European Community on Computational Methods in Applied Sciences

ECCOMAS

COMMUNICATED BY CARSTEN CARSTENSEN, CHAIRMAN OF THE ECCOMAS TECHNICAL COMMITTEE "SCIENTIFIC COMPUTING"

Rate Optimality of Adaptive Algorithms

The overwhelming practical success of adaptive mesh-refinement in computational sciences and engineering has recently obtained a mathematical foundation with a theory on optimal convergence rates. This article first explains an abstract adaptive algorithm and its marking strategy. Secondly, it elucidates the concept of optimality in nonlinear approximation theory for a general audience. It thirdly outlines an abstract framework with fairly general hypotheses (A1)—(A4), which imply such an optimality result. Various comments conclude this state of the art overview.

All details and precise references are found in the open access article [C. Carstensen, M. Feischl, P. Page, D. Praetorius, Comput. Math. Appl. 67 (2014)] at http://dx.doi.org/10.1016/j.camwa.2013.12.003.

The Algorithm

The geometry of the domain Ω in some boundary value problem (BVP) is often specified in numerical simulations in terms of a triangulation $\mathcal{T}$ (also called mesh or partition) which is a set of a large but finite number of cells (also called element-domains) $T_1, \dots, T_N$. Based on this mesh $\mathcal{T}$, some discrete model (e.g., finite element method (FEM)) leads to some discrete solution $U(\mathcal{T})$ which approximates an unknown exact solution u to the BVP. Usually, a posteriori error estimates motivate some computable error estimator

$$\eta(\mathcal{T})^2 = \sum_{j=1}^{N} \eta_{T_j}(\mathcal{T})^2.$$

The local contributions $\eta_{Tj}(\mathcal{T})$ serve as refinement-indicators in the adaptive mesh-refining algorithm, where the marking is the essential decision for refinement and written as a list of $\mathcal{M}$ cells (i.e. $\mathcal{M} \subseteq \mathcal{T}$) with some larger refinement-indicator. The refinement procedure then computes the smallest admissible refinement $\mathcal{T}'$ of the mesh $\mathcal{T}$ (see Section 3) such that at least the marked cells are refined.

The successive loops of those steps lead to the following adaptive algorithm, where the coarsest mesh $\mathcal{T}_0$ is an input data.

Adaptive Algorithm
Input: initial mesh $\mathcal{T}_0$
Loop: for $\ell = 0, 1, 2, \dots$ do steps 1-4:
1. Solve: Compute discrete approximation $U(\mathcal{T}_\ell)$.
2. Estimate: Compute refinement indicators $\eta_T(\mathcal{T}_\ell)$ for all $T \in \mathcal{T}_\ell$.
3. Mark: Choose set of cells to refine $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ (see Section 4 for details).
4. Refine: Generate new mesh $\mathcal{T}_{\ell+1}$ by refinement of at least all cells in $\mathcal{M}_\ell$ (see Section 3 for details).
Output: Meshes $\mathcal{T}_\ell$, approximations $U(\mathcal{T}_\ell)$, and estimators $\eta(\mathcal{T}_\ell)$.

The Optimality

Figure 1 displays a typical mesh for some adaptive 3D mesh-refinement of some L-shaped cylinder into tetrahedra with some global refinement as well as some local mesh-refinement towards the vertical edge along the re-entrant corner. The question whether this is a good mesh or not is an important issue in the mesh-design with many partially heuristic answers and approaches. We merely mention the coarsening techniques as in [Binev et al., 2004] when applied to the adaptive hp-FEM with the crucial decision about h- or p-refinement.

For the optimality analysis of the adaptive algorithm of Section 1, the

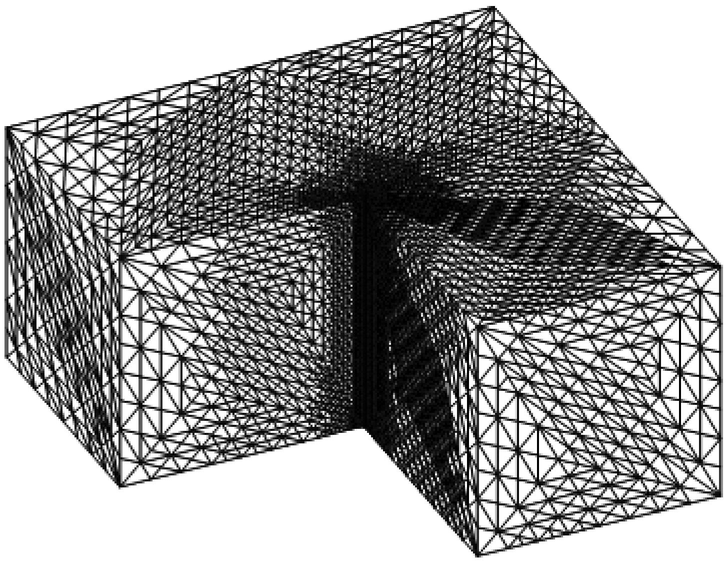


Figure 1: Strongly adaptively refined surface triangulation

natural comparison for optimality is with respect to the estimator η. The underlying class $\mathbb{T}$ of simplicial meshes is based, e.g., on newest vertex bisection (NVB) of an initial mesh $\mathcal{T}_0$; see, e.g., [Stevenson, 2008]. Since the typical work load is proportional to (and expected at least to be monotone increasing with) the number of tetrahedra $|\mathcal{T}|$ in the admissible mesh $\mathcal{T}$, given any non-negative integer N define

$$\mathbb{T}(N) := \{\mathcal{T} \in \mathbb{T} : |\mathcal{T}| \leq |\mathcal{T}_0| + N\}.$$

Then, the accuracy is measured in terms of the estimator η, and the optimal value for meshes with $\leq N$ extra tetrahedra, namely

$$\eta(\mathbb{T}(N)) := \min_{\mathcal{T}\in\mathbb{T}(N)} \eta(\mathcal{T}),$$

is studied as a function of N = 0,1,2,... and may be compared with $\eta(\mathcal{T}_\ell)$ of the computed solution and the number $N_\ell = |\mathcal{T}_\ell|$-$|\mathcal{T}_0|$ of extra tetrahedra in the computed triangulation of level ℓ = 0,1,2,...

Figure 2 presents a schematic scenario in a log-log scale which is written explicitly near the axes. The entries (N, $\eta(\mathbb{T}(N))$) are shown in red as a decreasing sequence for N =0,1,2,... of points as well as the corresponding entries (N_ℓ, $\eta(\mathcal{T}_\ell)$) for ℓ = 0,1,2,... in blue. By definition,

$$\eta(\mathbb{T}(N_\ell)) \leq \eta(\mathcal{T}_\ell)$$

for all ℓ = 0,1,2,... The converse estimate is unclear and expected to fail in general. However, rate optimality of the adaptive algorithm leads to an asymptotic comparison: Suppose there exists a convergence rate 0 < σ < ∞ in the sence that

$$M := \min_{N=0,1,\ldots} (N+1)^{\sigma} \eta(\mathbb{T}(N)) < \infty.$$

Then, optimality means that the adaptively computed solutions $U(\mathcal{T}_\ell)$ with corresponding estimators allow for the same rate in the sense that

$$\sup_{\ell=0,1,\ldots} (N_\ell+1)^{\sigma} \eta(\mathcal{T}_\ell) \leq C_{\text{qop}} M.$$

The visualization in Figure 2 shows two parallel straight lines of slope - σ in the log-log scaling. In fact, the log-transform of the above rate condition shows that

$$\log(\eta(\mathbb{T}(N))) \leq \log(M) - \sigma \log(N+1)$$

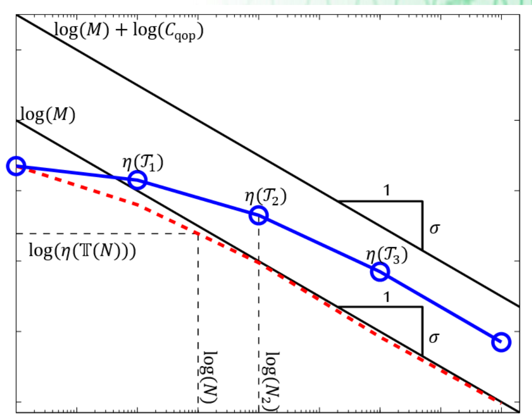
for all N =0,1,2,.... In other words, this straight line is an upper bound of the entries of the optimal meshes with an additive constant log(M) and multiplicative factor - σ. The constant $C_{\text{qop}} < \infty$ leads to a shift of the upper bound for the computed entries. The parallel straight line with an additive constant log(M) + log(C_{qop}) (and the same slope - σ) is in fact an upper bound. Stated explicitly, rate optimality of the adaptive algorithm means that the computed values (N_ℓ, $\eta(\mathcal{T}_\ell)$) will asymptotically below a curve parallel to the optimal curve (N, $\eta(\mathbb{T}(N))$) .

The optimality results from [Binev et al., 2004] and [Stevenson, 2007] show that (under some conditions) the same rate holds for the computed value in the sense that the generic constant $C_{\text{qop}} < \infty$ depends only on $\mathcal{T}_0$ and on the optimal rate σ as well as on the marking parameter θ from Section 4. The constants in the axioms of Section 5 below determine the quasi-optimality constant C_{qop}.

From the view of computational efficiency, not only the convergence rate σ is important, but also the number of adaptive steps. An adaptive algorithm could refine only a few elements in each step and, despite converging with optimal rate, may turn out to be extremely inefficient (with respect to CPU time).

The above considerations are formalized in the following two main results, which hold under the axioms (A1)—(A4) of Section 5.

Main Result 1: The adaptive algorithm guarantees linear convergence in the sense of



Fig

Figure 2: Schematic visualization of rate optimality

$$\eta(\mathcal{T}_{\ell+n})^2 \leq \frac{q^n}{1-q} \eta(\mathcal{T}_\ell)^2$$

for all ℓ= 0,1,2,... with some constant 0 < q < 1.

Main Result 2: For some sufficiently small adaptivity parameter θ from Section 4, the adaptive algorithm is quasi-optimal in the sense that it reveals the optimal rate of convergence.

The Refinement

The most fundamental property of NVB are the γ-*shape regularity,* which ensures that the cells of the meshes do not degenerate (i.e., the interior angles are bounded below), and the K -*mesh property*, which ensures that neighbouring cells in the mesh are of comparable size. This is why refinement of a cell may enforce refinements of additional cells and hence generates more refined cells $\mathcal{T}\backslash\mathcal{T}'$ than marked cells $\mathcal{M}$. However, both properties are only implicitly necessary for the validation of the axioms of Section 5.

Since the number of refined cells is the only factor which distinguishes adaptive refinement from uniform refinement, it is important to control this overhead. To that end, the analysis builds on the *closure estimate from [Binev et al., 2004]*

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \leq C_{\text{BDV}} \sum_{j=0}^{\ell-1} |\mathcal{M}_j|,$$

where the constant $C_{BDV} \geq 1$ depends only on $\mathcal{T}_0$ and $\mathcal{T}_{j+1}$, for all $j = 1, \ldots, \ell$-1, is the coarsest admissible refinement of $\mathcal{T}_j$, where all cells $T \in \mathcal{M}_j$ are bisected.

Moreover, a cell of $\mathcal{T}_{\ell-1}$ is refined into at most $C + 1$ son cells in $\mathcal{T}_\ell$. Counterexamples in the literature (even 1D bisection) show that

$$|\mathcal{T}_\ell| - |\mathcal{T}_{\ell-1}| \leq C|\mathcal{M}_{\ell-1}|,$$

cannot hold with some ℓ-independent constant $C > 0$ for refinement strategies which satisfy the K -mesh property mentioned above.

The optimality analysis relies on the comparison of different meshes. To that end, it is important that each two meshes $\mathcal{T}$, $\mathcal{T}'$ which are refinements of $\mathcal{T}_0$ have a coarsest common refinement $\mathcal{T} \oplus \mathcal{T}'$ in the sense that $\mathcal{T} \oplus \mathcal{T}'$ is a refinement of both $\mathcal{T}$ and $\mathcal{T}'$ and has less or equal cells than $|\mathcal{T}|+|\mathcal{T}'|-|\mathcal{T}_0|$.

Finally, a necessary requirement is that NVB has no blind ends. This means that for each refinement $\mathcal{T}$ of $\mathcal{T}_0$ and all $\varepsilon > 0$, there exists a refinement $\mathcal{T}'$ of $\mathcal{T}$ such that

$$\|u - U(\mathcal{T}')\| \leq \varepsilon.$$

Instead of NVB, any othere mesh-refinement could be used which guarantees the aforementioned properties.

THE MARKING

In each adaptive step, the error estimator $\eta_T(\mathcal{T}_\ell)$ gives a *heuristic* measure of the error on each cell $T \in \mathcal{T}_\ell$. If the adaptive algorithm is supposed to reduce the error sufficiently fast (i.e. by a factor 0 < q < 1 each step), it is sufficient and even necessary (see [Carstensen et al. 2014] for a proof) to apply the following Dörfler marking criterion [Dörfler, 1996] to identify the cells $\mathcal{M}_\ell$ to be refined: Given some fixed adaptivity parameter $0 < \theta \leq 1$, find a set $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of (almost) minimal cardinality with

$$\theta\eta(\mathcal{T}_\ell)^2 \leq \sum_{T\in\mathcal{M}_\ell}\eta_T(\mathcal{T}_\ell)^2.$$

The naive approach to find the set $\mathcal{M}_\ell$ sorts the list of cells such that

$$\eta_{T_1}(\mathcal{T}_\ell) \geq \cdots \geq \eta_{T_N}(\mathcal{T}_\ell)$$

and defines $\mathcal{M}_\ell := \{\mathcal{T}_1, \ldots, \mathcal{T}_j\}$ with the smallest index $1 \leq j \leq |\mathcal{T}_\ell| = N$ such that the Dörfler marking is satisfied. However, sorting requires at least $N \log(N)$ operations and therefore renders a (theoretical) bottleneck. An approximate bin sort algorithm [Stevenson, 2007] determines $\mathcal{M}_\ell$ with almost minimal cardinality (up to the multiplicative factor two) in $O(N)$ operations.

THE AXIOMS

If one aims at optimal asymptotic error reduction, the error estimator should satisfy the following four axioms. For simplicity, we abbreviate

$$\eta_{\mathcal{R}}(\mathcal{T})^2 := \sum_{T\in\mathcal{R}} \eta_T(\mathcal{T})^2$$

for the error estimator on any subset $\mathcal{R} \subseteq \mathcal{T}$ of cells. The generic constants $0 < q_{red} < 1$ and $C_{stab}, C_{red}, C_{orth}, C_{rel} \geq 1$ depend only on $\mathcal{T}_0$.

(A1) Stability: For a refinement $\mathcal{T}'$ of $\mathcal{T}$, $\mathcal{T} \cap \mathcal{T}'$ denotes the non-refined cells. The difference of the corresponding error indicators should be bounded by the difference of the solutions $U(\mathcal{T})$ and $U(\mathcal{T}')$ in the sense

$$\eta_{\mathcal{T}\cap\mathcal{T}'}(\mathcal{T}) - \eta_{\mathcal{T}\cap\mathcal{T}'}(\mathcal{T}') \leq C_{\text{stab}}\|U(\mathcal{T}) - U(\mathcal{T}')\|.$$

(A2) Reduction: For a refinement $\mathcal{T}'$ of $\mathcal{T}$, $\mathcal{T}\backslash\mathcal{T}'$ contains the refined cells of $\mathcal{T}$, whereas $\mathcal{T}'\backslash\mathcal{T}$ contains the newly generated cell of $\mathcal{T}'$. Each refinement of a cell reduces the mesh-width and hence should reduce the estimator in the sense

$$\eta_{\mathcal{T}'\backslash\mathcal{T}}(\mathcal{T}')^2 \leq q_{\text{red}}\eta_{\mathcal{T}\backslash\mathcal{T}'}(\mathcal{T})^2 + C_{\text{red}}\|U(\mathcal{T}) - U(\mathcal{T}')\|^2.$$

(A3) Orthogonality: Each iteration of the adaptive Algorithm improves the solution by adding $U(\mathcal{T}_{\ell+1}) - U(\mathcal{T}_\ell)$ to the existing approximation $U(\mathcal{T}_\ell)$ The estimator $\eta(\mathcal{T}_\ell)$ should control all subsequent steps of the adaptive Algorithm in the sense that

$$\sum_{k=\ell}^{\infty}\|U(\mathcal{T}_{k+1}) - U(\mathcal{T}_k)\|^2 \leq C_{\text{orth}}\eta(\mathcal{T}_\ell)^2.$$

(A4) Discrete reliability: The adaptive algorithm controls only the error estimator $\eta(\mathcal{T})$. Since the governing quantity for approximation quality is the error $\|u - U(\mathcal{T})\|$, the error estimator should bound the error from above in the sense that any refinement $\mathcal{T}'$ of $\mathcal{T}$ satisfies

$$\|U(\mathcal{T}') - U(\mathcal{T})\| \leq C_{\text{rel}}\,\eta_{\mathcal{R}}(\mathcal{T})$$

for some set $\mathcal{T}\backslash\mathcal{T}' \subseteq \mathcal{R} \subseteq \mathcal{T}$ with

$$|\mathcal{R}| \leq C_{\text{ref}}|\mathcal{T} \setminus \mathcal{T}'|.$$

In many FEM applications, it suffices to consider $\mathcal{R} = \mathcal{T}\backslash\mathcal{T}'$. However, some more involved applications may require $\mathcal{R}$ to contain a certain number of cell layers around the refined cells $\mathcal{T}\backslash\mathcal{T}'$. This discrete reliability implies

$$\|u - U(\mathcal{T})\| \leq C_{\text{rel}}\,\eta(\mathcal{T}).$$

THE HISTORY

The Main Results 1—2 of Section 2 are the accumulation of the following seminal results. [Dörfler, 1996] introduced the marking criterion from Section 4 and proved linear convergence of the error for some FEM for the Poisson problem up to some tolerance. [Morin et al., 2000] extended the analysis and included data approximation to prove convergence of a practical adaptive algorithm. [Binev et al., 2004] first proved convergence with optimal rates in the sense of Section 2 for the Poisson problem. However, their analysis required an additional mesh-coarsening step in the adaptive algorithm. [Stevenson, 2007] removed this coarsening step and proved convergence with optimal rates for the adaptive algorithm of Section 1. [Cascon et al., 2008] included standard newest vertex bisection as mesh refinement into the mathematical analysis.

Until then, only variations of FEM for the Poisson model problem with homogeneous Dirichlet boundary conditions were analyzed in the literature. Independently, [Feischl et al., 2013] and [Gantumur, 2013] developed the analysis for integral equations and proved convergence with optimal rates for standard boundary element methods (BEMs). For his contributions to [Feischl et al., 2013] and the field of adaptive BEM, Michael Karkulik won the Dr. Körper award 2013 of GAMM (Gesellschaft für Angewandte Mathematik und Mechanik). [Aurada et al., 2013] proved optimal convergence rates for FEM for the Poisson problem with general boundary conditions. Finally, [Feischl et al., 2014] concluded the theory for general second-order linear elliptic PDEs.

The recent work [Carstensen et al. 2014] collects all the mentioned seminal works in a unifying and abstract framework. The work identifies the axioms (A1)—(A4) from Section 5 and proves optimal rates for any problem that fits in the abstract setting. The latter covers the existing literature on rate optimality for conforming FEM (also the known results for nonlinear problems) and BEM as well as nonconforming and mixed FEM. With some additional (resp. relaxed) axioms, the abstract framework of [Carstensen et al., 2014] covers also inexact solvers and other error estimators (e.g., ZZ-type averaging error estimators).

THE COMPLEXITY

The asymptotic optimality notion of Section 2 may be seen as a first and important step towards a most effective computation. The computational complexity involves the usage of the iterative solver in the adaptive algorithm. The above results hold under the underlying assumption that the discrete solution as well as the estimators is computed exactly, which is unrealistic once optimal solvers (e.g., multigrid or BPX pre-conditioned CG) are employed. Section 7 of [Carstensen et al., 2014] shows a way to modify the adaptive algorithm: One needs to control the termination error in terms of the estimator and to engage some perturbation of the arguments behind the analysis for exact solve.

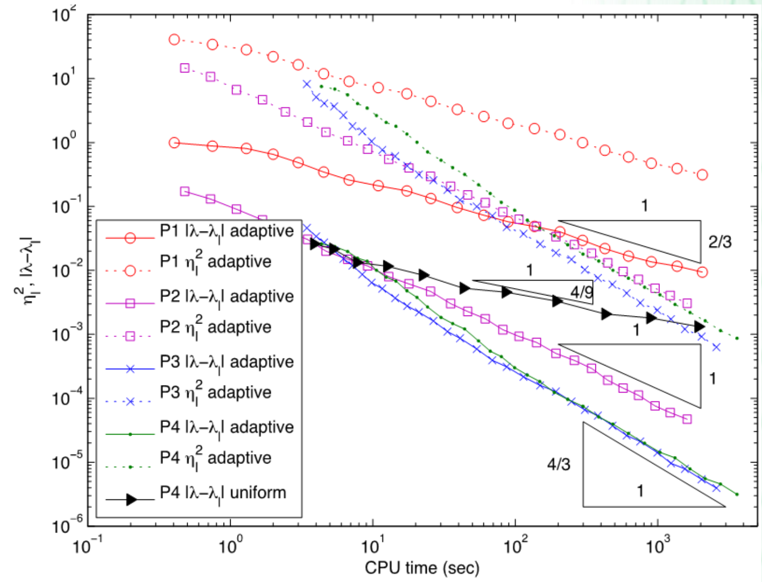


Figure 3: Convergence history plot of accuracy versus CPU time in a log-log scale proves optimal computational complexity

The situation is more dramatic for nonlinear problems, which always require an iterative solution procedure. Under realistic assumptions on the practical performance of the algebraic eigensolver (with multigrid preconditioning) [Carstensen et al., 2012] showed overall optimal complexity. Figure 4 displays numerical simulations for the first eigenvalue of the Laplacian in the 3D geometry of Figure 1. This numerical evidence suggests a practical optimal complexity and won Joscha Gedicke the SIAM student paper prize 2013.

Despite the first success, the overall proof of optimal computational complexity has to combine an analysis of optimal mesh-design with an analysis of the iterative solution process, and the emerging theory is still in its infancy for many important applications.

CARSTEN CARSTENSEN
DEPARTEMENT OF MATHEMATICS
HUMBOLDT-UNIVERSITÄT ZU BERLIN

MICHAEL FEISCHL
DIRK PRAETORIUS
INSTITUTE FOR ANALYSIS AND SCIENTIFIC COMPUTING
VIENNA UNIVERSITY OF TECHNOLOGY

CC@MATH.HU-BERLIN.DE

THE AUTHORS MF AND DP ACKNOWLEDGE FINANCIAL SUPPORT THROUGH THE AUSTRIAN SCIENCE FUND (FWF) UNDER GRANT P21732 AND W1245