_{1y} &\in [45\,\text{nm}; 205\,\text{nm}], \\ p_{2y} &\in [100\,\text{nm}; 210\,\text{nm}]. \end{aligned} \tag{20}$$

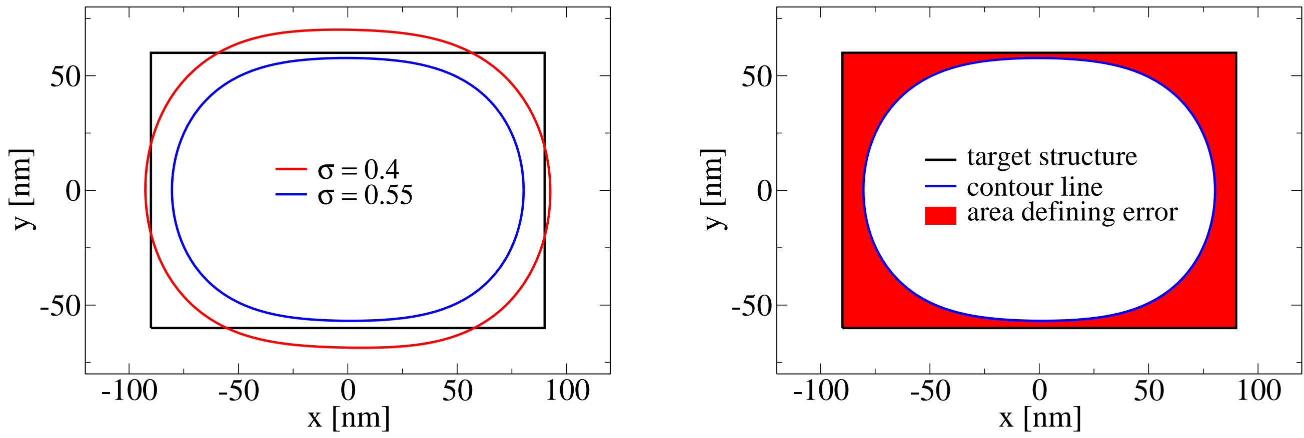
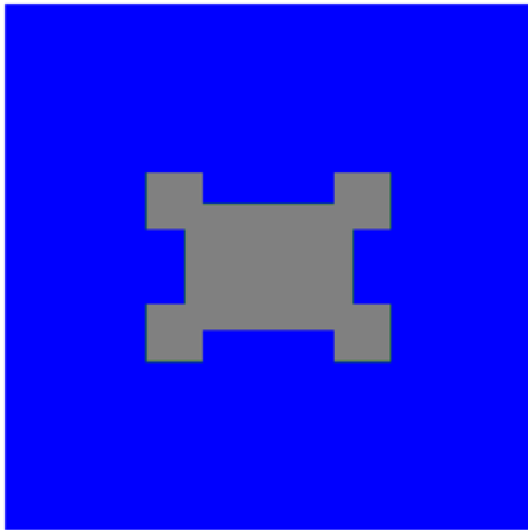


Figure 3. (a) Comparison of contour lines for different levels of σ. (b) Structure on wafer given as contour line of aerial image at $\sigma = 0.55$, i.e. 55% of maximum intensity, with target structure and area defining error functional.

(a)



(b)

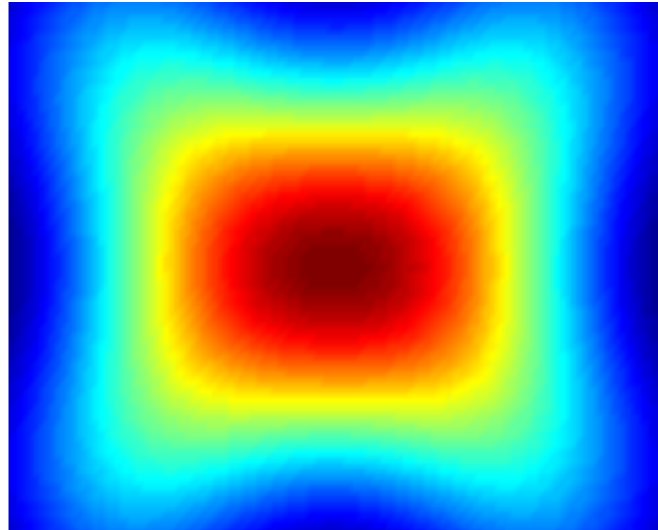


Figure 4. (a) Mask layout after OPC optimization and (b) corresponding aerial image (red: high intensity, blue: low intensity).

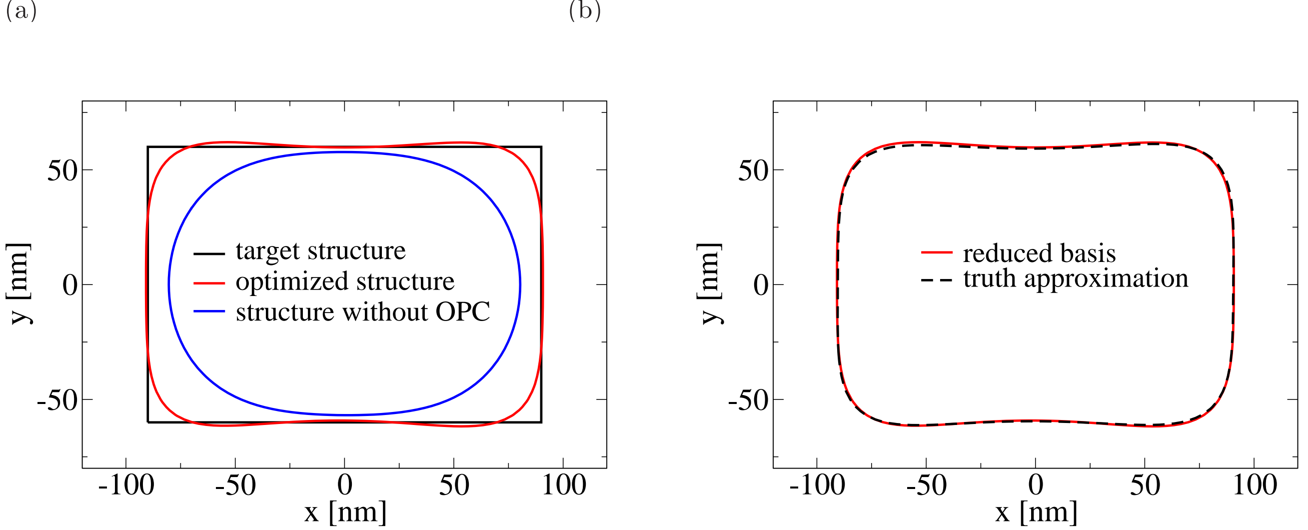


Figure 5. (a) Comparison of structure on wafer without and with optimized mask layout obtained from reduced basis computation. (b) Comparison of optimized reduced basis structure and corresponding structure obtained from truth approximation.

The dimension of the constructed reduced basis is $N = 40$.

Despite the large number of sources and outputs of interest the reduced basis computation only takes $1.1s$. This gives a speed up factor of about 8000 compared to the truth approximation. Hence we can perform the online part of the optimization at very low computational costs.

In order to perform the optimization we have to define a cost functional which is minimized. Let us denote by $\gamma(\nu)$ the shape of the contact hole on the wafer at a certain threshold intensity σ . Here we use a contour line of the aerial image for simplicity, see Fig. 3(a). We fix the threshold at $\sigma = 0.55$. Now we define $\Gamma(\nu)$ by the area enclosed by the target shape γ_T (the rectangle) and $\gamma_\sigma(\nu)$, as depicted in Fig. 3(b). The cost functional is then given by:

$$g(\nu) = ||1||_{L^1(\Gamma(\nu))} , \tag{21}$$

and we want to determine optimal parameters such that:

$$\nu_{\min} = \min_{\nu \in D} g(\nu). \tag{22}$$

Figure 4 shows the optimized geometry of the photomask and the corresponding aerial image.

The shape of the structure on the wafer without and with optimized serifs is depicted in Fig. 5(a). The optimized structure shows good agreement to the target structure. Of course corner rounding can not be avoided completely. Furthermore a comparison of the optimal structure computed with the reduced model and obtained from the truth approximation is given in Fig. 5(b). We observe very good agreement. Largest deviations are of the order of 1 nm below % of the hole CDs. These deviations will further decrease with higher reduced basis dimension.

6. CONCLUSIONS

We use the reduced basis method for fast solution of geometrically parametrized electromagnetic scattering problems. The reduced basis method is used to construct an approximative reduced system to a parameter dependent finite element problem. Error estimators thereby assure the reliability of the reduced basis solution.

A challenging numerical optimization example from optical proximity correction demonstrates the performance of the reduced basis method. A speed-up factor of about 8000 compared to the finite element simulation

was obtained, enabling a very fast optimization of the OPC structures of a contact hole. The obtained reduced basis result agrees very well with the exact finite element simulation.

REFERENCES

1. A. Patera and G. Rozza, *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations*, MIT Pappalardo Graduate Monographs in Mechanical Engineering, 1. ed., 2006.
2. G. Rozza, D. Huynh, and A. Patera, "Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics," *Archives of Computational Methods in Engineering* **15**(3), pp. 229–275, 2008.
3. Z. Zhu and F. Schmidt, "An efficient and robust mask model for lithography simulation," in *Optical Microlithography XXI*, H. J. Levinson and M. V. Dusa, eds., **6925**, p. 126, Proc. SPIE, 2008.
4. M. Ainsworth and J. T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, John Wiley and Sons, 1. ed., 2000.
5. J. Pomplun and F. Schmidt, "Reduced Basis Method for Electromagnetic Field Computations," in *Scientific Computing in Electrical Engineering, Proceedings of the SCEE-2008*, Springer, 2008.
6. Y. Chen, J. Hesthaven, Y. Maday, and J. Rodriguez, "A Monotonic Evaluation of Lower Bounds for inf-sup Stability Constants in the Frame of Reduced Basis Methods," *C. R. Acad. Sci. Paris* **1**, 2008.
7. P. Monk, *Finite Element Methods for Maxwell's Equations*, Oxford University Press, 2003.
8. M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, Academic Press, 2. ed., 1980.
9. J. Pomplun and F. Schmidt, "Reduced basis method for fast and robust simulation of electromagnetic scattering problems," in *Numerical Methods in Optical Metrology*, **7390**, Proc. SPIE, 2009.
10. Y. Maday, A. Patera, and D. Rovas, "A blackbox reduced-basis output bound method for noncoercive linear problems," in *Nonlinear Partial Differential Equations and their Applications*, *Proceedings of the College De France Seminars* , 2001.
11. D. B. P. Huynh, G. Rozza, S. Sen, and A. T. Patera, "A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants," *C. R. Acad. Sci. Paris* **I**(345), pp. 473–478, 2007.
12. J. Pomplun and F. Schmidt, "Accelerated a posteriori error estimation for the reduced basis method with application to 3d electromagnetic scattering problems," *SIAM J. Sci. Comput. (submitted)* .
13. H. Levinson, *Principles of Lithography*, SPIE, 2. ed., 2004.