Local error analysis of the interior penalty discontinuous Galerkin method for second order elliptic problems

G. Kanschat* R. Rannacher*

December 16, 2002

Summary. A local *a priori* and *a posteriori* analysis is developed for the Galerkin method with discontinuous finite elements for solving stationary diffusion problems. The main results are an optimal-order estimate for the point-wise error and a corresponding a posteriori error bound. The proofs are based on weighted L^2-norm error estimates for discrete Green functions as already known for the 'continuous' finite element method.

1 Introduction

This paper deals with the numerical approximation of second-order elliptic boundary value problems by *interior penalty discontinuous Galerkin finite element method* (in short *dG method*) on triangular or rectangular meshes; see Arnold [1] and Arnold et al. [2], and also [16]) In this Galerkin method the trial and test functions may be discontinuous across inter-element boundaries while continuity is enforced approximately via a penalty technique. We analyze the local convergence behavior of this method. Only weak assumptions are imposed on the meshes; they are allowed to be refined locally and the cells may be shifted against each other. The proof relies on the techniques for estimating discrete Green functions developed in Frehse and Rannacher [12] for the "continuous" Galerkin method (in short *cG method*. Finally, we apply the error estimation techniques in Becker and Rannacher [7] to obtain a posteriori error estimates.

For simplicity, we concentrate on the Poisson equation,

$$-\Delta u = f \quad \text{in } \Omega, \tag{1}$$

with boundary conditions

$$u_{|\partial\Omega_D} = u^D, \quad \partial_n u_{|\partial\Omega_N} = u^N. \tag{2}$$

on a polygonal domain $\Omega \subset \mathbb{R}^2$. The boundary is assumed to be decomposed like $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$, and the boundary data u^D, u^N are smooth. We allow for such

*Institut für Angewandte Mathematik, Universität Heidelberg, INF 293, D-69120 Heidelberg, Germany

general nonhomogeneous boundary conditions in order to highlight the particular flexibility of the dG method. However, in order to avoid technical difficulties, we will nevertheless assume that the weak solution of this problem possesses H^2 regularity which very much constrains the generality of this setting. We note that all results of this paper can be extended to more general equations in two and three dimensions with variable coefficients as well as to systems of such equations.

The main result of this paper is the local *a priori* error estimate

$$|J_a^\varepsilon(u-u_h)| \;\leqslant\; ch^2\,\ell(\varepsilon)\,\|\nabla^2 u\|_{L^\infty(B_a)} + ch^2\|\nabla^2 u\|_\Omega, \tag{3}$$

where $J_a^\varepsilon(u)$ is a local weighted average close to the point value $u(a)$, $\ell(\varepsilon) \approx \ln(1/\varepsilon)$, and B_a an $\mathcal{O}(1)$ neighborhood of the point $a \in \overline{\Omega}$. This result is complemented by an *a posteriori* error estimate of the form

$$|J_a^\varepsilon(u-u_h)| \;\leqslant\; \sum_{K\in\mathbb{T}_h} \left\{ \varrho_K^{(1)}\omega_K^{(1)} + \varrho_K^{(2)}\omega_K^{(2)} + \varrho_K^{(3)}\omega_K^{(3)} \right\}, \tag{4}$$

where $\varrho_K^{(i)}$ are cell and edge residuals computable from the data and the approximation u_h, and $\omega_K^{(i)}$ are mesh-dependent weights involving the solution z of the adjoint problem associated with the error functional $J_a^\varepsilon(\cdot)$. The asymptotic sharpness of these estimates is confirmed by numerical tests.

2 The interior penalty method

The polygonal domain $\Omega \subset \mathbb{R}^2$ is assumed to be regular, in the sense that the following regularity property holds:

Assumption 1. *Any weak solution* $v \in V := \{\phi \in H^1(\Omega),\ \phi_{|\partial\Omega_D} = 0\}$ *of the boundary value problem*

$$-\Delta v = f \ \text{ in } \ \Omega, \quad v_{|\partial\Omega_D} = 0, \ \ \partial_n v_{|\partial\Omega_N} = 0,$$

is in $H^2(\Omega)$ *and satisfies the a priori estimate*

$$\|v\|_{H^2(\Omega)} \leqslant c\|f\|_{L^2(\Omega)}. \tag{5}$$

Here and below, $H^s(\Omega)$ is the standard Sobolev space of order s on Ω. We will use the notation $\|v\| = \|v\|_{L^2(\Omega)}$ and $\|v\|_S = \|v\|_{L^2(S)}$ for the L^2 norm over Ω or any subset $S \subset \overline{\Omega}$, respectively.

Sufficient conditions for (5) on the interior angles of $\partial\Omega$ can be found, e.g., in Grisvard [14]. Under the same conditions, also Problem (1), (2), with nonhomogeneous boundary data, has a unique solution $u \in V \cap H^2(\Omega)$.

Let $\mathbb{T}_h$ be decompositions of Ω into (open) cells K, triangles or quadrilaterals, with boundaries ∂K. The cells are assumed to satisfy the uniform shape condition (see Brenner and Scott [10]), but are allowed to vary in size for local mesh adaptation. In particular, the meshes are not required to be "conforming", i.e., cells may be shifted

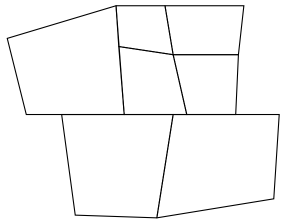

Figure 1: *An admissible mesh patch.*

against each other and possess "hanging" nodes. Figure 1 shows a generic situation for such a mesh. We set

$$h_K := \operatorname{diam}(K), \quad h := \max_{K \in \mathbb{T}_h} h_K.$$

The essential condition on the meshes is expressed in the following assumption.

Assumption 2. *The family of meshes $\{\mathbb{T}_h\}_{h>0}$ satisfies the uniform shape condition and the variation of the mesh size is smooth, i.e., for any two cells $K, K' \in \mathbb{T}_h$ with $K \cap K' \neq \emptyset$, the diameters are bounded by*

$$\frac{h_K}{h_{K'}} \leqslant c_{\mathbb{T}}. \tag{6}$$

with a constant $c_{\mathbb{T}}$ independent of h. In particular, this implies that the maximum number of neighboring cells of any $K \in \mathbb{T}_h$ sharing a non-trivial part of a boundary edge of $K \in \mathbb{T}_h$ is bounded by some number $\mu \geq 1$, uniformly for h.

On the mesh $\mathbb{T}_h$, we define the finite element space

$$V_h := \left\{ v_h : \overline{\Omega} \to \mathbb{R} \middle| \, v_{h|K} \in P(K), K \in \mathbb{T}_h \right\},$$

where $P(K)$ denotes the space of linear or (isoparametric) d-linear polynomials on K. The functions in V_h do not have to satisfy any boundary conditions. Additionally, we define the "broken Sobolev spaces"

$$H^s(\mathbb{T}_h) = \left\{ v \in L^2(\Omega) \mid \forall K \in \mathbb{T}_h : v|_K \in H^s(K) \right\}.$$

Boundary conditions as well as continuity between grid cells will be enforced weakly. To this end, we introduce the following notation for jumps and averages across inter-element boundaries: let Γ be the part of the boundary of K shared with the boundary of K'. Let v and v' be the restrictions of a function $v \in H^1(\mathbb{T}_h)$ to the current grid cell K and its neighbors K' with natural traces on ∂K and $\partial K'$, respectively. Then, we define for cell edges $\Gamma \subset \partial K$:

$$[v]_{|\Gamma} := \begin{cases} v_{|\Gamma} - v'_{|\Gamma}, & \text{if } \Gamma \subset \partial K \backslash \partial \Omega \\ 2v_{|\Gamma}, & \text{if } \Gamma \subset \partial \Omega_D \\ 0, & \text{if } \Gamma \subset \partial \Omega_N \end{cases},$$

and

$$\{v\}_{|\Gamma} := \begin{cases} \frac{1}{2}(v_{|\Gamma} + v'_{|\Gamma}), & \text{if } \Gamma \subset \partial K \backslash \partial \Omega \\ v_\Gamma, & \text{if } \Gamma \subset \partial \Omega_D \\ 0, & \text{if } \Gamma \subset \partial \Omega_N \end{cases}.$$

To define the interior penalty method for problem (1), (2), we introduce the symmetric bilinear form

$$a_h(v,w) := \sum_{K\in\mathbb{T}_h} \Big\{(\nabla v,\nabla w)_K + \frac{\kappa}{h_K}([v],[w])_{\partial K} - \frac{1}{2}(\{\partial_n v\},[w])_{\partial K} - \frac{1}{2}([v],\{\partial_n w\})_{\partial K}\Big\}.$$

on $H^2(\mathbb{T}_h)$. Remark that this bilinear form differs from the one found in the literature by increasing the penalty term at the boundary by a factor of four. Then, we seek $u_h \in V_h$ such that

$$a_h(u_h,\phi_h) = f_h(\phi_h) \quad \phi_h\in V_h, \tag{7}$$

where the linear form on the right hand side is

$$f_h(\phi) := \sum_{K\in\mathbb{T}_h} \Big\{(f,\phi)_K - (u^D,\partial_n\phi)_{\partial K_D} + (u^N,\phi)_{\partial K_N} + \frac{4\kappa}{h_K}(u^D,\phi)_{\partial K_D}\Big\},$$

with $\partial K_D := \partial K\cap\partial\Omega_D$ and $\partial K_N := \partial K\cap\partial\Omega_N$. With the form $a_h(\cdot,\cdot)$, we associate the "energy norm"

$$|||v|||_h \ := \ \Big(\sum_{K\in\mathbb{T}_h}\Big\{\|\nabla v\|_K^2 + \frac{\kappa}{h_K}\|[v]\|_{\partial K}^2\Big\}\Big)^{1/2},$$

defined on $H^2(\mathbb{T}_h)$, with "penalty parameter" $\kappa>0$. Let c_0 be the constant in the local inverse relation for polynomials

$$\|\partial_n\phi\|_{\partial K} \leqslant c_0 h_K^{-1/2}\|\nabla\phi\|_K, \quad \phi\in P(K),\, K\in\mathbb{T}_h. \tag{8}$$

Then, we estimate for $v_h\in V_h$ as follows :

$$\begin{aligned}|(\{\partial_n v_h\},[v_h])_{\partial K}| &\leqslant \frac{h_K}{2\kappa}\|\{\partial_n v_h\}\|_{\partial K}^2 + \frac{\kappa}{2h_K}\|[v_h]\|_{\partial K}^2\\ &\leqslant \frac{c_0^2}{2\kappa}\|\nabla v_h\|_{\tilde K}^2 + \frac{\kappa}{2h_K}\|[v_h]\|_{\partial K}^2,\end{aligned}$$

where

$$\tilde K = \bigcup_{K\cap K'\neq\emptyset} K'. \tag{9}$$

This implies that, with the parameter μ from Assumption 2, there holds

$$\begin{aligned}\Big|\sum_{K\in\mathbb{T}_h}(\{\partial_n v_h\},[v_h])_{\partial K}\Big| &\leqslant \sum_{K\in\mathbb{T}_h}\Big\{\frac{h_K}{2\kappa}\|\{\partial_n v_h\}\|_{\partial K}^2 + \frac{\kappa}{2h_K}\|[v_h]\|_{\partial K}^2\Big\}\\ &\leqslant \sum_{K\in\mathbb{T}_h}\Big\{\frac{c_0^2}{2\kappa}\|\nabla v_h\|_{\tilde K}^2 + \frac{\kappa}{2h_K}\|[v_h]\|_{\partial K}^2\Big\}\\ &\leqslant \sum_{K\in\mathbb{T}_h}\Big\{\frac{\mu c_0^2}{2\kappa}\|\nabla v_h\|_K^2 + \frac{\kappa}{2h_K}\|[v_h]\|_{\partial K}^2\Big\}.\end{aligned}$$

Hence, for

$$\kappa > c_0^2\mu, \tag{10}$$

we conclude that the bilinear form $a_h(.,.)$ is coercive on V_h:

$$\begin{aligned} a_h(v_h,v_h) &= \sum_{K\in\mathbb{T}_h} \Big\{ \|\nabla v_h\|_K^2 + \frac{\kappa}{h_K}\|[v_h]\|_{\partial K}^2 - (\{\partial_n v_h\},[v_h])_{\partial K} \Big\} \\ &\geq \sum_{K\in\mathbb{T}_h} \Big\{ \|\nabla v_h\|_K^2 + \frac{\kappa}{h_K}\|[v_h]\|_{\partial K}^2 - \frac{1}{2}\|\nabla v_h\|_{\tilde{K}}^2 - \frac{\kappa}{2h_K}\|[v_h]\|_{\partial K}^2 \Big\} \\ &\geq \frac{1}{2}|||v_h|||_h^2, \quad v_h \in V_h. \end{aligned} \tag{11}$$

This implies that problem (7) has a unique solution $u_h \in V_h$. For the "continuous" solution $u \in H^2(\Omega)$ of (1), (2), there holds on each interior cell edge:

$$[u]_{|\Gamma} = 0, \quad \{\partial_n u\}_{|\Gamma} = \partial_n u_{|\Gamma} \in L^2(\Gamma),$$

and along the boundary:

$$[u]_{|\Gamma} = u^D_{|\Gamma}, \quad \{\partial_n u\}_{|\Gamma} = u^N_{|\Gamma}.$$

This implies that

$$a_h(u,\phi_h) = \sum_{K\in\mathbb{T}_h} \Big\{ (\nabla u, \nabla\phi_h)_K - (\partial_n u, [\phi_h])_{\partial K} \Big\} = (f,\phi_h), \quad \phi_h \in V_h.$$

Consequently, the error $e := u - u_h$ satisfies the Galerkin orthogonality relation

$$a_h(e,\phi_h) = 0, \quad \phi_h \in V_h. \tag{12}$$

3 A priori error analysis

In the following, we will use the symbol c for a generic positive constant which may vary with the context, but is always independent of h and of the other currently free parameters. We prepare for the a priori error analysis by stating the local interpolation property of the trial spaces V_h. Let $I_h : H^k(\mathbb{T}_h) \to V_h$ be defined as the cell-wise L^2 projection satisfying

$$\|v - I_h v\|_K + h_K\|\nabla(v - I_h v)\|_K \leqslant c h_K^k \|\nabla^k v\|_K, \tag{13}$$

$$\|v - I_h v\|_{\partial K} + h_K\|\partial_n(v - I_h v)\|_{\partial K} \leqslant c h_K^{k-1/2} \|\nabla^k v\|_K, \tag{14}$$

for $v \in H^k(K)$ and $k \in \{1,2\}$. Then, we have the following well-known L^2 a priori error estimates (see Arnold [1]):

$$\|u - u_h\| \leqslant ch|||u - u_h|||_h \leqslant ch^2\, \|\nabla^2 u\|_{L^2(\Omega)}. \tag{15}$$

In the following, we will derive local error estimates. Since we want to allow for arbitrary local mesh refinement, it is appropriate to replace point values by local averages.

This is suggested by our experience with local a posteriori error estimation (see Becker and Rannacher [7]). For some point $a \in \overline{\Omega}$ and parameter $\varepsilon > 0$, let $\mathcal{U}_a^\varepsilon \subset \overline{\Omega}$ be a set, such that

$$a \in \mathcal{U}_a^\varepsilon, \quad \operatorname{diam}(\mathcal{U}_a^\varepsilon) = \varepsilon, \quad c\varepsilon^2 \leqslant |\mathcal{U}_a^\varepsilon| \leqslant \varepsilon^2. \tag{16}$$

Typically, we consider a ball of radius ε around a or a patch of mesh cells containing a. Further, let $\delta_a^\varepsilon \in L^\infty(U_a^\varepsilon)$ be a weighting function (regularized Dirac function) with support in $\mathcal{U}_a^\varepsilon$ and

$$\int_{\mathcal{U}_a^\varepsilon} \delta_a^\varepsilon(x)\,dx = 1, \quad \|\delta_a^\varepsilon\|_{L^\infty(\mathcal{U}_a^\varepsilon)} \leqslant c\varepsilon^{-2}.$$

Then, we define the evaluation functional

$$J_a^\varepsilon(u) = \int_{\mathcal{U}_a^\varepsilon} u(x)\delta_a^\varepsilon(x)\,dx, \tag{17}$$

which approximates the point evaluation $u(a)$.

Theorem 1. *Let, for some point $a \in \overline{\Omega}$, B_a be a ball around a with a fixed radius $\varrho > \varepsilon$. Then, for $h \leqslant \frac{1}{2}\varrho$, there holds*

$$|J_a^\varepsilon(u) - J_a^\varepsilon(u_h)| \leqslant ch^2\,\ell(\varepsilon)\|\nabla^2 u\|_{L^\infty(B_a)} + ch^2\|\nabla^2 u\|, \tag{18}$$

with $\ell(\varepsilon) := |\log\varepsilon| + 1$.

The proof of Theorem 1 will be given in the next section. For those readers who are experienced with pointwise error estimation in finite element methods, we remark on some technical aspects. The proof is by a duality technique employing regularized Green functions and weighted L^2-norms as developed in [12] for the standard cG scheme. This argument is, on the one hand, more complicated for the dG scheme because of the presence of additional edge-integral terms, but on the other hand also easier since the approximation properties of the dG method are more local than those of its "continuous" counterpart.

Corollary 1. *Assume that the mesh is not locally refined around the point $a \in \overline{\Omega}$, i.e., the diameter h_a of the cell K_a containing a is such that*

$$ch \leqslant h_a \leqslant h. \tag{19}$$

Then, (18) implies the point-wise error estimate

$$|u(a) - u_h(a)| \leqslant ch^2\,\ell(h)\|\nabla^2 u\|_{L^\infty(B_a)} + ch^2\|\nabla^2 u\|. \tag{20}$$

Proof. Let $\varepsilon = h_a$ and $\mathcal{U}_a^\varepsilon = K_a$ in (17). There exists a $\psi_a \in P(K_a)$, such that for any $q \in P(K_a)$ there holds

$$q(a) = \int_{K_a} q(x)\psi_a(x)\,dx.$$

Choosing $q \equiv 1$, we see that

$$\int_{K_a} \psi_a(x)\,dx = 1, \quad \max_{K_a} |\psi_a| \leqslant c|K_a|^{-1},$$

as is required for the application of Theorem 1. Using this construction, we derive the estimate

$$|e(a)| \leqslant 2 \max_{K_a} |u - I_h u| + \Big| \int_{K_a} e\psi_a \, dx \Big|.$$

Hence, by the approximation properties of Q_1 elements and the result of Theorem 1, we obtain

$$|e(a)| \leqslant ch^2 \, \|\nabla^2 u\|_{L^\infty(K_a)} + cch^2 \, \ell(h) \, \|\nabla^2 u\|_{L^\infty(B_a)} + ch^2 \, \|\nabla^2 u\|.$$

which completes the proof. □

4 Proof of Theorem 1

(i) First, we introduce some notation. Let

$$\sigma = \sigma_a := (|x-a|^2 + \varepsilon^2)^{1/2} \tag{21}$$

be the regularized distance function. The diameter of the cell K_a containing a is denoted by h_a. Furthermore, σ_h is the the cellwise constant L^2-projection of σ. In the following, we will use c as a generic constant which is independent of the mesh size h, the cell $K \in \mathbb{T}_h$, and the parameters ε and θ. By a simple calculation, using Assumption 2, we find the following relations (see [12]):

$$h_K \leqslant c\sigma_K, \quad K \in \mathbb{T}_h, \tag{22}$$

$$\max_K \sigma \leqslant c \min_K \sigma, \quad K \in \mathbb{T}_h, \tag{23}$$

$$\|\sigma^{-1}\| \leqslant c \, \ell(\varepsilon)^{1/2}. \tag{24}$$

We recall the local trace estimates

$$\|v\|_{\partial K} \leqslant ch_K^{-1/2} \|v\|_K + ch_K^{1/2} \|\nabla v\|_K, \tag{25}$$

$$\|\partial_n w\|_{\partial K} \leqslant ch_K^{-1/2} \|\nabla w\|_K + ch_K^{1/2} \|\nabla^2 w\|_K, \tag{26}$$

which hold for any functions $v \in H^1(K)$ and $w \in H^2(K)$, respectively, with constants independent of h_K. These estimates follow from their well known equivalents on the unit cell by a simple scaling argument.

(ii) Next, we define a 'regularized' Green function $g = g_a^\varepsilon \in V \cap H^2(\Omega)$ associated with the regularized Dirac function δ_a^ε by

$$(\nabla\phi, \nabla g) = (\phi, \delta_a^\varepsilon) \quad \forall \phi \in V. \tag{27}$$

Let $g_h \in V_h$ be the corresponding discontinuous Ritz projection defined by

$$a_h(\phi_h, g_h) = (\phi_h, \delta_a^\varepsilon) \quad \forall \phi_h \in V_h. \tag{28}$$

The following lemma provides the key estimates for the proof of the theorem.

Lemma 1. *With the notation introduced above, the following estimates hold:*

$$\|\nabla g\| + \|\sigma_h \nabla^2 g\| \leqslant c\,\ell(\varepsilon)^{1/2}, \tag{29}$$

$$\|g - g_h\| + h\|\nabla(g - g_h)\|_h \leqslant ch\,\ell(\varepsilon), \tag{30}$$

$$\|\sigma_h(g - I_h g)\| + h\|\sigma_h \nabla(g - I_h g)\|_h \leqslant ch\,\ell(\varepsilon), \tag{31}$$

where $\ell(\varepsilon) := |\log\varepsilon| + 1$ *and* $\|\cdot\|_h$ *denotes the cellwise semi-norm on* $H^1(\mathbb{T}_h)$ *.*

Proof. (i) We begin with proving (29). By definition, there holds

$$\|\nabla g\|^2 = (g, \delta_a^\varepsilon) \leqslant \max_{\mathcal{U}_a^\varepsilon} |g|. \tag{32}$$

Now, let $G(x, y)$ be the true Green function of the Laplacian on the domain Ω, for which we have the following bound (see [12]):

$$|G(x, y)| \leqslant c\big|\log|x - y|\big| + c$$

$$|g(x)| = |(G(x, \cdot), \delta)| \leqslant \frac{1}{|\mathcal{U}_a^\varepsilon|}\int_{\mathcal{U}_a^\varepsilon} |G(x, y)|dy \leqslant c\,\ell(\varepsilon).$$

This implies $\|\nabla g\| \leqslant c\,\ell(\varepsilon)^{1/2}$.
(ii) Next, we estimate using the regularity estimate (5),

$$\begin{aligned}
\|\sigma_h \nabla^2 g\|^2 &\leqslant c\|\sigma \nabla^2 g\|^2 = c\sum_{i=1}^d \|x_i \nabla^2 g\|^2 + \theta^2\varepsilon^2\|\nabla^2 g\|^2 \\
&\leqslant c\sum_{i=1}^d \|\nabla^2(x_i g)\|^2 + c\|\nabla g\|^2 + \theta^2\varepsilon^2\|\nabla^2 g\|^2 \\
&\leqslant c\sum_{i=1}^d \|\Delta(x_i g)\|^2 + \|\nabla g\|^2 + \theta^2\varepsilon^2(\|\Delta g\|^2 + \|\nabla g\|^2) \\
&\leqslant c\sum_{i=1}^d \|x_i \Delta g\|^2 + c\|\nabla g\|^2 + \theta^2\varepsilon^2\|\Delta g\|^2 \\
&\leqslant c\|\sigma \Delta g\|^2 + c\|\nabla g\|^2.
\end{aligned}$$

Consequently, by the definition of g,

$$\|\sigma_h \nabla^2 g\|^2 \leqslant c\|\sigma \delta_a^\varepsilon\|^2 + c\|\nabla g\|^2 \leqslant \ell(\varepsilon).$$

which implies the asserted bound (29).
(iii) For notational convenience, we introduce the following extended “energy norm”:

$$|||v|||_h^+ := \Big(\sum_{K\in\mathbb{T}_h} \Big\{\|\nabla v\|_K^2 + \frac{\kappa}{h_K}\|[v]\|_{\partial K}^2 + \frac{h_K}{\kappa}\|\{\partial_n v\}\|_{\partial K}^2\Big\}\Big)^{1/2}.$$

In virtue of the inverse estimate (8), for discrete functions $v_h \in V_h$, there holds

$$|||v_h|||_h \leqslant |||v_h|||_h^+ \leqslant c\,|||v_h|||_h. \tag{33}$$

Using the estimates (13) and (14) and the Sobolev inequality

$$\|\nabla v\| \leqslant c\,\|\nabla^2 v\|_{L^1(\Omega)}, \quad v \in V \cap H^2(\Omega),$$

we conclude for $v \in V \cap H^2(\Omega)$ the global approximation estimate

$$|||v - I_h v|||_h^+ \leqslant c \begin{cases} h\|\nabla^2 v\|_{L^2(\Omega)} \\ \|\nabla^2 v\|_{L^1(\Omega)} \end{cases}. \tag{34}$$

For $v, w \in V \oplus V_h$, there holds

$$|a_h(v,w)| \leqslant \sum_{K\in\mathbb{T}_h} \Big\{ \|\nabla v\|_K \|\nabla w\|_K + \kappa h_K \|[v]\|_{\partial K} \|[w]\|_{\partial K} + \|\{\partial_n v\}\|_{\partial K} \|[w]\|_{\partial K} + \|[v]\|_{\partial K} \|\{\partial_n w\}\|_{\partial K} \Big\},$$

and, consequently,

$$|a_h(v,w)| \leqslant c\,|||v|||_h^+\, |||w|||_h^+. \tag{35}$$

By the coercivity estimate (11) and Galerkin orthogonality, we conclude for the dual error $e^* := g - g_h$ that

$$\begin{aligned} |||I_h e^*|||_h^2 &\leqslant a_h(I_h e^*, I_h e^*) = a_h(e^*, I_h e^*) + a_h(g - I_h g, I_h e^*) \\ &= a_h(g - I_h g, I_h e^*). \end{aligned}$$

Consequently, by (35) and (33),

$$|||e^*|||_h \leqslant |||I_h e^*|||_h + |||g - I_h g|||_h \leqslant c\,|||g - I_h g|||_h^+.$$

Hence, the approximation estimate (34) yields

$$|||e^*|||_h \leqslant c\|\nabla^2 g\|_{L^1(\Omega)}. \tag{36}$$

Next, we employ a duality argument. Let $z \in V$ be the solution of

$$-\Delta z = e^* \|e^*\|^{-1} \ \text{ in } \Omega,$$

satisfying the a priori estimate $\|z\|_{H^2} \leqslant c$. By (34), we have

$$|||z - I_h z|||_h^+ \leqslant ch\,\|\nabla^2 z\| \leqslant ch.$$

Using Galerkin orthogonality, there holds

$$\begin{aligned} \|e^*\| = (e^*, -\Delta z) &= \sum_{K\in\mathbb{T}_h} \Big\{ (\nabla e^*, \nabla z)_K + (e^*, \partial_n z)_{\partial K} \Big\} \\ &= a_h(e^*, z) = a_h(e^*, z - I_h z). \end{aligned}$$

Consequently,

$$\|e^*\| \leqslant c\,|||e^*|||_h^+\, |||z - I_h z|||_h^+ \leqslant ch\,|||e^*|||_h^+ \leqslant ch\,\|\nabla^2 g\|_{L^1(\Omega)}.$$

It remains to bound the L^1 norm on the right. There holds

$$\|\nabla^2 g\|_{L^1(\Omega)} \leqslant \|\sigma^{-1}\|\,\|\sigma \nabla^2 g\|$$

which, in view of (24) and (29) yields the asserted L^2 error estimate (30).
(iv) Finally, we derive the interpolation estimate (31). In virtue of the local interpolation estimates (13), we have

$$\|\sigma_h(g - I_h g)\|^2 \leqslant \sum_{K\in\mathbb{T}_h} \sigma_{h,K}^2 \|\sigma_h(g - I_h g)\|_K^2 \\ \leqslant ch^4 \sum_{K\in\mathbb{T}_h} \sigma_{h,K}^2 \|\nabla^2 g\|_K^2 = ch^4 \|\sigma_h \nabla^2 g\|^2,$$

and analogously,

$$\|\sigma_h \nabla(g - I_h g)\|_h^2 \leqslant ch^2 \|\sigma_h \nabla^2 g\|^2.$$

Combining this with the a priori bound (29) completes the proof. □

(iii) Using the discrete Dirac function δ_a^ε, we have

$$|J_a^\varepsilon(e)| \leqslant c|(e, \delta_a^\varepsilon)| + c\|u - I_h u\|_{L^\infty(K_a)}.$$

Hence, the interpolation estimate

$$\|u - I_h u\|_{L^\infty(K_a)} \leqslant ch^2 \|\nabla^2 u\|_{L^\infty(K_a)},$$

yields

$$|J_a^\varepsilon(e)| \leqslant c|(e, \delta_a^\varepsilon)| + ch^2 \|\nabla^2 u\|_{L^\infty(K_a)}. \tag{37}$$

For abbreviation, we set $e^* := g - g_h$. Then, using Galerkin orthogonality, for e and e^*, we conclude

$$(e, \delta_a^\varepsilon) = a_h(e, g) = a_h(e, e^*) = a_h(\eta, e^*) \\ = \sum_{K\in\mathbb{T}_h} \Big\{ (\nabla\eta, \nabla e^*)_K + \frac{\kappa}{h_K}([\eta], [e^*])_{\partial K} - \frac{1}{2}(\{\partial_n \eta\}, [e^*])_{\partial K} \\ - \frac{1}{2}([\eta], \{\partial_n e^*\})_{\partial K} \Big\},$$

where $\eta := u - I_h u$ and abbreviate

$$|(e, \delta_a^\varepsilon)| \leqslant \Theta_1^{1/2} \Theta_2^{1/2},$$

where, with some parameter $\gamma \in (0, 1]$,

$$\Theta_1 = \sum_{K\in\mathbb{T}_h} \Big\{ \|\sigma_h^{-1} \nabla\eta\|_K^2 + \frac{c\kappa}{\gamma h_K}(\{\sigma_h^2\}^{-1}[\eta], [\eta])_{\partial K} \\ + \frac{h_K}{\kappa}(\{\sigma_h^2\}^{-1}\{\partial_n \eta\}, \{\partial_n \eta\})_{\partial K} \Big\},$$

$$\Theta_2 = \sum_{K\in\mathbb{T}_h} \Big\{ \|\sigma_h \nabla e^*\|_K^2 + \frac{\kappa}{h_K}(\{\sigma_h^2\}[e^*], [e^*])_{\partial K} \\ + \gamma h_K(\{\sigma_h^2\}\{\partial_n e^*\}, \{\partial_n e^*\})_{\partial K} \Big\}.$$

The two terms Θ_1 and Θ_2 will be estimated separately.

(iv) First, we estimate Θ_1 . For $\tilde{K}$ defined in (9), we conclude from the interpolation estimates (13) and (14), observing Assumption 2, that

$$\|\nabla\eta\|_K + h_K^{-1/2}\|[\eta]\|_{\partial K} + h_K^{1/2}\|\{\partial_n\eta\}\|_{\partial K} \leqslant c h_K\|\nabla^2 u\|_{\tilde{K}},$$

Now, we split the summation over $K \in \mathbb{T}_h$ as follows:

$$\Theta_1 = \sum_{K\in\mathbb{T}_h,\tilde{K}\subset B_a} \{\dots\} + \sum_{K\in\mathbb{T}_h,\tilde{K}\not\subset B_a} \{\dots\}.$$

On cells $K \not\subset B_a$, we have $\sigma_h^{-1} \leqslant c$ and therefore estimate

$$\begin{aligned}\Theta_1 &\leqslant \frac{c}{\gamma}\sum_{K\in\mathbb{T}_h,\tilde{K}\subset B_a} \frac{h_K^2}{\sigma_h^2}\|\nabla^2 u\|_K^2 + \frac{c}{\gamma}\sum_{K\in\mathbb{T}_h,\tilde{K}\not\subset B_a} h_K^2\|\nabla^2 u\|_K^2\\ &\leqslant \frac{c}{\gamma}h^2\|\nabla^2 u\|^2_{L^\infty(B_a)}\|\sigma_h^{-1}\|^2 + \frac{c}{\gamma}h^2\|\nabla^2 u\|^2.\end{aligned}$$

From this, we obtain

$$\Theta_1 \leqslant \frac{c}{\gamma}h^2\,\ell(\varepsilon)\|\nabla^2 u\|^2_{L^\infty(B_a)} + \frac{c}{\gamma}h^2\|\nabla^2 u\|^2. \tag{38}$$

(v) Next, we estimate Θ_2. Observing that σ_h is constant on each cell K , we have

$$\begin{aligned}a_h(\sigma_h^2 e^*, e^*) &= \sum_{K\in\mathbb{T}_h}\Big\{(\nabla(\sigma_h^2 e^*),\nabla e^*)_K + \frac{\kappa}{h_K}([\sigma_h^2 e^*],[e^*])_{\partial K}\\ &\qquad - \frac{1}{2}([\sigma_h^2 e^*],\{\partial_n e^*\})_{\partial K} - \frac{1}{2}(\{\partial_n(\sigma_h^2 e^*)\},[e^*])_{\partial K}\Big\}\\ &= \sum_{K\in\mathbb{T}_h}\Big\{(\sigma_h^2\nabla e^*,\nabla e^*)_K + \frac{\kappa}{h_K}(\{\sigma_h^2\}[e^*],[e^*])_{\partial K}\\ &\qquad + \frac{\kappa}{h_K}([\sigma_h^2]\{e^*\},[e^*])_{\partial K} - \frac{1}{2}([\sigma_h^2]\{e^*\},\{\partial_n e^*\})_{\partial K}\\ &\qquad - (\{\sigma_h^2\}[e^*],\{\partial_n e^*\})_{\partial K} - \frac{1}{8}([\sigma_h^2][\partial_n e^*],[e^*])_{\partial K}\Big\}.\end{aligned}$$

This leads us to

$$\Theta_2 = a_h(\sigma_h^2 e^*, e^*) + \Theta_3, \tag{39}$$

where

$$\begin{aligned}\Theta_3 = \sum_{K\in\mathbb{T}_h}\Big\{&\gamma h_K(\{\sigma_h^2\}\{\partial_n e^*\},\{\partial_n e^*\})_{\partial K} + \frac{1}{2}([\sigma_h^2]\{e^*\},\{\partial_n e^*\})_{\partial K}\\ &+ (\{\sigma_h^2\}[e^*],\{\partial_n e^*\})_{\partial K} + \frac{1}{8}([\sigma_h^2][\partial_n e^*],[e^*])_{\partial K}\\ &+ \frac{\kappa}{h_K}([\sigma_h^2]\{e^*\},[e^*])_{\partial K}\Big\}.\end{aligned}$$

(v) We proceed with the first term on the right in (39). Using Galerkin orthogonality of e^*, we obtain

$$a_h(\sigma_h^2 e^*, e^*) = a_h(\sigma_h^2 e^* - I_h(\sigma_h^2 e^*), e^*).$$

Since σ_h is constant on K and the projection I_h is local, we have $I_h(\sigma_h^2 e^*) = \sigma_h^2 I_h e^*$. Hence, with the notation $\eta^* = e^* - I_h e^* = g - I_h g$, it follows that

$$a_h(\sigma_h^2 e^*, e^*) = \sum_{K\in\mathbb{T}_h} \Big\{ (\nabla(\sigma_h^2 \eta^*), \nabla e^*)_K + \frac{\kappa}{h_K}([\sigma_h^2 \eta^*], [e^*])_{\partial K} - \frac{1}{2}([\sigma_h^2 \eta^*], \{\partial_n e^*\})_{\partial K} - \frac{1}{2}(\{\partial_n(\sigma_h^2 \eta^*)\}, [e^*])_{\partial K} \Big\}.$$

The four terms on the right hand side are now estimated separately using the trace inequalities (25), (26), and the interpolation estimates (13), (14). Furthermore, we employ the a priori bound (30) and the error estimate (30) for the Green function g and the dual error $\eta = g - g_h$. Note that in order to estimate the average values $\{\partial_n \eta\}|_{\partial K}$, it suffices to estimate each of the terms $\partial_n \eta|_{\partial K}$ separately. In virtue of Assumption 2, the weights σ_h can be estimated by

$$|[\sigma_h^2]|_{|\partial K} \leqslant c h_K \sigma_{h|K} \leqslant c h_K |\{\sigma_h\}|_{|\partial K},$$

and we use that $\nabla^2 g_{h|K} \equiv 0$. By (23), $\sigma_h \leqslant c\sigma$. We will use a free parameter $\gamma \in (0,1]$.

For the first term, we find

$$\Big| \sum_{K\in\mathbb{T}_h} (\nabla(\sigma_h^2 \eta^*), \nabla e^*)_K \Big| \leqslant c \sum_{K\in\mathbb{T}_h} \Big\{ \frac{1}{\gamma} h_K^2 \|\sigma \nabla^2 g\|_K^2 + \gamma \|\sigma_h \nabla e^*\|_K^2 \Big\} \leqslant \frac{c}{\gamma}\, \ell(\varepsilon)\, h^2 + c\gamma\Theta_2,$$

and, analogously, for the second term,

$$\begin{aligned} &\Big| \sum_{K\in\mathbb{T}_h} \frac{\kappa}{h_K}([\sigma_h^2 \eta^*], [e^*])_{\partial K} \Big| \\ &\leqslant c \sum_{K\in\mathbb{T}_h} \Big\{ \frac{\kappa}{\gamma h_K} \|\{\sigma_h\}^{-1}[\sigma_h^2 \eta^*]\|_{\partial K}^2 + \frac{\gamma\kappa}{h_K} \|\{\sigma_h\}[e^*]\|_{\partial K}^2 \Big\} \\ &\leqslant c \sum_{K\in\mathbb{T}_h} \Big\{ \frac{1}{\gamma} h_K^2 \|\sigma \nabla^2 g]\|_K^2 + \frac{\gamma\kappa}{h_K} \|\{\sigma_h\}[e^*]\|_{\partial K}^2 \Big\} \\ &\leqslant \frac{c}{\gamma}\, \ell(\varepsilon)\, h^2 + c\gamma\Theta_2. \end{aligned}$$

The third term is estimated by

$$
\begin{aligned}
&\Big| \sum_{K\in\mathbb{T}_h} ([\sigma_h^2\eta^*], \{\partial_n e^*\})_{\partial K} \Big| \\
&\qquad = \Big| \sum_{K\in\mathbb{T}_h} \Big\{ \big(2[\sigma_h]\{\sigma_h\}\{\eta^*\} + \frac{1}{4}[\sigma_h]^2\{\eta^*\} + \{\sigma_h\}^2[\eta^*], \{\partial_n e^*\}\big)_{\partial K} \Big\} \Big| \\
&\qquad \leqslant c \sum_{K\in\mathbb{T}_h} \|\sigma_h \eta^*\|_{\partial K} \|\sigma_h \partial_n e^*\|_{\partial K} \\
&\qquad \leqslant c \sum_{K\in\mathbb{T}_h} h_K^{3/2} \|\sigma\nabla^2 g\|_K \big(h_K^{-1/2} \|\sigma_h \nabla e^*\|_K + h_K^{1/2} \|\sigma\nabla^2 g\|_K \big) \\
&\qquad \leqslant c \sum_{K\in\mathbb{T}_h} \Big\{ \gamma \|\sigma_h \nabla e^*\|_K^2 + \frac{1}{\gamma} h_K^2 \|\sigma\nabla^2 g\|_K^2 \Big\} \\
&\qquad \leqslant c\gamma\Theta_2 + \frac{c}{\gamma}\, \ell(\varepsilon)\, h^2.
\end{aligned}
$$

Finally, for the fourth term, we find

$$
\begin{aligned}
\Big| \sum_{K\in\mathbb{T}_h} (\{\partial_n(\sigma_h^2\eta^*)\}, [e^*])_{\partial K} \Big| &\leqslant \sum_{K\in\mathbb{T}_h} \|\{\sigma_h\}^{-1}\{\partial_n(\sigma_h^2\eta^*)\}\|_{\partial K} \|\{\sigma_h\}[e^*]\|_{\partial K} \\
&\leqslant c \sum_{K\in\mathbb{T}_h} \Big\{ \frac{1}{\gamma}\big(\|\sigma_h\nabla\eta^*\|_K^2 + h_K^2 \|\sigma_h \nabla^2\eta^*\|_K^2 \big) + \gamma \frac{\kappa}{h_K} \|\{\sigma_h\}[e^*]\|_{\partial K}^2 \Big\} \\
&\leqslant c \sum_{K\in\mathbb{T}_h} \Big\{ \frac{1}{\gamma} h_K^2 \|\sigma\nabla^2 g\|_K^2 + \frac{\gamma\kappa}{h_K} \|\{\sigma_h\}[e^*]\|_{\partial K}^2 \Big\} \\
&\leqslant \frac{c}{\gamma}\, \ell(\varepsilon)\, h^2 + c\gamma\Theta_2.
\end{aligned}
$$

Combining these estimates yields

$$
|a_h(\sigma_h^2 e^*, e^*)| \leqslant c\gamma\Theta_2 + \frac{c}{\gamma}\, \ell(\varepsilon)\, h^2. \tag{40}
$$

(vi) Next, we estimate the five terms in Θ_3 separately again by using the local trace estimates (25), (26), and the a priori bound and error estimate (29) and (30). For the first term in Θ_3, we obtain

$$
\begin{aligned}
\Big| \sum_{K\in\mathbb{T}_h} \gamma h_K (\{\sigma_h^2\}\{\partial_n e^*\}, \{\partial_n e^*\})_{\partial K} \Big| &\leqslant c\gamma \sum_{K\in\mathbb{T}_h} \Big\{ \|\sigma_h \nabla e^*\|_K^2 + h_K^2 \|\sigma_h \nabla^2 g\|_K^2 \Big\} \\
&\leqslant c\gamma\Theta_2 + \frac{c}{\gamma}\, \ell(\varepsilon)\, h^2,
\end{aligned}
$$

and analogously for the second term,

$$
\begin{aligned}
&\Big| \sum_{K\in\mathbb{T}_h} \frac{1}{2}([\sigma_h^2]\{e^*\}, \{\partial_n e^*\})_{\partial K} \Big| \\
&\leqslant \sum_{K\in\mathbb{T}_h} \Big\{ \frac{c}{\gamma} h_K \|\{e^*\}\|_{\partial K}^2 + c\gamma h_K \|\{\sigma_h\}\{\partial_n e^*\}\|_{\partial K}^2 \Big\} \\
&\leqslant \sum_{K\in\mathbb{T}_h} \Big\{ \frac{c}{\gamma}\big(\|e^*\|_K^2 + h_K^2 \|\nabla e^*\|_K^2 \big) + c\gamma\big(\|\sigma_h \nabla e^*\|_K^2 + h_K^2 \|\sigma_h \nabla^2 g\|_K^2 \big) \Big\} \\
&\leqslant c\gamma\Theta_2 + \frac{c}{\gamma}\, \ell(\varepsilon)\, h^2.
\end{aligned}
$$

The third term in Θ_3 is the most critical one, since it does non contain a factor h_K. Therefore, we have to absorb this term into the other definite terms in Θ_2 using the stabilization parameter κ. We recall that for $v_h \in V_h$,

$$|(\{\partial_n v_h\}, [v_h])_{\partial K}| \leqslant \frac{c_0^2}{2\kappa}\|\nabla v_h\|_{\tilde{K}}^2 + \frac{\kappa}{2h_K}\|[v_h]\|_{\partial K}^2, \tag{41}$$

with $\tilde{K}$ defined in (9). Splitting the dual error like $e^* = \eta^* + I_h e^*$ with $\eta^* = g - I_h g$, we have

$$(\{\sigma_h^2\}[e^*], \{\partial_n e^*\})_{\partial K} = (\{\sigma_h^2\}[e^*], \{\partial_n \eta^*\})_{\partial K} + (\{\sigma_h^2\}[e^*], \{\partial_n I_h e^*\})_{\partial K}$$

The first term on the right is treated analogously as the other terms before leading to the estimate

$$\Big|\sum_{K \in \_H} (\{\sigma_h^2\}[e^*], \{\partial_n \eta^*\})_{\partial K}\Big| \leqslant \gamma\Theta_2 + \frac{c}{\gamma} h^2\, \ell(\varepsilon).$$

The second term is treated as follows:

$$\Big|\sum_{K \in \mathbb{T}_h} (\{\sigma_h^2\}[e^*], \{\partial_n I_h e^*\})_{\partial K}\Big| \leqslant \sum_{K \in \mathbb{T}_h} \Big\{\frac{\kappa}{2h_K}(\{\sigma_h^2\}[e^*], [e^*])_{\partial K} + \frac{h_K}{2\kappa}(\{\sigma_h^2\}\{\partial_n I_h e^*\}, \{\partial_n I_h e^*\})_{\partial K}\Big\}.$$

Using relation (41), we conclude by a lengthy but standard calculation:

$$\begin{aligned}
|(\{\sigma_h^2\}\{\partial_n I_h e^*\}, \{\partial_n I_h e^*\})_{\partial K}| &\leqslant (\sigma_{h;K}^2 + c h_K \sigma_{h;K})\|\{\partial_n I_h e^*\}\|_{\partial K}^2 \\
&\leqslant \frac{c_0^2}{h_K}(\sigma_{h;K}^2 + c h_K \sigma_{h;K})\|\nabla I_h e^*\|_{\tilde{K}}^2 \\
&\leqslant \frac{c_0^2}{h_K}\Big\{\Big(1+\frac{\alpha}{4}\Big)\|\sigma_h \nabla I_h e^*\|_{\tilde{K}}^2 + \frac{c}{\alpha} h_K^2 \|\nabla I_h e^*\|_{\tilde{K}}^2\Big\},
\end{aligned}$$

with an arbitrary constant $\alpha \in (0, 1]$. Consequently,

$$\begin{aligned}
&\sum_{K \in \mathbb{T}_h} \frac{h_K}{2\kappa}(\{\sigma_h^2\}\{\partial_n I_h e^*\}, \{\partial_n I_h e^*\})_{\partial K} \\
&\qquad \leqslant \sum_{K \in \mathbb{T}_h} \frac{c_0^2 \mu}{2\kappa}\Big\{\Big(1+\frac{\alpha}{4}\Big)\|\sigma_h \nabla I_h e^*\|_K^2 + \frac{c}{\alpha} h_K^2 \|\nabla I_h e^*\|_K^2\Big\}.
\end{aligned}$$

From

$$\|\nabla I_h e^*\|_K^2 \leqslant \Big(1+\frac{\alpha}{4}\Big)\|\nabla e^*\|_K^2 + \frac{c}{\alpha}\|\nabla \eta^*\|_K^2,$$

follows

$$\begin{aligned}
&\sum_{K \in \mathbb{T}_h} \frac{h_K}{2\kappa}(\{\sigma_h^2\}\{\partial_n I_h e^*\}, \{\partial_n I_h e^*\})_{\partial K} \\
&\qquad \leqslant \sum_{K \in \mathbb{T}_h} \frac{c_0^2 \mu}{2\kappa}(1+\alpha)\|\sigma_h \nabla e^*\|_K^2 \\
&\qquad\quad + c \sum_{K \in \mathbb{T}_h} \Big\{h_K^2 \|\nabla e^*\|_K^2 + \|\sigma_h \nabla \eta^*\|_K^2 + h_K^2 \|\nabla \eta^*\|_K^2\Big\}.
\end{aligned}$$

Hence, observing the results of Lemma 1, we obtain

$$\sum_{K\in\mathbb{T}_h}\frac{h_K}{2\kappa}(\{\sigma_h^2\}\{\partial_n I_h e^*\},\{\partial_n I_h e^*\})_{\partial K}\leqslant\frac{c_0^2\mu}{2\kappa}(1+\alpha)\Theta_2+\frac{c}{\alpha}\,\ell(\varepsilon)h^2.$$

Finally, the fourth and the fifth term are estimated in a similar way as before by

$$\begin{aligned}
&\Big|\sum_{K\in\mathbb{T}_h}([\sigma_h^2][\partial_n e^*],[e^*])_{\partial K}\Big|\\
&\qquad\leqslant\sum_{K\in\mathbb{T}_h}\Big\{\frac{c}{\gamma}h_K\|\{e^*\}\|_{\partial K}^2+c\gamma h_K\|\{\sigma_h\}\{\partial_n e^*\}\|_{\partial K}^2\Big\}\\
&\qquad\leqslant\sum_{K\in\mathbb{T}_h}\Big\{\frac{c}{\gamma}\big(\|e^*\|_K^2+h_K^2\|\nabla e^*\|_K^2\big)+c\gamma\big(\|\sigma_h\nabla e^*\|_K^2+h_K^2\|\sigma_h\nabla^2 g\|_K^2\big)\Big\}\\
&\qquad\leqslant c\gamma\Theta_2+\frac{c}{\gamma}\,\ell(\varepsilon)\,h^2,
\end{aligned}$$

and by

$$\begin{aligned}
&\Big|\sum_{K\in\mathbb{T}_h}\frac{\kappa}{h_K}([\sigma_h^2]\{e^*\},[e^*])_{\partial K}\Big|\\
&\qquad\leqslant\sum_{K\in\mathbb{T}_h}\frac{\kappa}{h_K}\Big\{\|[\sigma_h]\|\big(\frac{\gamma}{h_K}\|\{\sigma\}[e^*]\|_{\partial K}^2+\frac{h_K}{\gamma}\|\{e^*\}\|_{\partial K}^2\big)\Big\}\\
&\qquad\leqslant\sum_{K\in\mathbb{T}_h}\Big\{c\gamma\frac{\kappa}{h_K}(\{\sigma_h^2\}[e^*],[e^*])_{\partial K}+c\|e^*\|_K^2+ch_K^2\|\nabla e^*\|_K^2\Big\}\\
&\qquad\leqslant c\gamma\Theta_2+\frac{c}{\gamma}\,\ell(\varepsilon)\,h^2.
\end{aligned}$$

Collecting these results, we obtain

$$\Theta_3\leqslant\Big(c\gamma+\frac{c_0^2\mu}{2\kappa}(1+\alpha)\Big)\Theta_2+\Big(\frac{c}{\gamma}+\frac{c}{\alpha}\Big)\,\ell(\varepsilon)\,h^2, \tag{42}$$

and consequently, in virtue of (39) and (40),

$$\Theta_2\leqslant\Big(c\gamma+\frac{c_0^2\mu}{2\kappa}(1+\alpha)\Big)\Theta_2+\Big(\frac{c}{\gamma}+\frac{c}{\alpha}\Big)\,\ell(\varepsilon)\,h^2. \tag{43}$$

Now, we fix κ according to condition (10),

$$\kappa>c_0^2\mu,$$

and then choose α and γ sufficiently small, such that

$$\Big(c\gamma+\frac{c_0^2\mu}{2\kappa}(1+\alpha)\Big)\leqslant\theta<1,$$

for some $\theta\in(0,1)$. With this choice of the parameters, we obtain that

$$\Theta_2\leqslant c(\alpha,\gamma)\,\ell(\varepsilon)h^2. \tag{44}$$

which together with (38) completes the proof.

	$\mathbb{Q}_1$	
	$\|e_h\|_\infty$	$\tilde{e}_{h_L} - \tilde{e}_{h_{L-1}}$
0	5.0e-01	–
1	2.0e-01	-2.83e-01
2	6.3e-02	-2.27e-01
3	2.0e-02	-2.50e-01
4	5.7e-03	-2.04e-01
5	1.6e-03	-2.11e-01
6	4.6e-04	-2.15e-01
7	1.3e-04	-2.20e-01

Table 1: *L^∞-errors and their scaled differences for discontinuous $\mathbb{Q}_1$ elements.*

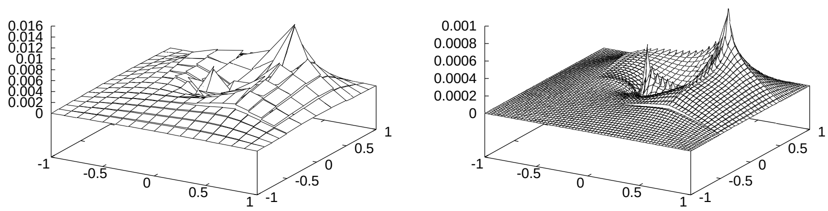

Figure 2: *Error function for refinement steps 2 and 4*

5 Numerical results 1

The result of Corollary 1 is asymptotically optimal. We show this using numerical results. The test problem is as follows: we solve the model problem (1), (2) on the square $\Omega = (-1,1)^2$ with homogeneous Dirichlet boundary values, i.e., $\partial\Omega = \partial\Omega_D$. The right hand side is chosen such that the solution is $u(x,y) = \cos(\pi/2x)\cos(\pi/2y)$. The stabilization parameter is $\kappa = 8$. The grids are generated by successive refinement of the coarse grid shown in Figure 3 in Section 7. Solutions are obtained using the multi-level algorithm described in Gopalakrishnan and Kanschat [13].

The scaled error $\tilde{e}_h := h^{-2}\|u-u_h\|_\infty$ admits the asymptotic estimate

$$\tilde{e}_{h_1} - \tilde{e}_{h_2} \approx c \log \frac{h_2}{h_1}.$$

Therefore, this difference should remain constant under refinement, if (20) is sharp. In Table 1, we display these values together with the L^∞-norm of the errors. This table clearly supports our theoretical result. As can be seen in Figure 2, the maxima of the error are located at the two irregular points of the mesh. Since the right graph is scaled up by a factor of 16 (i. e. h^2), the peaks are indeed growing by the logarithmic factor.

6 A posteriori error analysis

We will now derive *a posteriori* estimates for the interior penalty scheme (7). A similar analysis based on duality arguments has been developed in Becker, Hansbo and Larson [5] and Becker, Hansbo and Stenberg [6], but there the emphasis is on L^2-norm error bounds and on non-matching meshes in the context of domain decomposition. Let $J(\cdot)$ be an arbitrary *linear* functional on V with respect to which the error $e = u - u_h$ is to be estimated. Examples are local averages as considered above, contour integrals or integrals over subdomains:

$$J(u) = |B_\varepsilon|^{-1} \int_{B_\varepsilon} u(x)\,dx, \quad J(u) = \int_\Gamma u(x)\,ds, \quad J(u) = \int_{\Omega_0} u(x)\,dx;$$

for more examples see Becker and Rannacher [8]. With the functional $J(\cdot)$, we associate a *dual solution* $z \in V$ as the solution of the auxiliary problem

$$(\nabla\phi, \nabla z) = J(\phi) \quad \forall \phi \in V. \tag{45}$$

If the functional $J(\cdot)$ has an L^2 representation j, then the dual problem can be written in strong form as

$$-\Delta z = j \quad \text{in } \Omega, \quad z_{|\partial\Omega_D} = 0, \;\; \partial_n z_{|\partial\Omega_N} = 0. \tag{46}$$

Recalling that

$$J(\phi) = a_h(\phi, z), \quad \phi \in V \oplus V_h,$$

we have by Galerkin orthogonality

$$J(e) = a_h(e, z) = a_h(e, z - \psi_h),$$

for arbitrary $\psi_h \in V_h$. Integrating cell-wise by parts and reordering terms, we conclude by elementary but tedious calculation that, for $\xi := z - \psi_h$,

$$\begin{aligned} J(e) &= a_h(e, z - \zeta_h) \\ &= \sum_{K \in \mathbb{T}_h} \Big\{ (-\Delta e, \xi)_K + (\partial_n e_K, \xi_K)_{\partial K} - \frac{1}{2}(\{\partial_n e\}, [\xi])_{\partial K} \\ &\qquad - \frac{1}{2}([e], \{\partial_n \xi\})_{\partial K} + \frac{\kappa}{h_K}([e], [\xi])_{\partial K} \Big\} \\ &= \sum_K \Big\{ (f + \Delta u_h, \xi)_K - \frac{1}{2}([\partial_n u_h], \{\xi\})_{\partial K} \\ &\qquad - \frac{1}{2}([u_h], \{\partial_n \xi\})_{\partial K} + \frac{\kappa}{h_K}([u_h], [\xi])_{\partial K} \Big\}. \end{aligned}$$

This is rewritten in the form

$$\begin{aligned} J(e) = \sum_K \eta_K^{(0)} = \sum_K \Big\{ (R(u_h), \xi)_K + \frac{1}{2}(\partial r(u_h), \{\xi\}) \\ - \frac{1}{2}(r(u_h), \{\partial_n \xi\} + \frac{\kappa}{h_K}[\xi])_{\partial K} \Big\}, \end{aligned} \tag{47}$$

with the following notation of cell and edge residuals:

$$
\begin{aligned}
R(u_h)_{|K} &:= f + \Delta u_h, \\
r(u_h)_{|\Gamma} &:= \left\{ \begin{array}{ll} \frac{1}{2}[u_h], & \text{if } \Gamma \subset \partial K \backslash \partial\Omega, \\ u^D - u_h, & \text{if } \Gamma \subset \partial\Omega_D, \\ 0, & \text{if } \Gamma \subset \partial\Omega_N, \end{array} \right\}, \\
\partial r(u_h)_{|\Gamma} &:= \left\{ \begin{array}{ll} \frac{1}{2}[\partial_n u_h], & \text{if } \Gamma \subset \partial K \backslash \partial\Omega, \\ 0, & \text{if } \Gamma \subset \partial\Omega_D, \\ u^N - \partial_n u_h, & \text{if } \Gamma \subset \partial\Omega_N, \end{array} \right\}.
\end{aligned}
$$

From the error representation (47), we obtain the following error estimate:

Theorem 2. *For the error $e = u - u_h$ in the interior penalty scheme (7), there holds the a posteriori error estimate*

$$
J(e) = \sum_{K \in \mathbb{T}} \left\{ \varrho_K^{(1)} \omega_K^{(1)} + \varrho_K^{(2)} \omega_K^{(2)} + \varrho_K^{(3)} \omega_K^{(3)} \right\}, \tag{48}
$$

with the cell residuals $\varrho_K^{(i)}$ and weight factors $\omega_K^{(i)}$ being defined by

$$
\begin{aligned}
\varrho_K^{(1)} &= \|R(u_h)\|_K, & \omega_K^{(1)} &= \|z - \psi_h\|_K, \\
\varrho_K^{(2)} &= h_K^{-1/2} \|\partial r(u_h)\|_{\partial K}, & \omega_K^{(2)} &= h_K^{1/2} \|z - \psi_h\|_{\partial K}, \\
\varrho_K^{(3)} &= h_K^{-3/2} \|r(u_h)\|_{\partial K}, & \omega_K^{(3)} &= h_K^{3/2} \|\{\partial_n(z - \psi_h)\} - \kappa h_K^{-1}[z - \psi_h]\|_{\partial K},
\end{aligned}
$$

for arbitrary $\psi_h \in V_h$.

Theorem 2 provides a posteriori estimates for arbitrary functionals of the error. This also includes the L^2-error estimates. To see this, we take the special functional

$$
J(\phi) := (e, \phi)\|e\|^{-1}.
$$

The corresponding dual solution $z \in V$ satisfies $u \in H^2(\Omega)$ and the a priori bound

$$
\|\nabla^2 z\| \leqslant c_S, \tag{49}
$$

where the *stability constant* c_S only depends on the domain Ω. From (48) and the interpolation estimates (13), (14), we infer that

$$
\|e\| \leqslant c_I c_S \sum_{K \in \mathbb{T}} \left\{ h_K^2 \varrho_K^{(1)2} + h_K^2 \varrho_K^{(2)2} + h_K \varrho_K^{(3)2} \right\}^{1/2}. \tag{50}
$$

This a posteriori error estimate is asymptotically optimal, too.

Next, we state an a posteriori error estimate for the locally averaged error as considered in our a priori error analysis. In this case the dual solution is just the regularized Green function, $z = g_a^\varepsilon$ introduced in the proof of Theorem 1. We have the estimate

$$
|J_\varepsilon^a(e)| \leqslant \eta_1(u_h) := \sum_{K \in \mathbb{T}} \left\{ \varrho_K^{(1)} \omega_K^{(1)} + \varrho_K^{(2)} \omega_K^{(2)} + \varrho_K^{(3)} \omega_K^{(3)} \right\}, \tag{51}
$$

where the residual terms $\varrho_K^{(i)}$ are as defined above and the weights $\omega_K^{(i)}$ can be estimated as follows:

$$\omega_K^{(1)} = \|g_a^\varepsilon - \psi_h\|_K,$$
$$\omega_K^{(2)} = h_K^{1/2}\|g_a^\varepsilon - \psi_h\|_{\partial K},$$
$$\omega_K^{(3)} = h_K^{3/2}\|\{\partial_n(g_a^\varepsilon - \psi_h)\} - \frac{\kappa}{h_K}[g_a^\varepsilon - \psi_h]\|_{\partial K}.$$

7 Numerical results 2

The a posteriori error estimate in Theorem 2 is tested at the same configuration as before; see Figure 3. The error is evaluated at the point $a = (0.5, 0.5)$, which is the point with maximum error in Figure 2 (the value is taken from the solution in the upper right cell adjacent to this point) . We compare three types of "error estimators". The first one, $\eta_0(u_h)$, is obtained directly from the error representation (47), avoiding the use of triangle and Hölder inequalities:

$$J(e) = \eta_0(u_h) := \sum_K \eta_K^{(0)}. \tag{52}$$

The second estimator uses the local refinement indicators $|\eta_K^{(0)}|$:

$$\eta_1(u_h) := \sum_K \eta_K^{(1)} := \sum_K |\eta_K^{(0)}|.$$

The third one, $\eta_2(u_h)$, is given by Theorem 2 as described above:

$$\eta_2(u_h) := \sum_K \eta_K^{(2)} := \sum_K \{\varrho_K^{(1)}\omega_K^{(1)} + \varrho_K^{(2)}\omega_K^{(2)} + \varrho_K^{(3)}\omega_K^{(3)}\}.$$

For practical evaluation of the error estimators, we solve the dual problem (45) on the current mesh with bi-quadratic polynomials obtaining $\tilde{z} \in \tilde{V}_h$. We decided for exact computation in the higher order space to avoid additional error contributions. For a more efficient computation, $\tilde{z}$ can be obtained by post-processing; see [8] for such strategies and their influence on the estimator.

The quality of the resulting approximate error estimators $\tilde{\eta}_i(u_h),\ i = 0, 1, 2$, is measured by the "effectivity index":

$$I_{\text{eff}} := \frac{|\tilde{\eta}_i(u_h)|}{|J_a^\varepsilon(e)|}.$$

Mesh adaptation is based on "error indicators" $\eta_K^{(1)}$ and $\eta_K^{(2)}$, respectively. For mesh refinement, a fixed fraction (here 20%) of the grid cells with largest indicator are refined.

L	$e(a)$	$\tilde\eta_0(u_h)$	I_{eff}	$\tilde e(a)$	$\eta_1(u_h)$	I_{eff}
3	4.306e-3	4.288e-3	0.996	1.82e-5	6.127e-3	1.42
4	1.655e-3	1.652e-3	0.998	2.75e-6	2.296e-3	1.39
5	6.540e-4	6.528e-4	0.998	1.27e-6	9.431e-4	1.44
6	2.920e-4	2.917e-4	0.999	2.66e-7	4.182e-4	1.43
7	1.402e-4	1.401e-4	0.999	8.60e-8	2.000e-4	1.43
8	6.756e-5	6.751e-5	0.999	5.47e-8	9.678e-5	1.43
9	3.376e-5	3.373e-5	0.999	2.47e-8	4.875e-5	1.44
10	1.743e-5	1.742e-5	1.000	3.10e-9	2.522e-5	1.45
11	9.069e-6	9.068e-6	1.000	1.74e-9	1.324e-5	1.46

Table 2: *Point error $e(a)$ and $\tilde e(a) = e(a) + \tilde\eta_0(u_h)$ and the corresponding a posteriori estimators $\tilde\eta_0(u_h)$ and $\tilde\eta_1(u_h)$ (L is maximum refinement level).*

L	$e(a)$	$\tilde\eta_0(u_h)$	I_{eff}
3	3.877e-03	2.021e-02	5.211
4	1.459e-03	7.310e-03	5.009
5	6.508e-04	3.050e-03	4.686
6	3.146e-04	1.430e-03	4.544
7	1.608e-04	6.985e-04	4.344
8	8.205e-05	3.574e-04	4.355
9	4.203e-05	1.834e-04	4.365
10	8.205e-05	3.574e-04	4.355
11	4.203e-05	1.834e-04	4.365

Table 3: *Point error $e(a)$ and a posteriori estimator $\tilde\eta_2(u_h)$*

Table 2 presents results for estimators η_0 and η_1 obtained by adaptive refinement based on indicators $\eta_K^{(1)}$. Estimator $\eta_0(u_h)$ is asymptotically optimal while estimator $\eta_1(u_h)$ appears to be off by a factor of about $3/2$. The error representation (52) suggests to consider

$$\tilde J(u_h) := J(u_h) + \tilde\eta_0(u_h)$$

as new approximation. This post-processing step can improve accuracy dramatically, as shown in Table 2. The mesh obtained in the eighth step of this iteration is shown in Figure 3.

Table 2 presents the results obtained by the adaptive refinement process with a refined mesh shown in Figure 3. The error estimator $\eta_0(u_h)$ is asymptotically optimal while the estimator $\eta_1(u_h)$ appears to be only suboptimal.

In Table 3, we show results for estimator $\eta_2(u_h)$. Since this estimator involves Hölder and triangle inequalities on each cell, there is no cancellation between the different terms of the estimator. Therefore, the error is over-estimated by a factor between 4 and 5. The errors resulting from adaptive refinement —based on the estimator $\eta_2(u_h)$ here— are comparable to those of Table 2. Therefore, we conclude that both estimators are suited as refinement criteria.

Finally, in Table 4, we compare the “error/mesh-ratio” $e(a) * N$ for uniform and

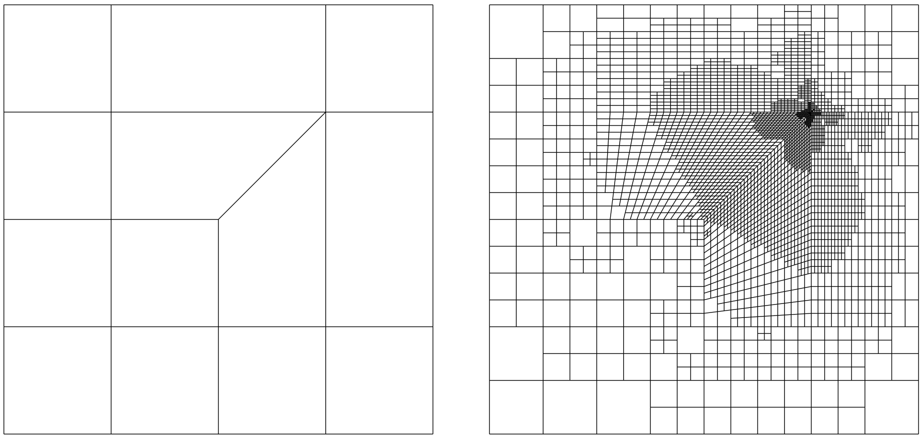

Figure 3: *Initial coarse mesh and an adapted mesh (level $L = 7$) for the point value computation.*

	uniform refinement		adaptive refinement	
L	$e(a)$	$e(a) * N$	$e(a)$	$e(a) * N$
1	3.811e-02	1.985e+00	3.811e-02	1.985e+00
2	1.126e-02	2.348e+00	1.448e-02	1.433e+00
3	3.212e-03	2.677e+00	4.306e-03	8.404e-01
4	8.956e-04	2.983e+00	1.655e-03	6.261e-01
5	2.464e-04	3.281e+00	6.540e-04	4.804e-01
6	6.716e-05	3.577e+00	2.920e-04	4.238e-01
7	1.818e-05	3.872e+00	1.402e-04	3.945e-01
8	–	–	6.756e-05	3.664e-01
9	–	–	3.376e-05	3.507e-01

Table 4: *Efficiency of computation of point value $u(a)$ on uniformly and adaptively refined meshes (L refinement level, N number of cells).*

adaptive refinement. On uniformly refined meshes, due to the asymptotic behavior $e(a) \approx h^2\,\ell(h)$ as demonstrated in Table 1, we expect $e(a) * N$ to grow like $\ell(1/N)$ while the mesh refinement at the point a should suppress this defect. This is demonstrated by Figure 4.

Extensions

On the basis of experience with the *continuous* Galerkin method, we believe that the results of this paper can be extended in several ways to more complex situations. Some of these extensions which are discussed below are obvious others seem more difficult.

1. More general second-order problems: The results of this paper can certainly be extended to more general elliptic problems of the form

$$-\nabla\cdot\{a\nabla u+bu\}+cu=f\quad\text{in }\Omega,\tag{53}$$

$$u_{|\partial\Omega}=u^D,\ (\partial_n u+\alpha u)_{|\partial\Omega_N}=u^N,\tag{54}$$

with regular coefficients $a>0,\ c\geq 0,\ \alpha\geq 0$, as well as to elliptic systems.

2. Irregular domains: In the presence of boundary irregularities, e.g., due to reentrant corners, the assumed H^2 regularity of weak solutions is generally lost. Since this property is crucial at several stages of the proof of Theorem 1 the local error estimate (18) does not so easily carry over to this situation. However, assuming that all corner singularities are properly resolved by mesh refinement, such that optimal-order of the energy-norm error estimate is preserved (see Heinrich and Nicaise [15]),

$$|||e|||_h\leqslant c(u)h,$$

one can prove the following analogue of (18):

$$|J_a^\varepsilon(e)|\leqslant Ch^2\,\ell(\varepsilon)\|\nabla^2 u\|_{L^\infty(B_a)}+C(u)h^2,\tag{55}$$

as long as a is not located at one of the corners. Since the proof is very technical and not really related to the particular aspects of pointwise error estimation, we have omitted this possible generalization.

3. Higher-oder approximation: Analogously as for the cG method, our analysis could be extended to higher-order finite elements. The argument of proof directly carries over to this situation with the obvious modifications in assumptions and results. We note that in the case of elements of order $m\geq 3$ (i.e. quadratics and higher) the logarithmic factor $\ell(\varepsilon)$ in the local estimate (18) can be dropped; see Nitsche [17] for the corresponding result for the cG method.

4. Estimates for gradient error: Using the local inverse and approximation properties of finite elements, the result of Theorem 1 implies the corresponding result, with reduced order, for local averages of the gradient error:

$$|J_\varepsilon^a(\nabla e)|\leqslant Ch\,\ell(\varepsilon)\,\|\nabla^2 u\|_{L^\infty(B_a)}+Ch\,\|\nabla^2 u\|.$$

Whether the logarithmic term $\ell(\varepsilon)$ can be dropped (as was established in Rannacher and Scott [18] for the cG method on quasi-uniform meshes) is not clear in the present case of possibly refined meshes.

5. Nonlinear problems: Once the L^∞-error analysis is established for general linear problems, it can be used to obtain corresponding results also for nonlinear equations and systems by linearization. Following the method in Dobrowolski and Rannacher [11] for the cG method, the a priori result of Theorem 1 could be proven also for nonlinear problems. The corresponding extension of the a posteriori analysis, at least in a formal sense, may be done by the abstract approach described in Becker and Rannacher [8].

Acknowledgments

This work has been supported by "Deutsche Forschungsgemeinschaft" (DFG) through "Schwerpunktprogramm ANumE" and "Sonderforschungsbereich 359 Reaktive Strömungen, Diffusion und Transport". Numerical computations have been conducted based on the `deal.II` finite element library [3, 4].

References

[1] D. N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19(4):742–760, 1982.

[2] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39:1749–1779, 2001.

[3] W. Bangerth, R. Hartmann, and G. Kanschat. `deal.II` *Differential Equations Analysis Library, Technical Reference*, 2002. first edition 1999.

[4] W. Bangerth and G. Kanschat. Concepts for object-oriented finite element software – the `deal.II` library. Preprint 43, SFB 359, Heidelberg, 1999.

[5] R. Becker, P. Hansbo, and M. G. Larson. Energy norm a posteriori error estimation for discontinuous Galerkin methods. *Calcolo*, submitted.

[6] R. Becker, P. Hansbo and R. Stenberg. A finite element method for domain decomposition with non-matching grids. *M2AN*, submitted.

[7] 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.

[8] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. In *Acta Numerica 2001* (A. Iserles, ed.), Cambridge University Press, Cambridge, 2001.

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

[10] S.C. Brenner and R.L. Scott. *The Mathematical Theory of Finite Element Methods*, Springer, Berlin-Heidelberg-New York, 1994.

[11] M. Dobrowolski and R. Rannacher. Finite element methods for nonlinear elliptic systems of second order. *Math. Nachr.*, 95:155-172, 1980.

[12] J. Frehse and R. Rannacher. Eine L^1-Fehlerabschätzung für diskrete Grundlösungen in der Methode der finiten Elemente. Tagungsband Finite Elemente, *Bonner Math. Schriften*, 89:92–114, 1976.

[13] J. Gopalakrishnan and G. Kanschat. A multilevel discontinuous Galerkin method. *Numer. Math.*, to appear.

[14] P. Grisvard. *Singularities in Boundary Value Problems*. Masson, Paris, and Springer, Berlin, 1992.

[15] B. Heinrich and S. Nicaise. Nitsche mortar finite element method for transmission problems with singularities. Technical Report 01-10, SFB 393, Chemnitz, 2001. 3.

[16] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei der Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Sem. Univ. Hamburg*, 36:9–15, 1971.

[17] J. Nitsche. Über L^∞-Abschätzungen von Projektionen auf Finite Elemente. Tagungsband Finite Elemente, *Bonner Math. Schriften*, 89:13–30, 1976.

[18] R. Rannacher and R. Scott. Some optimal error estimates for piecewise linear finite element approximations. *Math. Comp.*, 38:437–445, 1982.

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