Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

MARTIN HAMMERSCHMIDT, JAN POMPLUN, SVEN BURGER, FRANK SCHMIDT

Adaptive sampling strategies for efficient parameter scans in nano-photonic device simulations

This paper is made available as an electronic preprint with permission of SPIE and will be published in Proc. SPIE 8980, 89801O (2014). The full citation reads:
Martin Hammerschmidt, Jan Pomplun, Sven Burger and Frank Schmidt, "Adaptive sampling strategies for efficient parameter scans in nano-photonic device simulations", Physics and Simulation of Optoelectronic Devices XXII, Bernd Witzigmann, Marek Osinski, Fritz Henneberger, Yasuhiko Arakawa, Editors, Proc. SPIE 8980, 89801O (2014).

http://dx.doi.org/10.1117/12.2036363

Herausgegeben vom
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustraße 7
D-14195 Berlin-Dahlem

Telefon: 030-84185-0
Telefax: 030-84185-125

e-mail: bibliothek@zib.de
URL: http://www.zib.de

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

Adaptive sampling strategies for efficient parameter scans in nano-photonic device simulations

Martin Hammerschmidt[a], Jan Pomplun[b], Sven Burger[a,b], Frank Schmidt[a,b]

[a]Konrad-Zuse-Zentrum für Informationstechnik Berlin, Takustraße 7, 14195 Berlin, Germany
[b]JCMwave GmbH, Bolivarallee 22, 14050 Berlin, Germany

This paper is made available as an electronic preprint with permission of SPIE and will be published in Proc. SPIE 8980, 89801O (2014). The full citation reads:
Martin Hammerschmidt, Jan Pomplun, Sven Burger and Frank Schmidt, "Adaptive sampling strategies for efficient parameter scans in nano-photonic device simulations", Physics and Simulation of Optoelectronic Devices XXII, Bernd Witzigmann, Marek Osinski, Fritz Henneberger, Yasuhiko Arakawa, Editors, Proc. SPIE 8980, 89801O (2014).

http://dx.doi.org/10.1117/12.2036363

ABSTRACT

Rigorous optical simulations are an important tool in optimizing scattering properties of nano-photonic devices and are used, for example, in solar cell optimization. The finite element method (FEM) yields rigorous, time-harmonic, high accuracy solutions of the full 3D vectorial Maxwell's equations[1] and furthermore allows for great flexibility and accuracy in the geometrical modeling of these often complex shaped 3D nano-structures. A major drawback of frequency domain methods is the limitation of single frequency evaluations. For example the accurate computation of the short circuit current density of an amorphous silicon/micro-crystalline multi-junction thin film solar cell may require the solution of Maxwell's equations for over a hundred different wavelengths if an equidistant sampling strategy is employed. Also in optical metrology, wavelength scans are frequently used to reconstruct unknown geometrical and material properties of optical systems numerically from measured scatterometric data. In our contribution we present several adaptive numerical integration and sampling routines and study their efficiency in the context of the determination of generation rate profiles of solar cells. We show that these strategies lead to a reduction in the computational effort without loss of accuracy. We discuss the employment of tangential information in a Hermite interpolation scheme to achieve similar accuracy on coarser grids. We explore the usability of these strategies for scatterometry and solar cell simulations.

Keywords: finite element method, optical simulations, adaptive sampling, metrology, parameter scans, solar cells

1. INTRODUCTION

Determination of parameter dependent profiles such as the generation rate profile of a solar cell or the angular dependence of the scatterometric signal is oftentimes a tedious task if the underlying problem has to be solved numerically, as computation times for a single parameter value can easily require many hours of CPU time if 3D simulations are required. In other cases, like parameter optimization or reconstruction, the computation and evaluation of many of such spectra may be necessary. Usually one choses a regular, equidistant parameter spacing to sample the parameter interval and interpolates in between. However depending on the actual profile this strategy might not be optimal as it will probably miss narrow width peaks of the profile if the spacing is too broad or invest to much effort in areas where the profile changes smoothly with the parameter.

Further author information: (Send correspondence to Martin Hammerschmidt)
Martin Hammerschmidt: E-mail: hammerschmidt@zib.de, Telephone: +49 30 84185-149

2. NUMERICAL METHODS

As this contribution aims at comparing different sampling strategies in order to reduce the computational effort in high accuracy parameter scans within rigorous optical simulations, we present some of the numerical methods employed in this comparison.

2.1 The Finite Element Method

The finite element method (FEM) yields rigorous, time-harmonic, high accuracy solutions of the full 3D vectorial Maxwell's equations[1] and furthermore allows for great flexibility and accuracy in the geometrical modeling of these often complex shaped 3D nano-structures. In the following we describe in brief how scattering problems for the time-harmonic Maxwell's equations are investigated within this contribution. For this class of problems a monochromatic field is incident to the computational domain and the scattered field inside the computational domain is sought after. Using the finite element method (FEM) to discretize the arising differential operators we solve the second order curl-curl-equation for the electric field:

$$\nabla \times \mu^{-1} \nabla \times E(r) - \omega^2 \varepsilon E(r) = 0. \tag{1}$$

The complete mathematical formulation of a scattering problem with a detailed discussion of the arising boundary conditions has been published by Pomplun et al.[2]

Time-harmonic Maxwell solvers, like the FEM-implementation use for this work, allow the use of optical material properties as measured and do not require fitting of the complex refractive index to a physical model. In addition, FEM offers great geometric flexibility, i.e. variable mesh element sizes to resolve arbitrarily shaped material interfaces with sharp edges which is of interest especially in the solar cell model problem.

2.2 Sampling Strategies

The goal in determining parameter dependent profiles is to reconstruct the profile $f(p)$ of a scalar parameter $p \in P \subset \mathbb{R}$ within a parameter interval P as closely as possible with a minimal amount of (possibly expensive) evaluations of f. We assume no a priori knowledge of the underlying profile. In our sample problems 3.1 and 3.2 these are derived from solutions of Maxwell's equations computed with a FEM-solver. In general we assume f to be a real valued (nonlinear) map $p_i \in P \to f_i = f(p_i) \in \mathbb{R}$. Interpolation errors are usually measured as the difference between the real profile f and its interpolation $\tilde{f}$ over the parameter interval measured for example in the L^1 norm

$$\left\| f - \tilde{f} \right\|_1 = \int_P |f(p) - \tilde{f}(p)| \, \mathrm{d}p$$

or the L^2 norm

$$\left\| f - \tilde{f} \right\|_2 = \left(\int_P \left(f(p) - \tilde{f}(p) \right)^2 \, \mathrm{d}p \right)^{\frac{1}{2}}.$$

The L^2 norm penalizes differences $f - \tilde{f}$ stronger than the L^1 norm. Throughout this work we use the L^2 as an error measure.

In order to find the best $\tilde{f}$ we make use of the extensive mathematical research available[3] for the computation of integrals $\int_P f \mathrm{d}p$. The sequence of sampling points $p_i \in P$ used in computing $\int_P f \mathrm{d}p$ provides an interpolation $\tilde{f}$ of f that minimizes the L^1 error. There are a number of different numerical quadrature methods available within numerical packages such as `MATLAB`.[4] This includes error controlled adaptive schemes such as `quad`, which combines integral values computed on intervals using Simpson's rule and the trapezoidal rule. Except for the widely used equidistant routine, all of the following methods estimate and control the L^1 error:

equidistant A routine that samples f regularly with N sampling points in the parameter interval. No adaptive control is executed or required and the sampling points are spaced at $\frac{L}{N-1}$ distance apart where L is the length of the parameter interval P.

ode45 `ode45` is the `MATLAB` implementation of an explicit Runge-Kutta (4,5) formula. In this case we utilize the solver that is designed to solve initial value problems for ordinary differential equations as an adaptive integration routine.

quad `MATLAB` offers `quad` as a general purpose numerical quadrature tool. It implements an adaptive Simpson quadrature routine and is suggested to be used for low accuracy estimates with nonsmooth integrands.

quadl `quadl` is offerd by `MATLAB` as well. It is an implementation of an adaptive Lobatto quadrature routine and is suggested to be used at higher accuracies for smooth integrands.

quadgk `MATLAB` also implements an adaptive Gauss-Kronrod quadrature routine `quadgk` designed to be used for high accuracies and oscillatory integrands. Gauss-Kronrod quadrature rules provide very high order interpolation polynomials. Its standard settings are quite conservative as it uses at least 150 function evaluations.

adaptive The integration routine `adaptive` is a reimplementation with the Simpson rule similar to `MATLAB` 's `quad`. It makes use of all of its available function evaluations and increases the sampling point density locally in subintervals where the desired accuracy is not met, if other subintervals have the accuracy and the maximum number of function evaluations allowed is not yet reached. This design is meant to distribute the function evaluations more evenly across the subintervals of the integration interval in cases where the maximum function count will be reached. In contrast to `quad`, this should avoid situations where all the allowed function evaluations are used to resolve only one subinterval to meet the desired accuracy whereas others are left unrefined.

adaptive Hermite The adaptive Hermite routine is designed identical to the `adaptive` one. In addition to functional evaluations it also uses parameter derivatives to approximate the function and thus the integral. The adaptive error control takes this higher order interpolation into account.

multigrid The `multigrid` implementation follows the textbook description.[3] The general idea is to start with a very coarse grid and refinine locally only if the desired accuracy is not met. This yields a sequence of grids or refinement levels and error estimates that have to be accounted for. If a regular bisection is used to refine the grids, error estimates based on Simpson's and trapezoidal rule can be used which makes the error estimation similar to for example the `adaptive` routine or `MATLAB` 's `quad`.

multigrid Hermite Similar to the adaptive Hermite implementation this algorithm uses a Hermite interpolation and parameter derivatives instead of only function evaluations in combination with a multigrid scheme.

3. MODEL PROBLEMS

3.1 A metrology problem

The following metrology model problem has been taken from.[5] The geometry and finite element mesh of the 1D periodic line grating with a pitch of $100\,nm$ is shown in Figure 1. In the referenced paper the authors computed the dependence of the scatterometric signal strength I_0 on the azimuthal incidence angle θ at $\lambda_0 = 197\,nm$ for both polarization and studied the employment of physical parameter sensitivities in parameter reconstrutions. We use the same model parameters for our model problem in this contribution. The investigated profile is shown on the right in Figure 1. The shown reference solution is based on 2269 functional evaluations (in this case: FEM simulations of the scattering of the periodic line grating). It has been computed using a multigrid method. The different levels used by the sampling method can be observed in the same plot where the interval length is shown. The smallest interval is 8.4877e-05 *deg* used near the singularity at 68.4346 *deg*. At 31.2891 *deg* the largest grid interval is 0.3477 *deg*, 2^{12} times larger than the smallest. The three intervals used for the detailed analysis in section 4.1 roughly coincide with the refinement strategy used here: a coarser discretization for the first interval, a extremely fine sampling intervals for the middle and a fine grid for the right interval.

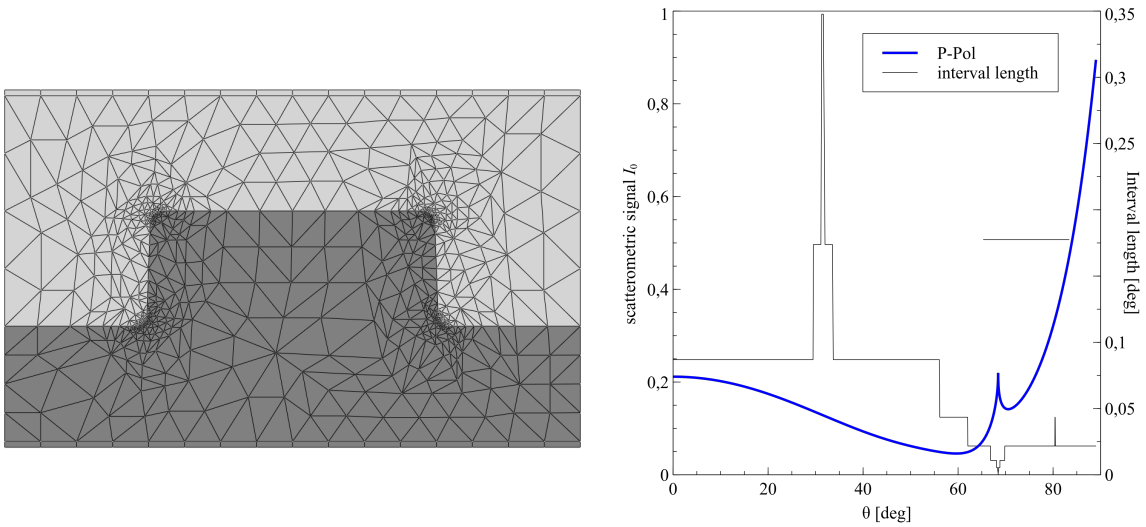

Figure 1. *Left:* Finite-element mesh for spatial discretization of the investigated 1D-periodic line grating. *Right:* Dependence of the scatterometric signal I_0 on the azimuthal angle of incidence of the illuminating plane waves, for P-polarization. The interval length between the sampling points of the reference solution demonstrates the different levels of grid densities used.

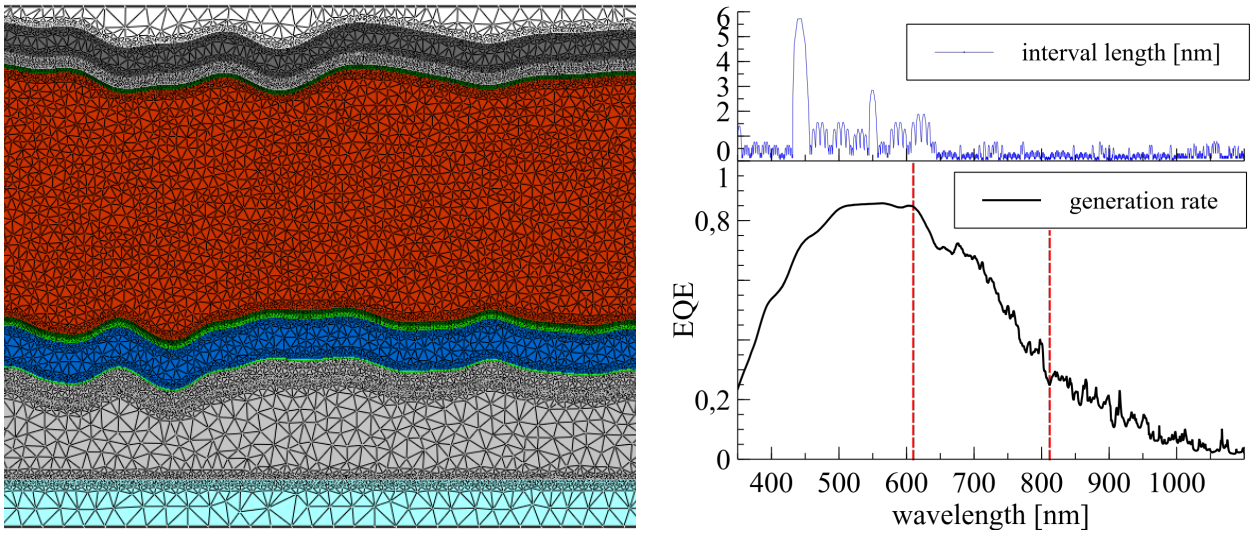

Figure 2. *Left:* Excerpt of the finite-element mesh for spatial discretization of the investigated thin-film silicon tandem solar cell. *Right:* Combined generation rate of both subcells over the investigated spectrum (bottom graph). The dotted red lines mark the three intervals with smooth and oscillatory behavior and a transition in between. The interval length between the sampling points of the reference solution is shown as a separate graph on top. The transition of the bottom graph to an oscillatory behavior is clearly reflected here by a more dense grid.

3.2 A solar cell problem

The thin-film silicon tandem solar cell model presented here has been adapted from[6] using the same material data. In the cited reference we investigated the converge behavior of the 3D numerical model. In the following we employ the same algorithm described in[6] to generate a random interface texture in 2D and look at the combined generation rate of both the a-Si and μc-Si subcells as a function of wavelength. As the model includes thick layers (730 nm SnO_2:F, 270 nm a-Si ,1600 nm μc-Si layer and 200 nm silver) as well as very thin doped layers (two *p*-type a-Si layers of 5 nm, 30 nm of n-type a-Si, 30 nm of p- and n-type μc-Si each and a 80 nm ZnO layer)

many narrow width peaks arise in the generation rate profile of the reference solution shown in Figure 2. The reference profile shown here was computed using 2940 sampling points separated by intervals ranging from 0.1 to 5.72 nm. The dotted red lines mark the three intervals with smooth and oscillatory behavior and a transition in between. Similarly to the metrology model problem in section 3.1 the different behavior of the generation rate profile is reflected in the sampling point density of the reference profile shown here. As this reference solution was computed with the `quadl` routine the density levels are not distinct as in Figure 1(right), but the transition of the reference profile to an oscillatory behavior is clearly reflected by a denser grid (graph on top in Figure 2(right)).

4. RESULTS AND DISCUSSION

For each of the two model problems we implemented a `MATLAB` function that returns the figure of merit (scatterometric signal strength and combined generation rate of the subcells) derived from the rigorous electromagnetic field solution upon input of the parameter varied (azimuthal incidence angle θ and vacuum wavelength λ_0). This was necessary to make use of the numerical quadrature functions provided by `MATLAB` itself.

4.1 Scatterometry

The results for the metrology model problem are shown in Figure 3. The graphs show the logarithm of the interpolation error $\log_{10}(\|f - \tilde{f}\|_2)$ measured in the L^2 norm for three different subintervals of the parameter space. These subintervals are indicated by a thicker line in profile plots shown in the insets. The choice of the boundaries of these three subintervals is arbitrary but can be motivated by the behavior of the characteristics of the integrand. In the first interval the scatterometric profile varies smoothly with the angle. This is reflected by the refinement strategy employed by the `multigrid` scheme used to compute the reference solution in Figure 1. The second interval includes the singularity and the remaining part of the parameter space forms the third.

We focus our discussion on the subintervals shown in (ii) to (iv). The most interesting interval is the second one (Figure 3(iii)). To resolve the singularity accurately a very fine mesh is needed. The `quad` routine (green line with cross markers) adaptively bisects its sampling intervals embedding the singularity and demonstrates a convergence behavior with a very steep convergence rate. The other `MATLAB` routines `quadl` and `quadgk` (green with star marker and dotted green) use more sampling points per interval resulting in a higher number of function evaluations. Due to the lack of regularity of the profile this expense cannot be recovered with a higher interpolation order. The `multigrid` routine uses an error estimate similar to `quad`'s based on Simpson's rule which lets it perform on par with `quad` for smooth integrands. The poor performance in this interval is due to a different starting grid it does not "see" the singularity until a fairly high accuracy (10^{-4} in L^1) is demanded. This explains the poor behavior in this interval. The good performance of the equidistant or regular sampling (black with square markers) underlines the need for fine resolution of the singularity. The Hermite routines (red and blue with circular markers) have a similarly steep convergence rate in this interval up to an error level of $10^{-3.5}$ where the accuracy of the reference solution seems to limit further gains. The routine named `adaptive` (dotted blue line) shows very little convergence due to the fact that it allows to invest computational effort not needed in other parts of the integration interval in subintervals, where the desired accuracy hasn't been met so far. This results in a high local accuracy in this particular interval whereas the global accuracy is not as high. In this case 37 of the 85 function evaluations (43.53%) lie in this interval compared to 5 used by `quad` on the same global accuracy setting of 10^{-3} in L^1.

In the smoother intervals (shown in (ii) and (iv)) the characteristics of the convergence properties of the different routines are different. A significant gain in accuracy is obtained by using derivatives in Hermite interpolation schemes in both intervals. The rate of convergence in general is determined by a regular global refinement. Especially in the first interval (ii) this can be observed. Both multigrid routines exhibit similar converge rates beyond $N = 100$ compared to the regular refinement as does the *adaptive* routine. `quad`, `ode45` and `quadgk` have a similar converge rate which is only surpassed by using derivative information in the Hermite schemes.

The last interval shown in (iv) exhibits a similar pattern with the convergence rate determined by the regular refinement at large Ns. The steepness of the profile leads to a finer sampling for all the routines, the exception

again being the adaptive Hermite scheme. Similar to before, `quad` performs as good or better as all remaining adaptive schemes and the equidistant sampling.

In (i) the results of the subintervals are combined and can be described more briefly. All adaptive schemes reach a higher accuracy at less functional evaluations resulting in a overall reduction of computational effort and thus an improvement in run times. The exception in this case is the ordinary different equation solver `ode45` which estimates errors conservatively and cannot sample the model profile as well as for example `quad`. As the `adaptive` routine is not designed to use the least amount of functional evaluations, but to perform better than a regular sampling at a similar number of function evaluations, it performs just as intended. As mentioned before the profile generated by `quad` is only surpassed in accuracy by using additional derivative information. The `multigrid` method performs analogously to `quad` in the smooth intervals, but it suffers from the lack of the detection of the singularity at low accuracies. This can also be seen in Figure **??** where the number of function evaluations required by the different algorithms to reach certain global error levels in the L^2 interpolation error is shown which makes readability easier than Figure 3 (i).

4.2 Solar Cell

The results for the solar cell model problem are shown in Figure 5 and again we plot the logarithm of the L^2 norm of the interpolation error for three different subintervals of the wavelength interval. The boundaries of the subintervals were chosen at points were the profile exhibits a distinct change in features. In the first interval the generation rate profile varies smoothly with the wavelength. In the third interval the profile is highly oscillatory with a large number of very narrow width singularities. The profile changes in the middle interval from a smooth curve to the oscillatory behavior observed in the last interval.

The convergence characteristics shown in the subplots (iii) and (iv) are very similar. The oscillatory profile demands a high sampling density throughout the whole interval as provided by the equidistant sampling. The adaptivity does not yield an advantage here as it results in a global regular refinement. The subplots focus on the interesting ranges of $N < 200$ and $N < 700$. The `MATLAB` routine `quadl` uses 1611 functional evaluations in the third interval alone and another 904 in the second. `quad` and `quadgk` similarly use more than 500 and 1000 function evaluations in the third interval. In contrast to smooth intervals the use of derivative information does not yield a better interpolation as it tends to overshoot at steep flanks present in the highy oscillatory part.

In the first interval shown in (ii) the Hermite schemes again yield a better interpolation at a lower number of functional evaluations. However the advantage is not as pronounced as in the metrology example. The steep flank of the profile at the beginning of the interval and the small changes in inclination at 400 and 450 nm lead again to fine almost equidistant sampling by the adaptive routines. The exception in this case is `quadgk` with its Gauss-Kronrod quadrature rule which is designed to cope with this sort of integrand. The routine stays at 58 function evaluations in this interval with estimated error of less than 10^{-2}.

5. CONCLUSION

To summarize, we have demonstrated that adaptive numerical integration and sampling routines are highly efficient if the investigated profile is sufficiently smooth and only occasional singularities occur. Furthermore the employment of derivatives yield a significant advantage provided the profile is again smooth. In the case of solar cells and the computation of short circuit current densities where the generation rate profile is not smooth at all the adaptive schemes cannot gain an advantage over a regular sampling strategy using equally spaced sampling points except for the smooth parts of the profile.

ACKNOWLEDGMENTS

This work was supported by Masdar PV GmbH and the German Federal Ministry for Environment, Nature Conservation and Nuclear Safety (BMU) in the framework of the program “Design und Demonstration der Technologie für Silizium-Dünnschichtsolarzellen mit 14% Zell- und 13% stabilem Modulwirkungsgrad (Demo 14)” as well as the DFG Research Center Matheon “Mathematics for key technologies” in Berlin. We further acknowledge funding withing the EMRP Joint Research Project IND 17 Scatterometry. The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union. The authors thank

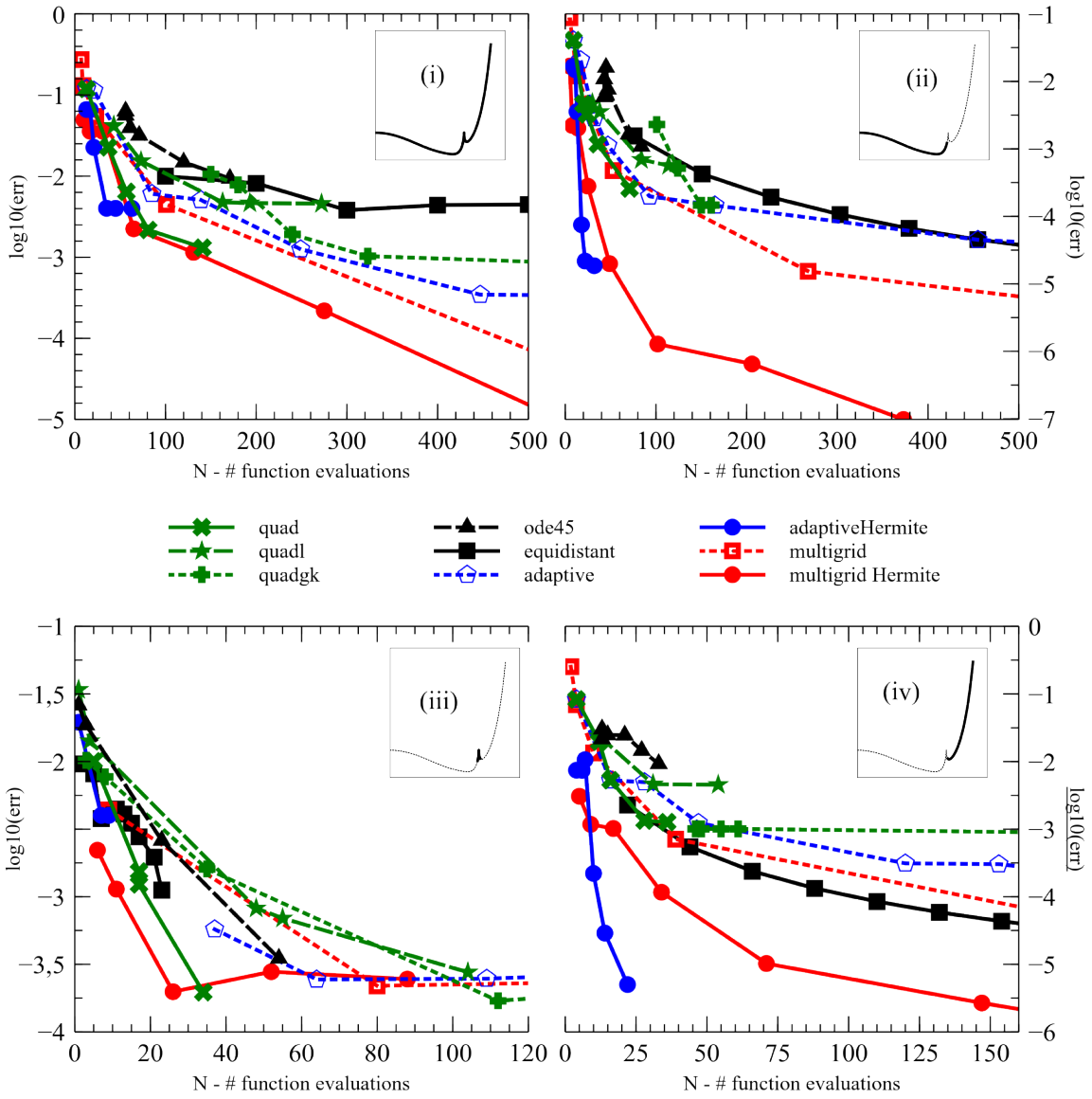

Figure 3. The graphs show the logarithm of the interpolation error $\log_{10}(\|f - \tilde{f}\|_2)$ for different intervals of the parameter space indicated by the insets where the active interval is marked with a thicker line. The legend shown in the center is valid for all four graphs. The Hermite integration routines show a distinct advantage in intervals where the profile is smooth ((ii) and (iv)). The other algorithms perform generally as well as a regular sampling in resolving the singularity (iii) and may show an overall reduction in computational effort (i) due to advantages in the smooth regions.

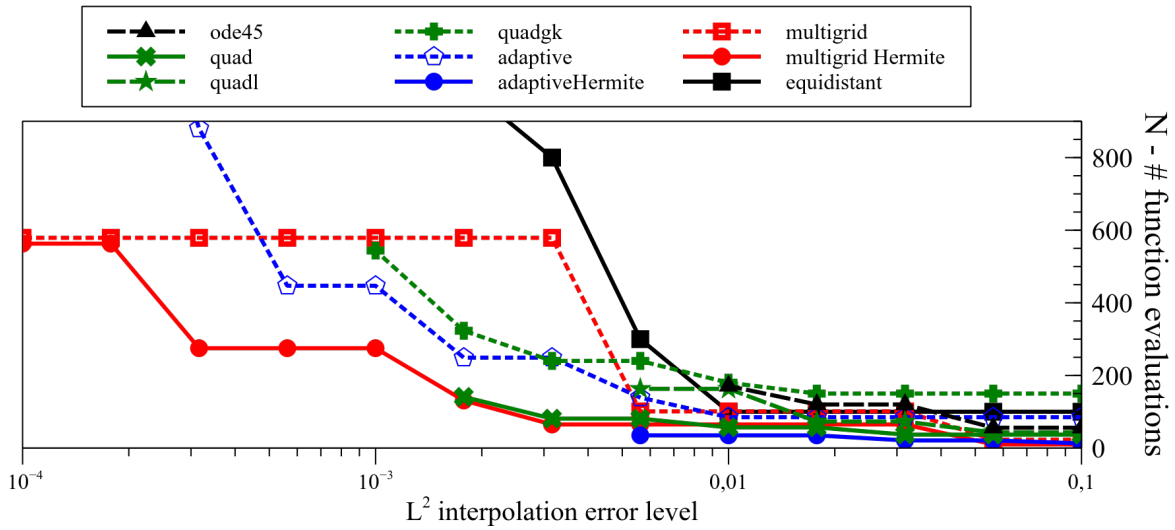

Figure 4. The number of function evaluations required by the different algorithms to reach certain global error levels in the L^2 interpolation error. The Hermite integration routines and `quad` require less function evaluations than the equidistant for all error levels and demonstrate a gentle incline at higher accuracies.

L.Zschiedrich and D. Lockau for many useful discussions concerning finite element based modeling and their continued support.

REFERENCES

[1] Burger, S., Zschiedrich, L., Pomplun, J., and Schmidt, F., "Jcmsuite: An adaptive fem solver or precise simulations in nano-optics," in [*Integrated Photonics and Nanophotonics Research and Applications*], *Integrated Photonics and Nanophotonics Research and Applications* , ITuE4, Optical Society of America (2008).

[2] Pomplun, J., Burger, S., Zschiedrich, L., and Schmidt, F., "Adaptive finite element method for simulation of optical nano structures," *phys. stat. sol. (b)* **244**, 3419 – 3434 (2007).

[3] Deuflhard, P. and Hohmann, A., [*Numerische Mathematik. I*], 3 ed. (2002).

[4] MATLAB, [*version 7.14.0.739 (R2012a)*], The MathWorks Inc., Natick, Massachusetts (2013).

[5] Burger, S., Zschiedrich, L., Pomplun, J., Schmidt, F., and Bodermann, B., "Fast simulation method for parameter reconstruction in optical metrology," *Proc. SPIE* **8681**, 868119–868119–7 (2013).

[6] Hammerschmidt, M., Lockau, D., Burger, S., Schmidt, F., Schwanke, C., Kirner, S., Calnan, S., Stannowski, B., and Rech, B., "FEM-based optical modeling of silicon thin-film tandem solar cells with randomly textured interfaces in 3D," in [*Proc. SPIE 8620, Physics, Simulation, and Photonic Engineering of Photovoltaic Devices II*], 86201H (2013).

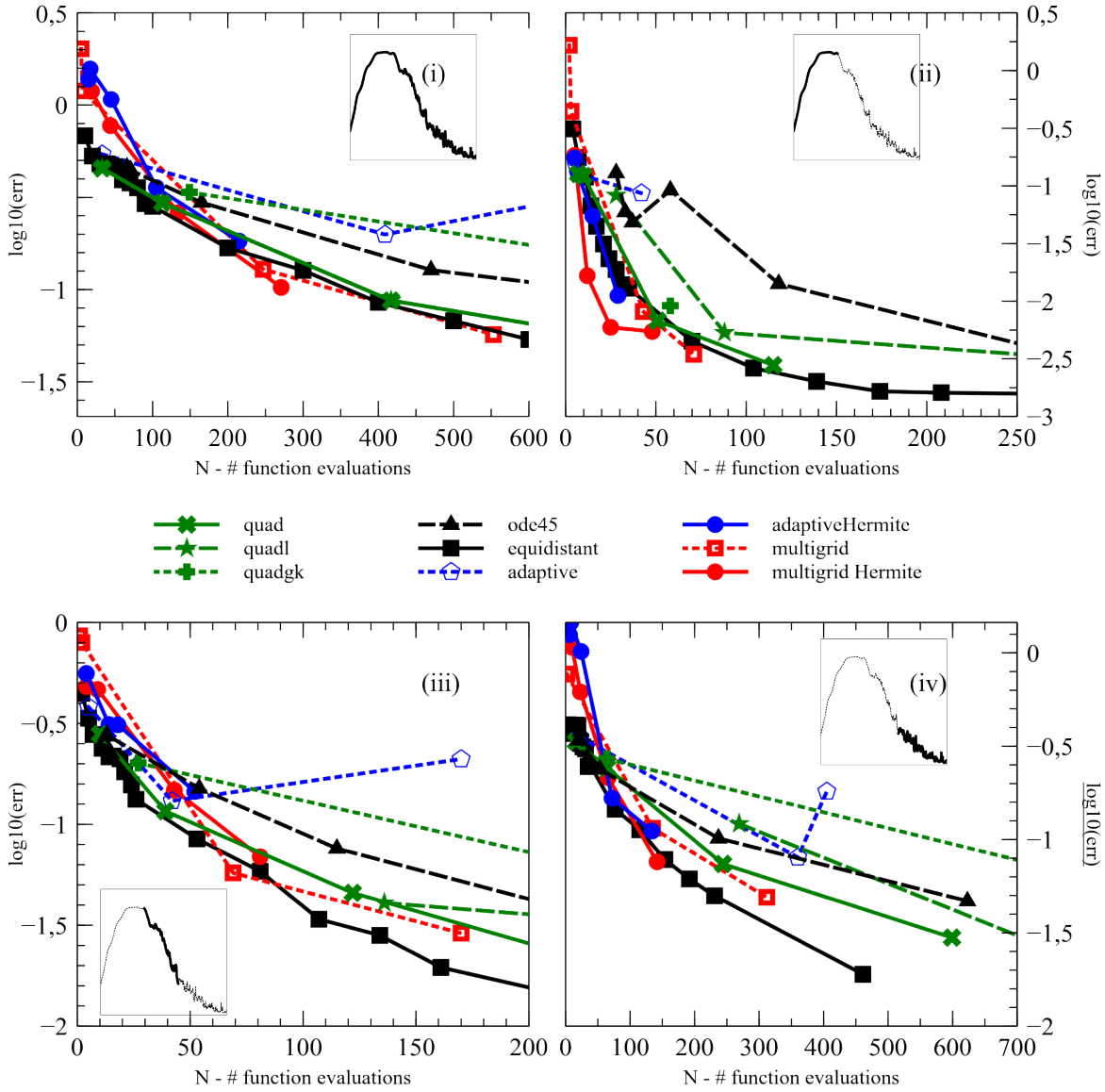

Figure 5. The graphs show the logarithm of the interpolation error $\log_{10}(\|f - \tilde{f}\|_2)$ for different subintervals of the wavelength interval indicated by the insets where the active interval is marked with a thicker line. The legend shown in the center is valid for all four graphs. The Hermite integration routines have a slight advantage in the interval where the profile is smooth (upper right), but cannot outperform the equidistant sampling routines in the highly oscillatory parts of the spectrum (lower row). The other algorithms perform generally as well as a regular, equidistant sampling routine. The `ode45` interpolation error in the smooth interval (upper right) deviates significantly from the others in a smooth region where the polynomial interpolation should yield a reasonable estimate. The strange behavior of the dotted blue graph in the lower right graph can be explained by the way the routine distributes functional evaluations in the parameter interval.