Mathematisches Forschungsinstitut Oberwolfach

Report No. 44/2016

DOI: 10.4171/OWR/2016/44

Adaptive Algorithms

Organised by
Carsten Carstensen, Berlin
Rob Stevenson, Amsterdam

18 September – 24 September 2016

Abstract. Overwhelming empirical evidence in computational science and engineering proved that self-adaptive mesh-generation is a must-do in real-life problem computational partial differential equations. The mathematical understanding of corresponding algorithms concerns the overlap of two traditional mathematical disciplines, numerical analysis and approximation theory, with computational sciences. The half workshop was devoted to the mathematics of optimal convergence rates and instance optimality of the Dörfler marking or the maximum strategy in various versions of space discretisations and time-evolution problems with all kind of applications in the efficient numerical treatment of partial differential equations.

Mathematics Subject Classification (2010): 65N30, 65N12, 65N50, 65M12, 41A25.

Introduction by the Organisers

The efficient numerical solution of PDEs requires an adaptive solution process in the sense that the discretization that is employed should depend in a proper, or rather optimal way on the solution. Since this solution is a priorily unknown, the adaptive algorithm has to extract the information where to refine from a sequence of increasingly more accurate computed numerical approximations, in a loop that is commonly denoted as **solve–estimate–mark–refine**. This loop has already been introduced by Babuska and co-workers in the seventies (eg. [1]), but is was not before 1997 that in [5], for linear elliptic PDEs in more than one dimensions, an adaptive finite element method (AFEM) was *proven* to converge to the exact solution.

Although this result meant an important step forward, it did not show that adaptive methods provide advantages over non-adaptive methods. These advantages are apparent from practical results, which often even show that adaptivity is paramount to be able to solve certain classes of complicated problems. This theoretical gap was closed some years later in [2, 6, 4], where it was shown that asymptotically AFEM converges with the *best possible rate* among all possible choices of the underlying partitions.

These initial results were restricted to model second order linear elliptic PDEs, discretized with standard conforming finite elements of a fixed polynomial degree, but they were a starting point for many researchers for proving similarly strong results for much larger classes of problems and discretizations. An abstract framework based on four assumptions ('axioms') (A1)-(A4), which identifies the basic mechanisms that underly the convergence and rate optimality proofs, and with that foster further applications was recently given in [3].

It was the purpose of this workshop to bring together experts that contributed to the recent developments in this field in order to exchange ideas, to stimulate further collaborations, and to identify or to make first steps in tackling remaining open challenges as *hp*-adaptivity, and optimal adaptive methods for instationary problems to name two important ones.

Moving to the contents of the in total 20 presentations, the aforementioned axioms do not not cover AFEMs with so-called separate marking for controlling data oscillation, as required for mixed and least squares problems. In her talk, Hella Rabus presented an even more general framework that applies to both collective and separate marking strategies.

The AFEM convergence theory was initially developed for symmetric positive definite (elliptic) problems, whereas later it was shown that possibly non-symmetric compact perturbations could be added. In his talk, Michael Feischl presented his recent ideas about how the framework could be extended to include strongly non-symmetric problems.

The application of the axioms for examples with non-conforming finite element schemes was discussed for the discrete reliability (A3) by Dietmar Gallistl and for boundary element schemes by Markus Melenk in terms of inverse estimates for the stability (A1) and (A2).

The Simons-professor Jun Hu presented three recent results on mixed finite element methods: an innovative class of novel mixed finite element methods with pointwise symmetric stress approximations in elasticity, an abstract framework for the convergence of adaptive algorithms, and preconditioners for the resulting system. He pointed out that there exist no convergence results for the Hellinger-Reissner formulation even in linear elasticity.

Joscha Gedicke presented the first robust a posteriori error estimators towards such an adaptive strategy and some numerical illustrations, where –to the surprise of the experienced audience– the bulk parameter in the Dörfler marking needed to be significantly smaller than 1/2. This and the size of the crucial bulk parameter in the framework of the aforementioned axioms showed the limitations of our

current understanding. Theoretical estimates in a worst case scenario suggest that it needs to be unrealistically small in comparison with the overall empirical knowledge of the workshop's participants. Speculations arose whether this implies that the current understanding of the optimal convergence rates leave out some undiscovered essentials. In the discussion, Carsten Carstensen also pointed out some difficulties in the treatment of vertices at the traction boundary in elasticity caused by the fact that the new and older symmetric stress approximations have nodal degrees of freedom whose naive approximation leads to contradictions even in simple benchmark examples and enforce a pre-asymptotic effect. The relaxed atmosphere in Oberwolfach allowed discussions amongst the experts and indeed sorted out this difficulty during the week, which will be visible in future research e.g. in ongoing projects of the SPP 1748 of the German research foundation.

Using techniques developed in the context of AFEM for eigenvalue problems, Alan Demlow demonstrated his proof of convergence and optimality of AFEM for harmonic forms. A major technical difficulty was the non-nestedness of the trial spaces under refinements. Another interesting class of non-standard problems is that of linear operator equations posed on Banach spaces. For those problems, Kris van der Zee presented a generalization of the Petrov-Galerkin methods with optimal test spaces, which methods in the generalized setting become nonlinear.

The class optimality result proven for AFEM says that if a solution can be approximated at a certain algebraic rate s, then the sequence of approximations produced by AFEM converges with that rate. Gantumur Tsogtgerel spoke about the question which functions can be approximated at rate s, characterized in terms of their smoothness. He presented new subtle results on the relation between approximation classes and classical smoothness spaces as Besov spaces.

Nearly all convergence results of AFEM are for bulk chasing, also known as the Dörfler marking strategy. An exception is given by a recent paper of Diening, Kreuzer and Stevenson in which, for the Poisson problem in two dimensions, even instance optimality is proven for an AFEM with a modified maximum marking strategy. In her talk, Mira Schedensack presented a generalization of this result to both Poisson and Stokes problems discretized with nonconforming Crouzeix Raviart finite elements.

Parameter dependent PDEs with a possibly infinite number of parameters are nowadays studied intensively. They arise for example by replacing a random coefficient field in a PDE by its Karhunen-Loeève expansion. In order to cope with the curse of dimensionality, one considers sparse (polynomial) expansions, or low rank approximations. In his talk, Wolfgang Dahmen presented an optimally converging adaptive solver in either of both formats. He showed that depending on the model either of the formats can realize the desired tolerance with the smallest number of terms.

For linear elliptic PDEs, it is known that AFEM does not only converge optimally in terms of the number of unknowns, but also in terms of the computational cost. For the latter it is needed that the exact solution of the arising Galerkin problems is replaced by an optimal iterative solution within a sufficiently small

relative tolerance. Dirk Praetorius spoke about a generalization of this result to nonlinear strongly monotone operators. As a first step he showed convergence of an AFEM with an iterative solver based on a Picard iteration. A related topic was discussed by Lars Diening. He presented a new algorithm of Kačanov type that can be used as an iterative solver of the nonlinear Galerkin problems that arise from the p-Laplacian.

So far most results about convergence of AFEM are for stationary (elliptic) problems. It can be foreseen that convergence and perhaps optimality of adaptive methods for time-dependent problems, in particular those of parabolic type, will an important topic in the coming years. In his talk, Omar Lakkis presented a posteriori error bounds for fully discrete Galerkin time-stepping methods using elliptic and time reconstruction operators.

The solution of a nonlinear parabolic problem may blow up in finite time. Emmanuil Georgoulis presented a conditional a posteriori bound for a fully-discrete first order in time implicit-explicit interior penalty discontinuous Galerkin in space discretization of a non self-adjoint semilinear parabolic PDE with quadratic nonlinearity. When used in a space-time adaptive algorithm to control the time step lengths and the spatial mesh modifications, the method detects and converges to the blow-up time without surpassing it.

Adaptive methods based on piecewise polynomial approximation of fixed degree can give at best algebraic convergence rates. Spectral- or hp finite element methods can yield even exponential convergence rates. Claudio Canuto showed how a convergent hp-afem can be turned into an instance optimal hp-afem by the addition of coarsening. An hp-adaptive tree algorithm that is perfectly suited for this task was presented by Peter Binev.

Having a p-robust a posteriori estimator is instrumental for the design of an hp-afem. In his talk, Martin Vohralík showed that the so-called equilibrated flux estimator, that is known to p-robust in two dimensions, is also p-robust in three dimensions. Serge Nicaise showed that this kind of error estimator is reliable and efficient for a magnetodynamic harmonic formulation of Maxwell's system.

Marco Verani considered adaptive spectral methods. He showed that exponential convergence rates can be achieved even without coarsening by the application of an 'aggressive' dynamic marking strategy.

References

[1] I. Babuška and M. Vogelius. Feedback and adaptive finite element solution of one-dimensional boundary value problems. *Numer. Math.*, 44(1):75–102, 1984.

[2] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219 – 268, 2004.

[3] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. *Comput. Math. Appl.*, 67(6):1195–1253, 2014.

[4] J.M. Cascon, Ch. Kreuzer, R.H. Nochetto, and K.G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.

[5] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996.
[6] R.P. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1049268, "US Junior Oberwolfach Fellows". Moreover, the MFO and the workshop organizers would like to thank the Simons Foundation for supporting Professor Jun Hu in the "Simons Visiting Professors" program at the MFO.

Workshop: Adaptive Algorithms

Table of Contents

Abstracts

A guaranteed equilibrated error estimator for the $\mathbf{A}-\varphi$ and $\mathbf{T}-\Omega$ magnetodynamic harmonic formulations of Maxwell's system

Serge Nicaise
(joint work with E. Creusé, R. Tittarelli)

In this talk a guaranteed equilibrated error estimator for the harmonic magnetodynamic formulation of Maxwell's system is proposed [Creusé, S. Nicaise and R. Tittarelli, A guaranteed equilibrated error estimator for the $\mathbf{A}-\varphi$ and $\mathbf{T}-\Omega$ magnetodynamic harmonic formulations of the Maxwell system, IMA Journal of Numerical Analysis, 2016]. First of all, since the estimator is based on two dual problems, these two equivalent potential formulations of Maxwell's system are presented. Secondly, the proof of the two key properties for an *a posteriori* equilibrated error estimator are proved: the reliability and efficiency results without generic constants. Finally, two numerical simulations validate the applicability of our estimator.

Let D be an open simply connected bounded polyhedral domain with a simply connected boundary Γ and let $D_c \subset D$ be the conductor domain supposed to be simply connected with a simply connected boundary Γ_c. The model of interest is given by the quasi-static approximation of Maxwell's equations in the magnetoharmonic regime, completed by the constitutive laws:

$$\mathbf{B} = \mu\,\mathbf{H} \quad \text{in the whole domain } D$$

and

$$\mathbf{J}_e = \sigma\,\mathbf{E} \quad \text{in the conductor domain } D_c\,.$$

Here $\mathbf{B}$, $\mathbf{H}$, $\mathbf{J}_e$ and $\mathbf{E}$ represent respectively the magnetic flux density, the magnetic field, the eddy current density and the electric field. Moreover, for the magnetic permeability $\mu \in L^\infty(D)$ and the electrical conductivity $\sigma \in L^\infty(D)$, we assume that $\sigma \equiv 0$ on $D\backslash D_c$ and that there exist $\mu_0 > 0$ and $\sigma_0 > 0$ such that $\mu > \mu_0$ and $\sigma > \sigma_0$ in D_c.

The *a posteriori* error estimator is built starting from the numerical solutions of two classical potential formulations. On one hand, the original system can be rewritten by a magnetic vector potential $\mathbf{A}$, defined in D, as well as an electrical scalar potential φ, defined in D_c. The finite element method resolution of that $\mathbf{A}-\varphi$ formulation, based on Nédélec and Lagrange elements, provides the following numerical solutions:

$$\mathbf{B}_h = \operatorname{curl} \mathbf{A}_h \ \text{ in } D \ \text{ and } \ \mathbf{E}_h = -i\omega\mathbf{A}_h - \nabla\varphi_h \ \text{ in } D_c\,.$$

On the other hand, the original system can be also rewritten by an electric vector potential $\mathbf{T}$, defined in D_c, as well as a magnetic scalar potential Ω, defined in

D. The finite element method resolution of that $\mathbf{T}-\Omega$ formulation, still based on Nédélec and Lagrange elements, provides the numerical solutions:

$$\mathbf{H}_h = \mathbf{H}_s + \mathbf{T}_h - \nabla\Omega_h \ \text{ in } D \ \text{ and } \ \mathbf{J}_h = \text{curl } \mathbf{T}_h \ \text{ in } D_c\,,$$

where $\mathbf{J}_s = \text{curl } \mathbf{H}_s$ denotes the source term.

We then design an error estimator η for the following energy norm of the error e

$$\begin{aligned} e \ = \ & (|| \, \mu^{-1/2}(\mathbf{B}-\mathbf{B}_h) \, ||^2_{L^2(D)} + || \, \mu^{1/2}(\mathbf{H}-\mathbf{H}_h) \, ||^2_{L^2(D)} \\ & + || \, (\omega\,\sigma)^{-1/2}(\mathbf{J}-\mathbf{J}_h) \, ||^2_{L^2(D_c)} + || \, \omega^{-1/2}\,\sigma^{1/2}\,(\mathbf{E}-\mathbf{E}_h) \, ||^2_{L^2(D_c)})^{1/2}. \end{aligned}$$

The estimator η measures the non-verification of the constitutive laws of the numerical fields. Indeed it is defined as

$$\eta^2 \ = \ \sum_{T\in\mathcal{T}_h} \eta^2_{m,T} \ + \ \sum_{T\in\mathcal{T}_h, T\subset D_c} \eta^2_{e,T}\,,$$

where $\mathcal{T}_h$ represents a tetrahedral and regular mesh and the element estimator contibutions are defined by

$$\eta_{m,T} = ||\mu^{1/2}(\mathbf{H}_h - \mu^{-1}\mathbf{B}_h)||_T \text{ and } \eta_{e,T} = || \, (\omega\sigma)^{-1/2}\,(\mathbf{J}_h - \sigma\mathbf{E}_h) \, ||_T.$$

The main result consists in proving the global equivalence between the error e and the estimator η, that is

$$\eta^2 = e^2 + r\,,$$

where r is a higher order term. The effort with respect to the magnetostatic case stays on the estimation of the remainder r, which vanishes in the static case. Thus it is estimated with the help of some functional analysis tools and using a sort of Aubin-Nitsche's trick. In addition, for adaptivity purposes, the local efficiency property is also proved, namely we show that

$$\eta_T = (\eta^2_{m,T} + \eta^2_{e,T})^{1/2} \ \leq \ \sqrt{2}\, e_{|\,T}, \forall T \in \mathcal{T}_h.$$

Finally, two benchmark tests are exihibed: an analytical one that illustrates the obtained theoretical results and a physical one that illustrates the estimator performance in an adaptive mesh refinement context.

Axioms of adaptivity for separate marking

Hella Rabus
(joint work with Carsten Carstensen)

The design of efficient numerical algorithms is one of the main challenges in computational mathematics. Various applications lead to boundary value problems with partial differential equations, whose discrete weak version is solved with finite element methods on the computer. It is well known, that uniform mesh-refinement eventually leads to suboptimal convergence rates. Thus, the design of optimal convergent adaptive finite element methods and its analysis is of particular interest.

Mixed finite element methods with flux errors in $H(\mathrm{div})$-norms and div-least-squares finite element methods require the separate marking strategy for data approximation reduction in obligatory adaptive mesh-refining. The refinement indicator $\sigma_\ell^2(K) = \eta_\ell^2(K) + \mu^2(K)$ of a finite element domain K in a triangulation $\mathcal{T}_\ell$ on the level ℓ consists of some residual-based error estimator η_ℓ with some reduction property under local mesh-refining and some data approximation error μ_ℓ. Separate marking (Safem) means either Dörfler marking if $\mu_\ell^2 \leq \kappa\eta_\ell^2$ or otherwise an optimal data approximation algorithm run with controlled accuracy as established in [CR11, Rab15] and reads as follows

for $\ell = 0, 1, \ldots$ **do**
  Compute $\eta_\ell(K)$, $\mu(K)$ for all $K \in \mathcal{T}_\ell$
  if $\mu_\ell^2 := \mu^2(\mathcal{T}_\ell) \leq \kappa\eta_\ell^2(\mathcal{T}_\ell, \mathcal{T}_\ell)$ **then**
    $\mathcal{T}_{\ell+1} := \texttt{Dörfler\_marking}(\theta_A, \mathcal{T}_\ell(K) : K \in \mathcal{T}_\ell)$
  else
    $\mathcal{T}_{\ell+1} := \mathcal{T}_\ell \oplus \texttt{approx}(\rho_B\mu_\ell^2, \mu(K) : K \in \mathcal{T}_\ell)$.

The enfolded set of axioms simplifies [CFPP14] for collective marking (with $\sigma^2 = \eta^2 + \mu^2$ for Case A and $\mu^2 \equiv 0$ for Case B), treats separate marking established for the first time in an abstract framework, generalizes [CP15] for least-squares schemes, and extends [CR11] to the mixed FEM with flux error control in $H(\mathrm{div})$.

The axioms (A1)–(A4) involve $\rho_2 < 1, \Lambda_k < \infty$, estimators σ, η, μ and distances $0 \leq \delta(\mathcal{T}, \hat{\mathcal{T}}) < \infty$ for all $\in \mathcal{T} \in \mathbb{T}$ and $\hat{\mathcal{T}} \in \mathbb{T}(\mathcal{T})$ and are sufficient for optimal asymptotic convergence rates. There exists $\mathcal{R}(\mathcal{T}, \hat{\mathcal{T}}) \subset \mathcal{T}$ such that $\mathcal{T} \setminus \hat{\mathcal{T}} \subseteq \mathcal{R}(\mathcal{T}, \hat{\mathcal{T}}) \wedge |\mathcal{R}(\mathcal{T}, \hat{\mathcal{T}})| \leq \Lambda_3 |\mathcal{T} \setminus \hat{\mathcal{T}}|$ and

(A1) $$|\eta(\hat{\mathcal{T}}, \mathcal{T} \cap \hat{\mathcal{T}}) - \eta(\mathcal{T}, \mathcal{T} \cap \hat{\mathcal{T}})| \leq \Lambda_1 \delta(\mathcal{T}, \hat{\mathcal{T}}),$$

(A2) $$\eta(\hat{\mathcal{T}}, \hat{\mathcal{T}} \setminus \mathcal{T}) \leq \rho_2 \eta(\mathcal{T}, \mathcal{T} \setminus \hat{\mathcal{T}}), + \Lambda_2 \delta(\mathcal{T}, \hat{\mathcal{T}})$$

(A3) $$\delta(\mathcal{T}, \hat{\mathcal{T}}) \leq \Lambda_3 \left(\eta^2(\mathcal{T}, \mathcal{R}) + \mu^2(\mathcal{T})\right) + \hat{\Lambda}_3 \eta(\hat{\mathcal{T}}),$$

(A4) $$\sum_{k=\ell}^{\infty} \delta^2(\mathcal{T}_k, \mathcal{T}_{k+1}) \leq \Lambda_4 \sigma_\ell^2 \quad \text{for all } \ell \in \mathbb{N}_0,$$

$\forall\,\mathrm{Tol} > 0$ $\mathcal{T}_{\mathrm{Tol}} = \texttt{data\_approx}(\mathrm{Tol}, \mu(K) : K \in \mathcal{T}_\ell) \in \mathbb{T}$ satisfies $\mu^2(\mathcal{T}_{\mathrm{Tol}}) \leq \mathrm{Tol}$ and

$$|\mathcal{T}_{\mathrm{Tol}}| - |\mathcal{T}_0| \leq \Lambda_5\,\mathrm{Tol}^{-1/(2s)}, \tag{B1}$$

$$\mu^2(\hat{\mathcal{T}}) \leq \Lambda_6 \mu^2(\mathcal{T}). \tag{B2}$$

Quasimonotonicity reads

$$\sigma^2(\hat{\mathcal{T}}) \leq \Lambda_7 \sigma^2(\mathcal{T}). \tag{QM}$$

Remark. $(\Lambda_1^2 + \Lambda_1^2)\hat{\Lambda}_3 < 1$ implies (QM) with Λ_7 depending on Λ_1, Λ_2, Λ_3, $\hat{\Lambda}_3$ and Λ_6.

Theorem. Suppose (A1)-(A4), (B1)-(B2) and (QM). Then, there exists $\kappa_0 > 0$ (which is $+\infty$ if $\Lambda_6 = 1$), such that for any choice of κ, θ_A, $0 < \rho_B < 1$ with $0 < \theta_A < \theta_0 := (1 - \kappa\Lambda_1^2\Lambda_3)/(1 + \Lambda_1^2\Lambda_3)$ and $\kappa < \kappa_1 := \min\left\{\frac{1-\rho_A}{\Lambda_6 - 1}, \Lambda_1^{-1}\Lambda_3^{-1}\right\}$ the output $(\mathcal{T}_\ell)_\ell$, $(\sigma_\ell)_\ell$ of SAFEM satisfy

$$\sup_{N\in\mathbb{N}_0} (1+N)^s \min_{\mathcal{T}\in\mathbb{T}(N)} \sigma(\mathcal{T}) \approx \sup_{\ell\in\mathbb{N}_0} (1 + |\mathcal{T}_\ell| - |\mathcal{T}_0|)^s \sigma_\ell.$$

In particular, if for some $s > 0$ the right-hand side is bounded and (B1) holds with the same s, then the left-hand side is bounded.

Examples. Separate marking for data approximation reduction is necessary for least-squares FEM and mixed FEM with convergence rates in $H(\mathrm{div}, \Omega) \times L^2(\Omega)$. Besides from natural a posteriori error control with residuals in least squares functional, [CP15] establishes an a posteriori error estimator in $H(\mathrm{div}, \Omega)$, namely $\sigma^2(\mathcal{T}, K) := \eta^2(\mathcal{T}, K) + \mu^2(K)$ for $K \in \mathcal{T}$ and $\mu^2(K) := \|f - \Pi_0 f\|^2_{L^2(K)}$. Since μ does not satisfy an estimator reduction SAFEM has to be applied instead of collective marking.

Let $(p_{LS}, u_{LS}) \in RT_K(\mathcal{T}) \times S_0^k(\mathcal{T})$ and $(\hat{p}_{LS}, \hat{u}_{LS}) \in RT_K(\hat{\mathcal{T}}) \times S_0^k(\hat{\mathcal{T}})$ be the discrete solution on $\hat{\mathcal{T}} \in \mathbb{T}(\mathcal{T})$ and $\mathcal{T} \in \mathbb{T}$, respectively. The proof of discrete reliability (A3) (for $k = 0$) still leaves the extra term

$$\begin{aligned}\|\hat{p}_{LS} - p_{LS} - \nabla(\hat{u}_{LS} - \hat{u}_{LS})\|^2_{L^2(\Omega)} + \|\mathrm{div}(\hat{p}_{LS} - p_{LS})\|^2_{L^2(\Omega)} \\ \lesssim \eta^2(\mathcal{T} \setminus \hat{\mathcal{T}}) + \|(1 - \Pi_0)\,\mathrm{div}\,\hat{p}_{LS}\|^2_{L^2(\Omega)}\,,\end{aligned}$$

which is not covered in [CFPP14]. The presented set of generalized axioms covers this special application [CR15].

Another example, the general second-order linear elliptic problem solved with MFEM with error estimation in $H(\mathrm{div})$ is analyzed in [CDR16].

The benefit of the new set of axioms is, that it guarantees rate optimality for AFEMs based on collective, as well as on separate marking. Thus, the existing literature of rate optimality of adaptive FEM is covered – to prove optimal rates, only the axioms (A1)-(A4), (B1)-(B2) and (QM) have to be verified.

References

[CDR16] C. Carstensen, A. Dond, H. Rabus. Quasi-optimality of adaptive mixed FEMs for general second-order linear elliptic problems. *in preparation*.

[CFPP14] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. *Comput. Methods Appl. Math.*, 67(6):1195–1253, 2014.

[CP15] C. Carstensen and E.-J. Park. Convergence and optimality of adaptive least squares finite element methods. *SIAM J. Numer. Anal.*, 53:43–62, 2015.

[CR11] C. Carstensen and H. Rabus. An optimal adaptive mixed finite element method. *Math. Comp.*, 80(274):649–667, 2011.

[CR15] C. Carstensen and H. Rabus. Axioms of adaptivity for separate marking. *arXiv:1606.02165 [math.NA], 2016*.

[Rab15] H. Rabus. Quasi-optimal convergence of AFEM based on separate marking – Part I and II. *Journal of Numerical Analysis*, 23(2):137–156, 57–174, 2015.

[Ste08] R. Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77:227–241, 2008.

Adaptive and Multilevel Mixed Finite Element Methods

Jun Hu

This talk consists of three parts. In the first part, we developed a new framework to design and analyze the mixed FEM for elasticity problems by establishing the following three main results:

(1) A crucial structure of the discrete stress space: on simplicial grids, the discrete stress space can be selected as the symmetric matrix-valued Lagrange element space, enriched by a symmetric matrix-valued polynomial H(div) bubble function space on each simplex; a corresponding choice applies to product grids.

(2) Two basic algebraic results: (1) on each simplex, the symmetric matrices of rank one produced by the tensor products of the unit tangent vectors of the (n+1)n/2 edges of the simplex, form a basis of the space of the symmetric matrices; (2) on each simplex, the divergence space of the above H(div) bubble function space is equal to the orthogonal complement space of the rigid motion space with respect to the corresponding discrete displacement space (Asimilar result holds on a macroelement for the product grids).

These define a two-step stability analysis which is new and different from the classic one in literature. As a result, on both simplicial and product grids, we were able to define the first families of both symmetric and optimal mixed elements with polynomial shape functions in any space dimension. Furthermore, the discrete stress space has a simple basis which essentially consists of symmetric matrix-valued Lagrange element basis functions.

On the simplicial grids, in order to avoid enriching $\boldsymbol{H}(\mathbf{div}, \Omega; \mathbb{S})$-$P_{n+1}$ face-bubble functions of piecewise polynomials or $\boldsymbol{H}(\mathbf{div})$-$P_k$ nonconforming face-bubble spaces for each $n-1$ dimensional simplex for the cases $1 \leq k \leq n$, we also designed two classes of stabilized mixed finite element methods. In the first class of elements, we use $\boldsymbol{H}(\mathbf{div}, \Omega; \mathbb{S})$-$P_k$ and $\boldsymbol{L}^2(\Omega; \mathbb{R}^n)$-$P_{k-1}$ to approximate the

stress and displacement spaces, respectively, for $1 \leq k \leq n$, and employ a stabilization technique in terms of the jump of the discrete displacement over the faces of the triangulation under consideration; in the second class of elements, we use $\boldsymbol{H}_0^1(\Omega;\mathbb{R}^n)$-$P_k$ to approximate the displacement space for $1 \leq k \leq n$, and adopt the stabilization technique suggested by Brezzi, Fortin, and Marini.

In the second part, we developed a block diagonal preconditioner with the minimal residual method and a block triangular preconditioner with the generalized minimal residual method for the aforementioned mixed finite element methods of linear elasticity. They are based on a new stability result of the saddle point system in mesh-dependent norms. The mesh-dependent norm for the stress corresponds to the mass matrix which is easy to invert while the displacement it is spectral equivalent to Schur complement. A fast auxiliary space preconditioner based on the H^1 conforming linear element of the linear elasticity problem is then designed for solving the Schur complement. For both diagonal and triangular preconditioners, it is proved that the conditioning numbers of the preconditioned systems are bounded above by a constant independent of both the crucial Lamé constant and the mesh-size. Numerical examples are presented to support theoretical results.

In the third part, we studied the convergence and optimality of adaptive mixed finite element methods (AMFEMs) for the Poisson equations and Stokes equations in an abstract setting. We generalized 6 Hypotheses for the finite element subspaces no matter the dimension $d = 2$ or 3. Under these hypotheses, the convergence and optimality can be obtained. As an application, We showed that the Raviart-Thomas elements and the Brezzi-Douglas-Martins elements satisfy these hypotheses.

References

[1] L. Chen, J. Hu, and X. - H. Huang,*Stabilized mixed finite element methods for linear elasticity on simplicial grids in* $\mathbb{R}^n$, arXiv:1512.03998, (2015).

[2] L. Chen, J. Hu, and X. - H. Huang, *Fast Auxiliary Space Preconditioner for Linear Elasticity in Mixed Form*, arXiv:1604.02568,2016.

[3] J. Hu, *Finite element approximations of symmetric tensors on simplicial grids in* $\mathbb{R}^n$*: the higher order case*, J. Comput. Math., 33 (2015), pp. 283–296.

[4] J. Hu, *A new family of efficient conforming mixed finite elements on both rectangular and cuboid meshes for linear elasticity in the symmetric formulation*, SIAM J. Numer. Anal., 53 (2015), pp. 1438–1463.

[5] J. Hu and R. Ma, *Conforming mixed triangular prism and nonconforming mixed tetrahedral elements for the linear elasticity problem*, arXiv:1604.07903, 2016.

[6] J. Hu, H. Man, and S. Zhang, *A simple conforming mixed finite element for linear elasticity on rectangular grids in any space dimension*, J. Sci. Comput., 58 (2014), pp. 367–379.

[7] J. Hu, H. Man, J. -Y. Wang, and S. Zhang *The simplest nonconforming mixed finite element method for linear elasticity in the symmetric formulation on n-rectangular grids*, Comput. Math. Appl., 71(2016), pp. 1317–1336.

[8] J. Hu and Z.-C. Shi, *Lower order rectangular nonconforming mixed finite elements for plane elasticity*, SIAM J. Numer. Anal., 46 (2007), pp. 88–102.

[9] J. Hu and G.-Z Yu, *A unified analysis of quasi-optimal convergence for adaptive mixed finite element methods*, arXiv:1601.00.00132, 2016.

[10] J. Hu and S. Zhang, *A family of conforming mixed finite elements for linear elasticity on triangular grids*, arXiv:1406.7457, (2015).

[11] J. Hu and S. Zhang, *A family of symmetric mixed finite elements for linear elasticity on tetrahedral grids*, Sci. China Math., 58 (2015), pp. 297–307.
[12] J. Hu and S. Zhang, *Finite element approximations of symmetric tensors on simplicial grids in $\mathbb{R}^n$: the lower order case*, M3AS 26 (2016), pp. 1649–1669.
[13] H.-Y. Man, J. Hu, and Z.-C. Shi, *Lower order rectangular nonconforming mixed finite element for the three-dimensional elasticity problem*, Math. Models Methods Appl. Sci., 19 (2009), pp. 51–65.

On Adaptivity in Tree Approximation

Peter Binev

The tree approximation is a convenient way to describe an adaptive approximation involving partitioning of a domain in which the next finer partition is formed by subdividing some elements of the current partition into several subsets in a prescribed way. In our considerations we replace "several" by "two" mentioning that this does not restrict the generality since a subdivision into several elements can be replaced by a sequence of binary subdivisions. The process of the adaptive partitioning is then described by a full binary tree (a graph in which exactly one node $\mathcal{R}$, called *root*, has degree 2 and the other nodes have either degree 1 and are called *leaves*, or degree 3 and are called *internal nodes*). Naturally, the domain is related to the root $\mathcal{R}$ and the current partition is composed by the subsets corresponding to the leaves $\mathcal{L}(T)$ of the tree T. A subdivision of an element Δ is performed by adding two new leaf nodes, Δ' and Δ'', to the graph and connecting them to Δ which becomes an internal node. We call Δ' and Δ'' *children* of Δ and refer to Δ as their *parent*. The set $\mathcal{A}(\Delta)$ of *ancestors* of Δ, including Δ itself, is composed of all nodes on the unique path connecting Δ with the root $\mathcal{R}$.

A generic adaptive approximation is defined via error functionals $e(\Delta)$ related to the nodes Δ of a tree T. Each of these functionals represents (an estimate of) the local error of approximation on the subset related to Δ. For example, if we approximate a function f by piecewise polynomials in L_q, $0 < q < \infty$ and P_Δ is the polynomial approximating f on the subset related to Δ, then one can choose $e(\Delta) = \int_\Delta |f - P_\Delta|^q dx$.

We choose no particular form of the error functional but require that it satisfies the following *subadditivity* property:

$$e(\Delta) \geq e(\Delta') + e(\Delta'') \qquad \text{for each node } \Delta \text{ with children } \Delta' \text{ and } \Delta''\,. \tag{1}$$

It is easy to see that this property is satisfied by (the q-th power of) any integral norm. However, in the case of local error indicators it should be replaced by its weak variant discussed in [4]. Here we use (1) for simplicity and to derive results with small constants.

The total error of approximation related to a given tree T is defined by $\mathcal{E}(T) := \sum_{\Delta\in\mathcal{L}(T)} e(\Delta)$. The property (1) guarantees that the process of growing the tree T decreases the error $\mathcal{E}(T)$. To assess how effective this process is, we should compare $\mathcal{E}(T)$ with the smallest possible such error for a tree with similar complexity.

Designating the complexity $n := \#\mathcal{L}(T)$ as the number of elements in the partition given by T, we define the best n-term approximation by

$$\sigma_n := \min_{T\,:\,\#\mathcal{L}(T)\le n} \mathcal{E}(T)\,.$$

The goal of the adaptive approximation is to design a course-to fine process of making adaptive decisions based only on the information from the errors at the nodes in the current tree. It was proven in [4] that for a particular algorithm, the error of approximation is compatible to the best one given by σ_n. Here we present the variant of this algorithm provided in [1, 3]. The direct greedy approach, namely subdividing the leaf Δ with the largest error $e(\Delta)$, while being computationally efficient, does not give an optimal performance. Instead, we define a modified error $\tilde{e}(\Delta)$, setting $\tilde{e}(\mathcal{R}) := e(\mathcal{R})$ and then inductively

$$\tilde{e}(\Delta) := \frac{\tilde{e}(\Delta^*) + e(\Delta)}{\tilde{e}(\Delta^*)e(\Delta)} = \left(\sum_{\nabla\in\mathcal{A}(\Delta)} \frac{1}{e(\nabla)}\right)^{-1}, \tag{2}$$

where Δ^* is the parent of Δ. The adaptive algorithm is very simple:

subdivide the leaf Δ with the largest $\tilde{e}(\Delta)$.

This type of algorithms are often referred to as *h-adaptive* since they use the same fixed approximation tool at each node and the improvements come from decreasing of the diameters of the elements of the partition sometimes denoted by h. The following is the main result about the above algorithm [1, 3].

Theorem 1 *Let the tree T_N be received by applying a greedy refinement strategy with respect to the quantities $\tilde{e}(\Delta)$ defined by (2). Then the tree T_N provides a near-best h-adaptive approximation*

$$\mathcal{E}(T_N) \le \frac{N}{N-n+1}\ \sigma_n \tag{3}$$

for any integer $n \le N$. The complexity of the algorithm for obtaining T_N is $\mathcal{O}(N)$ safe for the sorting of $\tilde{e}(\Delta)$ that requires $\mathcal{O}(N \log N)$ operations.

The sorting can be avoided by binning the values of $\tilde{e}(\Delta)$ into binary bins and choosing for subdivision any of the Δ from the largest nonempty bin. This increases the constant in (3) by 2 but the total complexity of the algorithm is $\mathcal{O}(N)$.

The approximation can be further adapted by varying the number of degrees of freedom $p = p(\Delta)$ used at each individual element Δ of the partition. For every $p = 1, 2, ...$ we consider the corresponding error functional at Δ denoted by $e_p(\Delta)$. It is reasonable to assume that these errors decrease as p increases, namely

$$e_p(\Delta) \ge e_{p+1}(\Delta)\,, \qquad p \ge 1\,.$$

In addition, we assume that the subadditivity property (1) also holds but only for the error functionals $e_1(\Delta)$.

The process of adapting both the partition and the number of degrees of freedom $p(\Delta)$ at each element Δ of the partition is known as *hp-adaptivity*. The approximation is described by the pair $(T, \mathcal{P}(T))$, where $\mathcal{P}(T) := \{p(\nabla) : \nabla \in \mathcal{L}(T)\}$ is the

list of the assigned degrees of freedom. We define $|\mathcal{P}(T)| := \sum_{\nabla\in\mathcal{L}(T)} p(\nabla)$ and use it as a measure of complexity of this approximation. The total error is calculated as $\mathcal{E}^{\rm hp}(T,P(T)) = \sum_{\nabla\in\mathcal{L}(T)} e_p(\nabla)$ and the best hp-adaptive approximation is

$$\sigma_n^{\rm hp} := \min_{T} \min_{\mathcal{P}(T)\,:\,|\mathcal{P}(T)|\le n} \mathcal{E}^{\rm hp}(T,P(T))\,.$$

An alternative description of the best approximation is using a "ghost tree" $\mathcal{G}$ with $\#\mathcal{L}(\mathcal{G}) = |\mathcal{P}(T)|$ to distribute the degrees of freedom and for any node $\Delta\in\mathcal{G}$ defines $p(\Delta,\mathcal{G}) := \#\{\nabla\in\mathcal{L}(\mathcal{G}) : \Delta\in\mathcal{A}(\nabla)\}$ to match the assignments given by $\mathcal{P}$. In this way, the pair $(\mathcal{G},T)$ defines an hp-adaptive approximation and it is easy to see that

$$\sigma_n^{\rm hp} = \min_{\mathcal{G}\,:\,\#\mathcal{L}(\mathcal{G})\le n} \min_{T\subset\mathcal{G}} \mathcal{E}^{\rm hp}(\mathcal{G},T)\,.$$

The following hp-adaptive algorithm describes how to find the ghost tree $\mathcal{G}_N$ of complexity $N = N_{\max}$ and the corresponding tree T_N:

(1) set $\mathcal{G}_1 := \{\mathcal{R}\}$, $T_1 := \{\mathcal{R}\}$, $\tilde e(\mathcal{R}) := e(\mathcal{R})$, $E_1(\mathcal{R}) := e(\mathcal{R})$, $\tilde E_1(\mathcal{R}) := \tilde e(\nabla)$, $q(\mathcal{R}) := \tilde e(\mathcal{R})$, $s(\mathcal{R}) := \mathcal{R}$, $p(\mathcal{R}) := 1$, and $N = 2$;
(2) set $\Delta := s(\mathcal{R})$ and expand the current tree $\mathcal{G}_{N-1}$ to $\mathcal{G}_N$ by subdividing Δ and adding two children nodes Δ'_N and Δ''_N to it;
(3) define T_N as the minimal full binary tree containing T_{N-1}, Δ'_N, and Δ''_N;
(4) for $\nabla = \Delta'_N$ and $\nabla = \Delta''_N$ calculate the quantities: $\tilde e(\nabla) := \frac{e(\nabla)\tilde e(\Delta)}{e(\nabla)+\tilde e(\Delta)}$, $E_1(\nabla) := e(\nabla)$, $\tilde E_1(\nabla) := \tilde e(\nabla)$, $q(\nabla) := \tilde e(\nabla)$, $s(\nabla) := \nabla$, $p(\nabla) := 1$;
(5) set $p(\Delta) := p(\Delta) + 1$ and calculate $e_{p(\Delta)}(\Delta)$;
(6) set Δ' and Δ'' to be the children of Δ;
(7) set $E_{p(\Delta)}(\Delta) := \min\{E_{p(\Delta')}(\Delta') + E_{p(\Delta'')}(\Delta'')\ ,\ e_{p(\Delta)}(\Delta)\}$;
(8) **if** $E_{p(\Delta)}(\Delta) = e_{p(\Delta)}(\Delta)$, **then** trim T at Δ making Δ a leaf node of T;
(9) set $\tilde E_{p(\Delta)}(\Delta) := \frac{E_{p(\Delta)}(\Delta)\tilde E_{p(\Delta)-1}(\Delta)}{E_{p(\Delta)}(\Delta)+E_{p(\Delta)-1}(\Delta)}$;
(10) set $D := \operatorname{argmax}\{q(\Delta'), q(\Delta'')\}$ and update $q(\Delta) := \min\left\{q(D), \tilde E_{p(\Delta)}(\Delta)\right\}$, $s(\Delta) := s(D)$;
(11) **if** $\Delta \ne \mathcal{R}$, **then** replace Δ with its parent and **go to** (5)
(12) **else** set $N := N + 1$ and **go to** (2) or **exit** if $N > N_{\max}$

A near-best performance of the above algorithm is established in [2, 3].

Theorem 2 *The pair $(\mathcal{G}_N, T_N)$ provides a near-best hp-adaptive approximation*

$$E^{\rm hp}(\mathcal{G}_N,T_N) \le \frac{2N-1}{N-n+1}\ \sigma_n^{\rm hp} \tag{4}$$

for any integer $n \le N$.

The algorithm performs $\sum_{\Delta\in\mathcal{G}_N} p(\Delta,\mathcal{G}_N)$ steps, a quantity that varies between $\mathcal{O}(N\log N)$ for well balanced trees to $\mathcal{O}(N^2)$ for highly unbalanced ones. This complexity can be significantly decreased if $e_p(\Delta)$ changes only at a lacunary sequence of indices p, e.g. at $p = 2^k$ setting $e_p(\Delta) := e_{2^k}(\Delta)$ for $2^k \le p < 2^{k+1}$.

This *hp*-adaptive algorithm was used by Canuto, Nochetto, Stevenson, and Verani in a recent paper [5] to find optimal convergence rates in the univariate case for a model problem.

References

[1] P. Binev, *Adaptive Methods and Near-Best Tree Approximation*, Oberwolfach Reports **29** (2007), 1669–1673.

[2] P. Binev, *Instance optimality for hp-type approximation*, Oberwolfach Reports **39** (2013), 2192–2194.

[3] P. Binev, *Tree approximation for hp-adaptivity*, IMI Preprint Series 2015:07

[4] P. Binev and R. DeVore, *Fast computation in adaptive tree approximation*, Numer. Math. **97** (2004), 193–217.

[5] C. Canuto, R. H. Nochetto, R. Stevenson, and M. Verani, *Convergence and Optimality of the hp-AFEM*, Numer. Math. (2016). doi:10.1007/s00211-016-0826-x

Convergence and Optimality of *hp*-AFEM

Claudio Canuto

(joint work with Ricardo H. Nochetto, Rob Stevenson, and Marco Verani)

We describe a recently proposed adaptive *hp*-type finite element algorithm, termed *hp-AFEM*, for the solution of operator equations such as, e.g., elliptic boundary value problems; the details can be found in [4]. The algorithm produces a sequence of *hp*-partitions of the domain (i.e., conforming finite element partitions equipped with a distribution of polynomial degrees over the elements) and corresponding Galerkin discrete solutions, with the following properties: the error in the energy norm decays at a fixed rate, and is instance optimal, meaning that it is comparable to the best *hp*-approximation error that can be obtained using a comparable number of degrees of freedom. In particular, if the solution admits *hp*-type approximations for which the error decays exponentially fast in the number of activated degrees of freedom, the same exponential behavior occurs for the Galerkin approximations built by our algorithm.

A result of this type (exponential convergence through a judicious selection of *h*-refinements and *p*-enrichments) was first obtained by Gui and Babuška [7, 8] for functions with localized singularities. However, these pioneering results implicitly exploit the particular structure of the considered singularities, thus avoiding the risk of making 'wrong' choices in the adaptive procedure. Unfortunately, this is not the generic situation; indeed, it is not difficult to build examples of functions (see again [4] for details) for which early choices between *h*-refinement and *p*-enrichment, based on the currently available information, must be subsequently corrected in order to stay close to the optimal *hp*-approximation. In other words, an *hp*-adaptive algorithm which works for a large class of functions (solutions and data) should incorporate the possibility of stepping-back, i.e., creating a new *hp*-partition starting from the available information on the current one, in view of guaranteeing optimality in the forthcoming steps. This stage is often referred to as 'coarsening'.

Our algorithm *hp-AFEM* hinges on two routines, termed *REDUCE* and *hp-NEARBEST*. The former, starting from the current *hp*-partition on which data are already replaced by suitable polynomial approximations (thus in a situation of no 'data oscillation') reduces the energy norm of the Galerkin error by a prescribed amount. The latter performs the coarsening as indicated above, in order to guarantee instance optimality of the coarsened approximations; it relies on the adaptive *hp*-tree approximation routine recently introduced by Binev ([1, 2], see also his contribution in this Report).

Our algorithm consists of a repetition of calls of *hp-NEARBEST* and *REDUCE* with decreasing error tolerances. Each call of *hp-NEARBEST* takes in input the data and the current Galerkin solution and produces a (nonconforming) *hp*-partition and a piecewise polynomial approximation of the input functions on this partition, in such a way that a specific error functional is less than a prescribed tolerance. Such a coarsening procedure, however, may increase the energy norm of the Galerkin error by up to a constant factor. This must be compensated by a judicious choice of the reduction factor of *REDUCE* so that the concatenation of the two routines produces a converging sequence.

The routine *REDUCE* is implemented as an AFEM consisting of the usual loop over *SOLVE*, *ESTIMATE*, *MARK*, and *REFINE*. In dimension 1, we construct an estimator that is reliable and discretely efficient, uniformly in p. Consequently, the number of iterations to achieve some fixed error reduction is independent on the maximal polynomial degree. In dimension 2, one may employ in *ESTIMATE* the residual-based a posteriori error estimator analyzed by Melenk and Wohlmuth [9], which however turns out to be p-sensitive. We show that with this choice, in order to achieve a fixed error reduction, it suffices to grow the number of iterations more than quadratically with respect to the maximal polynomial degree. This introduces an optimality degradation at stages intermediate between two consecutive calls of *hp-NEARBEST*.

As an alternative, one may resort to the equilibrated flux estimator, which has been shown to be p-robust in [3] (see also [6]). We use this estimator to mark stars for refinement, namely patches of elements around vertices, and to execute p-refinements only. In [5], we provide numerical evidence that a uniform saturation property holds provided the local polynomial degree p in each marked star is increased by any quantity proportional to p. Under this choice, the number of iterations in *REFINE* needed to reduce the Galerkin error by a fixed amount is uniformly bounded in h and p, thus guaranteeing overall optimality.

References

[1] P. Binev. Instance optimality for hp-type approximation. In *Oberwolfach Reports*, volume 39, pages 14–16, 2013.

[2] P. Binev. Tree approximation for *hp*-adaptivity. (in preparation).

[3] D. Braess, V. Pillwein, and J. Schöberl. Equilibrated residual error estimates are p-robust. *Comput. Methods Appl. Mech. Engrg.*, 198(13-14):1189–1197, 2009.

[4] C. Canuto, R.H. Nochetto, R. Stevenson, and M. Verani. Convergence and optimality of *hp*-AFEM. *Numer. Math.*, 10.1007/s00211-016-0826-x, 2016

[5] C. Canuto, R.H. Nochetto, R. Stevenson, and M. Verani. p-robust saturation property for hp-AFEM, in preparation.
[6] A. Ern and M. Vohralík. Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations. *SIAM J. Numer. Anal.*, 53(2):10581081, 2015.
[7] W. Gui and I. Babuška. The h, p and h-p versions of the finite element method in 1 dimension. II. The error analysis of the h- and h-p versions. *Numer. Math.*, 49(6):613–657, 1986.
[8] W. Gui and I. Babuška. The h, p and h-p versions of the finite element method in 1 dimension. III. The adaptive h-p version. Numer. Math., 49(6):659–683, 1986.
[9] J. M. Melenk and B. I. Wohlmuth. On residual-based a posteriori error estimation in hp-FEM. *Adv. Comput. Math.*, 15(1-4):311–331, 2002.

Robust residual-based a posteriori Arnold-Winther mixed finite element analysis in elasticity

Joscha Gedicke
(joint work with Carsten Carstensen)

This talk presents a residual-based *a posteriori* error estimator for the Arnold-Winther mixed finite element that utilises a post-processing for the skew-symmetric part of the strain [1]. Numerical experiments verify the proven reliability and efficiency for suitable approximation of the skew-symmetric deformation gradient. Numerical evidence supports that the L^2-stress error estimator is robust in the Poisson ratio and allows stable error control even in the incompressible limit.

The problem in linear elasticity considers the connected reference configuration of the elastic body $\Omega \subset \mathbb{R}^2$ with polygonal boundary $\partial\Omega = \Gamma_D \cup \Gamma_N$ with closed and connected Γ_D of positive surface measure and $\Gamma_N = \partial\Omega \backslash \Gamma_D$ for applied tractions. Given a volume force $f : \Omega \to \mathbb{R}^2$, a displacement $u_D : \Gamma_D \to \mathbb{R}^2$, and a traction $g : \Gamma_N \to \mathbb{R}^2$, find a displacement $u : \Omega \to \mathbb{R}^2$ and a stress tensor $\sigma : \Omega \to \mathbb{S} := \{\tau \in \mathbb{R}^{2\times 2} : \tau = \tau^T\}$ such that

$$
\begin{aligned}
-\operatorname{div}\sigma = f, \quad \sigma = \mathbb{C}\varepsilon(u) \quad \text{in } \Omega,\\
u = u_D \quad \text{on } \Gamma_D, \quad \sigma\nu = g \quad \text{on } \Gamma_N,
\end{aligned}
$$

where, $\mathbb{C}$ denotes the bounded and positive definite fourth-order elasticity tensor for isotropic linear elasticity.

The *a posteriori* error analysis for the symmetric Arnold-Winther mixed finite element method [2] may follow the ideas of [3] to derive a stress error control

$$
\begin{aligned}
||\sigma - \sigma_{AW}||^2_{\mathbb{C}^{-1}} \leq \min_{v\in V} ||\mathbb{C}^{-1}\sigma_{AW} - \varepsilon(u_D + v)||^2_{\mathbb{C}}\\
+ C_1 \mathrm{osc}^2(f, \mathcal{T}) + C_2 \mathrm{osc}^2(g, \mathcal{E}(\Gamma_N))
\end{aligned}
$$

for the stress error $\sigma - \sigma_{AW}$ even with a rather explicit estimate of the constant in front of the oscillations and the (unwritten) multiplicative constant 1 in front of the first term that measures the quality of the approximation $\mathbb{C}^{-1}\sigma_{AW}$ of symmetric

gradients $\varepsilon(v) := (Dv + D^T v)/2$ for $v \in V$. The space V consists of all square-integrable displacements with homogeneous boundary conditions along Γ_D and with a square-integrable functional matrix Dv.

A severe additional difficulty of this approximation is that only the symmetric part is approximated and not the full gradient Dv. Other mixed finite element schemes like PEERS [4] involve some additional variable to approximate the asymmetric part of the gradient. This talk presents an explicit error estimate which involves an arbitrary asymmetric approximation γ_h and provides an abstract *a posteriori* error control of the residual type, which is useful for adaptive mesh-refining algorithms,

$$\begin{aligned}\eta_\ell^2 &= \operatorname{osc}^2(f,\mathcal{T}) + \operatorname{osc}^2(g,\mathcal{E}(\Gamma_N)) \\ &\quad + \sum_{T\in\mathcal{T}} h_T^2 \|\operatorname{Curl}(\mathbb{C}^{-1}\sigma_{AW} + \gamma_h)\|_{L^2(T)}^2 \\ &\quad + \sum_{E\in\mathcal{E}(\Omega)} h_E \|[\mathbb{C}^{-1}\sigma_{AW} + \gamma_h]\tau_E|_{L^2(E)}^2 \\ &\quad + \sum_{E\in\mathcal{E}(\Gamma_D)} h_E \|(\mathbb{C}^{-1}\sigma_{AW} + \gamma_h - Du_D)\tau\|_{L^2(E)}^2.\end{aligned}$$

For any (piecewise smooth) choice of γ_h, this *a posteriori* error estimator is reliable in the sense that

$$\|\sigma - \sigma_{\mathrm{AW}}\|_{\mathbb{C}^{-1}} \le C_{\mathrm{rel}}\eta_\ell$$

with some λ-independent constant $C_{\mathrm{rel}} \approx 1$. One opportunity to ensure efficiency is a global minimisation over all piecewise polynomial γ_h of the error estimator η_ℓ. The bubble function technique shows that the particular choice of γ_h enters the efficiency estimates with some λ-independent constant $C_{\mathrm{eff}} \approx 1$,

$$\eta_\ell \le C_{\mathrm{eff}} \left(\|\sigma - \sigma_{\mathrm{AW}}\|_{\mathbb{C}^{-1}} + \|\operatorname{skew}(Du) - \gamma_h\|_{L^2(\Omega)} \right).$$

Hence, one efficient choice for γ_h is to choose it as a sufficiently accurate polynomial approximation of the asymmetric gradient $\operatorname{skew}(Du) := (Du - D^T u)/2$. Since a global approximation or even minimisation may be too costly, this talk proposes to apply a post-processing step to compute such a sufficiently accurate approximation $\gamma_h = \operatorname{skew}(Du^*_{AW})$ for the post-processed displacement u^*_{AW} in the spirit of Stenberg [5]. The approximation $\gamma_h = \operatorname{skew}(Du^*_{AW})$ is proven to be robust in the Poisson ratio $\nu \to 1/2$ for sufficiently smooth functions. For domains with re-entrant corners or incompatible boundary conditions, numerical experiments confirm that the proposed computation of γ_h leads empirically to reliable and efficient *a posteriori* error control independent of the Poisson ratio $\nu \to 1/2$.

References

[1] C. Carstensen and J. Gedicke, *Robust residual-based a posteriori Arnold-Winther mixed finite element analysis in elasticity*, Comput. Methods Appl. Mech. Engrg., **300** (2016), 245–264.

[2] D. N. Arnold and R. Winther, *Mixed finite elements for elasticity*, Numer. Math. **92** (2002), 401–419.
[3] C. Carstensen and G. Dolzmann, *A posteriori error estimates for mixed FEM in elasticity*, Numer. Math. **81(2)** (1998), 187–209.
[4] D. N. Arnold, F. Brezzi, and J. Douglas, *PEERS: A new mixed finite element for plane elasticity*, Jpn. J. Appl. Math. **1** (1984), 347–367.
[5] R. Stenberg, *A family of mixed finite elements for the elasticity problem*, Numer. Math. **53(5)** (1988), 513–538.

Towards optimal adaptivity of strongly non-symmetric problems

Michael Feischl

Rate optimality of adaptive algorithms has been an active research field since the seminal works [3, 4]. Initially, the research focused on simple symmetric PDEs like the Poisson equation, but nowadays, many interesting positive definite problems like (in general non-symmetric) second-order elliptic PDEs, non-local boundary element formulations of PDEs, and even certain non-linear problems are understood. For a good overview on the known results, we refer to [2]. In stark contrast to this is the development of the theory for strongly non-symmetric problems, i.e., problems which are more than a compact perturbation of a symmetric equation. The task of proving rate optimality of adaptive algorithms for positive definite non-symmetric problems is equivalent to proving rate optimality of indefinite symmetric problems. Hence, the goal of this talk is to understand how non-symmetric or indefinite problems fit into the theory of rate optimality.

1. Rate optimal adaptivity

Let X be a separable Hilbert space of functions on some domain $\Omega \subseteq \mathbb{R}^d$. Let $\mathcal{T}_0$ denote a triangulation of Ω into compact simplices $T \in \mathcal{T}_0$. Given a strongly positive definite and bounded linear operator $\mathcal{A}\colon X \to X^\star$, i.e., $\langle \mathcal{A}x\,,\,x\rangle \geq c\|x\|_X^2$ for all $x \in X$, we consider the equation

$$\mathcal{A}x = f \in X^\star. \tag{1}$$

Moreover, given a refinement (e.g., generated by newest-vertex-bisection) $\mathcal{T}_\ell$ of $\mathcal{T}_0$, let $X_\ell \subset X$ denote the associated finite dimensional subspace of functions (e.g., piecewise polynomials on $\mathcal{T}_\ell$). We consider the finite-dimensional system: Find $x_\ell \in X_\ell$ such that

$$\langle \mathcal{A}x_\ell\,,\,y\rangle = \langle f\,,\,y\rangle \quad \text{for all } y \in X_\ell. \tag{2}$$

With this, the adaptive algorithm reads:
Input: $\mathcal{T}_0$. For $\ell = 0, 1, 2, \dots$ do

a) solve (2),
b) compute error estimator to determine elements $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$,
c) refine at least the elements in $\mathcal{M}_\ell$ to design new triangulation $\mathcal{T}_{\ell+1}$.

By definition, an adaptive algorithm of the above structure is said to be rate optimal, if the adaptively generated triangulations $\mathcal{T}_\ell$ satisfy

$$\operatorname{err}(\mathcal{T}_\ell) \le C(\#\mathcal{T}_\ell)^{-s} \quad \text{for all } \ell \in \mathbb{N} \tag{3}$$

for all $s > 0$ which allow for a sequence $\mathcal{T}_\ell^{\mathrm{opt}}$ of optimal triangulations (with $\mathcal{T}_0^{\mathrm{opt}} = \mathcal{T}_0$), generated by the refinement rule of step (c) of the algorithm, such that $\operatorname{err}(\mathcal{T}_\ell^{\mathrm{opt}}) \le C(\#\mathcal{T}_\ell^{\mathrm{opt}})^{-s}$. This means that the adaptive algorithm performs at least as good as any other algorithm using the same mesh-refinement. The error term $\operatorname{err}(\mathcal{T}_\ell)$ is usually an upper bound for the error $\|x - x_\ell\|_X$. The work [2] specifies four sufficient (and partly necessary) requirements of the different parts of the adaptive algorithm to prove (3). For non-symmetric (or indefinite) problems, the major difficulty appears to be the so called *general quasi-orthogonality*.

2. General quasi-orthogonality

The most simple form of general quasi-orthogonality reads

$$\sum_{k=\ell}^{\infty} \|x_{k+1} - x_k\|_X^2 \le C\|x - x_\ell\|_X^2 \quad \text{for all } \ell \in \mathbb{N}. \tag{4}$$

There exist several generalizations of (4), e.g., for sufficiently small $\varepsilon \ge 0$

$$\sum_{k=\ell}^{\infty} \|x_{k+1} - x_k\|_X^2 - \varepsilon\chi_\ell^2 \le C\operatorname{err}(\mathcal{T}_\ell)^2 \quad \text{for all } \ell \in \mathbb{N}, \tag{5}$$

where $\chi_\ell \in \{\operatorname{err}(\mathcal{T}_\ell), \|x - x_\ell\|_X\}$. This general quasi-orthogonality is the key tool to prove *linear convergence* of the error in the sense that there exist $0 < q < 1$ and $C > 0$ such that

$$\operatorname{err}(\mathcal{T}_{\ell+k}) \le Cq^k \operatorname{err}(\mathcal{T}_\ell) \quad \text{for all } \ell, k \in \mathbb{N}. \tag{6}$$

Furthermore, linear convergence is one of the key tools to prove optimality (3). Moreover, with linear convergence, one can even improve the rate optimality to optimal complexity in the sense that given some $\gamma > 0$, even

$$\operatorname{err}(\mathcal{T}_\ell) \le C(\sum_{k=0}^{\ell} \#\mathcal{T}_k^\gamma)^{-s} \quad \text{for all } \ell \in \mathbb{N} \tag{7}$$

holds for all possible $s > 0$. Since $\sum_{k=0}^{\ell} \#\mathcal{T}_k^\gamma$ is equivalent to the the overall cost of the adaptive algorithm (for the right choice of γ, i.e., if each *solve*-step requires solving one sparse linear system, we choose $\gamma = 1$), this shows that the algorithm is optimal with respect to computational work. Finally, general quasi-orthogonality is a direct consequence linear convergence [2, Proposition 4.11], and thus proves to be a central concept of the theory.

3. Towards a proof of general quasi-orthogonality

For symmetric operators $\mathcal{A}$, general quasi-orthogonality (4) follows immediately from the equivalence $\|\cdot\|_{\mathcal{A}} := \langle \mathcal{A}(\cdot)\,,\,(\cdot)\rangle^{1/2} \simeq \|\cdot\|_X$, Galerkin orthogonality

$$\|x_{k+1} - x_k\|_{\mathcal{A}}^2 = \|x - x_k\|_{\mathcal{A}}^2 - \|x - x_{k+1}\|_{\mathcal{A}}^2, \tag{8}$$

and convergence $\lim_{\ell\to\infty} x_\ell = x$. For strongly non-symmetric (or indefinite) problems, however, (8) does not hold and new ideas are required. The basic approach can be sketched as follows. Assume that $\mathcal{A} \in \mathbb{R}^{N\times N}$ is a lower triangular matrix and let $X_\ell = \mathbb{R}^\ell \subset X := \mathbb{R}^N$. Then, we have $x_\ell = (\mathcal{A}|_{\mathbb{R}^\ell\times\mathbb{R}^\ell})^{-1} f|_{\mathbb{R}^\ell}$ for all $\ell \in \mathbb{N}$ and hence (up to identification of $\mathbb{R}^\ell$ as a subset of $\mathbb{R}^N$)

$$x_{k+1} - x_k = (\mathcal{A}|_{\mathbb{R}^{k+1}\times\mathbb{R}^{k+1}})^{-1} f|_{\mathbb{R}^{k+1}} - (\mathcal{A}|_{\mathbb{R}^k\times\mathbb{R}^k})^{-1} f|_{\mathbb{R}^k} = \alpha_{k+1} e_{k+1},$$

for the $(k+1)$-st unit vector $e_{k+1} \in \mathbb{R}^N$ and some $\alpha_{k+1} \in \mathbb{R}$. This shows general quasi-orthogonality (4) by $\sum_{k=\ell}^N \|x_{k+1} - x_k\|_X^2 = \sum_{k=\ell}^N \alpha_{k+1}^2 = \|x - x_\ell\|_X^2$.

For a general operator $\mathcal{A}$, the strategy is to find a suitable Riesz basis $(w_k)_{k\in\mathbb{N}}$ such that $X_\ell := \operatorname{span}\{w_1, \dots, w_{N_\ell}\}$ for some $N_\ell \in \mathbb{N}$ and all $\ell \in \mathbb{N}$. Then, the problems (1)–(2) are equivalent to

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{f} \quad \text{and} \quad \boldsymbol{A}|_{\{1,\dots,N_\ell\}\times\{1,\dots,N_\ell\}} \boldsymbol{x}_\ell = \boldsymbol{f}|_{\{1,\dots,N_\ell\}},$$

where $\boldsymbol{A} \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$, $\boldsymbol{A}_{ij} := \langle \mathcal{A} w_i\,,\, w_j\rangle$, $\boldsymbol{x}_\ell \in \mathbb{R}^{N_\ell}$, and $\boldsymbol{x}, \boldsymbol{f} \in \mathbb{R}^{\mathbb{N}}$. The Riesz basis property guarantees that $\boldsymbol{A}\colon \ell_2 \to \ell_2$ is a bounded and strongly elliptic operator. Moreover, there holds $x_\ell = \sum_{j=1}^{N_\ell} \boldsymbol{x}_{\ell,j} w_j$ and $x = \sum_{j=1}^\infty \boldsymbol{x}_j w_j$. As the above motivation for the finite dimensional case might suggest, general quasi-orthogonality (4) holds for the original problem (1) if $\boldsymbol{A}$ is sufficiently close to a triangular matrix.

Theorem 1. General quasi-orthogonality (5) holds if $\boldsymbol{A}$ admits a stable (block) LU-factorization, that is, $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ for matrices $\boldsymbol{L}, \boldsymbol{U} \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ which are (block) lower resp. upper triangular

$$\big(\exists \ell \in \mathbb{N},\, j > N_\ell \geq i\big) \quad \Longrightarrow \quad \boldsymbol{L}_{ij} = \boldsymbol{U}_{ji} = 0$$

and satisfy that for all $\varepsilon > 0$ there exists $\boldsymbol{U}_\varepsilon \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ such that $\boldsymbol{U}_\varepsilon, \boldsymbol{U}_\varepsilon^{-1}\colon \ell_2 \to \ell_2$ are bounded operators and $\|\boldsymbol{U}_\varepsilon - \boldsymbol{U}\|_{\ell_2\to\ell_\infty} \leq \varepsilon$. □

This theorem leads to several interesting questions.

- *When does an infinite matrix posses a stable LU-factorization in the above sense?* Little is known for non-symmetric matrices, see e.g. [1]. Numerical evidence suggests that not every strongly positive definite matrix does have a stable LU-factorization. Hence, this class of matrices might be too big.
- *How to characterize the closure* $\overline{L(\ell_2,\ell_2)} \subset L(\ell_2,\ell_\infty)$*?* It is well-known that the space of linear operators $L(\ell_2,\ell_2)$ is not dense in $L(\ell_2,\ell_\infty)$.
- *Can we find a Riesz basis such that* $\boldsymbol{A}$ *has a simple structure?* Wavelet techniques might allow to obtain quasi-sparse matrices $\boldsymbol{A}$. This can help to prove the existence of a stable LU-factorization.

- *What is a general characterization of matrices* $\boldsymbol{A}$ *which have a stable LU-factorization?* This would allow to characterize the class of operators $\mathcal{A}$ for which linear convergence (6) and thus optimal complexity (7) is achievable.

So far, none of the above questions is completely solved. It is the author's opinion that answering one of those questions can significantly advance the theory of rate optimality for adaptive algorithms.

References

[1] K. T. Andrews, J. D. Ward, *LU-factorization of order bounded operators on Banach sequence spaces*, J. Approx. Theory **48** (1986), 169–180.

[2] C. Carstensen, M. Feischl, M. Page, D. Praetorius, *Axioms of adaptivity*, Comput. Math. Appl. **67** (2014), 1195–1253.

[3] J. M. Cascon, C. Kreuzer, R. Nochetto, K. G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal. **46** (2008), 2524–2550.

[4] R. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math. **7** (2007), 245–269.

Rate optimal adaptive FEM with inexact solver for strongly monotone operators

Dirk Praetorius

(joint work with Gregor Gantner, Alexander Haberl, and Bernhard Stiftner)

Meanwhile, the mathematical understanding of adaptive FEM has reached a mature state; see [9, 14, 3, 15, 6, 11] for some milestones for linear elliptic PDEs, [17, 8, 2, 13] for non-linear problems, and [4] for some general framework. Optimal adaptive FEM with inexact solvers has already been addressed in [15, 1, 4] for linear PDEs and in [5] for eigenvalue problems. However, for problems involving nonlinear operators, optimal adaptive FEM with inexact solvers has not been analyzed yet. Our work [12] aims to close the gap between convergence analysis (e.g. [4]) and empirical evidence (e.g. [10]) by analyzing an algorithm from [7].

Model problem. We follow [4] and present our results from [12] in an abstract framework, while precise examples for our setting are given, e.g., in [2, 13]. Let $\mathcal{H}$ be a separable Hilbert space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ with norm $\|\cdot\|$ and scalar product $(\cdot,\cdot)_{\mathcal{H}}$. With the duality pairing $\langle\cdot,\cdot\rangle$ between $\mathcal{H}$ and its dual $\mathcal{H}^*$, let $A : \mathcal{H} \to \mathcal{H}^*$ be a nonlinear operator which satisfies the following assumptions:

(O1) $\boldsymbol{A}$ **is strongly monotone:** There exists $\alpha > 0$ such that

$$\alpha \|u - v\|^2 \le \operatorname{Re} \langle Au - Av, u - v\rangle \quad \text{for all } u, v \in \mathcal{H}.$$

(O2) $\boldsymbol{A}$ **is Lipschitz continuous:** There exists $L > 0$ such that

$$\|Au - Av\|_* := \sup_{w \in \mathcal{H}\setminus\{0\}} \frac{\langle Au - Av, w\rangle}{\|w\|} \le L \|u - v\| \quad \text{for all } u, v \in \mathcal{H}.$$

(O3) ***A* has a potential:** There exists a Gâteaux differentiable function $P : \mathcal{H} \to \mathbb{K}$ with Gâteaux derivative $\mathrm{d}P = A$, i.e.,

$$\langle Au, v\rangle = \lim_{r\to 0} \frac{P(u+rv)-P(u)}{r} \quad \text{for all } u, v \in \mathcal{H}.$$

Let $F \in \mathcal{H}^*$. According to the main theorem on strongly monotone operators [18], (O1)–(O2) imply the existence and uniqueness of $u^\star \in \mathcal{H}$ such that

$$\langle Au^\star, v\rangle = \langle F, v\rangle \quad \text{for all } v \in \mathcal{H}. \tag{1}$$

To sketch the proof, let $\mathcal{I} : \mathcal{H} \to \mathcal{H}^*$ denote the Riesz mapping defined by $\langle \mathcal{I}u, v\rangle = (u, v)_{\mathcal{H}}$ for all $u, v \in \mathcal{H}$. Then, $\Phi : \mathcal{H} \to \mathcal{H}$, $\Phi(u) := u - \frac{\alpha}{L^2}\mathcal{I}^{-1}(Au - F)$ satisfies

$$\|\Phi(u) - \Phi(v)\| \le q\,\|u - v\| \text{ for all } u, v \in \mathcal{H}, \text{ where } q := (1 - \alpha^2/L^2)^{1/2} < 1. \tag{2}$$

Hence, the Banach fixpoint theorem proves the existence and uniqueness of $u^\star \in \mathcal{H}$ with $\Phi(u^\star) = u^\star$ which is equivalent to (1). In particular, the Picard iteration $u^n := \Phi(u^{n-1})$ with arbitrary initial guess $u^0 \in \mathcal{H}$ converges to $u^\star$, and it holds

$$\|u^\star - u^n\| \le \frac{q}{1-q}\,\|u^n - u^{n-1}\| \le \frac{q^n}{1-q}\,\|u^1 - u^0\| \quad \text{for all } n \ge 1. \tag{3}$$

With (O3), (1) (resp. (4) below) is equivalent to energy minimization, and $\|v-u^\star\|^2$ is equivalent to the energy difference. This guarantees the quasi-orthogonality [4].

FEM with iterative solver based on Picard iteration. Let $\mathcal{X}_\bullet \subset \mathcal{H}$ be a discrete subspace. As in the continuous case, there is a unique $u^\star_\bullet \in \mathcal{X}_\bullet$ such that

$$\langle Au^\star_\bullet, v_\bullet\rangle = \langle F, v_\bullet\rangle \quad \text{for all } v_\bullet \in \mathcal{X}_\bullet. \tag{4}$$

According to (O1)–(O2), it holds the Céa-type quasi-optimality $\|u^\star - u^\star_\bullet\| \le \frac{L}{\alpha}\,\|u^\star - v_\bullet\|$ for all $v_\bullet \in \mathcal{X}_\bullet$. To solve the nonlinear system (4), we use the Picard iteration (applied in $\mathcal{X}_\bullet$): Given $u^{n-1}_\bullet \in \mathcal{X}_\bullet$, we compute $u^n_\bullet = \Phi_\bullet(u^{n-1}_\bullet)$ as follows:

- Solve the linear system $(w_\bullet, v_\bullet)_{\mathcal{H}} = \langle Au^{n-1}_\bullet - F, v_\bullet\rangle$ for all $v_\bullet \in \mathcal{X}_\bullet$.
- Define $u^n_\bullet := u^{n-1}_\bullet - \frac{\alpha}{L^2}\,w_\bullet$.

Applying (3) on the discrete level, we infer from [7] that

$$\|u^\star - u^n_\bullet\| \le \|u^\star - u^\star_\bullet\| + \frac{q}{1-q}\,\|u^n_\bullet - u^{n-1}_\bullet\| \le \frac{L}{\alpha}\min_{v_\bullet\in\mathcal{X}_\bullet}\|u^\star - v_\bullet\| + \frac{q^n}{1-q}\,\|u^1_\bullet - u^0_\bullet\|. \tag{5}$$

***A posteriori* error estimator.** We suppose that all considered discrete spaces $\mathcal{X}_\bullet \subset \mathcal{H}$ are associated with a conforming triangulation $\mathcal{T}_\bullet$ of a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \ge 2$. For all $T \in \mathcal{T}_\bullet$ and all $v_\bullet \in \mathcal{X}_\bullet$, we suppose an *a posteriori* computable refinement indicator $\eta_\bullet(T, v_\bullet) \ge 0$. We then define

$$\eta_\bullet(v_\bullet) := \eta_\bullet(\mathcal{T}_\bullet, v_\bullet) \quad \text{and} \quad \eta_\bullet(\mathcal{U}_\bullet, v_\bullet)^2 := \sum_{T\in\mathcal{U}_\bullet} \eta_\bullet(T, v_\bullet)^2 \text{ for } \mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet. \tag{6}$$

We suppose that there exist constants $C_{\mathrm{ax}} > 0$ and $0 < \varrho_{\mathrm{ax}} < 1$ such that for all $\mathcal{T}_\bullet$ and all refinements $\mathcal{T}_\circ$ of $\mathcal{T}_\bullet$, the following properties (A1)–(A3) from [4] hold:

(A1) Stability on non-refined element domains:

$$|\eta_\circ(\mathcal{T}_\bullet \cap \mathcal{T}_\circ, v_\circ) - \eta_\bullet(\mathcal{T}_\bullet \cap \mathcal{T}_\circ, v_\bullet)| \le C_{\mathrm{ax}}\,\|v_\circ - v_\bullet\| \text{ for all } v_\bullet \in \mathcal{X}_\bullet, v_\circ \in \mathcal{X}_\circ.$$

(A2) Reduction on refined element domains:

$$\eta_\circ(\mathcal{T}_\circ\backslash\mathcal{T}_\bullet, v_\circ) \le \varrho_{\rm ax}\,\eta_\bullet(\mathcal{T}_\bullet\backslash\mathcal{T}_\circ, v_\bullet) + C_{\rm ax}\,\|v_\circ - v_\bullet\| \text{ for all } v_\bullet \in \mathcal{X}_\bullet, v_\circ \in \mathcal{X}_\circ.$$

(A3) Discrete reliability:

$$\|u_\circ^\star - u_\bullet^\star\| \le C_{\rm ax}\,\eta_\bullet(\mathcal{T}_\bullet\backslash\mathcal{T}_\circ, u_\bullet^\star).$$

Note that (A1)–(A2) are required for all discrete functions (and follow from inverse estimates), while (A3) is only required for the discrete solutions $u_\circ^\star$ resp. $u_\bullet^\star$ of (4).

Adaptive algorithm. With adaptivity parameters $0 < \theta \le 1$, $\lambda > 0$, and $C_{\rm mark} \ge 1$, an initial conforming triangulation $\mathcal{T}_0$, and an initial guess $u_0^0 \in \mathcal{X}_0$, our adaptive algorithm iterates the following steps (i)–(iii) for all $\ell = 0, 1, 2, \dots$

(i) Repeat (a)–(b) for all $n = 1, 2, 3, \dots$, until $\|u_\ell^n - u_\ell^{n-1}\| \le \lambda\,\eta_\ell(u_\ell^n)$.
  (a) Compute discrete Picard iterate $u_\ell^n \in \mathcal{X}_\ell$.
  (b) Compute refinement indicators $\eta_\ell(T, u_\ell^n)$ for all $T \in \mathcal{T}_\ell$

(ii) Define $u_\ell := u_\ell^n$ and determine a set $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of minimal cardinality, up to the multiplicative factor $C_{\rm mark}$, such that $\theta\,\eta_\ell(u_\ell) \le \eta_\ell(\mathcal{M}_\ell, u_\ell)$.

(iii) Employ newest vertex bisection [16] to generate the coarsest conforming refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ such that $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell\backslash\mathcal{T}_{\ell+1}$ (i.e., all marked elements have been refined) and define $u_{\ell+1}^0 := u_\ell \in \mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1}$.

In step (iii), we suppose that mesh-refinement leads to nested discrete spaces.

Lucky break-down of adaptive algorithm. First, if the repeat loop in step (i) does not terminate, it holds $u^\star \in \mathcal{X}_\ell$. Moreover, there exists $C > 0$ with

$$\|u^\star - u_\ell^n\| + \eta_\ell(u_\ell^n) \le C\,q^n \xrightarrow{n\to\infty} 0. \tag{7}$$

Second, if the repeat loop in step (i) terminates with $\mathcal{M}_\ell = \emptyset$ in step (ii), then $u^\star = u_k$ as well as $\mathcal{M}_k = \emptyset$ for all $k \ge \ell$. Overall, we may thus suppose that the repeat loop in step (i) terminates and that $\#\mathcal{T}_\ell < \#\mathcal{T}_{\ell+1}$ for all $\ell \ge 0$.

Bounded number of Picard iterations in step (i). There exists $C > 0$ such that nested iteration $u_\ell^0 := u_{\ell-1} \in \mathcal{X}_{\ell-1} \subseteq \mathcal{X}_\ell$ guarantees

$$u_\ell = u_\ell^n \quad \text{with} \quad n \le C\Big[1 + \log\Big(\max\Big\{1, \frac{\eta_{\ell-1}(u_{\ell-1})}{\eta_\ell(u_\ell)}\Big\}\Big)\Big] \quad \text{for all } \ell \ge 1. \tag{8}$$

Linear convergence. For $0 < \theta \le 1$ and all sufficiently small $\lambda > 0$, there exist constants $0 < \varrho < 1$ and $C > 0$ such that

$$\eta_{\ell+n}(u_{\ell+n}) \le C\varrho^n\,\eta_\ell(u_\ell) \quad \text{for all } \ell, n \ge 0. \tag{9}$$

In particular, there exists $C' > 0$ such that

$$\|u^\star - u_\ell\| \le C'\,\eta_\ell(u_\ell) \le C'C\varrho^\ell\eta_0(u_0) \xrightarrow{\ell\to\infty} 0. \tag{10}$$

Optimal algebraic convergence rates. For sufficiently small $0 < \theta \ll 1$, sufficiently small $\lambda > 0$, and all $s > 0$, there exists $C > 0$ such that

$$\eta_\ell(u_\ell) \le C\,\big(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\big)^{-s} \quad \text{for all } \ell \ge 0, \tag{11}$$

provided that the rate s is possible with respect to certain nonlinear approximation classes [4, 6].

Optimal computational complexity. Currently, the proof of optimal computational complexity is open, but it is observed in numerical experiments.

References

[1] Arioli, Georgoulis, Loghin. *SIAM J. Sci. Comput.*, 35(3):A1537–A1559, 2013.
[2] Belenki, Diening, Kreuzer. *IMA J. Numer. Anal.*, 32(2):484–510, 2012.
[3] Binev, Dahmen, DeVore. *Numer. Math.*, 97(2):219–268, 2004.
[4] Carstensen, Feischl, Page, Praetorius. *Comput. Math. Appl.*, 67(6):1195–1253, 2014.
[5] Carstensen, Gedicke. *SIAM J. Numer. Anal.*, 50(3):1029–1057, 2012.
[6] Cascon, Kreuzer, Nochetto, Siebert. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.
[7] Congreve, Wihler. *J. Comput. Appl. Math.*, 311:457–472, 2017.
[8] Diening, Kreuzer. *SIAM J. Numer. Anal.*, 46(2):614–638, 2008.
[9] Dörfler. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.
[10] Ern, Vohralík. *SIAM J. Sci. Comput.*, 35(4):A1761–A1791, 2013.
[11] Feischl, Führer, Praetorius. *SIAM J. Numer. Anal.*, 52(2):601–625, 2014.
[12] Gantner, Haberl, Praetorius, Stiftner. *in preparation*, 2016.
[13] Garau, Morin, Zuppa. *Numer. Math. Theory Methods Appl.*, 5(2):131–156, 2012.
[14] Morin, Nochetto, Siebert. *SIAM J. Numer. Anal.*, 38(2):466–488, 2000.
[15] Stevenson. *Found. Comput. Math.*, 7(2):245–269, 2007.
[16] Stevenson. *Math. Comp.*, 77(261):227–241, 2008.
[17] Veeser. *Numer. Math.*, 92(4):743–770, 2002.
[18] Zeidler. *Part II/B*, Springer, 1990.

Convergence and optimality of adaptive FEM for harmonic forms

Alan Demlow

Harmonic forms arise in computational electromagnetics. More broadly they are an important part of the finite element exterior calculus (FEEC) framework for approximating solutions to problems of Hodge-Laplace types posed on domains possibly having nontrivial topology. In the three-dimensional setting, they consist of curl- and divergence-free vector fields satisfying certain boundary conditions. In this talk we prove convergence and rate optimality of adaptive FEM for controlling L_2 (equivalently, energy) errors in the approximation of harmonic forms. Admissible finite element methods include standard mixed methods based on de Rham complex-conforming finite element complexes in all space dimensions.

Proving AFEM convergence results for harmonic forms requires overcoming several hurdles. The first is non-nestedness, that is, that is, there is no subset relationship between spaces of harmonic forms on adjacent mesh levels even if the meshes themselves are nested. We instead use properties of broader portions of the Hodge decomposition, resulting in sufficient nestedness and orthogonality properties to obtain AFEM convergence. A second challenge is nonalignment of bases for the spaces of discrete harmonic forms on nested meshes. This is a technical problem that occurs more generally when approximating multidimensional invariant-type spaces. Typical computational methods do not produced fixed alignments of the bases, so these alignments can change from mesh level to mesh level. We use recent techniques developed in the context of eigenvalue AFEM in order to meet this challenge [1]. Finally, computation of harmonic forms may be either a linear or

mildly nonlinear problem, depending on the method used to compute them. The most general viewpoint, which we take, views the problem as a mildly nonlinear problem. The result is that some constants in our estimates depend on the degree of overlap between the space of discrete harmonic forms on the initial mesh and the (target) space of continuous harmonic forms. We however are able to show that the dependence of these constants on the initial overlap disappears as the mesh is refined.

We finally obtain rate optimality. First we prove a local discrete upper bound. Using essentially standard techniques we then obtain rate optimality using a new argument based on the commuting and locally defined projecting quasi-interpolant of Falk and Winther [2]. As when proving that AFEM contracts, we also discuss dependence of various constants on the initial overlap between the continuous and discrete spaces of harmonic forms.

References

[1] A. Bonito and A. Demlow, *Convergence and optimality of higher-order adaptive finite element methods for eigenvalue clusters*, SIAM J. Numer. Anal. **54** (2016), 2379–2388.

[2] R. Falk and R. Winther *Local bounded cochain projections*, Math. Comp. **83** (2014), 2631–2656.

Adaptivity and blowup detection for nonlinear parabolic problems

Emmanuil H. Georgoulis

(joint work with Andrea Cangiani, Irene Kyza, Stephen Metcalfe)

In this work, we derive a practical conditional a posteriori bound for a fully-discrete first order in time implicit-explicit (IMEX) interior penalty discontinuous Galerkin (dG) in space discretization of a non self-adjoint semilinear parabolic PDE with quadratic nonlinearity. The choice of an IMEX discretization and, in particular, the explicit treatment of the nonlinearity, offers advantages in the context of finite time blow-up – this is highlighted below via the discretization of the related ODE problem with various time-stepping schemes. The choice of a dG method in space offers stability of the spatial operator in convection-dominated regimes on coarse meshes; we stress, however, that the theory presented below is directly applicable to the case of conforming finite element approximations in space. The conditional a posteriori error bounds are derived in the $L^\infty(L^2) + L^2(H^1)$-type norm. The derivation is based on energy techniques combined with the Gagliardo-Nirenberg inequality while retaining the key idea introduced in [6] – judicious usage of Gronwall's lemma. A key novelty of our approach is the use of a *local-in-time* continuation argument in conjunction with a space-time reconstruction. Global-in-time continuation arguments have been used to derive conditional a posteriori error estimates for finite element discretizations of PDEs with globally bounded solutions, cf. [1, 4, 5]. A useful by-product of the local continuation argument used in this work is that it gives a natural stopping criterion for approach towards the blow-up time. The use of space-time reconstruction,

introduced in [7, 8] for conforming finite element methods and in [2, 3] for dG methods, allows for the derivation of a posteriori bounds in norms weaker than $L^2(H^1)$ and offers great flexibility in treating general spatial operators and their respective discretizations.

Furthermore, a space-time adaptive algorithm is proposed which uses the conditional a posteriori bound to control the time step lengths and the spatial mesh modifications. The adaptive algorithm is a non-trivial modification of typical adaptive error control procedures for parabolic problems. In the proposed adaptive algorithm, the tolerances are adapted in the run up to blow-up time to allow for larger absolute error in an effort to balance the relative error of the approximation. The space-time adaptive algorithm is tested on three numerical experiments, two of which exhibit point blow-up and one which exhibits regional blow-up. Each time the algorithm appears to detect and converge to the blow-up time without surpassing it.

References

[1] S. Bartels, *A posteriori error analysis for time-dependent Ginzburg-Landau type equations*, Numer. Math., 99 (2005), pp. 557–583.

[2] A. Cangiani, E. H. Georgoulis, and S. Metcalfe, *Adaptive discontinuous Galerkin methods for nonstationary convection-diffusion problems*, IMA J. Numer. Anal., 34 (2014), pp. 1578–1597.

[3] E. H. Georgoulis, O. Lakkis, and J. M. Virtanen, *A posteriori error control for discontinuous Galerkin methods for parabolic problems*, SIAM J. Numer. Anal., 49 (2011), pp. 427–458.

[4] E. H. Georgoulis and C. Makridakis, *On a posteriori error control for the Allen-Cahn problem*, Math. Method. Appl. Sci., 37 (2014), pp. 173–179.

[5] D. Kessler, R. H. Nochetto, and A. Schmidt, *A posteriori error control for the Allen-Cahn problem: circumventing Gronwall's inequality*, ESAIM Math. Model. Numer. Anal., 38 (2004), pp. 129–142.

[6] I. Kyza and C. Makridakis, *Analysis for time discrete approximations of blow-up solutions of semilinear parabolic equations*, SIAM J. Numer. Anal., 49 (2011), pp. 405–426.

[7] O. Lakkis and C. Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp., 75 (2006), pp. 1627–1658.

[8] C. Makridakis and R. H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 1585–1594.

Linearization of the p-Poisson equation

Lars Diening

(joint work with M. Wank, M. Fornasier, R. Tomasi)

The p-Poisson problem consists in finding the solution $u \in W_0^{1,p}(\Omega)$ of

$$-\operatorname{div}\left(|\nabla u|^{p-2}\nabla u\right) = f$$

for $1 < p < \infty$ and suitable f. The function u may be scalar or vectorial. For $p = 2$ this is the standard Poisson problem, but for $p \neq 2$ the problem is non-linear.

Our goal is to approximate the solution by means of the finite element method. Many results have appeared in this direction. The a priori analysis goes back to

[1, 7, 6]. In [8, 4] a first notion of error estimators has been introduced. In [9] a convergence theorem for the adaptive finite element method has been proved without any rates. Later in [2] rate optimality has been proved for new error estimators. However, all of these results assume that the corresponding discrete finite element solution can be calculated exactly. Since the problem is non-linear, this is a serious obstacle. The purpose of this paper is to introduce an iterative algorithm based on linear problems, that approximates the solutions of the p-Poisson problem. The only result in this direction can be found in [3], where the super-linear case $p \geq 2$ has been studied using a wavelet stabilization. However, no rates of convergence have been of obtained. In contrast, we study the sub-linear case $p \leq 2$ and we are able to derive rates of convergence. We are interested in this paper in the outer non-linear loop, so the results apply to an iteration in the space $W_0^{1,p}(\Omega)$ as well for any conformal finite element approximation.

From now on let $1 < p \leq 2$. Starting from $v_0 \in W_0^{1,p}(\Omega)$ a first idea would be to define v_{n+1} iteratively by

$$-\operatorname{div}\left(|\nabla v_n|^{p-2} v_{n+1}\right) = f.$$

Each of this problem is linear, but the weight $|\nabla v_n|^{p-2}$ may degenerate. This cannot be avoided, since $|\nabla v_n| = 0$ at any critical point of v_n. This will generate matrices of arbitrary high stiffness. This problem can only be avoided by truncation of the weight, so we could consider for $0 < \varepsilon_- \leq \varepsilon_+ < \infty$ the problem

$$-\operatorname{div}\left((\varepsilon_- \vee |\nabla v_n| \wedge \varepsilon_+)^{p-2} v_{n+1}\right) = f.$$

Now the question arises how to adapt ε during the iteration to ensure that the algorithm converges to the correct solution. This idea is only adhoc, but we will derive a similar algorithm in the following by means of a variational problem.

The solution u of our p-Poisson problem is the minimizer of the energy

$$\mathcal{J}(v) := \int \tfrac{1}{p} |\nabla v|^p \, dx - \int v f \, dx.$$

Now we relax the energy by an additional function $a : \Omega \to (0, \infty)$ and define

$$\mathcal{J}(v, a) := \int \tfrac{1}{2} |a|^{p-2} |\nabla v|^2 + (\tfrac{1}{p} - \tfrac{1}{2}) |a|^p \, dx - \int v f \, dx.$$

The basic idea now is to minimize alternately with respect to v and a. While the minimization w.r.t. to v is a linear problem, the minimization w.r.t. a is explicitly given. To overcome the degeneracy we introduce $\varepsilon_n = (\varepsilon_{n,-}, \varepsilon_{n,+})$ and restrict the minimization of a to the functions with values in $(\varepsilon_{n,-}, \varepsilon_{n,+})$. We propose the following algorithm of Kačanov type.

Algorithm

Start with $v_0 \in W_0^{1,p}(\Omega)$ and $\varepsilon = (\varepsilon_{0,-}, \varepsilon_{0,+}) = (1, 1)$. Repeat

(1) Choose new truncation interval $\varepsilon_{n+1} \supset \varepsilon_n$.
(2) $a_{n+1} := \operatorname{argmin}_{a \in (\varepsilon_{n,-}, \varepsilon_{n,+})} \mathcal{J}(a, v_n)$, i.e. $a_{n+1} = \varepsilon_{n,-} \vee |\nabla v_n| \wedge \varepsilon_{n,+}$.
(3) $v_{n+1} := \operatorname{argmin}_v \mathcal{J}(a_{n+1}, v)$, i.e. $-\operatorname{div}(a_{n+1}^{p-2} \nabla v_{n+1}) = f$.

We consider two questions: First, how does the minimizer u_ε of the energy $\mathcal{J}_\varepsilon(v) := \mathcal{J}(v, \varepsilon_- \vee |\nabla v| \wedge \varepsilon_+)$ converge to u? We show that $\mathcal{J}_\varepsilon(u_\varepsilon) - \mathcal{J}(u) \leq c\,\varepsilon_-^p + \varepsilon_+^{-d/(p-1)}$. The proof is based on the Lipschitz truncation technique, see [5]. Note that $\mathcal{J}_\varepsilon$ is defined on $W_0^{1,2}(\Omega)$ but $\mathcal{J}$ only on $W_0^{1,p}(\Omega)$.

Second, how does the algorithm converge for fixed ε to u_ε. We prove that $\mathcal{J}_\varepsilon(v_n) - \mathcal{J}_\varepsilon(u_\varepsilon) \leq (1 - \delta(\varepsilon))^n(\mathcal{J}_\varepsilon(v_0) - \mathcal{J}_\varepsilon(u_\varepsilon))$. This is linear (exponential) convergence, but the speed depends on the truncation ε, since $\delta(\varepsilon) \eqsim (\varepsilon_-/\varepsilon_+)^{2-p}$.

Combining these two results, we obtain a convergence result with specific rates. We can show that the choice $\varepsilon = (n^{-\alpha}, n^\beta)$ with $\alpha, \beta > 0$ ensures that the energy error $\mathcal{J}_{\varepsilon_n}(v_n) - \mathcal{J}(u)$ decreases with $\mathcal{O}(n^{-\gamma})$ for some $\gamma > 0$.

We have also performed numerical experiments including adaptivity. We developed error estimators that indicate when to reduce $\varepsilon_{n,-}$, increase $\varepsilon_{n,+}$, refine the grid, or just recalculate a_n and v_n. These experimental results suggest that our adaptive algorithm performs rate optimally in the sense of energy accuracy vs. costs.

References

[1] J. W. Barrett and W. B. Liu. Finite element approximation of the p-Laplacian. *Math. Comp.*, 61(204):523–537, 1993.

[2] L. Belenki, L. Diening, and Ch. Kreuzer. Optimality of an adaptive finite element method for the p-Laplacian equation. *IMA J. Numer. Anal.*, 32(2):484–510, 2012.

[3] C. Canuto and K. Urban. Adaptive optimization of convex functionals in Banach spaces. *SIAM J. Numer. Anal.*, 42(5):2043–2075 (electronic), 2005.

[4] C. Carstensen, W. Liu, and N. Yan. A posteriori error estimates for finite element approximation of parabolic p-Laplacian. *SIAM J. Numer. Anal.*, 43(6):2294–2319, 2006.

[5] L. Diening, Ch. Kreuzer, and E. Süli. Finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *SIAM J. Numer. Anal.*, 51(2):984–1015, 2013.

[6] L. Diening and M. Růžička. Interpolation operators in Orlicz Sobolev spaces. *Num. Math.*, 107(1):107–129, July 2007.

[7] C. Ebmeyer and WB. Liu. Quasi-norm interpolation error estimates for the piecewise linear finite element approximation of p-Laplacian problems. *Numer. Math.*, 100(2):233–258, 2005.

[8] W. Liu and N. Yan. On quasi-norm interpolation error estimation and a posteriori error estimates for p-Laplacian. *SIAM J. Numer. Anal.*, 40(5):1870–1895 (electronic), 2002.

[9] A. Veeser. Convergent adaptive finite elements for the nonlinear Laplacian. *Numer. Math.*, 92(4):743–770, 2002.

Aposteriori analysis of discontinuous Galerkin timestepping methods

Omar Lakkis
(joint work with E.H. Georgoulis and T.P. Wihler)

1. Set-up

1.1. **Time discontinuous Galerkin.** We are interested in adaptive numerical time-discontinuous Galerkin approximation of a function $u : [0,T] \to \mathcal{X}$ satisfying the *parabolic initial value problem*

$$u' + \mathcal{A}u = f \text{ and } u(0) = u_0, \tag{1}$$

where $\mathcal{A} : \mathcal{X} \to \mathcal{X}'$ is an elliptic operator, with respect to a Gelfand triple of Hilbert spaces a partition of the time interval $\mathcal{I} := \{I_n : n = 1, \dots, N\}$, $I_n := (t_{n-1}, t_n]$ and $I_0 := \{0\}$, with $t_0 < \cdots < t_N$. A *discontinuous Galerkin method in time* for equation (1) consists in finding a function $\mathsf{U} : [0,T] \to \mathcal{X}$ that is $\mathcal{I}$-wise continuous and such that

$$\mathsf{U}(0) := \Pi_n 0 u_0, \tag{2}$$

$$\begin{aligned} \int_{I_n} [(\mathsf{U}', \mathsf{V})_{\mathcal{H}} + \langle \mathcal{A}\mathsf{U} \,|\, \mathsf{V}\rangle] + \left(\llbracket \mathsf{U} \rrbracket_{n-1}, \mathsf{V}(t_{n-1}^+)\right)_{\mathcal{H}} \\ = \int_{I_n} \langle f \,|\, \mathsf{V}\rangle \quad \forall \mathsf{V} \in \mathbb{P}^{r_n}(I_n; \mathbb{X}_n), \end{aligned} \tag{3}$$

where for each $n = 1, \dots, N$, $\mathbb{X}_n$ is a finite dimensional subspace of $\mathcal{X}$, $\Pi_n n$ indicates the $\mathcal{H}$-orthogonal projection from $\mathcal{X}'$ onto $\mathbb{X}_n$, $(\phi, \psi)_{\mathcal{H}}$ is the $\mathcal{H}$-inner product of $\phi, \psi \in \mathcal{H}$ and $(l, \phi)_{\mathcal{H}}$ is the duality action of $l \in \mathcal{X}'$ onto $\phi \in \mathcal{X}$.

The derivation of adaptive methods may be rigorously based on *aposteriori error bounds*. Establishing such bounds is this work's object, which we approach using the following two main technical tools

(i) the elliptic reconstruction for conforming methods as described in [MN03, LM06, GLM13],
(ii) the *time lifting operator* introduced and studied by [SW10, MN06], in the context of time-discretizations of (1).

Our main contribution is the establishment of general aposteriori error bounds for the *fully discrete schemes* of the form (2)–(3).

1.2. **Elliptic reconstruction.** For each $n = 0, \dots, N$, denoting the Ritz projector by $\mathsf{R}_n : \mathcal{X} \to \mathbb{X}_n$, the *elliptic reconstructor* is defined as $\mathcal{R}_n := \mathcal{A}^{-1}\Pi_n\mathcal{A}$ (that is, a preinverse of R_n, $\mathsf{R}_n\mathcal{R}_n = \mathrm{id}_{\mathbb{X}_n}$). We *assume existence of aposteriori elliptic error estimators* $\mathcal{E}_{\mathcal{Z}}$, $\mathcal{Z} \in \{\mathcal{X}, \mathcal{H}, \mathcal{X}'\}$ such that

$$\|\mathcal{R}_n \mathsf{w} - \mathsf{w}\|_{\mathcal{Z}} \le \mathcal{E}_{\mathcal{Z}}[\mathsf{w}, \mathcal{A}\mathcal{R}_n\mathsf{w}] \quad \forall \mathsf{w} \in \mathbb{X}_n, n = 1, \dots, N. \tag{4}$$

The *elliptic reconstruction of* U is now defined as $\tilde{U} : [0,T] \to \mathcal{X}$ such that

$$\tilde{U}(t) := \mathcal{R}_n \mathsf{U}(t) \text{ for each } t \in I_n \text{ and } n = 0, \dots, N. \tag{5}$$

1.3. **Time reconstruction.** To analyze the effects of time discretization, we use a *time reconstructor* defined, for each $\mathcal{Z} \subseteq \mathcal{H}$ Hilbert subspace, and each $W \in \mathbb{P}^{r_n}(I_n;\mathcal{Z})$, for $n = 1,\dots,N$ with $\hat{W}$ defined from W by the weak initial value problem

$$\int_{I_n} \left(\hat{W}', V\right)_{\mathcal{H}} = \left(\llbracket W \rrbracket_{n-1}, V(t_{n-1}^+)\right)_{\mathcal{H}} + \int_{I_n} (W', V)_{\mathcal{H}} \quad \forall V \in \mathbb{P}^{r_n}(I_n;\mathcal{X}), \tag{6}$$

$$\hat{W}(t_{n-1}^+) = W(n-1). \tag{7}$$

Following [MN06, SW10] there hold the *time-reconstruction error identities*

$$\left\| W - \hat{W} \right\|_{\mathrm{L}_\infty(I_n;\mathcal{Z})} = \| \llbracket W \rrbracket_{n-1} \|_{\mathcal{Z}}, \tag{8}$$

$$\left\| W - \hat{W} \right\|_{\mathrm{L}_2(I_n;\mathcal{Z})} = \left(\frac{k(r+1)}{(2r+1)(2r+3)} \right)^{1/2} \| \llbracket W \rrbracket_{n-1} \|_{\mathcal{Z}}. \tag{9}$$

2. Analysis

2.1. **Splitting the error.** Starting from U, we may therefore apply the elliptic reconstructor to obtain $\tilde{U}$ and then the time reconstructor to obtain the *space-time reconstruction* $\hat{U} := \hat{\tilde{U}} \in \mathrm{C}^{0,1}(0,T;\mathcal{A})$. This allows the splitting of the full error as follows

$$e := \mathsf{U} - u = \mathsf{U} - \hat{U} + \hat{U}u - u =: -\epsilon + \rho, \tag{10}$$

where the error ρ satisfies

$$\rho' + \mathcal{A}\rho = \xi := \xi^1 + \xi^2, \text{ with } \xi^i := \sum_{n=1}^{N} 1_{I_n} \xi_n^i, \qquad i = 1,2, \tag{11}$$

and 1_A denoting the indicator function of a generic set A, where

$$\begin{aligned} \xi_n^1 &:= \left(\tilde{U}' - \mathsf{U}'\right) + \chi_n\left(\left[\!\left[\tilde{U} - \mathsf{U}\right]\!\right]_{n-1}\right) + \mathcal{A}\left(\hat{U} - \tilde{U}\right), \\ \xi_n^2 &:= \Pi_n f - f + \chi_n\left(\llbracket \mathsf{U} - \pi_n \mathsf{U} \rrbracket_{n-1}\right). \end{aligned} \tag{12}$$

The core part of the analysis consists in using energy techniques on (11) to bound ρ and the elliptic error esimates (4) to bound ϵ. The following two results are crucial to close the analysis.

2.2. **Time reconstruction error estimate.** We have the bound

$$\begin{aligned} &\int_{I_n} \left\| \mathcal{A}(\hat{U} - \tilde{U}) \right\|_{\mathcal{X}'}^2 \le (\beta \eta_n^{\mathrm{time}})^2, \text{ with } \beta := \sup_{w \in \mathcal{X}} \|\mathcal{A}w\|_{\mathcal{X}'} / \|w\|_{\mathcal{X}} \\ &\eta_n^{\mathrm{time}} := \left(\frac{k(r+1)}{(2r+1)(2r+3)} \right)^{1/2} \left((\mathcal{E}_{\mathcal{X}}[\llbracket \mathsf{U} \rrbracket_{n-1}, \llbracket \mathsf{A}_n \mathsf{U} \rrbracket_{n-1}])^2 + \left\| \llbracket \mathsf{U} \rrbracket_{n-1} \right\|_{\mathcal{X}} \right). \end{aligned} \tag{13}$$

2.3. **Elliptic reconstruction error.** For $t \in I_n$, $n = 1, \dots, N$, we have

$$\left\| \tilde{U}' + \chi_n \left[\!\left[\tilde{U} \right]\!\right]_{n-1} - \mathsf{U}' - \chi_n \left[\!\left[\mathsf{U} \right]\!\right]_{n-1} \right\|_{\mathcal{X}'} \le \eta_n^{\text{space}}, \tag{14}$$
$$\text{with } \eta_n^{\text{space}} := \mathcal{E}_{\mathcal{X}'}[\mathsf{U}' + \chi_n([\![\mathsf{U}]\!]_{n-1}), \mathsf{A}_n\mathsf{U}' + \chi_n([\![\mathsf{A}_n\mathsf{U}]\!]_{n-1})].$$

3. Main results

Denote by α the coercivity of $\mathcal{A}$ and by $C_{\text{PF},\mathcal{X}}$ the Poincaré–Friedrichs constant associated with $\mathcal{H}$ and $\mathcal{X}$, $\lambda_n := \min\{1, t_n^{-1}\}$. The following are valid.

3.1. $\mathrm{L}_2(\mathcal{X})$**-norm aposteriori bound.** For each $n = 1, \dots, N$, we have the bound

$$\begin{aligned} \|u - \mathsf{U}\|^2_{\mathrm{L}_2(0,t_n;\mathcal{X})} \le{}& \frac{6}{\alpha} \|u_0 - \pi_0 u_0\|^2_{\mathcal{H}} + \Big(\frac{27}{2\alpha^2} + 3\Big) \sum_{j=1}^{n} (\eta_j^{\text{time}})^2 \\ &+ 3\sum_{j=1}^{n} \big(\mathcal{E}_{\mathcal{X}}[\mathsf{U}, \mathsf{A}_n\mathsf{U}]\big)^2 + 3\sum_{j=1}^{n} \int_{I_j} \big((\beta\eta_j^{\text{space}})^2\big) \\ &+ \frac{27}{\alpha}\lambda_n \Big(\sum_{j=1}^{n} \int_{I_j} \eta_j^{\text{mesh}} \Big)^2 + \frac{27}{2\alpha^2} C^2_{\text{PF},\mathcal{X}} (1-\lambda_n)^2 \sum_{j=1}^{n} (\eta_j^{\text{mesh}})^2. \end{aligned} \tag{15}$$

3.2. $\mathrm{L}_\infty(\mathcal{H})$**-norm aposteriori bound.** For $n = 1, \dots, N$, we have the bound

$$\begin{aligned} \|u - \mathsf{U}\|_{\mathrm{L}_\infty(0,t_n;\mathcal{H})} \le{}& \Big(2\|u_0 - \pi_0 u_0\|^2_{\mathcal{H}} + \frac{3}{\alpha} \sum_{j=1}^{n} \Big((\eta_j^{\text{time}})^2 + \int_{I_j} (\eta_n^{\text{space}})^2 \Big) \\ &+ 2\lambda_n \Big(\sum_{j=1}^{n} \int_{I_j} \eta_j^{\text{mesh}} \Big)^2 + \frac{3}{\alpha} C^2_{\text{PF},\mathcal{X}} (1-\lambda_n)^2 \sum_{j=1}^{n} (\eta_j^{\text{mesh}})^2 \Big)^{\frac{1}{2}} \\ &+ \max_{j=1,\dots,n} \Big(\mathcal{E}_{\mathcal{H}}[[\![\mathsf{U}]\!]_{n-1}, [\![\mathsf{A}_n\mathsf{U}]\!]_{n-1}] + \big\|[\![\mathsf{U}]\!]_{n-1}\big\|_{\mathcal{H}} \Big) + \max_{0 \le t \le t_n} \mathcal{E}_{\mathcal{H}}[\mathsf{U}, \mathsf{A}_n\mathsf{U}]. \end{aligned} \tag{16}$$

References

[GLM13] Emmanuil H. Georgoulis, Omar Lakkis, and Charalambos Makridakis, *A posteriori $L^\infty(L^2)$-error bounds for finite element approximations to the wave equation*, IMA J. Numer. Anal. **33** (2013), no. 4, 1245–1264. MR 3119715

[LM06] Omar Lakkis and Charalambos Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp. **75** (2006), no. 256, 1627–1658. MR 2240628 (2007e:65122)

[MN03] Charalambos Makridakis and Ricardo H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal. **41** (2003), no. 4, 1585–1594. MR 2034895 (2004k:65157)

[MN06] ———, *A posteriori error analysis for higher order dissipative methods for evolution problems*, Numer. Math. **104** (2006), no. 4, 489–514. MR 2249675 (2008b:65114)

[SW10] Dominik Schötzau and Thomas P. Wihler, *A posteriori error estimation for hp-version time-stepping methods for parabolic partial differential equations*, Numer. Math. **115** (2010), no. 3, 475–509. MR 2640055

Adaptive sparse and low-rank methods for parametric PDEs

Wolfgang Dahmen
(joint work with Markus Bachmayr, Albert Cohen)

Parameter dependent families of partial differential equations (PDEs) arise in a diversity of contexts. For instance, Karhunen-Loève, polynomial chaos or other types of expansions of random coefficient fields in a PDE, in principle, give rise to even infinitely many parameters. Therefore, we consider a class of parametric operator equations where the involved parameters could either be of deterministic or stochastic nature. Specifically, we focus on scenarios involving a large number of parameters in the context of elliptic boundary value problems of the type

$$-\mathrm{div}(a(y)\nabla u) = f \quad \text{in } \Omega, \quad u = 0 \text{ on } \partial\Omega, \quad y \in \mathcal{Y} := [-1,1]^{\mathcal{I}}, \tag{1}$$

where the coefficient field is given as an expansion

$$a(x,y) = a_0(x) + \sum_{j\in\mathcal{I}} y_j \theta_j(x) \tag{2}$$

subject to suitable conditions on the θ_j that ensure convergence of the expansion as well as uniform ellipticity of the operator. Thus, (1) has a unique solution (as a function of the spatial variables x as well as of the parametric variables $y \in \mathcal{Y} = [-1,1]^{\mathcal{I}}$) in $L_2(\mathcal{Y}, H_0^1(\Omega)) = H_0^1(\Omega) \otimes L_2(\mathcal{Y})$. As mentioned before, $\#(\mathcal{I}) = \infty$ is permitted in which case $u(y) = u(\cdot, y)$ can be understood as the limit of solutions to finitely parametrized problems obtained by truncating the expansion (2). Hence, in this case the *anisotropy* of the parameter dependence, induced by the decay conditions on the expansion (2), is essential.

From a practical point of view a typical "forward" simulation task concerns the approximate computation of a quantity of interest such as the expectation of the solution in the stochastic case or local spatial averages as functions of the parameters. In either case one is confronted in one way or the other with what is commonly referred to as "curse of dimensionality". Therefore, understanding the whole parameter-to-solution map $y \mapsto u(y)$ is of fundamental importance, see e.g. [6, 7, 11] and the references therein. On the theoretical side the vast majority of contributions to this problem area are concerned with *sparse polynomial* parametric expansions. One such strategy is based on performing sparse best n-term approximations of the solution map in an *a priori* chosen system of tensor product form in the parametric variables. This approach has been extensively analyzed in the case of tensor product Legendre polynomial bases, for which approximation rates have been established, see e.g. [3, 3]. On the practical side, a prominent approach is to combine such expansions with collocation, see e.g. [4]. The complexity analysis of such methods is essentially based on a priori estimates.

An alternative strategy for addressing the challenges posed by high dimensionality is to use low-rank approximations of solutions based on a separation of spatial and parametric variables. In this respect we review here some recent results from [2] which, in particular, ask what can be gained by exploiting further low rank structures, in particular using optimized systems of basis functions obtained by

singular value decomposition techniques. Based on the concepts from [1], a generic adaptive solver is proposed that is guaranteed to achieve a given target accuracy after finitely many steps without requiring any a priori knowledge about the solution. Moreover, it can be specified so as to produce approximate solutions to (1) in each of the following approximation formats: (a) sparse n-term approximation with respect a fixed background basis for $H_0^1(\Omega)\otimes L_2(\mathcal{Y})$, (b) Hilbert-Schmidt type low-rank approximations separating spatial and parametric variables, as well as (c) approximate solutions in hierarchical tensor formats separating also all parametric variables. The latter variant is suited best for a possibly large but *fixed* finite number of parametric variables with "equal importance". Although it involves a similar format as [10] it is conceptually different and can be shown to exhibit near-optimal performance with respect to relevant approximation classes, see [2]. Therefore, we confine the subsequent discussion to the two former variants (a), (b) focusing on the case of ininitely many parameters of "anisotropic importance", remarking that for variant (a) the scheme agrees in essence with the one proposed in [5] in the wavelet context and extended to parametric PDEs in [12, 13].

On the theoretical side, inspired by the findings in [3], several model scenarios are presented which show that optimized low-rank expansions can either bring significant or no improvement over sparse polynomial expansions, regarding the number of terms needed to realize a given target accuracy. On the computational side, the complexity of the adaptive solver is analyzed for its respective specification to (b) low-rank structure as well as (a) sparse background basis expansions. First, guided by the above mentioned model scenarios, benchmark classes are introduced which quantifiy the expceted approximability properties of solutions with respect to the decay of singular values, sparsity of the (lower-dimensional) Hilbert-Schmidt factors, or of expansions in terms of the fixed background basis for $H_0^1(\Omega) \otimes L_2(\mathcal{Y})$. Moreover, it is indicated that specific types of parametric expansions (2) have a significant effect on the performance on algorithms based on a posteriori residual evaluations. Specifically, in contrast to Karhunen-Loève type representations, expansion functions θ_j of multilevel or wavelet type allow one to fully exploit the approximability properties given in the benchmark classes. In fact, the algorithm is shown to realize then near-best convergence and complexity rates for both variants (a) and (b) improving somewhat on the results in [13] with regard to (a). Adaptive solvers based on finite element discretizations in the spatial variable have been proposed also in [8, 9] where, however, the complexity with respect to the parametric components does not seem to be controled. An essential difference between the algorithm discussed here from those in [10, 8, 9] lies in: (i) a new adaptive operator application scheme using a posteriori information through a judicious decomposition of the current aproximation and a careful compressibility analysis of the operator components; (ii) threshold based tensor-recompression and coarsening schemes in the spirit of [5, 1]. Some numerical experiments quantify the theoretical findings and support the predictions derived from the above

mentioned model scenarios. On the other hand, the price for the obtained near-optimal performance is a high level of intrusiveness, i.e., standard numerical tools cannot easily be incorporated.

References

[1] M. Bachmayr, W. Dahmen, *Adaptive near-optimal rank tensor approximation for high-dimensional operator equations*, Found. Comput. Math., **15** (No. 4) (2015), 839–898, DOI 10.1007/s10208-013-9187-3.

[2] M. Bachmayr, A. Cohen, W. Dahmen, *Parametric PDEs: Sparse or low-rank approximations?* July 2016, http://arxiv.org/abs/1607.04444 [math.NA].

[3] M. Bachmayr, A. Cohen, G. Migliorati, *Sparse polynomial approximation of parametric elliptic PDEs. Part I: affine coefficients*, 2015, arXiv:1509.07045 [math.NA],

[4] J. Beck, R. Tempone, F. Nobile, L. Tamellini, *On the optimal polynomial approximation of stochastic PDEs by Galerkin and collocation methods*, Math. Models Methods Appl. Sci., **22** (No.9), (2012), DOI 10.1142/S0218202512500236,

[5] A. Cohen, W. Dahmen, R. DeVore, *Adaptive wavelet methods II – beyond the elliptic case*, Foundations of Computational Mathematics, **2** (No. 3) (2002), 203–245.

[6] A. Cohen, R. DeVore, *Approximation of high-dimensional parametric PDEs*, Acta Numer., **24**, (2015), 1–159, DOI 10.1017/S0962492915000033.

[7] A. Cohen, R. DeVore, Ch. Schwab, *Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs*, Found. Comput. Math., **10**, (No.6) (2010), 615–646, DOI 10.1007/s10208-010-9072-2.

[8] M. Eigel,C.J. Gittelson, Ch. Schwab, E. Zander, *Adaptive stochastic Galerkin FEM*, Comput. Methods Appl. Mech. Engrg., **270** (2014), 247–269, DOI10.1016/j.cma.2013.11.015.

[9] M. Eigel, C.J. Gittelson, Ch. Schwab, E. Zander, *A convergent adaptive stochastic Galerkin finite element method with quasi-optimal spatial meshes*, ESAIM: M2AN, **49**, (No.5) (2015), 1367–1398.

[10] M. Eigel,M. Pfeffer, R. Schneider, *Adaptive stochastic Galerkin FEM with hierarchical tensor representions*, Preprint 2153, 2015, WIAS Berlin,

[11] C.J. Gittelson, Ch. Schwab, *Sparse tensor discretization of high-dimensional parametric and stochastic PDEs*, Acta Numerica, **20** (2011),

[12] C.J. Gittelson, *An adaptive stochastic Galerkin method for random elliptic operators*, Mathematics of Computation, **82** (2013), 1515–1541.

[13] C.J. Gittelson, *Adaptive wavelet methods for elliptic partial differential equations with random operators*, Numer. Math., **126** (2014), 471–513.

On the discrete reliability for nonconforming finite element methods

Dietmar Gallistl

(joint work with Carsten Carstensen, Mira Schedensack)

Let $\mathcal{T}$ be a simplicial partition of the bounded Lipschitz polytope $\Omega \subseteq \mathbb{R}^n$, $n \geq 2$. Classical a posteriori error estimates for finite element discretizations of elliptic partial differential equations (PDEs) posed in the Sobolev space $H^1_0(\Omega)$ (for instance Poisson's equation) usually take the form

$$\|\nabla(u - u_{\mathcal{T}})\|^2_{L^2(\Omega)} \leq C_{\mathrm{rel}} \sum_{T \in \mathcal{T}} \eta^2_T(u_{\mathcal{T}})$$

where u is the solution to the PDE and $u_{\mathcal{T}}$ is its finite element approximation with respect to the mesh $\mathcal{T}$. The bound on the right hand side consists of the sum of

computable a posteriori error estimator contributions $\eta_T^2(u_{\mathcal{T}})$ for each cell $T \in \mathcal{T}$. Such an error estimate is referred to as *reliability* estimate. It was observed in [4] that, for conforming finite elements, the comparison of $u_{\mathcal{T}}$ with the finite element solution $u_{\widehat{\mathcal{T}}}$ with respect to a refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}$ leads to the so-called *discrete reliability*

$$\|\nabla(u_{\widehat{\mathcal{T}}} - u_{\mathcal{T}})\|_{L^2(\Omega)}^2 \leq C_{\mathrm{drel}} \sum_{T \in \mathcal{T} \setminus \widehat{\mathcal{T}}} \eta_T^2(u_{\mathcal{T}}).$$

Only the error estimator contributions on those elements of $\mathcal{T}$ that are refined and therefore not contained in $\widehat{\mathcal{T}}$ enter the upper bound. The constant C_{drel} is universal and only depends on the shape-regularity of the triangulations. It does not depend on the mesh-size or the grade of refinement from $\mathcal{T}$ to $\widehat{\mathcal{T}}$. This localization argument is a key argument in the proof of optimal convergence rates of adaptive finite element methods [4].

In the case of nonconforming finite element methods, proofs of the discrete reliability are not obvious because the discrete spaces on refined triangulations are not nested. This contribution presents a proof of the discrete reliability [1] for the nonconforming P_1 finite element method (Crouzeix-Raviart method). For a given triangulation $\mathcal{T}$, the discrete space $V(\mathcal{T})$ consists of piecewise affine functions that are continuous in the midpoints of the interior $(n-1)$-dimensional hyper-faces and vanish in the midpoints of the $(n-1)$-dimensional hyper-faces on the boundary $\partial\Omega$. Existing results in the literature were based on discrete Helmholtz decompositions and, thus, restricted to simply-connected planar domains. The new result is valid in general Lipschitz polytopes in $\mathbb{R}^n$. The crucial estimate is the *discrete distance control*

$$\inf_{v_{\widehat{\mathcal{T}}} \in V(\widehat{\mathcal{T}})} \|v_{\mathcal{T}} - v_{\widehat{\mathcal{T}}}\|_{L^2(\Omega)}^2 \leq C_{\mathrm{ddc}} \sum_{F \in \mathcal{F}(\mathcal{T} \setminus \widehat{\mathcal{T}})} h_F^{-1} \, \|[v_{\mathcal{T}}]_F\|_{L^2(F)}^2$$

for any $v_{\mathcal{T}} \in V(\mathcal{T})$. Here, $\mathcal{F}(\mathcal{T} \setminus \widehat{\mathcal{T}})$ is the set of all $(n-1)$-dimensional hyper-faces that belong to the simplices of $\mathcal{T}$ but not to those of $\widehat{\mathcal{T}}$. For any hyper-face F, its diameter is denoted by h_F and the bracket $[\cdot]_F$ denotes the jump across F (if F belongs to $\partial\Omega$, it denotes the trace). The stated version was presented in [3] and is slightly sharper than the result of [1]. The proof of discrete distance control is based on fine properties [2] of the employed mesh-refinement procedure (Newest Vertex Bisection).

References

[1] C. Carstensen, D. Gallistl, M. Schedensack, *Discrete reliability for Crouzeix-Raviart FEMs*, SIAM J. Numer. Anal., **51** (2013), 2935–2955.

[2] D. Gallistl, M. Schedensack, R. Stevenson, *A remark on newest vertex bisection in any space dimension*, Comput. Methods Appl. Math. **14** (2014), 317–320.

[3] Ch. Kreuzer, M. Schedensack, *Instance optimal Crouzeix-Raviart adaptive finite element methods for the Poisson and Stokes problems*, IMA J. Numer. Anal., **36** (2016), 593–617.

[4] R. Stevenson, *The completion of locally refined simplicial partitions created by bisection*, Math. Comp., **77** (2008), 227–241.

Instance optimal Crouzeix-Raviart adaptive FEM for the Poisson and Stokes problems

Mira Schedensack
(joint work with C. Kreuzer)

Recently, the seminal contribution [5] proved instance optimality for an adaptive algorithm based on a (modified) maximum marking strategy. The considered problem is the Poisson problem in $\mathbb{R}^2$ approximated with P_1 conforming finite elements.

This talk is based on [7] and considers an adaptive algorithm with the same modified maximum marking strategy for the Poisson problem and the Stokes equations in $\mathbb{R}^2$ approximated with P_1 nonconforming finite elements, also named after Crouzeix and Raviart [4]. Let $I_{\mathcal{T}}$ denote the nonconforming interpolation operator which satisfies [2]

$$\frac{1}{\Lambda}\|h_{\mathcal{T}}^{-1}(v - I_{\mathcal{T}}v)\|_{L^2(T)} \leq \|\nabla_{\mathrm{NC}}(v - I_{\mathcal{T}}v)\|_{L^2(T)} \leq \|\nabla_{\mathrm{NC}}v\|_{L^2(T)},$$

where ∇_{NC} denotes the piecewise gradient and $h_{\mathcal{T}}$ denotes the piecewise constant mesh-size function which equals $\sqrt{\mathrm{area}(T)}$ on each triangle T. For a fixed $\gamma > 2\Lambda^2 > 0$, we define the energy

$$\mathcal{G}(\mathcal{T}) := -\int_\Omega \left(\tfrac{1}{2}|\nabla_{\mathrm{NC}}u_{\mathcal{T}}|^2 - fu_{\mathcal{T}}\right) dx + \gamma\|h_{\mathcal{T}}f\|^2_{L^2(\Omega)}.$$

Then we have for any admissible refinement $\mathcal{T}_\star$ of $\mathcal{T}$ the equivalence of discrete errors and energy differences in the sense that

$$\mathcal{G}(\mathcal{T}) - \mathcal{G}(\mathcal{T}_\star) \approx \|\nabla_{\mathrm{NC}}(u_{\mathcal{T}_\star} - u_{\mathcal{T}})\|^2_{L^2(\Omega)} + \|h_{\mathcal{T}}f\|^2_{L^2(\Omega(\mathcal{T}\setminus\mathcal{T}_\star))}, \tag{1}$$

where $\Omega(\mathcal{T}\setminus\mathcal{T}_\star)$ denotes the union of all triangles that are refined in $\mathcal{T}_\star$ and $A \approx B$ abbreviates that there exist constants C_1 and C_2 that are independent of the mesh size such that $C_1A \leq B \leq C_2A$. This in particular implies that the sequence of energies is monotonically decreasing on a sequence of refined triangulations. Define the local error estimator on each side S of $\mathcal{T}$ by

$$\mathcal{E}^2_{\mathcal{T}}(S) := \|h_{\mathcal{T}}f\|^2_{L^2(\omega_S)} + h_S\|[\nabla_{\mathrm{NC}}u_{\mathcal{T}}\cdot\tau]_S\|^2_{L^2(S)},$$

where ω_S denotes the patch around S, $[\bullet]_S$ denotes the jump across S and τ denotes the tangent of S. Let $\mathcal{S}(\mathcal{T})$ denote the set of sides of $\mathcal{T}$ and define for any subset $\tilde{\mathcal{S}} \subseteq \mathcal{S}(\mathcal{T})$

$$\mathcal{E}^2_{\mathcal{T}}(\tilde{\mathcal{S}}) := \sum_{S\in\tilde{\mathcal{S}}} \mathcal{E}^2_{\mathcal{T}}(S).$$

Then we have for any admissible refinement $\mathcal{T}_\star$ of $\mathcal{T}$ discrete efficiency and discrete reliability in the sense that

$$\mathcal{E}^2_{\mathcal{T}}(\mathcal{S}(\mathcal{T})\setminus\mathcal{S}(\mathcal{T}_\star)) \approx \|\nabla_{\mathrm{NC}}(u_{\mathcal{T}_\star} - u_{\mathcal{T}})\|^2_{L^2(\Omega)} + \|h_{\mathcal{T}}f\|^2_{L^2(\Omega(\mathcal{T}\setminus\mathcal{T}_\star))}. \tag{2}$$

The proof of the discrete reliability employs a transfer operator from [1].

The third important property is the lower diamond estimate. A lower diamond $(\mathcal{T}^\wedge, \mathcal{T}_\vee; \mathcal{T}_1, \dots, \mathcal{T}_m)$ consists of regular triangulations $\mathcal{T}^\wedge, \mathcal{T}_\vee, \mathcal{T}_1, \dots, \mathcal{T}_m$ such that $\mathcal{T}^\wedge$ is the finest common coarsening of $\mathcal{T}_1, \dots, \mathcal{T}_m$, and $\mathcal{T}_\vee$ is the coarsest common refinement of $\mathcal{T}_1, \dots, \mathcal{T}_m$ and the areas of coarsening $\Omega(\mathcal{T}_j \setminus \mathcal{T}_\vee)$ are pairwise disjoint [5]. Then the lower diamond estimate holds

$$\mathcal{G}(\mathcal{T}^\wedge) - \mathcal{G}(\mathcal{T}_\vee) \approx \sum_{j=1}^{m} \big(\mathcal{G}(\mathcal{T}_j) - \mathcal{G}(\mathcal{T}_\vee)\big). \tag{3}$$

A similar assertion follows directly from the definition of a lower diamond for the difference of the discrete solution on the finest triangulation $\mathcal{T}_\vee$ and its nonconforming interpolant on the triangulations $\mathcal{T}^\wedge$ and $\mathcal{T}_1, \dots, \mathcal{T}_m$. Then a discrete best-approximation result, which can be regarded as a discrete version of the results from [6, 3] and which is not immediate due to the nonconformity of the method, allows to conclude the lower diamond estimate in terms of the energy differences (3).

The three ingredients (1)–(3) allow for the proof of instance optimality of the adaptive algorithm up to data oscillations in the following sense. There exist constants $C, \tilde{C} \geq 1$ such that the sequence of triangulations $(\mathcal{T}_k)_{k\in\mathbb{N}}$ created by the adaptive algorithm satisfies

$$\|\nabla_{\mathrm{NC}}(u - u_{\mathcal{T}_k})\|_{L^2(\Omega)} + \operatorname{osc}(f, \mathcal{T}_k) \leq \tilde{C}\big(\|\nabla_{\mathrm{NC}}(u - u_{\mathcal{T}})\|_{L^2(\Omega)} + \operatorname{osc}(f, \mathcal{T})\big)$$

for all admissible refinements $\mathcal{T}$ of the initial triangulation $\mathcal{T}_0$ with

$$C\#(\mathcal{T} \setminus \mathcal{T}_0) \leq \#(\mathcal{T}_k \setminus \mathcal{T}_0).$$

This instance optimality can be generalised to the nonconforming P_1 discretisation of the Stokes equations. The key ingredients for this generalisation are the two conservation properties

$$\operatorname{div}_{\mathrm{NC}} I_{\mathcal{T}} v_{\mathcal{T}_\star} = 0 \qquad \text{for all } v_{\mathcal{T}_\star} \in \mathrm{CR}^1_0(\mathcal{T}_\star) \text{ with } \operatorname{div}_{\mathrm{NC}} v_{\mathcal{T}_\star} = 0,$$

$$\int_T \operatorname{div}_{\mathrm{NC}}(\mathcal{K}_{\mathcal{T}_\star} v_{\mathcal{T}})\, dx = 0 \qquad \text{for all } v_{\mathcal{T}} \in \mathrm{CR}^1_0(\mathcal{T}) \text{ with } \operatorname{div}_{\mathrm{NC}} v_{\mathcal{T}} = 0,$$

where $\mathcal{T}_\star$ is an admissible refinement of $\mathcal{T}$ and $\mathcal{K}_{\mathcal{T}_\star}$ maps P_1 nonconforming functions on the coarse triangulation to P_1 nonconforming functions on the refinement $\mathcal{T}_\star$.

References

[1] C. Carstensen, D. Gallistl and M. Schedensack, *Discrete reliability for Crouzeix–Raviart FEMs*, SIAM J. Numer. Anal., **51** (2013), 2935–2955.

[2] C. Carstensen and J. Gedicke, *Guaranteed lower bounds for eigenvalues*, Math. Comp., **83** (2014), 2605–2629.

[3] C. Carstensen, D. Peterseim and M. Schedensack, *Comparison results of finite element methods for the Poisson model problem*, SIAM J. Numer. Anal., **50** (2012), 2803–2823.

[4] M. Crouzeix and P.-A. Raviart, *Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I*, Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge, **7** (1973), 33–75.

[5] L. Diening, C. Kreuzer, and R. Stevenson, *Instance optimality of the adaptive maximum strategy*, Found. Comput. Math., **16** (2016), 33–68.
[6] T. Gudi, *A new error analysis for discontinuous finite element methods for linear elliptic problems*, Math. Comp., **79** (2010), 2169–2189.
[7] C. Kreuzer and M. Schedensack, *Instance optimal Crouzeix-Raviart adaptive finite element methods for the Poisson and Stokes problems*, IMA J. Numer. Anal., **36** (2016), 593–617.

On Adaptive Spectral Galerkin Methods with Dynamic Marking

Marco Verani
(joint work with Claudio Canuto, Ricardo H. Nochetto and Rob Stevenson)

The convergence and optimality theory of adaptive Galerkin methods for linear elliptic PDEs is almost exclusively based on the Dörfler marking. This entails a fixed marking parameter and leads to a contraction constant bounded below away from zero. More precisely, the convergence is proven to be *linear*, i.e., a certain expression controlling the error (a norm, or a combination of norm and estimator) contracts with some fixed parameter $\rho < 1$ from one iteration to the next one, e.g., $\|u - u_{k+1}\| \leq \rho \|u - u_k\|$. This is typically achieved if the adaptation strategy is based on some form of *Dörfler marking* (or *bulk chasing*) with fixed parameter $\theta < 1$: assuming that $\sum_{i\in\mathcal{I}} \eta_i^2$ is some additive error estimator at iteration k, one identifies a minimal subset $\mathcal{I}' \subset \mathcal{I}$ such that $\sum_{i\in\mathcal{I}'} \eta_i^2 \geq \theta^2 \sum_{i\in\mathcal{I}} \eta_i^2$ and utilizes $\mathcal{I}'$ for the construction of the new discretization at iteration $k+1$. For wavelet or h-type fem discretizations, optimality is guaranteed by performing cautious successive adaptations, i.e., by choosing a moderate value of θ, say $0 < \theta \leq \theta_{\max} < 1$ [4]. This avoids the need of cleaning-up the discrete solution from time to time, by subjecting it to a *coarsening* stage.

On the other hand, the resulting contraction factor $\rho = \rho(\theta)$ turns out to be bounded from below by a positive constant, say $0 < \rho_{\min} \leq \rho < 1$ (related to the 'condition number' of the exact problem), regardless of the choice of θ. This is not restrictive for fixed-order methods [4, 3], but for spectral Galerkin methods it is a severe limitation which affects performance. Indeed, when a method of spectral type is used, one expects a fast (possibly, exponentially fast) decay of the discretization error for smooth solutions. In such a situation, a slow convergence of the iterations of the adaptive algorithm would spoil the overall performance of the method; from this perspective, it is useful to be able to make the contraction factor as close to 0 as desired.

Yet, linear convergence of the adaptive iterations is not enough to guarantee the optimality of the method. In order for the marking strategy to guarantee super-linear convergence, one needs to adopt a dynamic choice of Dörfler's parameter θ, which pushes its value towards 1 as the iterations proceed. We accomplish this requirement by equating the quantity $1 - \theta_k^2$ to some function of the dual norm of the residual r_k, which is monotonically increasing and vanishing at the origin. This defines our *dynamic marking strategy* introduced in [2]. The order of the root

at the origin dictates the exponent q in the super-linear convergence estimate of our adaptive algorithm.

In the following, using an idealized setting for the ease of presentation, we explain why linear convergence does not guarantee the optimality of the algorithm, and why a *super-linear* convergence is preferable. We refer to [2] for more details.

As customary in Nonlinear Approximation, we consider the *best N-term approximation error* $E_N(u)$ of the exact solution u, in a suitable norm, using combinations of at most N functions taken from a chosen basis. We prescribe a decay law of $E_N(u)$ as N increases, which classically for fixed-order approximations is *algebraic* and reads

$$\sup_N N^s E_N(u) < \infty, \tag{1}$$

for some positive s. However, for infinite-order methods such as spectral approximations an *exponential* law is relevant that reads

$$\sup_N \mathrm{e}^{\eta N^\alpha} E_N(u) < \infty \tag{2}$$

for some $\eta > 0$ and $\alpha \in (0, 1]$, where $\alpha < 1$ accommodates the inclusion of C^∞-functions that are not analytic. This defines corresponding algebraic and exponential *sparsity classes* for the exact solution u. These classes are related to Besov and Gevrey regularity of u respectively.

We now assume the ideal situation that at each iteration of our adaptive algorithm [1]

$$\|u - u_k\| \eqsim N_k^{-s} \qquad \text{or} \qquad \|u - u_k\| \eqsim e^{-\eta N_k^\alpha}, \tag{3}$$

where N_k is the cardinality of the discrete solution u_k, i.e., the dimension of the approximation space activated at iteration k. We assume in addition that the error decays linearly from one iteration to the next, i.e., it satisfies precisely

$$\|u - u_{k+1}\| = \rho \, \|u - u_k\|. \tag{4}$$

If u belongs to a sparsity class of algebraic type, then one easily gets $N_k \eqsim \rho^{-k/s}$, i.e., cardinalities grow exponentially fast and

$$\Delta N_k := N_{k+1} - N_k \eqsim N_k \eqsim \|u - u_k\|^{-1/s},$$

i.e., the increment of cardinality between consecutive iterations is proportional to the current cardinality as well as to the error raised to the power $-1/s$. The important message stemming from this ideal setting is that for a *practical* adaptive algorithm one should be able to derive the estimates $\|u - u_{k+1}\| \le \rho \, \|u - u_k\|$ and $\Delta N_k \lesssim \|u - u_k\|^{-1/s}$, because they yield, employing a geometric-series argument,

$$N_n = \sum_{k=0}^{n-1} \Delta N_k \lesssim \sum_{k=0}^{n-1} \|u - u_k\|^{-1/s} \le \|u - u_n\|^{-1/s} \sum_{k=0}^{n-1} \rho^{(n-k)/s} \lesssim \|u - u_n\|^{-1/s}.$$

[1] We write $A_k \lesssim B_k$ to indicate that A_k can be bounded by a multiple of B_k, independently of the iteration counter k and other parameters which A_k and B_k may depend on; $A_k \eqsim B_k$ means $A_k \lesssim B_k$ and $B_k \lesssim A_k$.

The performance of a practical adaptive algorithm is thus quasi-optimal.

If u belongs to a sparsity class of exponential type, instead, the situation changes radically. In fact, assuming (3) and (4), one has $e^{-\eta N_k^\alpha} \eqsim \rho^k$, and so

$$\lim_{k\to\infty} k^{-1/\alpha} N_k = \big(|\log\rho|/\eta\big)^{1/\alpha},$$

i.e., the cardinality N_k grows polynomially. One of the key points of [2] is the observation that if the convergence of the adaptive algorithm is super-linear, then one is back to the simpler case of exponential growth of cardinalities which is ameanable to a sharper performance analysis. To see this, let us assume a super-linear relation between consecutive errors:

$$\|u - u_{k+1}\| = \|u - u_k\|^q \tag{5}$$

for some $q > 1$. If additionally u_k satisfies (3), then one infers that $\mathrm{e}^{-\eta N_{k+1}^\alpha} \eqsim \mathrm{e}^{-\eta q N_k^\alpha}$, whence

$$\lim_{k\to\infty} \frac{\Delta N_k}{N_k} = q^{1/\alpha} - 1, \qquad \lim_{k\to\infty} \frac{|\log\|u-u_k\||^{1/\alpha}}{N_k} = \eta^{1/\alpha},$$

the latter being just a consequence of (3). This suggests that the geometric-series argument may be invoked again in the optimality analysis of the adaptive algorithm.

This ideal setting does not apply directly to a *practical* adaptive algorithm. Indeed, in [2] we prove estimates that are consistent with the preceding derivation to some extent, namely

$$\|u - u_{k+1}\| \le \|u - u_k\|^q, \qquad \Delta N_k \le Q|\log\|u-u_k\||^{1/\bar\alpha},$$

with constants $Q > 0$ and $\bar\alpha \in (0, \alpha]$. Invoking $\|u - u_n\| \le \|u - u_k\|^{q^{n-k}}$, we then realize that

$$N_n = \sum_{k=0}^{n-1} \Delta N_k \le Q \sum_{k=0}^{n-1} \big|\log\|u-u_k\|\big|^{1/\bar\alpha} \le \frac{Qq^{1/\bar\alpha}}{q^{1/\bar\alpha}-1} \big|\log\|u-u_n\|\big|^{1/\bar\alpha}.$$

Setting $\bar\eta := \big(\frac{Qq^{1/\bar\alpha}}{q^{1/\bar\alpha}-1}\big)^{-\bar\alpha}$, we deduce the estimate

$$\sup_n \mathrm{e}^{\bar\eta N_n^{\bar\alpha}} \|u - u_n\| \le 1,$$

which is similar to (2), albeit with different class parameters. The most important parameter is $\bar\alpha$. Its possible degradation relative to α is mainly caused by the fact that the residual, the only computable quantity accessible to our practical algorithm, belongs to a sparsity class with a main parameter generally smaller than that of the solution u. This perhaps unexpected property is typical of the exponential class and has been elucidated in [1].

References

[1] C. Canuto, R. H. Nochetto, and M. Verani. *Adaptive Fourier-Galerkin methods.* Math. Comp., 83(288):1645–1687, 2014.

[2] C. Canuto, R. H. Nochetto, R. Stevenson, and M. Verani. *Adaptive Spectral Galerkin Methods with Dynamic Marking.* arXiv:1511.00233. Accepted for publication on SINUM.

[3] J. M. Cascón, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. *Quasi-optimal convergence rate for an adaptive finite element method.* SIAM J. Numer. Anal., 46(5):2524–2550, 2008.

[4] R. Stevenson. *Optimality of a standard adaptive finite element method.* Found. Comput. Math., 7(2):245–269, 2007.

Stable broken H^1 and $H(\mathrm{div})$ polynomial extensions

Martin Vohralík
(joint work with Alexandre Ern)

Braess *et al.* showed in [1] that equilibrated flux a posteriori error estimates lead to local efficiency and polynomial-degree robustness. This means that the estimators give local lower bounds for the error, up to a generic constant independent of the polynomial degree of the approximate solution. These results apply to conforming finite element methods in two space dimensions and require solving local homogeneous Neumann hat-function-weighted residual problems posed on vertex-centered patches of elements via the mixed finite element method. The proof of the p-robustness relies on two key components: uniform stability with respect to the polynomial degree of the right inverse of the divergence operator of Costabel and McIntosh [2, Corollary 3.4], and uniform stability with respect to the polynomial degree of the right inverse of the normal trace of Demkowicz *et al.* [4, Theorem 7.1]. In our contribution [5], we extended p-robustness to any numerical method satisfying a couple of clearly identified assumptions, including various nonconforming, discontinuous Galerkin, and mixed finite elements. Here one is led to solve additionally local homogeneous Dirichlet conforming finite element problem on each vertex-centered patch, with a hat-function-weighted piecewise continuous datum. The results of [5] still only hold in two space dimensions, and their proof proceeds through the same stability arguments as in [1], i.e. those of [2, 4]. In particular, advantage of the two-dimensional setting is taken to formulate the local conforming Dirichlet problems using rotated Raviart–Thomas mixed finite elements. We present in this contribution the extension of these results to three space dimensions, with necessarily a different concept for the local Dirichlet problems. We do so in an abstract setting not necessarily linked to a posteriori error analysis.

Let a shape-regular patch of simplicial elements sharing the given vertex (we treat both interior patches and patches around a vertex on the boundary of a computational domain) be given, together with a p-degree polynomial r_F associated with each interior face F of the patch in the H^1 setting. In the $\boldsymbol{H}(\mathrm{div})$ setting, let the data be a p-degree polynomial r_F per face F and a p-degree polynomial r_K per element K of the patch. These data need to satisfy some compatibility

conditions. We want to extend them to a piecewise polynomial of degree p over the patch such that the jumps over the interior faces are prescribed by r_F in the H^1 setting. In the $\boldsymbol{H}(\mathrm{div})$ setting, we want to find a piecewise Raviart–Thomas–Nédélec function of degree p such that its normal component jumps are given by r_F and its piecewise divergence is given by r_K. The goal is to show that these extensions are, up to a constant only depending on the patch shape regularity, subordinate in the broken energy norm to the best possible (minimizing the broken energy norm) extensions in the whole broken H^1 or $\boldsymbol{H}(\mathrm{div})$ spaces. This result can also equivalently be reformulated as the statement that 1) best-approximation of a discontinuous piecewise scalar polynomial is as good by continuous piecewise polynomials of the same degree as it is by all H^1_0 Sobolev functions on the patch; 2) best-approximation of a normal-trace-discontinuous piecewise vector polynomial is as good by $\boldsymbol{H}(\mathrm{div})$-conforming Raviart–Thomas–Nédélec functions of the same degree as it is by all $\boldsymbol{H}(\mathrm{div})$ functions on the patch.

In contrast to [1], where the work with dual norms was essential, our proofs only work with the (broken) energy norms. They are constructive and also indicate a possible practical replacement of the local Neumann mixed finite element / local Dirichlet conforming finite element problems by a single explicit run through the patch, with possibly a solve on each element only. The key ingredients on a single element are still the right inverse of the divergence [2, Corollary 3.4] and the right inverse of the normal trace [4, Theorem 7.1] in the $\boldsymbol{H}(\mathrm{div})$ setting, but this becomes the right inverse of the trace of Demkowicz *et al.* [3, Theorem 6.1] in the H^1 setting. Gluing the elemental contributions together turns out to be a rather involved ingredient of the proofs in three space dimensions, since the two-dimensional argument of turning around a vertex can no longer be invoked. To achieve a suitable enumeration of the mesh cells composing the vertex-centered patch in three dimensions, we crucially rely on the notion of shelling for polytopes, as presented in, e.g., [7]. Finally, parts of the proofs additionally rely on a graph result concerning the existence of two- and three-coloring of the vertices in the patch or a suitable sub-patch. All the details can be found in [6].

References

[1] D. Braess, V. Pillwein, and J. Schöberl, *Equilibrated residual error estimates are p-robust*, Comput. Methods Appl. Mech. Engrg. **198** (2009), 1189–1197.

[2] M. Costabel and A. McIntosh, *On Bogovskiĭ and regularized Poincaré integral operators for de Rham complexes on Lipschitz domains*, Math. Z. **265** (2010), 297–320.

[3] L. Demkowicz, J. Gopalakrishnan, and J. Schöberl, *Polynomial extension operators. Part I*, SIAM J. Numer. Anal. **46** (2008), 3006–3031.

[4] ———, *Polynomial extension operators. Part III*, Math. Comp. **81** (2012), 1289–1326.

[5] A. Ern and M. Vohralík, *Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations*, SIAM J. Numer. Anal. **53** (2015), 1058–1081.

[6] A. Ern and M. Vohralík, *Stable broken H^1 and $\boldsymbol{H}(\mathrm{div})$ polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions*, to be sumbitted, 2016.

[7] G. M. Ziegler, *Lectures on polytopes*, vol. 152 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1995.

On approximation classes of adaptive methods

Gantumur Tsogtgerel

A crucial notion in the theory of adaptive finite element methods is that of *approximation classes*, which we discuss here in a simple but very paradigmatic setting. Given a polygonal domain $\Omega \in \mathbb{R}^2$, a conforming triangulation P_0 of Ω, and a number $s > 0$, we say that a function $u \in H^1(\Omega)$ belongs to the approximation class $\mathscr{A}^s$ if

$$\min_{\{P \in \mathscr{P} : \#P \leq N\}} \inf_{v \in S_P} \|u - v\|_{H^1} \leq cN^{-s},$$

for all $N \geq \#P_0$ and for some constant c, where $\mathscr{P}$ is the set of conforming triangulations of Ω that are obtained by a sequence of newest vertex bisections from P_0, and S_P is the space of continuous piecewise affine functions subordinate to the triangulation P. In a certain sense, $\mathscr{A}^s$ is the collection of all functions that can be approximated with the error $\sim N^{-s}$ given the budget of N triangles.

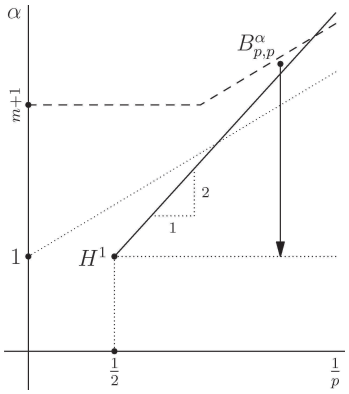
When $u \in \mathscr{A}^s$ is the solution of a boundary value problem, a natural question is if the convergence rate N^{-s} can be achieved by any practical algorithm, and it was answered in the seminal works [1, 6]: These papers established that the convergence rates of certain adaptive finite element methods are optimal, in the sense that if $u \in \mathscr{A}^s$ for some $s > 0$, then the method converges with the rate not slower than s. This approach was greatly clarified and improved upon in [3].

Having established that the smallest approximation class $\mathscr{A}^s$ in which the solution u belongs to essentially determines how fast adaptive finite element methods converge, the next issue is to determine how large these classes are and if the solution to a typical boundary value problem would belong to an $\mathscr{A}^s$ with large s. A first step towards addressing this issue is to characterize the approximation classes in terms of classical smoothness spaces, and the main work in this direction so far appeared is [2], which, upon tailoring to our situation and a slight simplification, tells that $B^\alpha_{p,p} \subset \mathscr{A}^s \subset B^\sigma_{p,p}$ for $\frac{2}{p} = \sigma < 1 + \frac{1}{p}$ and $\sigma < \alpha < \max\{2, 1 + \frac{1}{p}\}$ with $s = \frac{\alpha - 1}{2}$. Here $B^\alpha_{p,q}$ are the standard Besov spaces defined on Ω. This result has recently been generalized to higher order Lagrange finite elements by [5]. In particular, they show that the direct embedding $B^\alpha_{p,p} \subset \mathscr{A}^s$ holds in the larger range $\sigma < \alpha < m + \max\{1, \frac{1}{p}\}$, where m is the polynomial degree of the finite element space, see Figure 1(a). However, the restriction $\sigma < 1 + \frac{1}{p}$ on the inverse embedding $\mathscr{A}^s \subset B^\sigma_{p,p}$ cannot be removed, since for instance, any finite element function whose derivative is discontinuous cannot be in $B^\sigma_{p,p}$ if $\sigma \geq 1 + \frac{1}{p}$ and $p < \infty$. To get around this problem, Gaspoz and Morin proposed to replace the Besov space $B^\sigma_{p,p}$ by the approximation space $A^\sigma_{p,p}$ associated to uniform refinements. We call the spaces $A^\sigma_{p,p}$ *multilevel approximation spaces*. For the purposes of this abstract, and roughly speaking, the space $A^\sigma_{p,p}$ is the collection of functions $u \in L^p$ for which

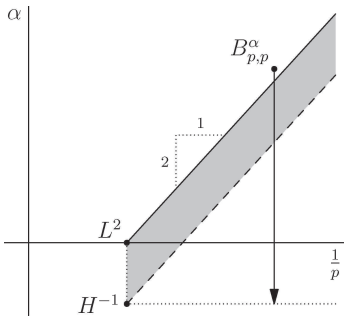
$$\inf_{v \in S_{P_k}} \|u - v\|_{L^p} \leq ch_k^\sigma,$$

where $\{P_k\} \subset \mathscr{P}$ is a sequence of triangulations such that P_{k+1} is the uniform refinement of P_k, and h_k is the diameter of a typical triangle in P_k. Note for instance that finite element functions are in every $A^\sigma_{p,p}$. With the multilevel approximation spaces at hand, the inverse embedding $\mathscr{A}^s \subset A^\sigma_{p,p}$ is recovered for all $\sigma = \frac{2}{p} > 0$.

In the recent paper [4], we prove the direct embedding $A^\alpha_{p,p} \subset \mathscr{A}^s$, so that the existing situation $B^\alpha_{p,p} \subset \mathscr{A}^s \subset A^\sigma_{p,p}$ is improved to $A^\alpha_{p,p} \subset \mathscr{A}^s \subset A^\sigma_{p,p}$. It is a genuine improvement, since $A^\alpha_{p,p}(\Omega) \supsetneq B^\alpha_{p,p}(\Omega)$ for $\alpha \geq 1 + \frac{1}{p}$. Moreover, as one stays entirely within an approximation theory framework, one can argue that the link between $\mathscr{A}^s$ and $A^\alpha_{p,p}$ is more natural than the link between $\mathscr{A}^s$ and $B^\alpha_{p,p}$. Once the link between $\mathscr{A}^s$ and $A^\alpha_{p,p}$ has been established, one can then invoke the well known relationships between $A^\alpha_{p,p}$ and $B^\alpha_{p,p}$. It seems that this two step process offers more insight into the underlying phenomenon.



(A) If the space $B^\alpha_{p,p}$ is located above the solid line and below the dashed line, then $B^\alpha_{p,p} \subset \mathscr{A}^s$ with $s = \frac{\alpha-1}{2}$. The inverse embeddings $\mathscr{A}^{s+\varepsilon} \subset B^\alpha_{p,p}$ hold on the solid line and below the (slanted) dotted line.



(B) If the space $B^\alpha_{p,p}$ is located above or on the solid line, and if $u \in \mathscr{A}^s$ and $\Delta u \in B^\alpha_{p,p}$ with $s = \frac{\alpha+1}{2}$, then $u \in \mathscr{A}^s_*$. It is as if the approximation of Δu is taking place in H^{-1}, with the proviso that the shaded area is excluded from all considerations.

FIGURE 1. Illustration of various embeddings. The point $(\frac{1}{p}, \alpha)$ represents the Besov space $B^\alpha_{p,p}$.

The approximation classes $\mathscr{A}^s$ are associated to measuring the error of an approximation in the H^1-norm. Of course, this can be generalized to other function space norms, such as L^p and $B^\alpha_{p,p}$, which we consider in [4]. However, we do not stop there, and consider more general approximation classes corresponding to ways of measuring the error between a general function u and a discrete function $v \in S_P$ by a quantity $\rho(u, v, P)$ that may depend on the triangulation P and is required to make sense merely for discrete functions $v \in S_P$. An example of such

an error measure is

$$\rho(u,v,P)=\left(\|u-v\|_{H^1}^2+\sum_{\tau\in P}(\operatorname{diam}\tau)^2\|f-\Pi_\tau f\|_{L^2(\tau)}^2\right)^{\frac{1}{2}},$$

where $f=\Delta u$, and $\Pi_\tau : L^2(\tau)\to\mathbb{P}_d$ is the $L^2(\tau)$-orthogonal projection onto $\mathbb{P}_d$, the space of polynomials of degree not exceeding d. It has been shown in [3] that if the solution u of the boundary value problem

$$\Delta u=f \quad \text{in } \Omega \qquad \text{and} \qquad u|_\Omega=0,$$

satisfies

$$\min_{\{P\in\mathscr{P}:\#P\le N\}}\inf_{v\in S_P}\rho(u,v,P)\le cN^{-s},$$

for all $N\ge\#P_0$ and for some constants c and $s>0$, then a typical adaptive finite element method converges with the rate not slower than s. Moreover, there are good reasons to consider that the approximation classes $\mathscr{A}_*^s$ defined by the preceding equation are more attuned to certain practical adaptive finite element methods than the standard approximation classes. Obviously, we have $\mathscr{A}_*^s\subset\mathscr{A}^s$ but we cannot expect the inclusion $\mathscr{A}^s\subset\mathscr{A}_*^s$ to hold in general. In [3], an effective characterization of $\mathscr{A}_*^s$ was announced as an important pending issue.

The general results in [4] imply a characterization of $\mathscr{A}_*^s$ in terms of memberships of u and $f=\Delta u$ into suitable approximation spaces, which in turn are related to Besov spaces. For instance, we show that if $u\in\mathscr{A}^s$ and $f\in B_{p,p}^\alpha$ with $\frac{\alpha}{2}\ge\frac{1}{p}-\frac{1}{2}$ and $s=\frac{\alpha+1}{2}$, then $u\in\mathscr{A}_*^s$, see Figure 1(b). Note that the approximation rate $s=\frac{\alpha+1}{2}$ is as if we were approximating f in the H^{-1}-norm, which is illustrated by the arrow downwards. However, the parameters must satisfy $\frac{\alpha}{2}\ge\frac{1}{p}-\frac{1}{2}$ (above or on the solid line), which is more restrictive compared to $\frac{\alpha+1}{2}>\frac{1}{p}-\frac{1}{2}$ (above the dashed line), the latter being the condition we would expect if the approximation was indeed taking place in H^{-1}. This situation cannot be improved in the sense that if $\frac{\alpha}{2}<\frac{1}{p}-\frac{1}{2}$ then $B_{p,p}^\alpha\not\subset L^2$, hence the quantity ρ would be infinite in general for $f\in B_{p,p}^\alpha$.

References

[1] P. Binev, W. Dahmen, and R. DeVore, *Adaptive finite element methods with convergence rates*, Numer. Math. **97** (2004), 219–268.

[2] P. Binev, W. Dahmen, R. DeVore, and P. Petrushev, *Approximation classes for adaptive methods*, Serdica Math. J., **28** (2002), 391–416.

[3] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal., **46** (2008), 2524–2550.

[4] T. Gantumur, *Convergence Rates of Adaptive Methods, Besov Spaces, and Multilevel Approximation*, Found. Comp. Math., (2016), 1–40.

[5] F. D. Gaspoz and P. Morin, *Approximation classes for adaptive higher order finite element approximation*, Math. Comp., **83** (2014), 2127–2160.

[6] R. P. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., **7** (2007), 245–269.

Adaptive discretization in Banach spaces with the nonlinear Petrov–Galerkin method

Kristoffer G. van der Zee
(joint work with Ignacio Muga)

This short note reviews some initial results and open problems for the adaptive discretization of linear problems in Banach-space settings with the *nonlinear Petrov–Galerkin* method. This is relevant to, for example, PDEs with rough data or nonsmooth solutions.

1. The nonlinear Petrov–Galerkin method

In the setting of Banach spaces, consider the linear problem:

$$\text{Find } u \in \mathbb{U}: \quad \big\langle Bu, v\big\rangle_{\mathbb{V}^*,\mathbb{V}} = \big\langle f, v\big\rangle_{\mathbb{V}^*,\mathbb{V}} \quad \forall v \in \mathbb{V}\,,$$

where $B \in \mathcal{L}(\mathbb{U};\mathbb{V}^*)$ is a linear continuous operator, $f \in \mathbb{V}^*$, $\mathbb{U}$ is a Banach space and $\mathbb{V}$ is a strictly-convex reflexive Banach space with strictly-convex dual $\mathbb{V}^*$.[1] In typical examples, $\mathbb{U}$ and $\mathbb{V}$ are distinct reflexive Sobolev spaces (e.g., L^p or $W^{1,p}$, $1 < p < \infty$).

To ensure well-posedness of the above problem (by the Banach closed-range theorem and open-mapping theorem), we assume that B is *bounded below*, i.e.,

$$\|Bw\|_{\mathbb{V}^*} \geq \gamma \|w\|_{\mathbb{U}} \quad \forall w \in \mathbb{U}\,,$$

for some $\gamma > 0$, and that its dual operator has a trivial kernel, i.e., $\mathrm{Ker}(B^*) = \{0\}$.

Given a discrete subspace $\mathbb{U}_n \subset \mathbb{U}$, the nonlinear Petrov–Galerkin (NPG) discretization, proposed in [11], is given by:

$$\text{Find } u_n \in \mathbb{U}_n: \quad \Big\langle Bw_n\,,\, J_{\mathbb{V}_m}^{-1}\big(f - Bu_n\big)\Big\rangle_{\mathbb{V}^*,\mathbb{V}} = 0 \quad \forall w_n \in \mathbb{U}_n\,, \tag{1}$$

where $\mathbb{V}_m \subseteq \mathbb{V}$ is a subspace, and $J_{\mathbb{V}_m} : \mathbb{V}_m \to \mathbb{V}_m^*$ is the bijective *duality map* for $\mathbb{V}_m$. The duality map for a Banach space $\mathbb{X}$ with strictly convex $\mathbb{X}^*$ is a (generally *nonlinear*) map defined by $\|J_{\mathbb{X}}(x)\|_{\mathbb{X}^*} = \|x\|_{\mathbb{X}}$ and $\big\langle J_{\mathbb{X}}(x), x\big\rangle_{\mathbb{X}^*,\mathbb{X}} = \|x\|_{\mathbb{X}}^2$ for all $x \in \mathbb{X}$. It is a linear map only in *Hilbert* spaces, in which case it coincides with the (linear) *Riesz map*.

2. Connections and tractability

The NPG method builds on recent ideas in *Petrov–Galerkin* and *residual minimization* methods: It extends the seminal optimal Petrov–Galerkin methodology of Demkowicz and Gopalakrishnan [7] to Banach spaces; it provides for a (monotone) nonlinear extension of the corresponding mixed formulation, cf. Dahmen et al [6]; and it extends the L^p residual-minimization method of Guermond [10] to arbitrary dual Banach spaces.

[1] A normed space $\mathbb{X}$ is strictly convex iff $\|\theta x + (1-\theta)y\|_{\mathbb{X}} < 1$ for all $\theta \in (0,1)$, where $x \neq y$, and $\|x\|_{\mathbb{X}} = \|y\|_{\mathbb{X}} = 1$.

The case $\mathbb{V}_m = \mathbb{V}$ corresponds to the *ideal* setting, and delivers the *best* approximation in $\mathbb{U}_n$ as measured in the norm $\|B(\cdot)\|_{\mathbb{V}^*}$. The NPG method becomes tractable however for *discrete* $\mathbb{V}_m$. In that case, the (discrete) well-posedness and the quasi-optimality of the NPG method are guaranteed under the following *Fortin compatibility condition* for the pair $(\mathbb{U}_n, \mathbb{V}_m)$: There exists a bounded projection $\Pi : \mathbb{V} \to \mathbb{V}_m$ such that

$$\big\langle Bw_n, v - \Pi(v)\big\rangle_{\mathbb{V}^*,\mathbb{V}} = 0 \quad \forall w_n \in \mathbb{U}_n\,, \tag{2}$$

for all $v \in \mathbb{V}$. Hence, the NPG method needs, in particular, that $\dim \mathbb{V}_m \geq \dim \mathbb{U}_n$.

Another generalization provided by the NPG method is the following: Upon choosing a special compatible pair $(\mathbb{U}_n, \mathbb{V}_m)$ for which $\dim \mathbb{V}_m = \dim \mathbb{U}_n$, one recovers the *standard* (linear) Petrov–Galerkin method! See [11] for more details.

3. Adaptive discretization: Challenges

The NPG method is a general discretization method which aims to deliver, with guaranteed stability, near-best approximations to problems in functional settings *beyond* Hilbert spaces, thereby opening up the possibility to address a range of new applications. From a theoretical point of view, further research is required, e.g., to clarify how to tackle the Fortin condition (which is likely to be problem dependent), and to devise good solvers for the nonlinear system contained in (1).

Significant research is also needed to develop and understand *adaptive discretization strategies* with the NPG method. The key ingredient in the efficient adaptive generation of (nested) subspaces $\{\mathbb{U}_n\}$ are *a posteriori error estimates*. Coincidentally, the NPG method has a natural a posteriori error analysis, which can be employed, in standard manners, to generate adaptively-refined subspaces (as in, e.g., adaptive finite element methods):

Theorem (A posteriori error analysis) *Assume that the Fortin condition* (2) *holds. Let $u_n \in \mathbb{U}_n$ denote the (well-defined) approximation given by the NPG method* (1)*, and let $r_m := J_{\mathbb{V}_m}^{-1}(f - Bu_n) \in \mathbb{V}_m$ denote its discrete residual representation. Then*

$$\frac{1}{\|B\|}\|r_m\|_{\mathbb{V}} \leq \|u - u_n\|_{\mathbb{U}} \leq \frac{\|\Pi\|}{\gamma}\|r_m\|_{\mathbb{V}} + \frac{1}{\gamma}\operatorname{osc}(f)\,,$$

where the data oscillation is defined by $\operatorname{osc}(f) := \big\|f \circ (I - \Pi)\big\|_{\mathbb{V}^*}$. □

The proof of this result can be found [11]. It is perhaps striking that the a posteriori result is exactly the same as the one in Hilbert-space settings [4, 1]; this is not true for the *a priori* error analysis, which simplifies in Hilbert spaces because of linearity and Kato's identity for projectors.

The above theorem is a first result that is needed to understand adaptivity. However, since the NPG method is a nonstandard, nonlinear discretization methodology, the established theories for the analysis of adaptive methods do not seem to apply. There are many open problems that need to be addressed; we close this note by mentioning some exciting questions:

- Can one guarantee that a sequence of adaptively-computed approximations $\{u_n\}$ *converges* to the solution u, and if so, is this convergence *optimal* with respect to a suitable adaptive-approximation class (cf. [12, 3, 2])?
- Is it possible to avoid the (intractable) computation of $\mathrm{osc}(f)$, e.g., by resorting to suitable (adaptive) data approximation (cf. [5]) or adaptive computation of $\mathbb{V}_m$ (cf. [6])?
- Can one design optimally-convergent *goal-oriented* adaptive NPG algorithms that aim to efficiently resolve output-quantities of interest (cf. [9])?
- How should one include *linearization errors* which results from the solution of the nonlinear discrete system (cf. [8])?

References

[1] C. Carstensen, L. Demkowicz, J. Gopalakrishnan, *A posteriori error control for DPG methods*, SIAM J. Numer. Anal., 52 (2014), pp. 1335–1353.

[2] C. Carstensen, M. Feischl, M. Page, D. Praetorius, *Axioms of adaptivity*, Comput. Math. Appl., 67 (2014), pp. 1195–1253.

[3] J. M. Cascon, C. Kreuzer, R. H. Nochetto, K. G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal., 46 (2008), pp. 2524–2550.

[4] A. Cohen, W. Dahmen, G. Welper, *Adaptivity and variational stabilization for convection-diffusion equations*, M2AN Math. Model. Numer. Anal., 46 (2012), pp. 1247–1273.

[5] A. Cohen, R. DeVore, R. H. Nochetto, *Convergence rates of AFEM with H^{-1} data*, Found. Comput. Math., 12 (2012), pp. 671–718.

[6] W. Dahmen, C. Huang, C. Schwab, G. Welper, *Adaptive Petrov–Galerkin methods for first order transport equations*, SIAM J. Numer. Anal., 50 (2012), pp. 2420–2445.

[7] L. Demkowicz, J. Gopalakrishnan, *An overview of the Discontinuous Petrov Galerkin method*, in Recent Developments in Discontinuous Galerkin Finite Element Methods for Partial Differential Equations: 2012 John H Barrett Memorial Lectures, X. Feng, O. Karakashian and Y. Xing (eds.), pp. 149–180, Springer, (2014)

[8] A. Ern, M. Vohralík, *Adaptive inexact Newton methods: A posteriori error control and speedup of calculations*, SIAM News, 46 (2013).

[9] M. Feischl, D. Praetorius, K. G. van der Zee, *An abstract analysis of optimal goal-oriented adaptivity*, SIAM J. Numer. Anal., 54 (2016), pp. 1523–1448.

[10] J. L. Guermond, *A finite element technique for solving first-order PDEs in L^p*, SIAM J. Numer. Anal., 42 (2004), pp. 714–737

[11] I. Muga, K. G. van der Zee, *Discretization of linear problems in Banach spaces: Residual minimization, nonlinear Petrov–Galerkin, and monotone mixed methods*, arXiv:1511.04400 [math.NA].

[12] R. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., 7 (2007), pp. 245–269.

Local error analysis of the boundary element method for polygonal/polyhedral domains

Jens Markus Melenk
(joint work with Markus Faustmann)

Energy projection methods such as many finite element methods (FEM) or boundary element methods (BEM) minimize the error in some global norm. Locally, i.e., on fixed subdomains, the convergence may be better, especially, if the sought solution is smoother there. In the context of the FEM, the numerical analyses of this setting goes back at least to [4]; subsequent generalizations include [9, 1] and the references there. These estimates take the following form: The error in the *local* H^1-norm is estimated by a *local* best approximation error in H^1 and an error in a weaker norm (L^2-norm or negative norm). Significantly fewer works are concerned with local error estimates for the BEM. The case of smooth curves is studied in [5, 8], and [7] analyze screen domains. For the case of piecewise smooth geometries, sharp local error estimates that exploit the maximum local regularity of the solution do not seem to be available. Additionally, the local estimates in [5, 8, 7] focus on the energy norm and do not provide sharp estimates in stronger norms. For the case of Symm's integral equation, Theorem 1 below presents sharp estimates in the local L^2-norm, which is stronger than the local energy norm. Our numerics illustrate the sharpness of these *a priori* bounds. Full proofs are given in the forthcoming paper [3].

Main results. We study polygonal ($d = 2$) or polyhedral ($d = 3$) Lipschitz domain $\Omega \subset \mathbb{R}^d$ with boundary $\Gamma := \partial\Omega$. An important fact is that for such domains an elliptic shift theorem for the interior *and* exterior Dirichlet problem holds in an extended range, which we denote by a parameter $\alpha_D \in (0, 1/2)$. Specifically, let the ball $B_{R_\Omega}(0)$ be so large that $\overline{\Omega} \subset B_{R_\Omega}(0)$. Then we require for every $0 < \epsilon \le \alpha_D$ and both $\widehat{\Omega} = \Omega$ and $\widehat{\Omega} = B_{R_\Omega}(0) \setminus \overline{\Omega}$ the validity of the shift theorem

$$\|u\|_{H^{3/2+\epsilon}(\widehat{\Omega})} \le C_\varepsilon \|f\|_{H^{-1/2+\epsilon}(\widehat{\Omega})}, \tag{1}$$

where u denotes the solution of $-\Delta u = f$ in $\widehat{\Omega}$, $u = 0$ on $\partial\widehat{\Omega}$. We mention that for $d = 2$ the parameter α_D can be determined explicitly terms of the largest interior and exterior angles of the polygon, viz., any α_D with $1/2 < \alpha_D + 1/2 < \min_i\{\pi/\omega_i, \pi/(2\pi - \omega_i)\}$, where the ω_i are the internal angles of the polygon.

We consider Symm's integral equation of finding $\phi \in H^{-1/2}(\Gamma)$ such that $V\phi = g$, where $g \in H^{1/2}(\Gamma)$ is given and $V : H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$ denotes the single-layer boundary integral operator (see [6, p. 119]).

On a regular quasi-uniform triangulation $\mathcal{T}_h$ of Γ with mesh size h, we consider a lowest order Galerkin discretization in the space of piecewise constant functions on $\mathcal{T}_h$, which we denote by $S^{0,0}(\mathcal{T}_h)$.

Theorem 1. *Let* (1) *hold. Let* $\phi \in H^{-1/2}(\Gamma)$ *and* $\phi_h \in S^{0,0}(\mathcal{T}_h)$ *satisfy the Galerkin orthogonality condition*

$$\langle V(\phi - \phi_h), \psi_h \rangle = 0 \qquad \forall \psi_h \in S^{0,0}(\mathcal{T}_h).$$

Let Γ_0, $\widehat{\Gamma}$ be open subsets of Γ with $\Gamma_0 \subset \widehat{\Gamma} \subsetneq \Gamma$ and $R := \operatorname{dist}(\Gamma_0, \partial\widehat{\Gamma}) > 0$. Assume $\phi \in L^2(\widehat{\Gamma})$. Then,

$$\|\phi - \phi_h\|_{L^2(\Gamma_0)} \leq C \left(\inf_{\chi_h \in S^{0,0}(\mathcal{T}_h)} \|\phi - \chi_h\|_{L^2(\widehat{\Gamma})} + \|\phi - \phi_h\|_{H^{-1-\alpha_D}(\Gamma)} \right),$$

where $C > 0$ depends only on Γ, Γ_0, d, R, and the γ-shape regularity of $\mathcal{T}_h$. The norm $\|\cdot\|_{H^{-1-\alpha_D}(\Gamma)}$ is the dual norm of $\|\cdot\|_{H^{1+\alpha_D}(\Gamma)}$, which is defined as $\|w\|_{H^{1+\alpha_D}(\Gamma)} := \inf\{\|W\|_{H^{3/2+\alpha_D}(\Omega)} : W|_\Gamma = w\}$.

The key step of the local error analysis of the FEM [4, 9, 1] is a Caccioppoli-type inequality for discrete harmonic functions. The corresponding estimate for the BEM is likewise key and stated in the following proposition:

Proposition 1 ([2, Lemma 3.9])**.** *Let $\zeta_h \in S^{0,0}(\mathcal{T}_h)$ and define the potential $u := \widetilde{V}\zeta_h \in H^1_{\ell oc}(\mathbb{R}^d)$. Let $B_1 \subset B_2$ be two concentric balls with radii $R_1 < R_2$. Then there is $C > 0$ (depending only on R_1, R_2, and Γ) such that for h sufficiently small the following holds: If $\langle V\zeta_h, \psi_h\rangle = 0$ for all $\psi_h \in S^{0,0}(\mathcal{T}_h)$ with $\operatorname{supp}\psi_h \subset B_2$, then*

$$\|\nabla u\|_{L^2(B_1)} \leq C \left[h\|\nabla u\|_{L^2(B_2)} + \|u\|_{L^2(B_2)} \right].$$

By assuming the solution ϕ of $V\phi = g$ to be in the Sobolev space $H^{-1/2+\alpha}(\Gamma)$, $\alpha > 0$, and to have locally even better Sobolev regularity H^β, $\beta \in [0,1]$, we can be explicit in the local convergence rates:

Corollary 1. *Let the assumptions of Theorem 1 be fulfilled. Let $\widetilde{\Gamma} \subset \Gamma$ be a subset with $\widehat{\Gamma} \subsetneq \widetilde{\Gamma}$ and $\operatorname{dist}(\widehat{\Gamma}, \partial\widetilde{\Gamma}) = R > 0$. Additionally, assume $\phi \in H^{-1/2+\alpha}(\Gamma) \cap H^\beta(\widetilde{\Gamma})$ with $\alpha \geq 0$, $\beta \in [0,1]$. Then, with a constant $C > 0$ depending only on $\Gamma, \widetilde{\Gamma}$, d, R, α, β, and the γ-shape regularity of $\mathcal{T}_h$,*

$$\|\phi - \phi_h\|_{L^2(\Gamma_0)} \leq C h^{\min\{1/2+\alpha+\alpha_D, \beta\}}.$$

The error term $\|\phi - \phi_h\|_{H^{-1-\alpha_D}(\Gamma)}$ in Theorem 1 is treated with a duality argument that utilizes the maximum amount of regularity provided by the shift theorem given in (1). Therefore, one does not expect that replacing the contribution $\|\phi - \phi_h\|_{H^{-1-\alpha_D}(\Gamma)}$ in Theorem 1 with an even weaker norm (which is possible) leads to improvements in the convergence rate. Indeed, the numerics below show that our estimates are sharp.

Numerical Example. In order to underline that our local estimates are optimal, we study the behavior of the local error on the L-shaped domain in the $L^2(\Gamma_0)$-norm and the $H^{-1/2}(\Gamma_0)$-norm, computed by $\|\phi\|^2_{H^{-1/2}(\Gamma_0)} \sim \langle V(\chi_0\phi), \chi_0\phi\rangle$, where χ_0 is the characteristic function for $\Gamma_0 \subset \Gamma$. Here, Γ_0 is fixed and a union of elements away from the reentrant corner at the origin. For the L-shaped domain, we may choose $\alpha_D = 1/6 - \epsilon$ for any $\epsilon > 0$. We prescribe the solution $u(r,\theta) = r^\alpha \cos(\alpha\theta)$ of Poisson's equation in polar coordinates. Then, the normal derivative $\phi = \partial_n u$ solves $V\phi = (K + 1/2)\gamma_0 u$ with the double layer operator K. We have $\phi \in H^{-1/2+\alpha}(\Gamma) \cap H^1(\Gamma_0)$, and with the choice $\alpha = 1/6$, Corollary 1 leads to an expected local rate of $O(h^{\alpha+1/2+\alpha_D}) = O(h^{5/6-\epsilon})$.

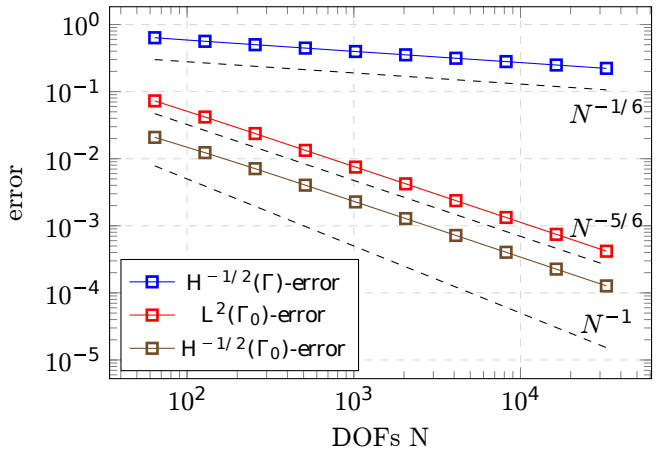


FIGURE 1. Local and global convergence of Galerkin-BEM for Symm's equation, L-shaped domain, $\alpha = 1/6$.

Figure 1 shows that the theoretical local rate is indeed obtained in the $L^2(\Gamma_0)$-norm. In the weaker $H^{-1/2}(\Gamma_0)$-norm we observe no improvement of the rate. This is a reflection of the fact that the global error term $\|\phi - \phi_h\|_{H^{-1-\alpha_D}(\Gamma)}$ in Theorem 1 is the limiting factor for the convergence in both norms.

REFERENCES

[1] A. Demlow, J. Guzmán, A.H. Schatz. *Local energy estimates for the finite element method on sharply varying grids*, Math. Comp., **80** (2011), no.273, 1–9.

[2] M. Faustmann, J.M. Melenk, D. Praetorius, *Existence of $\mathcal{H}$-matrix approximants to the inverse of BEM matrices: the simple- layer operator*, Math. Comp. **85** (2016), no.297, 119–152.

[3] M. Faustmann, J.M. Melenk, *Local convergence of the boundary element method on polygonal and polyhedral domains (in preparation)*, (2016).

[4] J.A. Nitsche, A.H. Schatz, *Interior estimates for Ritz-Galerkin methods*, Math. Comp. **28** (1974), 937–958.

[5] J. Saranen, *Local error estimates for some Petrov-Galerkin methods applied to strongly elliptic equations on curves*, Math. Comp. **48** (1987), no.178, 485–502.

[6] O. Steinbach. *Numerical Approximation Methods for Elliptic Boundary Value Problems: Finite and Boundary Elements*. Springer Verlag, 2008.

[7] E.P. Stephan, T. Tran, *Localization and post processing for the Galerkin bound- ary element method applied to three-dimensional screen problems*, J. Integral Equations Appl. **8** (1996), no. 4, 457–481.

[8] T. Tran, *The K-operator and the Galerkin method for strongly elliptic equations on smooth curves: local estimates*, Math. Comp. **64** (1995), no. 210, 501–513.

[9] L. Wahlbin, *Local behavior in finite element methods*, Handbook of numerical analysis. Volume II: Finite element methods (Part 1), North Holland (1991), 353–522.

Reporter: Rob Stevenson

Participants

Prof. Dr. Peter G. Binev
Department of Mathematics
University of South Carolina
Columbia, SC 29208
UNITED STATES

Prof. Dr. Zhiqiang Cai
Department of Mathematics
Purdue University
West Lafayette, IN 47907-1395
UNITED STATES

Prof. Dr. Claudio Canuto
Dipartimento di Matematica
Politecnico di Torino
Corso Duca degli Abruzzi, 24
10129 Torino
ITALY

Prof. Dr. Carsten Carstensen
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
GERMANY

Prof. Dr. Wolfgang Dahmen
Institut für Geometrie und
Praktische Mathematik
RWTH Aachen
Templergraben 55
52056 Aachen
GERMANY

Prof. Dr. Alan Demlow
Department of Mathematics
Texas A & M University
College Station, TX 77843-3368
UNITED STATES

Prof. Dr. Lars Diening
Fachbereich Mathematik/Informatik
Universität Osnabrück
Albrechtstrasse 28 A
49076 Osnabrück
GERMANY

Dr. Michael Feischl
School of Mathematics and Statistics
University of New South Wales
Sydney NSW 2033
AUSTRALIA

Dr. Dietmar Gallistl
Institut für Angewandte
und Numerische Mathematik
Karlsruher Institut für Technologie
(KIT)
76128 Karlsruhe
GERMANY

Dr. Joscha Gedicke
Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Dr. Emmanuil H. Georgoulis
Department of Mathematics
University of Leicester
University Road
Leicester LE1 7RH
UNITED KINGDOM

Friederike Hellwig
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10117 Berlin
GERMANY

Prof. Dr. Jun Hu
School of Mathematical Sciences
Peking University
No. 5 Yiheyuan Road
Beijing 100 871
CHINA

Dr. Omar Lakkis
Department of Mathematics
University of Sussex
Falmer Campus
Brighton BN1 9RH
UNITED KINGDOM

Prof. Dr. Jens M. Melenk
Institut für Analysis und Scientific Computing
Technische Universität Wien
Wiedner Hauptstrasse 8 - 10
1040 Wien
AUSTRIA

Prof. Dr. Neela Nataraj
Department of Mathematics
Indian Institute of Technology Bombay
Powai, Mumbai 400 076
INDIA

Prof. Dr. Serge Nicaise
ISTV, LAMAV
Université de Valenciennes
Le Mont Houy
59313 Valenciennes Cedex 9
FRANCE

Prof. Dr. Daniel Peterseim
Institut für Numerische Simulation
Rheinische
Friedrich-Wilhelms-Universität Bonn
Wegelerstrasse 6
53115 Bonn
GERMANY

Prof. Dr. Dirk Praetorius
Institut für Analysis und Scientific Computing
Technische Universität Wien
Wiedner Hauptstrasse 8 - 10
1040 Wien
AUSTRIA

Dr. Hella Rabus
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
GERMANY

Dr. Mira Schedensack
Institut für Numerische Simulation
Universität Bonn
Wegelerstraße 6
53115 Bonn
GERMANY

Prof. Dr. Rob Stevenson
Korteweg-de Vries Instituut
Universiteit van Amsterdam
Postbus 94248
1090 GE Amsterdam
NETHERLANDS

Prof. Dr. Gantumur Tsogtgerel
Department of Mathematics and Statistics
McGill University
805, Sherbrooke Street West
Montreal, P.Q. H3A 2K6
CANADA

Dr. Kris G. van der Zee
School of Mathematical Sciences
The University of Nottingham
University Park
Nottingham NG7 2RD
UNITED KINGDOM

Prof. Dr. Marco Verani
Dipartimento di Matematica
Politecnico di Milano
Via Bonardi 9
20133 Milano
ITALY

Prof. Dr. Martin Vohralik
INRIA Rocquencourt
Domaine de Voluceau
BP 105
78153 Le Chesnay Cedex
FRANCE