Variational Localizations of the Dual Weighted Residual Estimator

Thomas Richter [*]   Thomas Wick [†]

The dual weighted residual method (DWR) and its localization for mesh adaptivity applied to elliptic partial differential equations is investigated. The contribution of this paper is twofold: first, we introduce a novel localization technique based on the introduction of a partition of unity. This new technique is very easy to apply, as neither strong residuals nor jumps over element edges are required. Second, we compare and analyze (theoretically and numerically) different localization techniques used for mesh adaptivity with respect to their effectivity. Here, we focus on localizations in variational formulations that do not require the evaluation of the corresponding differential operator in the classical strong formulation. In our mathematical analysis, we show for different localization techniques (established methods and our new approach), that the local error indicators used for mesh adaptivity converge with proper order in the error functional. Several numerical tests substantiate our theoretical investigations.

1 Introduction

In this work, we investigate the dual weighted residual method (DWR) and its localization for mesh adaptivity applied to elliptic partial differential equations. Our goal is twofold: First, and most important, we introduce a new localization technique, given in weak form that avoids both the evaluation of strong residuals and jump terms over element edges. This method is easy to implement and therefore suitable for coupled multiphysics systems with many different equations. The second aim is then to analyze different localization techniques with respect to their effectivity. For some established localization techniques, this has not yet been accomplished.

The DWR method allows for estimating the error $u - u_h$ between the exact solution $u \in V$ (for a function space V) of the PDE and its Galerkin solution $u_h \in V_h \subset V$ in general (error) functionals $J : V \to \mathbb{R}$. These functionals can be norms but also more general expressions, like point-values, (local) averages or technical expressions like (in the case of fluid dynamics) lift- or drag-coefficients. Error estimators based on the DWR method always consist of residual evaluations, that are weighted by adjoint sensitivity measures. These sensitivities are the solution to adjoint problems that measure the influence of the error functional J.

---

[*]Institute for Applied Mathematics, University of Heidelberg, thomas.richter@iwr.uni-heidelberg.de

[†]Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, Texas 78712, United States twick@ices.utexas.edu

The DWR method goes back to Becker & Rannacher [8, 10] and is based on the pioneering work by Eriksson, Estep, Hansbo and Johnson [20]. It has been further developed by various researchers [32, 1, 23] and has been applied to a vast number of application problems including fluid-dynamics [7], structural dynamics, and further to complex multiphysics problems like chemically reactive flows [13] or fluid-structure interactions [24, 39, 34, 38, 37]. A completely different field, where the use of strong residuals is to be avoided is kinetic theory, see [26] for an application of goal oriented error estimation to Boltzmann-type equations.

In this contribution to the DWR method, we focus on more principle questions that arise in its application. First, the adjoint weights entering the error estimator usually must be approximated, as they involve the unknown exact solution $z \in V$ of an adjoint problem. Section 3 provides an overview of different approximation techniques commonly used. Second, if used for adaptive mesh refinement, the error estimator $\eta_h \approx J(u) - J(u_h)$ must be localized to positive error indicators

$$|\eta_h| \approx \sum_i |\eta_i|, \tag{1}$$

which describe the local error contribution $|\eta_i|$ of a mesh element or a mesh node, and that can be used to establish adaptive mesh refinement schemes. In the central Section 4, we describe different localization techniques for the DWR estimator and discuss their effectivity: a localization is called effective, if the sum of local indicators do not overestimate the error. Error indicators, that highly overestimate the error will lead to adaptive meshes, which do not fit to the problem. Usually, for adaptive methods, one aims at showing effectivity, such that the estimator bounds the error from below and above

$$c_1 \sum_i |\eta_i| \le \|\nabla(u - u_h)\| \le c_2 \sum_i |\eta_i|. \tag{2}$$

We cannot expect such a sharp result, as we are not looking at norms only, but at errors in general functionals $J(\cdot)$. The DWR estimator is an error approximation $\eta_h \approx J(u) - J(u_h)$, but usually not a rigorous estimate. Usually, it is straightforward to bound the indicators by the estimator from below

$$|\eta_h| \le \sum_i |\eta_i|. \tag{3}$$

The main contribution of this work is to provide insight to the opposite direction. We are not able to bound the sum of indicators by the estimator $|\eta_h|$ or even the functional error $|J(u) - J(u_h)|$ itself, but we can show, that the error $|J(u) - J(u_h)|$ and the indicators $\sum_i |\eta_i|$ satisfy a common upper bound. This has not been accomplished for some commonly used localization techniques. Finally, we introduce a novel localization technique, that is strikingly simple in its application and also permits a very simple proof to show the effectivity of indicators (within the limits just discussed).

In Section 5, several numerical test cases are presented to discuss the performance of the different localization strategies. Finally, in Section 6, we conclude with some further remarks.

Let us begin in the following second section by gathering the notation and shortly introducing the dual weighted residual method for error estimation.

2 The Dual Weighted Residual Method for Error Estimation

By $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ we denote a domain with polygonal or polyhedral domain. On Ω, we denote by $(\cdot, \cdot)$ the L^2-inner product and by $\|\cdot\|$ the corresponding L^2-norm. By $H^{r+1}(\Omega)$ we denote the space of Lebesgue functions with square integrable weak derivatives up to degree $r + 1$. In particular, by $V := H_0^1(\Omega)$ we denote the space of $H^1(\Omega)$ functions with trace zero on the boundary $\partial\Omega$.

2.1 DWR for the Poisson problem and linear goal functionals

Next, by $u \in V$ we denote the solution of the Poisson equation

$$(\nabla u, \nabla \phi) = (f, \phi) \quad \forall \phi \in V, \tag{4}$$

for a given right hand side function $f \in H^{-1}(\Omega)$. We consider the case of homogenous Dirichlet boundary conditions on $\partial\Omega$ only. Next, we denote by $V_h := V_h^{(r)} \subset V$ a finite dimensional, piece-wise polynomial of degree r finite element subspace and by $u_h \in V_h$ the finite element solution

$$(\nabla u_h, \nabla \phi_h) = (f, \phi_h) \quad \forall \phi_h \in V_h. \tag{5}$$

Here we only consider finite element spaces of polynomial degree $r \geq 1$ on shape-regular triangulations Ω_h, such that there exists an interpolation operator $i_h : V \to V_h$ on every element $K \in \Omega_h$

$$\|\nabla^k(u - i_h u)\|_K \leq c_{\text{in}} h_K^{r+1-k} \|\nabla^{r+1} u\|_K \quad \forall u \in H^{r+1}(K), \ k = 0, 1, 2, \tag{6}$$

with $h_K := \text{diam}(K)$ and where $\|\cdot\|_K$ is the L^2-norm on K. The interpolation constant c_{in} depends on the polynomial degree r and the triangulation Ω_h. Further, on element boundaries ∂K, we use the estimate

$$\|u - i_h u\|_{\partial K} \leq c_{\text{in}} h^{r+\frac{1}{2}} \|\nabla^{r+1} u\|_K, \tag{7}$$

with $h = \max_K h_K$. Adaptive meshes are realized with hanging nodes, see [25] for details on the construction.

We assume, that the problem data, e.g. right hand side f and domain Ω are such, that the following two a priori error estimates hold for the finite element solution

$$\|\nabla(u - u_h)\| \leq ch^r \|\nabla^{r+1} u\|, \quad \|u - u_h\| \leq ch^{r+1} \|\nabla^{r+1} u\|. \tag{8}$$

For linear elements with $r = 1$, this is given for $f \in L^2(\Omega)$ on polygonal convex domains or if the boundary is smooth (piece-wise C^2) with only convex corners [21]. By $J : V \to \mathbb{R}$ we denote a linear continuous functional and by $z \in V$ we denote the adjoint solution to (4)

$$(\nabla \phi, \nabla z) = J(\phi) \quad \forall \phi \in V. \tag{9}$$

Existence and uniqueness of this adjoint solution follows by standard arguments. The regularity of $z \in V$ depends on the regularity of the functional J. For $J \in H^{-1}(\Omega)$ it holds $z \in H^1(\Omega)$. Given a more regular functional like the L^2-error $J(\phi) = \|e_h\|^{-1}(e_h, \phi)$ with $J \in L^2(\Omega)^*$, it holds $z \in H^2(\Omega)$ on suitable domains (convex polygonal or smooth boundary with C^2-parametrization). By $z_h \in V_h \subset V$ we denote the corresponding adjoint finite element solution

$$(\nabla \phi_h, \nabla z_h) = J(\phi_h) \quad \forall \phi_h \in V_h. \tag{10}$$

Key relation in the context of the dual weighted residual method, see Becker & Rannacher [8, 10], is the following error identity which is only based on Galerkin orthogonality by plugging-in $\nabla i_h z$:

$$J(u) - J(u_h) = (f, z - i_h z) - (\nabla u_h, \nabla(z - i_h z)). \tag{11}$$

The error in the functional $J(u) - J(u_h)$ can be expressed in terms of a residual, that is weighted by (the local) adjoint sensitivity information $z - i_h z$. Further, a second adjoint error identity is directly given as by introducing ∇z_h; and Galerkin orthogonality of the adjoint equation can be used such that

$$J(u) - J(u_h) = J(u - i_h u) - (\nabla(u - i_h u), \nabla z_h). \tag{12}$$

Here, the residuals of the adjoint equation are weighted with the primal interpolation error.

The two error representations involve as unknown parts in the weights the primal solution $u \in V$ or adjoint $z \in V$ solution. Section 3 deals with approximation techniques for the evaluation of these interpolation weights. This approximation is required to obtain an usable error estimator. Then, in Section 4 we come to the localization of the error representations and the definition of different local error indicators.

For the following we collect some useful inequalities. All these inequalities are given for two-dimensional domains. First, on a mesh element $K \in \Omega_h$, we frequently use the trace-inequality

$$\|u\|_{\partial K} \leq c_{\mathrm{tr}} h_K^{-\frac{1}{2}} \left(\|u\|_K + h_K \|\nabla u\|_K \right), \tag{13}$$

and in addition, we recall the inverse estimate

$$0 \leq s \leq k : \quad \|\nabla^k u_h\|_K \leq c h_K^{s-k} \|\nabla^s u_h\|_K \quad \forall v_h \in V_h, \tag{14}$$

which - on shape regular meshes - easily follows by equivalence of norms in discrete spaces, see [15].

2.2 DWR for nonlinear problems and nonlinear functionals

The theory presented above is limited to linear partial differential equations and linear functionals. Here, we shortly recapitulate the full nonlinear DWR theory as presented by Becker & Rannacher [10]. Let $J : V \to \mathbb{R}$ be a differentiable error functional and let $a(\cdot)(\cdot)$ be a differentiable semilinear form, which is linear in the second argument. Let $u \in V$ be the solution to the nonlinear problem

$$a(u)(\phi) = (f, \phi) \quad \forall \phi \in V, \tag{15}$$

and let $z \in V$ be the solution to the linearized adjoint problem

$$a'(u)(\phi, z) = J'(u)(\phi) \quad \forall \phi \in V, \tag{16}$$

where by $a'(u)(\cdot, \cdot)$ we denote the Gâteaux derivative of $a(\cdot)(\cdot)$ in $u \in V$, and by $J'(u)(\cdot)$ the Gâteaux derivative of $J(\cdot)$ in $u \in V$. Then, it holds the following mixed error representation

$$J(u) - J(u_h) = \frac{1}{2} \left\{ (f, z - i_h z) - a(u_h)(z - i_h z) \right\}$$
$$+ \frac{1}{2} \left\{ J'(u_h)(u - i_h u) - a'(u_h)(u - i_h u, z_h) \right\} + \mathcal{R}^{(3)}(u - u_h, z - z_h), \tag{17}$$

where both primal and adjoint residual appear, each tested with interpolation weights coming from the other problem. The error identity (17) includes a remainder term $\mathcal{R}^{(3)}$, that is of third order in the errors $u - u_h$ and $z - z_h$ and stems from the application of the trapezoidal quadrature rule, see Becker & Rannacher [10]. We notice that the primal estimator (11) is still valid for nonlinear problems, it is however disturbed by a second order error term $\mathcal{R}^{(2)}(u - u_h, z - z_h)$.

If $a(\cdot)(\cdot)$ describes a bilinear form, e.g. if we consider linear equations and if the goal functional $J(\cdot)$ is linear, the mixed error identity (17) is exact with $\mathcal{R}^{(3)} = 0$ and the adjoint problem is defined by (9).

As primal (11) and dual (12) error identities are exact in the linear case, the mixed error formula (17) is an equivalent formulation. While all these error representations are exact for measuring the error, they lead to different error indicators which may produce adaptive meshes of different quality. See Sections 4 and 5 for discussions on this point.

In contrast to other classical a posteriori error estimators for the energy norm, the DWR method is not an estimator in the strict case. Considering linear problems, the DWR method is an error identity. However, as discussed in the following section, it is not possible to evaluate the error relation (even in the linear case), as the weights depend on the exact values of primal and dual solution. Hence, computational approaches deliver an approximation of the error identity η_h:

$$\eta_h(u_h, z_h) \approx J(u) - J(u_h).$$

The quality of this approximation procedure can be measured by the *effectivity index* eff_h, defined as

$$\mathrm{eff}_h := \frac{\eta(u_h, z_h)}{J(u) - J(u_h)}. \tag{18}$$

For $\mathrm{eff}_h \to 1$, the estimate is asymptotically exact. As functional errors carry a sign, no absolute values may be used in defining the effectivity index. The following discussion shows, that cancellation effects by different signs are the major cause for difficulties connected to the localization of the DWR method. Local indices to be used for adaptivity will have to be positive measures of the error contribution.

3 Approximation of the Weights

For evaluation of the error identities (11), (12) and (17), we need approximations of the interpolation errors $z - i_h z$ and/or $u - i_h u$. Obtaining such an approximation is the critical part in the DWR framework that stands in the way of strict reliability. Examples have been constructed [31], where the DWR estimator underestimated the error due to coarse approximation of the weights $z - i_h z$ and $u - i_h u$. A remedy is only given by spending sufficient effort on the estimation of these weights on fine meshes [10, 17] or an additional control of the approximation error in $z - i_h z$ and $u - i_h u$ [31]. For simplicity of presentation, we consider the case of linear goal functionals throughout this section.

3.1 Classical approximation of the DWR estimator

First, we consider a classical approach, that is based on applying Cauchy-Schwarz inequality locally on every element for getting strict upper bounds, see e.g. [8, 9, 6]. With integration by parts on every mesh element $K \in \Omega_h$ it holds:

$$J(u) - J(u_h) = \sum_{K \in \Omega_h} \left(f + \Delta u_h, z - i_h z \right)_K + \int_{\partial K} \partial_n u_h \cdot (z - i_h z) \, \mathrm{d}s. \tag{19}$$

Following the usual procedure for residual based error estimators [36], we combine each two boundary integrals over element edges to a normal jump and proceed with Cauchy Schwarz to get

$$|J(u) - J(u_h)| \leq \sum_{K \in \Omega_h} \underbrace{\left(\|f + \Delta u_h\|_K + \frac{1}{2} h_K^{-\frac{1}{2}} \|[\partial_n u_h]\|_{\partial K} \right)}_{=: \rho_K(u_h)} \underbrace{\left(\|z - i_h z\|_K + h_k^{\frac{1}{2}} \|z - i_h z\|_{\partial K} \right)}_{=: \omega_K(z)}, \tag{20}$$

where by $[\partial_n u_h]$ we denote the jump of the u_h derivative in normal direction. On the outer boundary $\partial \Omega$, we set $[\partial_n u_h] = 0$. The residual part ρ_K only contains the discrete solution u_h and the problem data. A similar estimator can be derived based on the adjoint form (12), where the residuals of the adjoint equation $\rho_K^*(u_h, z_h)$ are weighted with primal interpolation errors $\omega_K(u)$.

Remark 1 (Evaluation of the strong residuals)**.** *The evaluation of the classical error localization can bring along enormous computational costs, as for higher order finite elements the assembly of second order differential operators on general mesh elements is computationally expensive.*

Considering nonlinear problems, we further have to evaluate the adjoint residuals $\rho_K^*(u_h, z_h)$. For some complex problems, e.g. like fluid-structure interactions [24, 39, 34], the strong adjoint residual formulation has not even been derived yet. Such complex nonlinear problems are our key motivation in deriving localization techniques that do not require the evaluation of classical residuals.

Remark 2 (Strong residuals for C^1-continous approximations)**.** *Edge terms in strong residuals only appear for H^1-conforming finite elements. By using C^1-continuous approximations, it holds $[\partial_n u_h] = 0$ on the element edges in (19), such that no edge terms appear in (20). Traditionally, C^1-continuous elements found little usage due to their high computational effort. This however changes in the context of isogeometric analysis, see [19] or [27] with application to goal oriented error estimation.*

The error weights $\omega_K(z)$ involve the unknown adjoint solution $z \in V$, which has to be approximated. Here, two possible approaches exist: one can directly approximate the interpolation error $z - i_h z$ using available discrete quantities only. This approach is described in the following section. As an alternative, one could first apply an interpolation error estimate

$$\omega_K(z) \leq c_{\text{in}} h^{r+1} \|\nabla^{r+1} z\|_{P(K)}, \tag{21}$$

where $P(K)$ is the patch of all those elements $K' \in \Omega_h$ that share a common boundary with $K \in \Omega_h$, followed by an approximation of the $(r+1)$-th derivative

$$h^{r+1} \|\nabla^{r+1} z\|_{P(K)} \approx h^{r+1} \sqrt{|K|} \|\nabla_h^{r+1} z_h\|_{L^\infty(P(K))}, \tag{22}$$

which is based on discrete recovery concepts. This approach is typical for the gradient recovery error estimator and it is directly applicable to the DWR method. We refer to the literature [40, 41]

The drawback of the classical DWR-approach is due to the application of the Cauchy-Schwarz inequality that rules out possible cancellation effects by local orthogonality and interpolation error estimates, which both bring along unknown constants. Usually, error estimators based on this approach result in an over-estimation of the true error. We provide examples in Section 5.

3.2 Variational approximation of the DWR estimator

The second possibility for the evaluation of the weights is by a direct approximation of the interpolation error without prior estimates. If an approximation for $z - i_h z$ is available, it can be both used for the error estimate given in the classical form (20), but also directly for the error identities (11), (12) and (17), given in weak formulation. Here, every discrete approximation to $\psi_h \approx z - i_h z$ must be finer than the trial space V_h, as the residual is orthogonal on V_h. A first obvious possibility is to simply solve for an approximation z_h^* in a higher accurate space, e.g.,

$$z_{h/2} \in V_{h/2} : \quad (\nabla \phi_{h/2}, \nabla z_{h/2}) = J(\phi_{h/2}) \quad \forall \phi_{h/2} \in V_{h/2}, \tag{23}$$

where $z - i_h z \approx z_{h/2} - i_h z_{h/2}$. Alternatively, it would be possible to solve for $z_h^{(2)} \in V_h^{(2r)}$ in a finite element space of higher polynomial degree. Both approaches work very well in application and usually yield optimal error estimators with the effectivity index eff_h (18) going to one:

$$\text{eff}_h(u_h, z_h) = \frac{(f, z_h^{(2)} - i_h z_h^{(2)}) - (\nabla u_h, \nabla(z_h^{(2)} - i_h z_h^{(2)}))}{J(u) - J(u_h)} \to 1 \quad (h \to 0). \tag{24}$$

We refer the reader to Section 5 for an example. In the practical application, however, this approach is not feasible, since it means, that only for getting a reliable and effective error estimator we have to spend very high numerical effort [10]. A modification of this approach is to define local subproblems, that live on a small subset of elements each, and that can be solved efficiently and in parallel using higher order finite elements.

A third possibility that goes without the solution of additional higher order problems uses a reconstruction of the already computed discrete approximations $u_h \in V_h$ and $z_h \in V_h$. For this reconstruction, the discrete solutions are simply reinterpreted using a higher order basis. First, we assume, that the finite element space V_h is constructed in a patched manner, such that each element is part of four elements arising from one common father-element $P \in \Omega_{2h}$. Such a patched finite element set allows for a reinterpretation of the finite element basis by combining four r-th order elements to one element of order $2r$ (in 2d). As the finite element spaces $V_h^{(r)}$ and $V_{2h}^{(2r)}$ have the same number of unknowns in the same Lagrange points, a higher order reconstruction is directly given by an exchange of the basis:

$$z_h = \sum_i \mathbf{z}_i \phi_h^i \approx \sum_i \mathbf{z}_i \phi_{2h}^{(2r),i} =: z_h^{(2)}, \tag{25}$$

where $\mathbf{z} \in \mathbb{R}^N$ stands for the coefficient vector and by $\phi_{2h}^{(2r),i}$ we denote the basis functions of the finite element space of double degree on a mesh with double mesh size, see Fig. 1. Details on the application of this reinterpretation process on unstructured meshes are given by Carpio et al. [16]. This reconstruction strategy is highly reliable and effective for a large class of problems, see [13, 33, 35]. Similar to gradient recovery error estimators, it is based on super-approximation results obtained by error expansion techniques, see e.g. Blum and coworkers [12, 11]; however it can not be rigorously shown on adaptive meshes. The success in numerical examples however works in favor of this cost-efficient approach.

4 Localizations of the error identity and effectivity of localizations

In this section, we discuss with the localization of the error estimator. In most representations, the error estimator allows for a direct splitting into a sum $J(u) - J(u_h) \approx \eta_h = \sum_i \eta_i$. The absolute values of the local quantities $|\eta_i|$ are the indicators used for refinement. Our discussion follows two goals: the localization procedure should be simple in terms of implementation and numerical effort. Second, the localization should be effective, such that the sum of local error indicators does not heavily overestimate the error. Similar to the effectivity index (18), we define the *indicator index* to measure the quality of the localization process:

$$\mathrm{ind}_h := \frac{\sum_i |\eta_i|}{|J(u) - J(u_h)|}. \tag{26}$$

It is not possible to reach strict effectivity with $\mathrm{ind}_h \to 1$ in the context of goal-oriented errors. The functional error $J(u) - J(u_h)$ has a sign, and hence the error can vanish, although the solution shows a very large approximation error, for example by symmetry reasons. The local estimator values η_i might have changing sign, such that $\sum_i |\eta_i|$ may be a strong over-estimation. However, we aim at strategies, where ind_h is uniformly bounded in h. By a priori estimates, a worst case bound for the functional error is given by the product of primal and dual energy errors

$$|J(u) - J(u_h)| \leq c \|\nabla(u - u_h)\| \, \|\nabla(z - z_h)\|, \tag{27}$$

with a constant $c > 0$ that depends only on the continuity of the variational formulation, and with $c = 1$ in the case of the Poisson equation. This bound is not sharp, as it neglects possible orthogonality

of primal and dual errors. A functional error $J(u) - J(u_h)$ can be zero, even if the energy errors are substantial. In the following, we call error indicators *effective*, if this worst case estimate also holds for the indicators

$$\sum_i |\eta_i| \le c \|\nabla(u - u_h)\| \|\nabla(z - z_h)\|, \tag{28}$$

with a constant $c > 0$ that may depend on the bilinear form of the equation, the finite element space and the shape regularity of the mesh, but that should be robust with respect to the mesh size parameter h. For practical approximations of the error identity, the weights are given by interpolation errors. Hence, the concept of effectivity defined in (28) must be slightly altered to also allow for bounds in the interpolation error $\|\nabla(u - i_h u)\|$ and $\|\nabla(z - i_h z)\|$. Finally, whenever formulations based on strong residuals are considered, a pure H^1-estimate is not sufficient. Instead, we define a mesh-dependent norm, that is - assuming enough regularity $u \in H^2(\Omega)$ - equivalent in terms of convergence $\|\nabla(u - u_h)\| \sim \|u - u_h\|_h$ in h:

$$\|\phi\|_h := \left(\|\nabla\phi\|_\Omega^2 + \sum_{K \in \Omega_h} \left\{ h_K^2 \|\nabla^2\phi\|_K^2 + h_K^{-2} \|\phi\|_K^2 \right\} \right)^{\frac{1}{2}}. \tag{29}$$

Using this norm, we can now define our concept of effective error indicators:

$$\sum_i |\eta_i| \le c \max\{\|u - u_h\|_h, \|u - i_h u\|_h\} \max\{\|z - z_h\|_h, \|z - i_h z\|_h\}. \tag{30}$$

Every localization can only be as accurate, as the approximation of the weights permits as discussed in the previous section. For the following discussion, we assume that the approximation of the weights is sufficiently accurate. On coarse meshes, this is a simplification, as shown by the discussion of Nochetto et al. [31].

4.1 Localization based on the classical (strong) formulation

The typical localization procedure [8, 10] for residual based error estimators is based on the classical formulation of the error estimator (20) by defining local element-wise indicator values $\eta_K := \rho_K \omega_K$. Reliability of these indicators depends on the approximation properties of the interpolation weights $z - i_h z$. The question of effectivity is a bigger concern, as the Cauchy-Schwarz inequality has been used. We know, that a functional error $J(u) - J(u_h)$ can change its sign and pass through zero. This behavior cannot be represented by the classical localization. However, given the more subtle definition of effectivity (30), it holds:

Lemma 1 (Effectivity of the localization based on the classical residual)**.** *Let $u, z \in V \cap H^2(\Omega)$ be the solution and adjoint solution, respectively. Furthermore, let $u_h, z_h \in V_h = V_h^{(r)}$ be the corresponding finite element solutions of degree r. The classical error indicators $\eta_K = \rho_K \omega_K$ given by (20) are effective, i.e.,*

$$\sum_{K \in \Omega_K} \rho_K \omega_K \le c \|u - u_h\|_h \|z - i_h z\|_h, \tag{31}$$

with a constant $c > 0$.

Proof. (i) We split the residual part $\rho_K = \rho_K^i + \rho_K^e$ into inner part $\rho_K^i = \|f + \Delta u_h\|_K$ and edge part $\rho_K^e = h_K^{-\frac{1}{2}} \|[\partial_n u_h]\|_{\partial K}$. First, it holds by standard a priori analysis using $-\Delta u = f$:

$$\rho_K^i = \|f + \Delta u_h\|_K = \|\Delta(u - u_h)\|_K \le \|\nabla^2(u - u_h)\|_K. \tag{32}$$

On every edge, it holds for $u \in V \cap H^2(\Omega)$, that $[\partial_n u] = 0$ and hence with (13)

$$\rho_K^e = h_K^{-\frac{1}{2}} \|[\partial_n u_h]\|_{\partial K} = h_K^{-\frac{1}{2}} \|[\partial_n (u - u_h)]\|_{\partial K}$$
$$\leq c_{\mathrm{tr}} h_K^{-1} \left(\|\nabla(u - u_h)\|_K + h_K \|\nabla^2(u - u_h)\|_K \right). \quad (33)$$

Combining (32) and (33) gives the first part of the estimate.

(ii) Next, for the weights, we get by using (13)

$$\omega_K = \|z - i_h z\|_K + h_k^{\frac{1}{2}} \|z - i_h z\|_{\partial K} \leq c h_K \left(\|\nabla(z - i_h z)\|_K + h_K^{-1} \|z - i_h z\|_K \right). \quad (34)$$

(iii) The result follows by combining the estimates for ρ_K and ω_K using Hölder's inequality. $\qquad \square$

We can bound the sum of error indicators by the product of approximation error in the primal solution and interpolation error of the adjoint solution. Using the dual form of the error identity (12) we would get the opposite result. Starting with the mixed identity (17) results in the sum of both estimates, or - in terms of the concept of effectivity (30) - in the maximum value of approximation and interpolation error.

For an evaluation of this indicator (at a higher polynomial degree) one must assemble the strong residual of the equation, namely $f - L u_h$, where L is the second order differential operator in classical formulation. This evaluation can be very costly, when parametric finite elements of higher order are used, see Remark 1. Finally, having the full nonlinear case in mind, see Section 2.2, the error indicator consists of a primal and adjoint part, where $\eta_K^* = \rho_K^* \omega_K^*$, with

$$\rho_K^* = \|J'(u_h) - L'^*(u_h) z_h\|_K, \quad (35)$$

where $L'^*(u_h)$ is the linearized adjoint operator at u_h. For complex coupled problems it is sometimes not possible to assemble the adjoint operator in strong formulation, see e.g. [34].

4.2 Localization based on filtering the variational formulation

Braack and Ern [14] proposed a localization technique for the DWR estimator that is fully based on the variational formulation and firmly linked to the approximation of the weights using a higher order representation $z_h^* \in V_h^*$ of the adjoint solution $z_h \in V_h^{(r)}$, where $V_h^{(r)}$ is the finite element space of degree r, see Section 3.2. Given a patched mesh and finite element space setup, we can define the space $V_{2h}^{(r)} \subset V_h^{(r)}$ of double mesh spacing and introducing $i_{2h}(z - i_h z) = 0$. As the two interpolation operators commute $i_{2h} i_h = i_h i_{2h}$ (considering a standard nodal interpolation), it holds

$$z - i_h z - i_{2h}(z - i_h z) = (\mathrm{id} - i_{2h})(\mathrm{id} - i_h)z = (\mathrm{id} - i_h)(\mathrm{id} - i_{2h})z = (\mathrm{id} - i_h)\pi_{2h} z, \quad (36)$$

where the patch-wise *filtering operator* is $\pi_{2h} := \mathrm{id} - i_{2h}$; introduced by Braack and Ern. It remains to apply the approximation of the weights as described in Section 3.2:

$$J(u) - J(u_h) \approx \left(f, (i^* - \mathrm{id})\pi_{2h} z_h \right) - \left(\nabla u_h, \nabla \left((i^* - \mathrm{id})\pi_{2h} z_h \right) \right). \quad (37)$$

The operator $i^* : V_h^{(r)} \to V_h^* := V_{2h}^{(2r)}$ is the patch-wise interpolation into the space of double polynomial degree $V_{2h}^{(2r)}$ on the patch mesh , and for $\phi_h^i \in V_h^{(r)}$ it holds $i_h i^* \phi_h^i = \phi_h^i$. Let $P = \{K_1, \ldots, K_p\}$ be a patch of elements $K \in \Omega_h$ and $\phi_h^{i,*} := i^* \phi_h^i \in V_{2h}^{(2r)}$. Then it holds,

$$(i^* - \mathrm{id})\phi_h^i = \phi_h^{i,*} - i_h \phi_h^{i,*}, \quad (38)$$

such that for $k = 0$ and $k = 1$ the interpolation estimate gives

$$h_K^k \|\nabla^k (i^* - \mathrm{id})\phi_h^i\|_P = h_K^k \|\nabla^k (\phi_h^{i,*} - i_h \phi_h^{i,*})\|_P \leq c h_K \|\nabla \phi_h^{i,*}\|_P \leq c h_K. \tag{39}$$

Applied to z_h, the filtering operator $\pi_{2h} z_h$ is a strictly local algebraic process acting on the coefficient vector $\mathbf{z} \in \mathbb{R}^N$:

$$z_h - i_{2h} z_h = \sum_i \mathbf{z}_i (\phi_h^i - i_{2h} \phi_h^i) =: \sum_i (\pi_{2h} \mathbf{z})_i \phi_h^i, \tag{40}$$

as the interpolations of the finite element bases functions $i_{2h} \phi_h^i$ can be linearly combined by ϕ_h^j. Finally, Braack and Ern defined the local error indicators as

$$J(u) - J(u_h) \approx \sum_i \underbrace{\left\{ \left(f, (i^* - \mathrm{id})\phi_h^i \right) - \left(\nabla u_h, \nabla \left((i^* - \mathrm{id})\phi_h^i \right) \right) \right\}}_{=: \eta_i^\pi} (\pi_{2h} \mathbf{z})_i. \tag{41}$$

We refer to Fig. 1 for a sketch of the different interpolation operators i_{2h}, i^* and the filtering operator π_{2h} employed for this approach. We notice that these local indicators η_i^π are node-wise and not element-wise contributions. The error indicators inherit the patch structure and it holds, that $\eta_i = 0$ for every second degree of freedom (in a tensor-product way), e.g. for all degrees of freedom that belong to Lagrange points x_i of the finite element space V_h as well as the finite element space V_{2h}. If these indicators are to be used for mesh-refinement, a first step has to be a summation of all indicators associated to one patch. Then, refinement is carried out on the patch-mesh.

Lemma 2 (Effectivity of the algebraic filter-approach). *Let $u \in V \cap H^2(\Omega)$ and $z \in V$ be solution to primal and adjoint problem, $u_h \in V_h$ and $z_h \in V_h$ be the corresponding finite element solutions of degree r. The filtering indicator defined by (41) is effective*

$$\sum_{i=1}^N |\eta_i^\pi| \leq c\|u - u_h\|_h \left(\|\nabla(z - z_h)\| + \|\nabla(z - i_{2h} z)\| \right), \tag{42}$$

with a constant $c > 0$ and where $i_{2h} : H^2(\Omega) \to V_{2h}^{(r)}$ is the interpolation operator into the finite element space of the same degree r on the patched mesh with mesh-size $2h$.

Proof. We consider the two dimensional case only. The three dimensional case follows by similar arguments.

All indicators η_i belonging to Lagrange points $x_i \in \Omega_{2h}$ vanish. We must distinguish between indicators η_e belonging to degrees of freedom on edges of a patch $x_e \in \partial P$ and indicators η_m belonging to inner points $x_m \in P$, see Fig. 2. (In the three dimensional case we would have to add all indicators η_f belonging to faces of elements.)
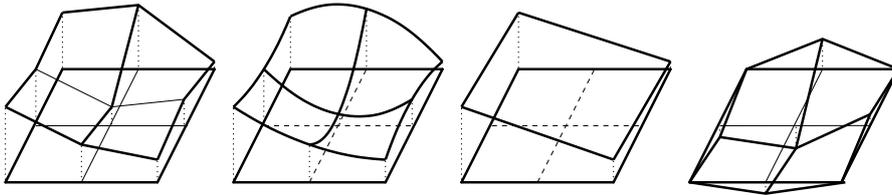

Figure 1: Patch of four elements. Discrete solution $u_h \in V_h^{(r)}$, reconstruction $i^* u_h \in V_{2h}^{(2r)}$, coarse-mesh interpolation $i_{2h} u_h \in V_{2h}^{(r)}$ and fluctuation operator $\pi_{2h} u_h$ (going from left to right).

(i) Inner points. Let $x_m \in P$ be the midpoint of a patch P. Then, the basis functions $\phi_h^i, i^*\phi_h^i$ and $i_{2h}\phi_h^i$ all have their support in P. Let $\bar{z} = \pi_{2h}z_h(x_m)$.

$$\eta_i^\pi = \bar{z}\left\{\left(f, (i^* - \mathrm{id})\phi_h^i\right)_P - \left(\nabla u_h, \nabla\left((i^* - \mathrm{id})\phi_h^i\right)\right)_P\right\}$$

$$= \bar{z}\left\{\left(f + \Delta u_h, (i^* - \mathrm{id})\phi_h^i\right)_P - \sum_{K \in P}\left\langle\partial_n u_h, (i^* - \mathrm{id})\phi_h^i\right\rangle_{\partial K}\right\}, \tag{43}$$

where $K \in P$ are all in the patch. On $\partial K \in \partial P$, the boundary integral vanishes. By $e_i \in P$ we denote the interiour edges between the elements. They appear twice, such that using the normal jumps it holds

$$\eta_i^\pi = \bar{z}\left\{\left(f + \Delta u_h, (i^* - \mathrm{id})\phi_h^i\right)_P - \sum_{e_i \in P}\left\langle[\partial_n u_h] - [\partial_n u], (i^* - \mathrm{id})\phi_h^i\right\rangle_{e_i}\right\}. \tag{44}$$

As in the proof to Lemma 1, we added $[\partial_n u] = 0$ on all inner edges $e_i \in P$:

$$|\eta_i^\pi| \leq |\bar{z}|\Big\{\|\Delta(u - u_h)\|_P \|(i^* - \mathrm{id})\phi_h^i\|_P$$

$$+ ch_K^{-1}\big(\|\nabla(u - u_h)\|_P + h_K\|\nabla^2(u - u_h)\|_K\big)$$

$$\cdot\big(\|(i^* - \mathrm{id})\phi_h^i\|_K + h_K\|\nabla(i^* - \mathrm{id})\phi_h^i\|_K\big)\Big\}. \tag{45}$$

We proceed with (39) to get

$$|\eta_i^\pi| \leq c|\bar{z}|\Big\{h_K\|\nabla^2(u - u_h)\|_P + \|\nabla(u - u_h)\|_P\Big\}. \tag{46}$$

It remains to estimate the discrete fluctuation $\bar{z}$. By inverse estimates, it holds

$$|\bar{z}| \leq \|\pi_{2h}z_h\|_{L^\infty(P)} \leq ch_K^{-1}\|\pi_{2h}z_h\|_P \leq c\|\nabla\pi_{2h}z_h\|_K. \tag{47}$$

This last inverse estimate works out, as for $\nabla\pi_{2h}z_h = 0$ it must follow that $\pi_{2h}z_h$ is constant on P and as $z_h \in V_{2h}$ finally $\pi_{2h}z_h = 0$. By introducing $\pm z$, and using the stability of the interpolation operator i_{2h}, the estimator sum gets

$$\sum_i |\eta_i^\pi| \leq c\left(\|\nabla(u - u_h)\|^2 + \sum_{K \in \Omega_h} h_K^2\|\nabla^2(u - u_h)\|_K^2\right)^{\frac{1}{2}}$$

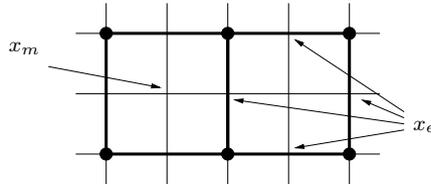$$\cdot\left(\|\nabla(z - z_h)\| + \|\nabla(z - i_{2h}z)\|\right).$$

Figure 2: The filter based error indicator on two patches $P \in \Omega_h$. It holds $\eta_i = 0$ for all outer points $x_i \in \Omega_{2h}$. Splitting of the estimator into inner nodes x_m and edge nodes x_e.

(ii) Boundary points. Let $x_e \in \partial P$ be a node on the edge of two patches P_1 and P_2. Then, it holds

$$\eta_i^\pi = \bar{\mathbf{z}} \left\{ \left(f, (i^* - \mathrm{id})\phi_h^i \right)_{P_1 \cup P_2} - \left(\nabla u_h, \nabla \left((i^* - \mathrm{id})\phi_h^i \right) \right)_{P_1 \cup P_2} \right\}, \tag{48}$$

where again, $\bar{\mathbf{z}} = \pi_{2h} z_h(x_e)$. Here, the test-function ϕ_h^i and its interpolants have their support in the joint patch $P_1 \cup P_2$ and all terms can be estimated as in step *(i)*. $\qquad\square$

Remark 3. *This estimate slightly differs from the effectivity concept (30). The patch-structure enters by the interpolation $\|\nabla(z - i_{2h}z)\|$ onto the coarse mesh. This term however is of the same order in h as $\|z - i_h z\|_h$, just carrying a larger constant.*

The benefit of this localization strategy is the simplicity of implementation, if a patch structured mesh is available. For evaluation of the estimator, two residuals must be calculated,

$$r_i := (f, \phi_h^i) - (\nabla u_h, \nabla \phi_h^i), \quad r_i^* := (f, \phi_h^{i,*}) - (\nabla u_h, \nabla \phi_h^{i,*}), \quad i = 1, \ldots, N, \tag{49}$$

the first using the standard basis, the latter with a higher order basis. Then, given the filtered coefficient vector $\pi_{2h}\mathbf{z}$, the estimator is given by the algebraic computation

$$\eta_i^\pi = (r_i^* - r_i)(\pi_{2h}\mathbf{z})_i, \quad i = 1, \ldots, N. \tag{50}$$

All these ingredients are usually available in standard finite element libraries. One drawback of this localization is its interpretation, as the indicator values η_i^π are neither given in an element-wise way, nor strictly in a node-wise manner, as $\eta_i^\pi = 0$ on all coarse Lagrange points. The indicators must instead be regarded in a patch-wise sense which comes at the cost of loosing granularity. This might be an issue regarding 3D simulations as it will lead to meshes, which are up to a factor of 8 more complex as the optimal ones.

4.3 Localization using partition of unity (PU)

Finally, we introduce a new localization approach based on the variational formulation that combines the simplicity of the filter based approach (as it is given in terms of variational residuals) with a very simple structure, which does not require patched meshes. Localization is simply based on introducing a partition of unity (PU) $\sum \psi_i \equiv 1$ into the error identity (11):

$$J(u) - J(u_h) = \sum_{i=1}^N \underbrace{\left\{ (f, (z - i_h z)\psi_i) - (\nabla u_h, \nabla((z - i_h z)\psi_i)) \right\}}_{=:\eta_i^{\mathrm{PU}}}. \tag{51}$$

The resulting error indicators η_i^{PU} are node-wise contributions of the error. Mesh adaptivity can directly be carried out in a node-wise fashion: if a node is picked for refinement, all elements touching this node will be refined. Alternatively, one could also first assemble element wise indicators by summing up all indicators belonging to nodes of the element and then carry out adaptivity in the usual element-wise way.

Lemma 3 (Effectivity of the PU localization)**.** *Let $u \in V$ be the solution to the Poisson equation, $z \in V$ be the adjoint solution. $u_h, z_h \in V_h = V_h^{(r)}$ their discrete counter-part. Further, let $\sum \psi_i \equiv 1$ be a PU with $\|\nabla \psi_i\|_\infty = O(h^{-1})$. The error indicators given by (51) are effective, i.e.,*

$$\sum_{i=1}^N |\eta_i^{PU}| \le c\|\nabla(u - u_h)\| \cdot \left(\|\nabla(z - i_h z)\|^2 + \sum_{K \in \Omega_h} h_K^{-2}\|z - i_h z\|_K^2 \right)^{\frac{1}{2}}, \tag{52}$$

with a constant $c > 0$.

Proof. Let $\text{supp}(\psi_i) \subset P_i = \cup K_j$ for some elements $K_j \in \Omega_h$. It holds with $(f, \phi) = (\nabla u, \nabla \phi)$ for all $\phi \in H_0^1(\Omega)$:

$$
\begin{aligned}
\eta_i^{\mathrm{PU}} &= (f, (z - i_h z)\psi_i) - (\nabla u_h, \nabla((z - i_h z)\psi_i)) \\
&= (\nabla(u - u_h), \nabla((z - i_h z)\psi_i))_{P_i} \\
&\leq \|\nabla(u - u_h)\|_{P_i} (\|\nabla(z - i_h z)\|_{P_i} \|\psi_i\|_{L^\infty(P_i)} + \|z - i_h z\|_{P_i} \|\nabla \psi_i\|_{L^\infty(P_i)}).
\end{aligned}
\tag{53}
$$

The result follows by using $\|\nabla \psi\|_{L^\infty(P_i)} = O(h^{-1})$. $\qquad\square$

In contrast to the classical localization and the filtering approach, the PU localization technique requires minimal regularity $u, z \in H_0^1(\Omega)$ only. A similar technique based on a PU has been used by Kuzmin & Korotov [28] to localize a strong residual formulation of the DWR estimator applied to 1D transport problems. The introduction of a partition of unity into the strong formulation of the residual is also the fundamental basis for the family of *flux-free* error estimators [29, 18]. Here, the PU is used to define local sub-problems that are used to construct robust error estimators. The construction with help of a PU directly yields a localized form of the estimator. This technique is not only accurate for energy norm estimates but also robustly applied in the context of linear output functionals [29]. Further, it is possible to design a convergent finite element method based on flux-free error estimators [30]. In contrast to flux-free error estimators, we simply insert a partition of unity to localizing the standard DWR estimator.

Realization   As PU, we consider the space of piece-wise bilinear elements $V_h^{(1)}$ (without restrictions on Dirichlet boundaries) with usual nodal basis $\{\psi_h^i, \ i = 1, \ldots, N^{(1)}\}$. The approximated local error indicator is then given by

$$
\tilde{\eta}_i^{\mathrm{PU}} := \sum_{j=1}^{N^{(r)}} \left\{ (f, (\phi_{2h}^{(2),j} - \phi_h^j)\psi_h^i)_\Omega - \left(\nabla u_h, \nabla((\phi_{2h}^{(2),j} - \phi_h^j)\psi_h^i)\right)_\Omega \right\} \mathbf{z}_j,
\tag{54}
$$

and it can be efficiently computed in an element-wise manner, as only few test-functions ϕ_h^j, $\phi_{2h}^{(2),j}$ and ψ_h^i overlap on every element $K \in \Omega_h$.

On adaptive meshes with hanging nodes, the evaluation of the PU indicator is straightforward: First, the partition of unity is assembled on basis functions ψ_h^i. In a second step, the contributions belonging to hanging nodes are condensed in the usual way by distribution to the neighboring indicators, see [3] for details on handling hanging nodes. The benefit of this localization technique is its simplicity and its accuracy according to Lemma 3 demonstrated in the numerical examples in Section 5. For the application, we only need evaluations of the right hand side and the residual with modified test-functions. As PU we can simply use the standard nodal Lagrange basis of the continuous finite element space of lowest polynomial degree. This localization technique can be readily applied to general meshes in two and three dimensions. In contrast to the filtering approach, we do not require special mesh structures, such as patches. In particular for three dimensional simulations, the use of patched meshes can substantially increases the problem size. However, the problem of obtaining good approximations to the weights $z - i_h z$ and $u - i_h u$ remains and here, using reconstruction of patches still is one of the most efficient strategies. The second advantage is the easy application of the localization to complex nonlinear systems, where the evaluation of strong residuals can be cumbersome. Once again we point out, that the adjoint operator in strong formulation is not even known for some complex multiphysics problems, see e.g. [22, 34].

Remark 4 (General elliptic problems)**.** *The results stated in Lemmata 1, 2 and 3 can all be transferred to the case of general elliptic problems like transport reaction diffusion problems. We have only used standard a priori results in the H^1-seminorm and the L^2-norm. Considering the general operator $Lu := -\Delta u + \beta \cdot \nabla u + \alpha u$, the adjoint operator reads $L^* z := -\Delta z - \beta \cdot \nabla z + \alpha z$ with opposite transport direction.*

5 Numerical Tests

In this final section, we substantiate our theoretical findings by three numerical tests with increasing complexity. In the first test, we use standard Poisson's problem with a regular goal functional on a regular domain. The second test case deals with a low regularity problem on a L-shaped domain with singular right hand side and functional. Finally, we consider a three-dimensional nonlinear elasticity system to demonstrate the great flexibility and simple realization of the PU approach. In every test-case, we analyze different forms of the error estimator (primal, dual, mixed) and different localization techniques and compare them with respect to estimator and indicator effectivities on both uniform and adaptive meshes. The computations use quadrilateral and hexadedra meshes and are performed with Gascoigne 3D [5] and deal.II [4].

#el	$J(u - u_h^{(1)})$	η_h	$J(u - u_h^{(2)})$	η_h	$J(u - u_h^{(3)})$	η_h
16	$8.11 \cdot 10^{-4}$	$8.08 \cdot 10^{-4}$	$2.23 \cdot 10^{-6}$	$2.23 \cdot 10^{-6}$	$5.16 \cdot 10^{-8}$	$5.14 \cdot 10^{-8}$
64	$2.04 \cdot 10^{-4}$	$2.04 \cdot 10^{-4}$	$1.77 \cdot 10^{-7}$	$1.77 \cdot 10^{-7}$	$3.20 \cdot 10^{-9}$	$3.19 \cdot 10^{-9}$
256	$5.11 \cdot 10^{-5}$	$5.11 \cdot 10^{-5}$	$1.34 \cdot 10^{-8}$	$1.34 \cdot 10^{-8}$	$0.20 \cdot 10^{-9}$	$0.20 \cdot 10^{-9}$
1024	$1.28 \cdot 10^{-5}$	$1.28 \cdot 10^{-5}$	$0.98 \cdot 10^{-9}$	$0.98 \cdot 10^{-9}$	$< TOL$	$< TOL$
order	2.00		3.78		4.00	

Table 1: Configuration 1: error and error estimator on uniform meshes. From left to right: linear, quadratic and cubic finite elements. The last line shows the estimated order of convergence.

5.1 Configuration 1: A regular Poisson example

In this first example, we consider Poisson's equation $-\Delta u = 1$ on a unit square $\Omega := (0,1)^2$ with a homogenous Dirichlet conditions on $\partial\Omega$. As target functional we evaluate the average of the solution

$$J(u) = \int_\Omega u \, dx. \tag{55}$$

This functional corresponds to the adjoint problem $-\Delta z = 1$, again with $z = 0$ on $\partial\Omega$. Hence it holds $u = z$ and the regularity of $u, z \in H^{3-\epsilon}(\Omega)$ for $\epsilon > 0$ is limited by the edges of the unit square [21]. For such a regular problem we expect the a priori estimate

$$|J(u) - J(u_h)| \leq c\|\nabla(u - u_h)\| \, \|\nabla(z - z_h)\| \leq ch^{\min\{2r, 4-2\epsilon\}}, \tag{56}$$

where $r \geq 1$ is the polynomial degree of the finite element space V_h. By accurate computations on very fine meshes using extrapolation we identify the reference value $\tilde{J} = 0.03514425375 \pm 10^{-10}$.

We compute primal and adjoint solution $z_h^{(r)}, u_h^{(r)} \in V_h^{(r)}$ by using finite elements of degree $r = 1$, $r = 2$ and $r = 3$. The interpolation weights are either approximated by using global finite element solutions $u_h^{(2r)}$ and $z_h^{(2r)} \in V_h^{(2r)}$ of double polynomial degree, or obtained by local patch-wise reconstruction $u_{2h}^{(2r)}, z_{2h}^{(2r)} \in V_{2h}^{(2r)}$, see Section 3.
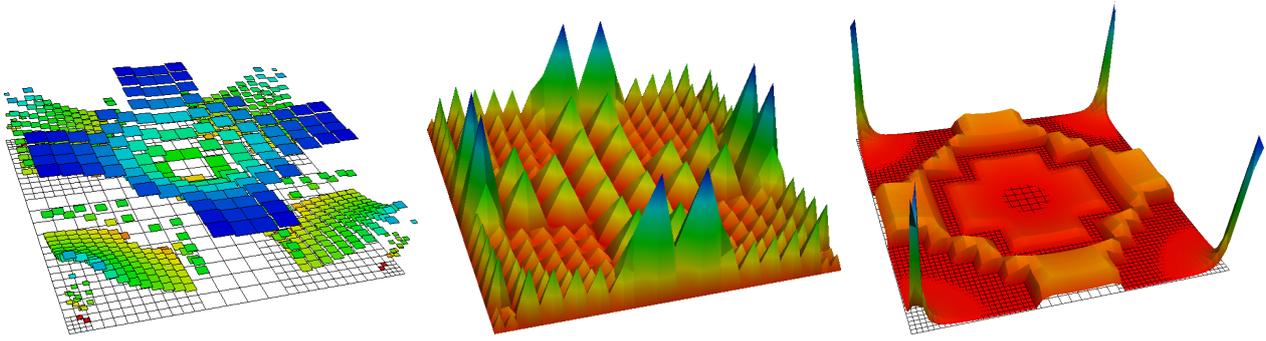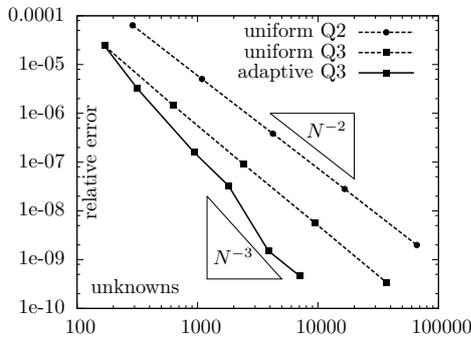
Figure 3: Configuration 1: Visualization of the local error indicators on similar meshes: (left) classical formulation, (middle) filtering approach, (right) PU. All computations with biquadratic finite elements.

# dof's	$J(u) - J(u_h^{(3)})$	η_h^{PU}	$\mathrm{eff}_h^{\mathrm{PU}}$
169	$8.51 \cdot 10^{-07}$	$8.47 \cdot 10^{-07}$	0.99
317	$1.12 \cdot 10^{-07}$	$1.37 \cdot 10^{-07}$	1.23
937	$5.56 \cdot 10^{-09}$	$7.54 \cdot 10^{-09}$	1.36
1 813	$1.14 \cdot 10^{-09}$	$1.41 \cdot 10^{-09}$	1.24
3 877	$5.28 \cdot 10^{-11}$	$8.05 \cdot 10^{-11}$	1.52
7 057	$1.61 \cdot 10^{-11}$	$2.07 \cdot 10^{-11}$	1.29

Figure 4: Configuration 1. Left: error slopes on uniform (Q2 and Q3 elements) vs. adaptive meshes (Q3 elements). Right: error, estimator and effectivity index for adaptive Q3-elements.

In Table 1 we show the functional error $J(u) - J(u_h^{(r)})$ and the estimated error η_h on a sequence of uniform meshes for different polynomial degrees. It can be seen, that the error estimator shows perfect effectivity $\mathrm{eff}_h \approx 1$ even on very coarse meshes. As the problem and the functional are linear, all three versions of the error identity; namely, primal, dual and mixed result in the same findings. This is also found numerically, hence, only one value η is given in the table. We find no difference, whether the weights are approximated using higher order simulations or by the reconstruction process. Hence, just one value is given.

Localization and adaptivity   Table 1 further shows, that going beyond second order finite elements does not result in an increased approximation order on uniform meshes. This is due to the limited regularity $u, z \in H^{3-\epsilon}(\Omega)$. Hence, we next consider localization of the error estimator and adaption of the meshes. For mesh adaption, we follow a simple equalization strategy that aims at balancing the element wise error indicators, such that a mesh element $K \in \Omega_h$ is being refined, if the error indicator $|\eta_K|$ is above average. While the classical localization technique described in Section 4.1 directly gives element-wise indicators η_K, we agglomerate the adjacent node-wise values to element wise values in the case of the PU approach in Section 4.3 and to patch-wise values for the filtering approach of Section 4.2.

In Fig. 3 we plot the error indicators η_K, η_i^{π} and η_i^{PU} as function over the domain Ω. By construction, the classical indicator values η_K are all positive, while the two variational settings η_i^{π} and η_i^{PU} show

both negative and positive values. Furthermore, the projection based indicators η_i^π show patch-wise fluctuations, while the PU approach yields smooth node-wise contributions.

Next, in Fig. 4, the error slopes obtained on uniform and adaptive meshes using quadratic and cubic finite elements are displayed. For refinement we used the PU localization. Further, we indicate error, estimator and effectivity index eff_h on a sequence of locally adapted meshes. Here optimal order of convergence in terms of unknowns with respect to the relation $N \sim h^{-2}$ is recovered. In addition, we observe good effectivities. For this simple and regular problems, all three localization techniques result in the same finite element meshes. Hence, we always considered the PU method only. Lastly, localizations based on the primal, dual or mixed formulation all result in the same adaptive meshes.

5.2 Configuration 2: Poisson problem with low regularity

As second test-case, we consider Poisson's equation on an L-shaped domain $\Omega_L = (-1, 1)^2 \setminus (-1, 0)^2$, where the right hand side is given by a Dirac in $x_0 = (-0.5, 0.5)$

$$-\Delta u = \delta_{x_0} \text{ in } \Omega_L, \quad u = 0 \text{ on } \partial\Omega_L. \tag{57}$$

As functional of interest, we consider the point evaluation in $x_1 = (0.5, -0.5)$ such that $J(\phi) = \phi(x_1)$. The adjoint problem corresponds to solving Poisson's equation $-\Delta z = \delta_{x_1}$ with a Dirac right hand side in x_1. Both the primal problem and the adjoint problem lack the required minimal regularity for the standard finite element theory, such that a regularization by averaging is required, e.g. by averaging over a small subdomain:

$$J_\epsilon(\phi) = \frac{1}{2\pi\epsilon^2} \int_{|x-x_1|<\epsilon} \phi(x)\,\mathrm{d}x, \tag{58}$$

where $\epsilon > 0$ is a small parameter not depending on h. As reference functional quantity we identify the value

$$\tilde{J} = 2.134929 \cdot 10^{-3} \pm 10^{-7}. \tag{59}$$

Due to limited regularity of primal and adjoint solution, we cannot expect high order convergence. Adaptivity based on good localization is important for an accurate approximation. We start by comparing the different localization techniques discussed in Section 4. Table 2 shows values obtained on a sequence of uniform meshes using piece-wise bilinear finite elements. Here, we provide the number of mesh elements, the error as well as effectivity index eff_h and indicator index ind_h, see (18) and (26), for the three different localization techniques based on the strong residual, the filtering approach and the PU method.

DoFs	$J(u) - J(u_h^{\mathrm{Q1}})$	eff_h^K	ind_h^K	eff_h^π	ind_h^π	$\mathrm{eff}_h^{\mathrm{PU}}$	$\mathrm{ind}_h^{\mathrm{PU}}$
48	$1.45 \cdot 10^{-4}$	1.80	1.80	1.06	2.92	1.06	1.74
192	$4.66 \cdot 10^{-5}$	2.04	2.04	1.24	3.27	1.24	2.19
768	$1.46 \cdot 10^{-5}$	1.98	1.98	1.08	2.80	1.08	2.03
3 072	$5.31 \cdot 10^{-6}$	1.83	1.83	1.02	2.28	1.02	1.77
12 288	$2.02 \cdot 10^{-6}$	1.66	1.66	1.02	1.85	1.02	1.50

Table 2: Configuration 2: calculations on uniform meshes, comparing the effectivity index and the indicator index for the three different localization techniques using the classical formulation η_h^K, the filtering approach η_h^π and the PU technique η_h^{PU}.

N	$\tilde{J} - J(u_h)$	$\mathrm{eff}_h^{\mathrm{PU}}$	$\mathrm{ind}_h^{\mathrm{PU}}$	N	$\tilde{J} - J(u_h)$	$\mathrm{eff}_h^{\mathrm{PU}}$	$\mathrm{ind}_h^{\mathrm{PU}}$
12	$7.98 \cdot 10^{-5}$	0.87	0.95	12	$4.15 \cdot 10^{-5}$	0.87	0.96
48	$1.25 \cdot 10^{-5}$	0.81	0.82	48	$6.39 \cdot 10^{-6}$	0.83	0.83
84	$4.98 \cdot 10^{-6}$	0.86	0.87	84	$2.43 \cdot 10^{-6}$	0.85	0.87
240	$1.85 \cdot 10^{-6}$	0.86	0.86	120	$8.91 \cdot 10^{-7}$	0.92	0.97
276	$6.67 \cdot 10^{-7}$	0.96	0.97	324	$2.80 \cdot 10^{-7}$	1.17	1.16

Table 3: Configuration 2: effectivity of the estimator and the indicators (partition of unity) on adaptive meshes with N elements. Left: discretization with piece-wise cubic, right with piece-wise quartic finite elements.
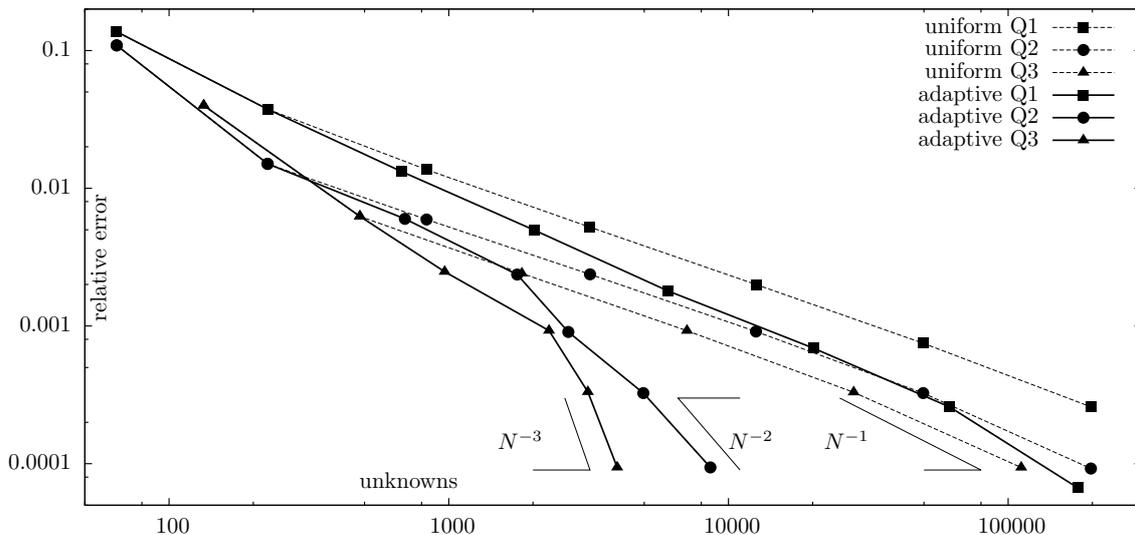

Figure 5: Configuration 2: relative error over number of unknowns for computations using linear, quadratic and cubic finite elements. Uniform vs. adaptive mesh refinement.

The two approximation techniques based on the variational formulation, η_h^π and η_h^{PU} result in a better effectivity index. Here, the formulation based on the strong residual and application of Cauchy-Schwarz shows a small overestimation of the error. For the local error indicators, we can only discover a small difference between the three techniques under investigation. The PU method yields a slightly better constant, due to the prevention of local oscillation, that is typical for the algebraic filtering approach.

Table 3 shows the functional error, the estimator effectivity and the indicator effectivity on a sequence of locally refined meshes using high order finite elements. Localization and refinement is based on the PU method. It shows, that the estimator and the localization are highly accurate for finite elements of high polynomial degree. In Fig. 5, we show the convergence of the functional values on sequence of uniform and adaptive meshes using linear, quadratic and cubic finite elements. First, we identify the necessity of local mesh adaptation, as an increasing polynomial degree does not result in better approximation order on uniform meshes. Using adaptivity and localization based on the PU method, we recover the optimal order of convergence (based on the relation $N = h^{-2}$) for all polynomial degrees. We however also see, that using isotropic adaptive meshes is not sufficiently able to resolve the singularities. The optimal order is only recovered on very fine meshes.
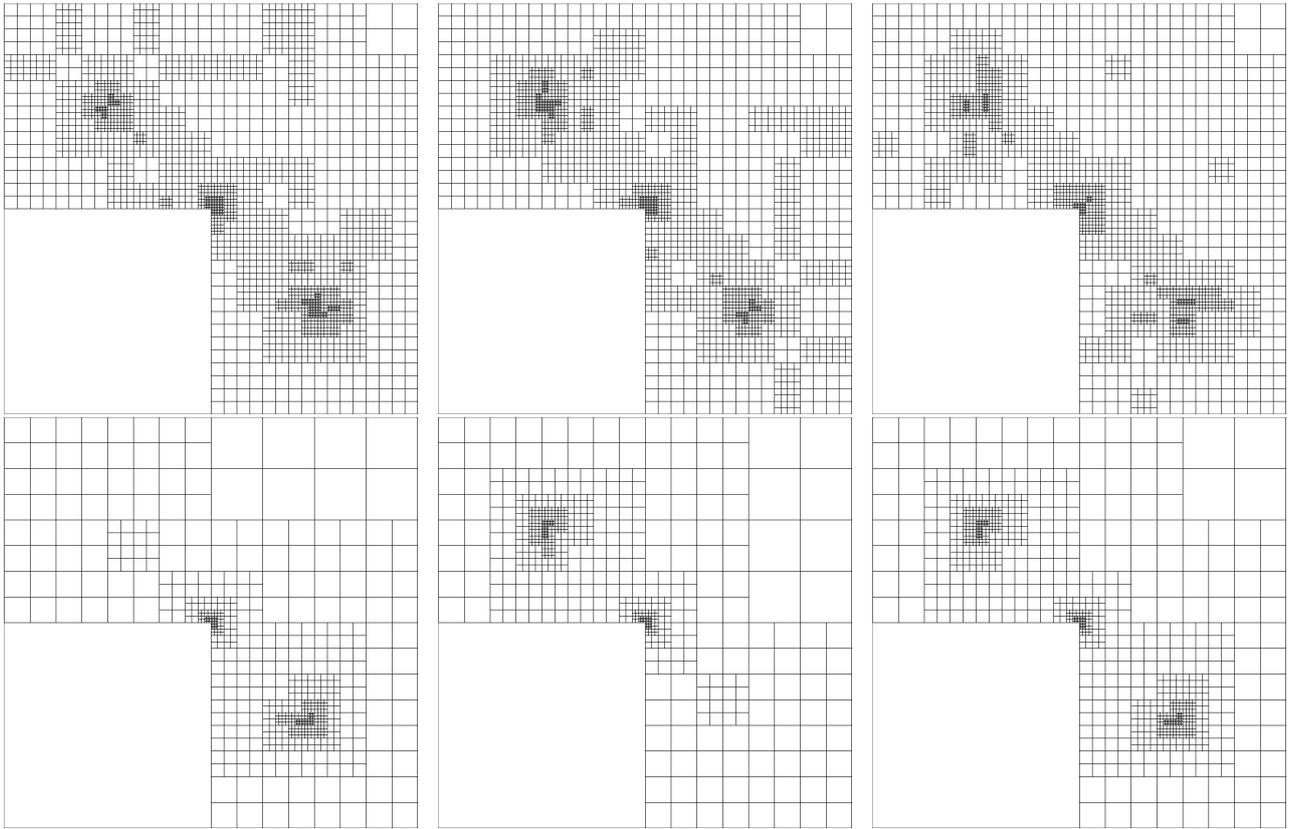
Figure 6: Configuration 2: Comparisons of patched meshes on refinement level 6 using primal (left), dual estimator (middle) and the mixed estimator (right). We compare from top to bottom the Q_1 and $Q4$ discretizations. In particular, the localization of the error estimator using higher order polynomials is observed.

Then, we use this test-case to study the differences in the localization behavior of primal, dual, and mixed estimators. For linear problems, all three versions of the error estimator (11), (12) and (17) are equivalent. Their localizations however will depend on a different weighting of approximation and interpolation errors as discussed in Section 4.

By the pollution effect, which is supposed to get stronger for higher order finite elements, see [2], we may experience different localizations. In Fig. 6 we show locally refined meshes obtained from localizations based on the three different error representation formulas; primal, adjoint and mixed - all meshes differ. The refinement of primal and adjoint formulation is mirrored at the line $x = y$ through the midpoint. The meshes corresponding to the mixed formulation can be regarded as a union of the two one-sided error representations, leading to symmetric meshes. The effect of different meshes gets stronger with higher polynomial degree, as is expected by the analysis of the pollution effect. Even though this example is atypical, we stress the importance of a correct balancing of primal and adjoint residuals for adaptive mesh refinement.
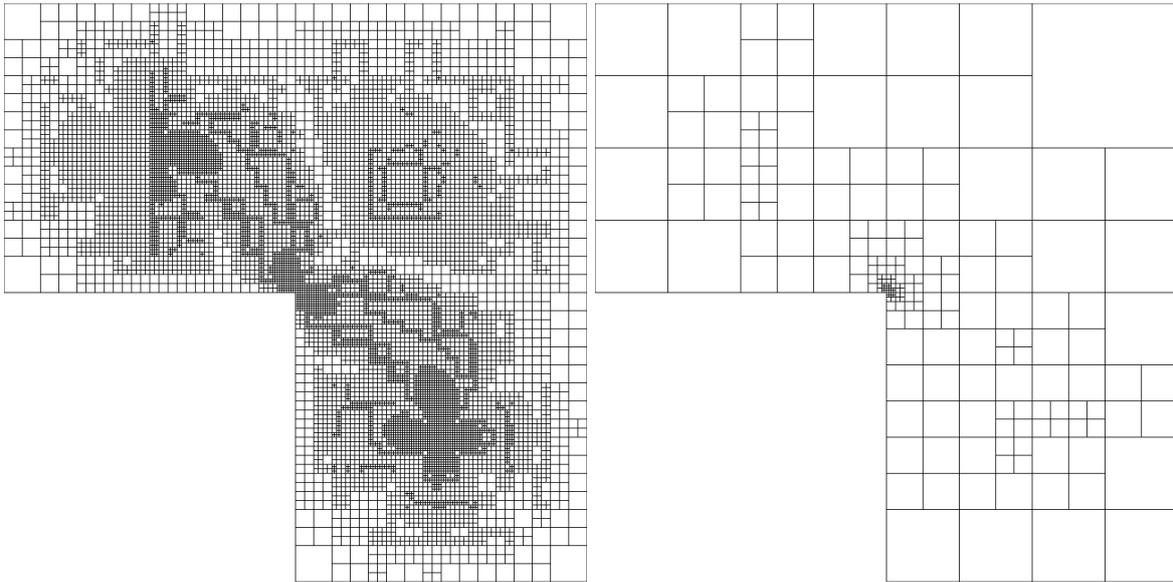
Figure 7: Configuration 2: Meshes on refinement level 6 without patch structure for Q_1 (left) and Q_4 (right) discretizations using the primal (non-symmetric) error estimator.
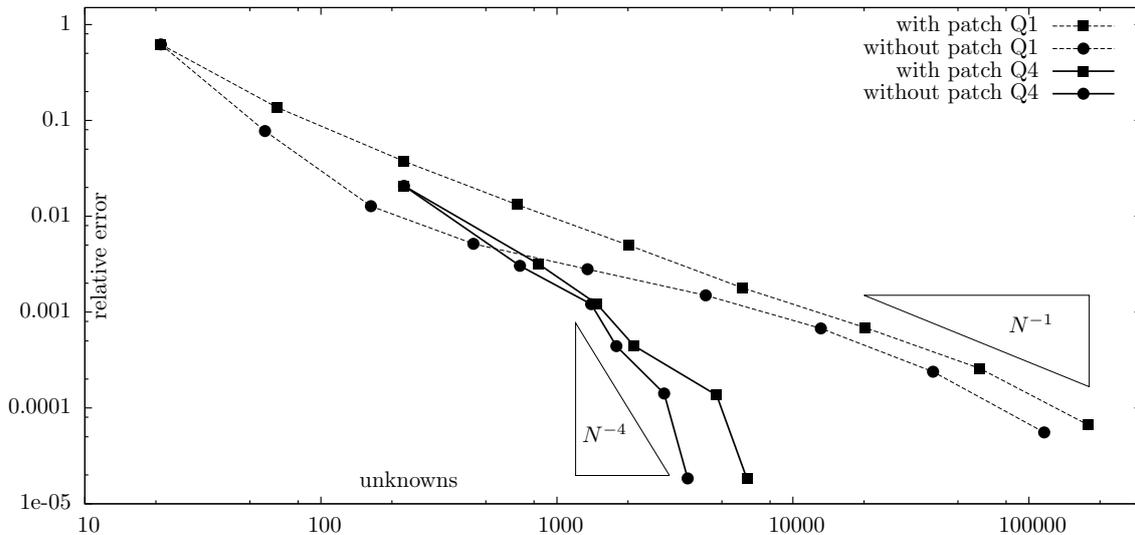

Figure 8: Configuration 2: Relative error over number of unknowns for the primal error estimator. The rate of convergence for the same order of finite element discretization is the same for patched and non-patched meshes. However, the error constant is slightly better using non-patched meshes.

Next, we demonstrate that the PU error estimator does not rely on a patched mesh structure. To see this, we run and compare four different settings; namely Q_1 and Q_4 discretizations using patched and non-patched meshes. For estimating the error, the primal formulation (11) is considered. Omiting the patch structure allows us to realize a sharper refinement towards singularities. On the other hand, without the patch structure, we cannot use the simple reconstruction technique for the weights that

DOF's	$J(u)$	error	η_h	eff_h	ind_h
3 072	136 070	4 230	730	0.17	3.00
8 448	134 626	2 786	215	0.08	2.63
22 224	133 487	1 647	556	0.33	2.54
57 504	132 808	968	523	0.54	2.48
152 592	132 358	518	325	0.63	2.56
378 384	132 152	312	235	0.76	2.30
871 800	132 023	183	152	0.83	2.31

Table 4: Configuration 3: convergence history on a sequence on locally refined meshes with number of unknowns, functional value $J(\mathbf{u})$, error $J(\mathbf{u}) - J(\mathbf{u}_h)$, effectivity $\text{eff}_h = \eta_h/|J(\mathbf{u}) - J(\mathbf{u}_h)|$ and indicator index $\text{ind}_h = \sum |\eta_i|/|J(\mathbf{u}) - J(\mathbf{u}_h)|$ for the PU localization.

has been described in Section 3.

In Fig. 7 we show meshes for $Q1$ and $Q4$ elements on refinement level 6. These meshes must be compared to the left column in Figure 6, where the same computations have been carried out - but here using patched meshes. In Fig. 8 we compare the error slopes for $Q1$ and $Q4$ elements on adaptive meshes with and without patches. We first identify the same order of convergence for both cases. The computations without patches give a slightly better error constant, as the strong singularities in this test-case can be better resolved. Comparing the meshes in Fig 7 with the left column in Fig. 6 shows, that by omiting the mesh structure, the regularity of the meshes in lessened. In particular the $Q1$ computations introduce many *islands* of refined cells, which lead to a large number of hanging ndoes. However, by consulting the error plots in Figure 8, we cannot detect a negative influence on the resulting error with respect to the number of unknowns (including all hanging nodes).

5.3 Configuration 3: Application to a nonlinear system

The benefit of a variational localization of the error identity is in particular given for complex nonlinear systems of partial differential equations, where assembling the strong formulation of the system is too costly. To exploit the localization technique we consider a nonlinear elasticity problem. For the construction shown in Fig. 9, we compute the deformation $\mathbf{u} : \Omega \to \mathbb{R}^3$ of an elastic beam $\Omega = (0, L) \times (0, D) \times (0, H)$ with length $L = 2$, depth $D = 1$ and height $H = 0.5$ under a given volume force $\mathbf{f} = -100\mathbf{e}_3$. The beam is attached $\mathbf{u} = 0$ on parts of the frontal boundary $\Gamma_D = (0 \times L/2) \times \{0\} \times (0, H)$. All other boundary parts are free. As quantity of interest, we measure the
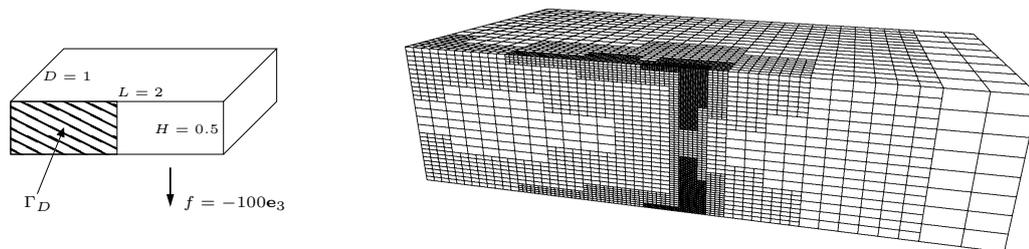

Figure 9: Configuration 3: Deformation of partially fixed elastic beam Ω_b under gravity. Left: sketch of the configuration. Right: locally adapted mesh.
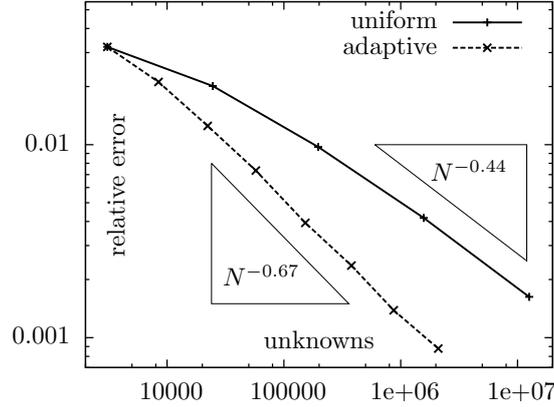
Figure 10: Configuration 3: comparison of the relative error on uniform and adaptive meshes.

total stress within the structure

$$J(\mathbf{u}) = \int_\Omega |\mathbf{F}_s \boldsymbol{\Sigma}_s|^2 \, \mathrm{d}x, \tag{60}$$

where $\mathbf{F}_s = \mathbf{I} + \nabla\mathbf{u}$ is the deformation gradient and $\boldsymbol{\Sigma}_s$ the 2nd Piola Kirchhoff stress tensor. The balance equation reads:

$$-\operatorname{div}\left(\mathbf{F}_s \boldsymbol{\Sigma}_s\right) = \mathbf{f}, \tag{61}$$

and the material is of St. Venant Kirchhoff type:

$$\boldsymbol{\Sigma}_s := 2\mu_s \mathbf{E}_s + \lambda_s \operatorname{tr}(\mathbf{E}_s)\mathbf{I}, \quad \mathbf{E}_s := \frac{1}{2}\left(\mathbf{F}_s^T \mathbf{F}_s - \mathbf{I}\right), \tag{62}$$

where $\mu_s = 5 \cdot 10^4$ and $\lambda_s = 10^5$. This corresponds to a very soft and compressible (Poisson's ratio $\nu_s = 1/3$) material. A reference value $\tilde{J} = 131\,840 \pm 0.1\%$ is obtained with the help of extrapolation on a very fine, initially graded mesh with about $4\,000\,000$ unknowns. The weak formulation of this problem with Dirichlet data on $\Gamma_D \subset \partial\Omega$ is to find $\mathbf{u} \in \mathcal{X} := H_0^1(\Omega; \Gamma_D)^3$

$$a(\mathbf{u}, \boldsymbol{\phi}) := (\mathbf{F}_s \boldsymbol{\Sigma}_s, \nabla\boldsymbol{\phi})_\Omega = (\mathbf{f}, \boldsymbol{\phi})_\Omega \quad \forall \boldsymbol{\phi} \in \mathcal{X}. \tag{63}$$

Here, we must consider the full nonlinear version of the DWR estimator (17) as introduced in Section 2.2. For its application, we must solve the adjoint problem and evaluate residuals of both primal and adjoint formulation. For the derivative $a'(\mathbf{u})(\mathbf{w}, \boldsymbol{\phi})$ it holds for $\mathbf{u}, \mathbf{w}, \boldsymbol{\phi} \in \mathcal{X}$

$$a'(\mathbf{u})(\mathbf{w}, \boldsymbol{\phi}) = \left(\frac{d\mathbf{F}_s(\mathbf{u})}{d\mathbf{u}}(\mathbf{w})\boldsymbol{\Sigma}_s + \mathbf{F}_s \frac{d\boldsymbol{\Sigma}_s(\mathbf{u})}{d\mathbf{u}}(\mathbf{w}), \nabla\boldsymbol{\phi}\right)_\Omega, \tag{64}$$

where $\partial_u \mathbf{F}_s(\mathbf{w}) = \nabla\mathbf{w}$ and

$$\begin{aligned}
\frac{d\boldsymbol{\Sigma}_s(\mathbf{u})}{d\mathbf{u}}(\mathbf{w}) &= 2\mu_s \frac{d\mathbf{E}_s(\mathbf{u})}{d\mathbf{u}}(\mathbf{w}) + \lambda \operatorname{tr}\left(\frac{d\mathbf{E}_s(\mathbf{u})}{d\mathbf{u}}(\mathbf{w})\right), \\
\frac{d\mathbf{E}_s(\mathbf{u})}{d\mathbf{u}}(\mathbf{w}) &= \frac{1}{2}\left(\nabla\mathbf{w}^T \mathbf{F}_s + \mathbf{F}_s^T \nabla\mathbf{w}\right).
\end{aligned} \tag{65}$$

While it is straightforward to set up a system matrix according to this derivative (or its transposed $a'(\mathbf{u})(\boldsymbol{\phi}, \mathbf{z})$); derivation of the strong formulation is cumbersome, mostly due to strongly coupled terms like

$$\left(\frac{\lambda_s}{2} \mathbf{F}_s \operatorname{tr} \left(\nabla \boldsymbol{\phi}^T \mathbf{F}_s + \mathbf{F}_s^T \nabla \boldsymbol{\phi} \right), \nabla \mathbf{z} \right). \tag{66}$$

How to separate and exempt the test-function $\boldsymbol{\phi}$ from all derivatives is not obvious. A fully variational evaluation of the error estimator and its localization has the advantage that strong formulations are not required. The functional derivative is given by

$$J'(\mathbf{u})(\boldsymbol{\phi}) = 2 \int_\Omega \mathbf{F}_s \boldsymbol{\Sigma}_s : \left(\nabla \mathbf{w} \boldsymbol{\Sigma}_s + \mathbf{F}_s \frac{d\boldsymbol{\Sigma}_s(\mathbf{u})}{d\mathbf{u}} \right) \, \mathrm{d}x. \tag{67}$$

In Table 4 we collect the functional value and error $|\tilde{J} - J(\mathbf{u}_h)|$, the error estimator as well as effectivity index and indicator index on a sequence of adaptive meshes. We get only a slight underestimation of the error even for this complex problem. The effectivity is close to one $\mathrm{eff}_h \approx 1$ for $h \to 0$. Localization with the PU method yields very good local approximations to the error. The indicator index ind_h is about two on all meshes. Once again we stress that this localization technique is not only very accurate, it also helps to realize an efficient application to complex nonlinear problems, where the evaluation of classical residuals in strong operator formulation has to be avoided.

In addition, the error slopes on uniform and on adapted meshes are shown in Fig. 10 . By using adaptivity, singularities appearing at edges are resolved and a higher order of convergence is obtained with respect to the number of unknowns. Finally, by choosing adaptive meshes, a relative error of 0.1% can be reached with less than $1\,000\,000$ unknowns compared to more than $10\,000\,000$ required unknowns on uniform meshes.

6 Conclusion

We have investigated several localization strategies for goal-oriented error estimators. In particular, we provided further insight to variational based localization techniques as the filtering approach by Braack & Ern [13] that is in particular useful for practical applications, where the evaluation of strong residuals is not possible due to the high effort or simply because a classical formulation of the adjoint problem is not available. Furthermore, we have introduced a new and very simple localization technique based on introducing a partition of unity (PU), that is locally effective for minimal regularity problems and that can be applied to general nonlinear problems in a straightforward manner. Specifically for high order finite elements this technique is highly accurate and the implementation is straightforward.

References

[1] M. Ainsworth and J. T. Oden. A posteriori error estimation in finite element analysis. *Computer Methods in Applied Mechanics and Engineering*, 142(1-2):1–88, 1997.

[2] I. Babuška, T. Strouboulis, A. Mathur, and C.S. Upadhyay. Pollution error in the *h*-version of the finite element method and the local quality of a-posteriori error estimators. Technical report, Institute for Physical Science and Technology, University of Maryland, 1994.

[3] W. Bangerth and O. Kayser-Herold. Data structures and requirements for hp finite element software. *ACM Trans. Math. Softw.*, 36:4:1–4:31, 2009.

[4] Wolfgang Bangerth, Timo Heister, and Guido Kanschat. *Differential Equations Analysis Library*, 2012.

[5] R. Becker, M. Braack, D. Meidner, T. Richter, and B. Vexler. The finite element toolkit GAS-COIGNE. HTTP://WWW.GASCOIGNE.UNI-HD.DE.

[6] R. Becker, M. Braack, and R. Rannacher. Numerical simulation of laminar flames at low Mach number with adaptive finite elements. *Combust. Theory Modelling*, 3:503–534, 1999.

[7] R. Becker, V. Heuveline, and R. Rannacher. An optimal control approach to adaptivity in computational fluid mechanics. *Int. J. Numer. Math. Fluids.*, 40:105–120, 2002.

[8] R. Becker and R. Rannacher. Weighted a posteriori error control in FE methods. In et al. H. G. Bock, editor, *ENUMATH'97*. World Sci. Publ., Singapore, 1995.

[9] R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: Basic analysis and examples. *East-West J. Numer. Math.*, 4:237–264, 1996.

[10] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. In A. Iserles, editor, *Acta Numerica 2001*, volume 37, pages 1–225. Cambridge University Press, 2001.

[11] H. Blum. Asymptotic error expansion and defect correction in the finite element method. Habilitationsschrift, Institut für Angewandte Mathematik, Universität Heidelberg, 1991. SFB-123 Preprint 640.

[12] H. Blum, Q. Lin, and R. Rannacher. Asymptotic error expansion and richardson extrapolation for linear finite elements. *Numer. Math.*, 49:11–37, 1986.

[13] M. Braack and A. Ern. Adaptive computation of reactive flows with local mesh refinement and model adaptation. In et. al. M. Feistauer, editor, *Numerical Mathematics and Advanced Applications, ENUMATH 2003*, pages 159–168. Springer, 2004.

[14] M. Braack and A. Ern. Coupling multimodeling with local mesh refinement for the numerical solution of laminar flames. *Combust. Theory Modelling*, 8(4):771–788, 2004.

[15] D. Braess. *Finite Elemente*. Springer, 1997.

[16] J. Carpio, H.L. Prieto, and R. Bermejo. Anisotropic "goal-oriented" mesh adaptivity for elliptic problems. *SIAM J. Sci. Comput.*, 35:861–885, 2013.

[17] C. Carstensen. Estimation of higher sobolev norm from lower order approximation. *SNA*, 42:2136–2147, 2004.

[18] C. Carstensen and S.A. Funken. Fully reliable localized error control in the FEM. *SIAM J. Sci. Comput.*, 21(4):1465–1484, 1999/00.

[19] J.A. Cottrell, T. Hughes, and Y. Bazilevs. *Isogeometric Analysis: Toward Integration of CAD and FEA*. Wiley, 2009.

[20] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson. Introduction to adaptive methods for differential equations. In A. Iserles, editor, *Acta Numerica 1995*, pages 105–158. Cambridge University Press., 1995.

[21] L.C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics. American Mathematical Society, 2010.

[22] P.W. Fick, E.H. van Brummelen, and K.G. van der Zee. On the adjoint-consistent formulation of interface conditions in goal-oriented error estimation and adaptivity for fluid-structure interaction. *Computer Methods in Applied Mechanics and Engineering*, 199:3369–3385, 2010.

[23] M.B. Giles and E. Süli. Adjoint methods for pdes: a posteriori error analysis and postprocessing by duality. *Acta Numerica 2002*, pages 145–236, 2002. A. Iserles, ed.

[24] T. Grätsch and K.J. Bathe. Goal-oriented error estimation in the analysis of fluid flows with structural interactions. *Computer Methods in Applied Mechanics and Engineering*, 2006:5673–5684, 195.

[25] V. Heuveline and F. Schieweck. h^1-interpolation on quadrilateral and hecahedral meshes with hanging nodes. *Computing*, 80:203–220, 2007.

[26] W. Hoitinga and E.H. van Brummelen. Goal-oriented adaptive methods for a boltzmann-type equation. volume 1333 of *AIP Conference Proceedings*, pages 81–86. American Institute of Physics, 2011.

[27] G. Kuru, C.V. Verhoosel, K.G. van der Zee, and E.H. van Brummelen. Goal-adaptive isogeometric analysis with hierarchical splines. *Comput. Methods Appl. Mech. Engrg.*, 270:270–292, 2014.

[28] D. Kuzmin and S. Korotov. Goal-oriented a posteriori error estimates for transport problems. *Math. Comp. Sim.*, 80(8):1674 – 1683, 2010.

[29] L. Machiels, Y. Maday, and A.T. Patera. A "flux-free" nodal neumann subproblem approach to output bounds for partial differential equations. *Comptes-Rendus de l'Acadmie des Sciences - Series I - Mathematics*, 330:249–254, 2000.

[30] P. Morin, N.H. Nochetto, and K.G. Siebert. Local problems on stars: a posteriori error estimators, convergence and performance. *Math. Comp.*, 72:1067–1097, 2003.

[31] R.H. Nochetto, A. Veeser, and M. Verani. A safeguarded dual weighted residual method. *IMA Journal of Numerical Analysis*, 29(1):126–140, 2009.

[32] J. Peraire and A.T. Patera. Bounds for linear-functional outputs of coercive partial differential equations: local indicators and adaptive refinement. In P. Ladeveze and J.T. Oden, editors, *Advances in Adaptive Computational Methods in Mechanics*, pages 199–215. Elsevier, Amsterdam, 1998.

[33] T. Richter. A posteriori error estimation and anisotropy detection with the dual weighted residual method. *Int. J. Numer. Math. Fluids.*, 62(1):90–118, 2010.

[34] T. Richter. Goal oriented error estimation for fluid-structure interaction problems. *Comput. Methods Appl. Mech. Engrg.*, 223-224:28–42, 2012.

[35] M. Schmich and B. Vexler. Adaptivity with dynamic meshes for space-time finite element discretizations of parabolic equations. *SIAM J. Sci. Compu.*, 30(1):369–393, 2008.

[36] R. Verfürt. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley/Teubner, New York-Stuttgart, 1996.

[37] T. Wick. *Adaptive Finite Element Simulation of Fluid-Structure Interaction with Application to Heart-Valve Dynamics*. PhD thesis, Heidelberg University, 2011.

[38] T. Wick. Goal-oriented mesh adaptivity for fluid-structure interaction with application to heart-valve settings. *Arch. Mech. Engrg.*, 59(6):73–99, 2012.

[39] K.G. van der Zee, E.H. van Brummelen, I. Akkerman, and R. de Borst. Goal-oriented error estimation and adaptivity for fluid-structure interaction using exact linearized adjoints. *Comput. Methods Appl. Mech. Engrg.*, 200:2738–2757, 2011.

[40] O. C. Zienkiewicz and J. Z. Zhu. The superconvergent patch recovery and a posteriori error estimates. part 1: The recovery technique. *Int. J. Numer. Meth. Engrg.*, 33(7):1331–1364, 1992.

[41] O.C. Zienkiewicz and J.Z. Zhu. The superconvergent patch recovery and a posteriori error estimates. part 2: Error estimates and adaptivity. *Int. J. Numer. Meth. Engrg.*, 33(7):1365–1382, 1992.