IMA Journal of Numerical Analysis (2018) **00**, 1–26
doi:10.1093/imanum/dry006

Adaptive vertex-centered finite volume methods for general second-order linear elliptic partial differential equations

CHRISTOPH ERATH*
Department of Mathematics, TU Darmstadt, Dolivostrasse 15, 64293 Darmstadt, Germany
*Corresponding author: erath@mathematik.tu-darmstadt.de

AND

DIRK PRAETORIUS
Institute for Analysis and Scientific Computing, TU Wien, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria

[Received on 21 September 2017; revised on 08 January 2018]

We prove optimal convergence rates for the discretization of a general second-order linear elliptic partial differential equation with an adaptive vertex-centered finite volume scheme. While our prior work Erath & Praetorius (2016, Adaptive vertex-centered finite volume methods with convergence rates. *SIAM J. Numer. Anal.*, **54**, 2228–2255) was restricted to symmetric problems, the present analysis also covers nonsymmetric problems and hence the important case of present convection.

Keywords: finite volume method; Céa-type quasi-optimality; *a posteriori* error estimators; adaptive algorithm; local mesh refinement; optimal convergence rates; nonsymmetric problems.

1. Introduction

We consider a general second-order linear elliptic partial differential equation (PDE) and approximate the solution with an adaptive vertex-centered finite volume method (FVM). FVM are well established in fluid mechanics, since they naturally preserve numerical flux conservation.

1.1 *Model problem*

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded Lipschitz domain with polygonal boundary $\Gamma := \partial\Omega$. As a model problem, we consider the following stationary diffusion problem: given $f \in L^2(\Omega)$, find $u \in H^1(\Omega)$ such that

$$\operatorname{div}(-\mathbf{A}\nabla u + \mathbf{b}u) + cu = f \quad \text{in } \Omega \qquad \text{and} \qquad u = 0 \quad \text{on } \Gamma. \tag{1.1}$$

We suppose that the diffusion matrix $\mathbf{A} = \mathbf{A}(x) \in \mathbb{R}^{d\times d}$ is bounded, symmetric and uniformly positive definite, i.e., there exist constants $\lambda_{\min}, \lambda_{\max} > 0$ such that

$$\lambda_{\min} |\mathbf{v}|^2 \le \mathbf{v}^{\mathrm{T}}\mathbf{A}(x)\mathbf{v} \le \lambda_{\max} |\mathbf{v}|^2 \quad \text{for all } \mathbf{v} \in \mathbb{R}^d \text{ and almost all } x \in \Omega. \tag{1.2}$$

Let $\mathscr{T}_0$ be a given initial triangulation of Ω; see Section 2.2 below. For convergence of FVM and well-posedness of the residual error estimator, we additionally require that $\mathbf{A}(x)$ is piecewise Lipschitz continuous, i.e.,

$$\mathbf{A} \in W^{1,\infty}(T)^{d\times d} \quad \text{for all } T \in \mathscr{T}_0. \tag{1.3}$$

We suppose that the lower-order terms satisfy the assumption

$$\mathbf{b} \in W^{1,\infty}(\Omega)^d \quad \text{and} \quad c \in L^\infty(\Omega) \quad \text{with} \quad \tfrac{1}{2}\operatorname{div}\mathbf{b} + c \geq 0 \quad \text{almost everywhere on } \Omega. \tag{1.4}$$

With $(\phi,\psi)_\omega = \int_\omega \phi(x)\psi(x)\,\mathrm{d}x$ being the L^2-scalar product on a subdomain $\omega \subseteq \Omega$, the weak formulation of the model problem (1.1) reads as follows: find $u \in H^1_0(\Omega)$ such that

$$\mathscr{A}(u,w) := (\mathbf{A}\nabla u - \mathbf{b}u, \nabla w)_\Omega + (cu,w)_\Omega = (f,w)_\Omega \quad \text{for all } w \in H^1_0(\Omega). \tag{1.5}$$

According to our assumptions (1.2)–(1.4), the bilinear form $\mathscr{A}(\cdot,\cdot)$ is continuous and elliptic on $H^1_0(\Omega)$. Existence and uniqueness of the solution $u \in H^1_0(\Omega)$ of (1.5) thus follow from the Lax–Milgram theorem. Moreover, the operator-induced quasi-norm $|||\cdot|||$ satisfies

$$C_{\mathrm{ell}}\|v\|^2_{H^1(\Omega)} \leq |||v|||^2 := \mathscr{A}(v,v) \leq C_{\mathrm{cont}}\|v\|^2_{H^1(\Omega)} \qquad \text{for all } v \in H^1_0(\Omega), \tag{1.6}$$

where $C_{\mathrm{ell}} > 0$ depends only on $\lambda_{\min}$ and Ω, whereas $C_{\mathrm{cont}} > 0$ depends only on $\lambda_{\max}$, $\|\mathbf{b}\|_{L^\infty(\Omega)}$ and $\|c\|_{L^\infty(\Omega)}$.

1.2 *Adaptive FVM*

In the past 20 years there have been major contributions to the mathematical understanding of adaptive mesh-refinement algorithms, mainly in the context of the finite element method (FEM). While the seminal works Dörfler (1996), Morin *et al.* (2000), Binev *et al.* (2004), Stevenson (2007) and Cascón *et al.* (2008) were restricted to symmetric operators, the recent works Mekchay & Nochetto (2005), Cascón & Nochetto (2012), Feischl *et al.* (2014) and Bespalov *et al.* (2017) proved convergence of adaptive FEM with optimal algebraic rates for general second-order linear elliptic PDEs. The work Carstensen *et al.* (2014) gives an exhaustive overview of the developments and it gains, in an abstract framework, a general recipe to prove optimal adaptive convergence rates of adaptive mesh-refining algorithms. Basically, the numerical discretization scheme, the *a posteriori* error estimator and the adaptive algorithm have to fulfill four criteria (called *axioms in* Carstensen *et al.*, 2014), namely, *stability on nonrefined elements*, *reduction on refined elements*, *general quasi-orthogonality* and *discrete reliability*. Building upon these findings, our recent work (Erath & Praetorius, 2016) provides the first proof of convergence of adaptive FVM with optimal algebraic rates for a symmetric model problem (1.1) with $\mathbf{b} = \mathbf{0}$ and $c = 0$.

1.3 *Contributions and outline*

In this work, we are in particular interested in the nonsymmetric model problem with $\mathbf{b} \neq \mathbf{0}$ in (1.1). The proofs of *stability on nonrefined elements*, *reduction on refined elements* and *discrete reliability* follow basically the proofs in Erath & Praetorius (2016); see Sections 3.3 and 3.4. Thus, the major contribution of the present work is the proof of the *general quasi-orthogonality* property

for the nonsymmetric problem, which is satisfied under some mild regularity assumptions on the dual problem. Similar assumptions are required in Mekchay & Nochetto (2005) and Cascón & Nochetto (2012) to prove convergence for an adaptive FEM procedure. Moreover, we note that Mekchay & Nochetto (2005) and Cascón & Nochetto (2012) require slightly more restrictions on the model data (namely, $\operatorname{div}(\mathbf{b}) = 0$) and on the mesh refinement (the so-called *interior node property*) for proving quasi-orthogonality, which are avoided in the present analysis.

At this point, we note that Feischl *et al.* (2014) and Bespalov *et al.* (2017) improve the FEM result of Mekchay & Nochetto (2005) and Cascón & Nochetto (2012) by a different approach. Instead of the duality argument, the analysis exploits the *a priori* convergence of FEM solutions (which follows from the classical Céa lemma) by splitting the operator into a symmetric and an elliptic part and a compact perturbation. In particular, there is no duality argument applied. Therefore, no additional regularity assumption is required. However, it seems to be difficult to transfer the analysis of Feischl *et al.* (2014) and Bespalov *et al.* (2017) to FVM due to the lack of the Céa lemma.

We also mention that unlike the FEM literature, a direct proof of *general quasi-orthogonality* is not available for FVM due to the lack of Galerkin orthogonality. Instead, the FVM work Erath & Praetorius (2016) first proves linear convergence which relies on a quasi-Galerkin orthogonality (see Erath & Praetorius, 2016, Lemma 11) for FVM. Unfortunately, this auxiliary result does not hold for nonsymmetric problems.

Hence, to handle the nonsymmetric case, the missing Galerkin orthogonality and the lack of an optimal L^2-estimate for FVM seem to be the bottlenecks. To overcome these difficulties, we first estimate the FVM error in the bilinear form by oscillations in Lemma 3.3. Then we provide a new L^2-type estimate in Lemma 3.4 which depends on the regularity of the corresponding dual problem plus oscillations. These two results provide the key arguments to prove a quasi-Galerkin orthogonality in Proposition 3.2. Unlike the literature, this estimate also includes a mesh-size weighted estimator term. With the aid of the previous results, we show linear convergence in Theorem 3.6, where the proof relies on the previous results. Finally, optimal algebraic convergence rates are guaranteed by Theorem 3.10 which follows directly from the literature.

We remark that the proposed Algorithm 3.1 additionally marks oscillations to overcome the lack of classical Galerkin orthogonality. Note that this is not required for adaptive FEM. However, since FVM is not a best approximation method, the proposed approach appears to be rather natural. In practice, however, this additional marking is negligible (see also Erath & Praetorius, 2016, Remark 12). Furthermore, if problem (1.1) is slightly convection dominated, Algorithm 3.1 and thus our analysis can be used with caution. We discuss the difficulties for such model problems in Sections 4.3 and 5 in more detail. An extension of our analysis to PDEs with nonlinearities appears to be difficult and is thus beyond the scope of this work. Overall, the present work seems to be the first that proves convergence with optimal rates of an adaptive FVM algorithm for the solution of general second-order linear elliptic PDEs.

2. Preliminaries

This section introduces the notation and the discrete scheme, as well as the residual *a posteriori* error estimator. In particular, we fix our notation used throughout this work.

2.1 *General notation*

Throughout, $\mathbf{n}$ denotes the unit normal vector to the boundary pointing outward from the respective domain. In the following, we mark the mesh dependency of quantities by appropriate indices, e.g., u_ℓ is

the solution on the triangulation $\mathcal{T}_\ell$. Furthermore, $\lesssim$ abbreviates $\leq$ up to some (generic) multiplicative constant which is clear from the context.

2.2 *Triangulations*

The FVM relies on two partitions of Ω: the *primal mesh* $\mathcal{T}_\times$ and the associated *dual mesh* $\mathcal{T}_\times^*$. The primal mesh $\mathcal{T}_\times$ is a regular triangulation of Ω into nondegenerate closed triangles/tetrahedra $T \in \mathcal{T}_\times$, where the possible discontinuities of the coefficient matrix $\mathbf{A}$ are aligned with $\mathcal{T}_\times$. Define the local mesh-size function

$$h_\times \in L^\infty(\Omega), \quad h_\times|_T := h_T := |T|^{1/d} \quad \text{for all } T \in \mathcal{T}_\times. \tag{2.1}$$

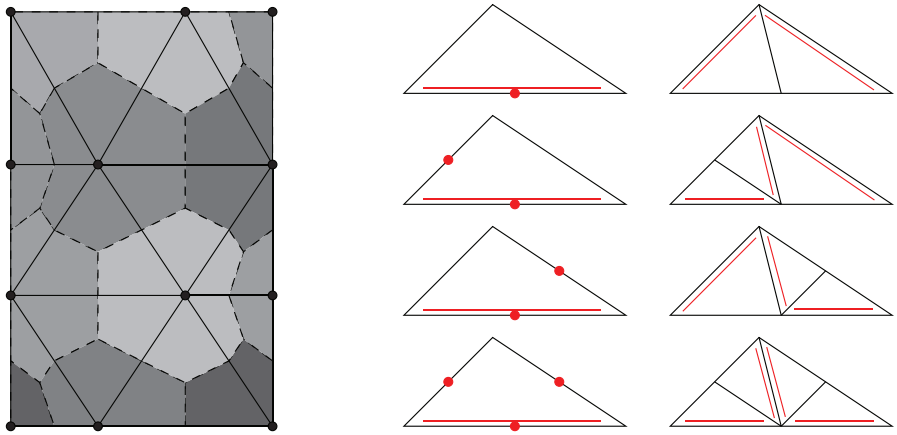
Let $\operatorname{diam}(T)$ be the Euclidean diameter of T. Suppose that $\mathcal{T}_\times$ is σ-shape regular, i.e.,

$$\max_{T\in\mathcal{T}_\times} \frac{\operatorname{diam}(T)}{|T|^{1/d}} \leq \sigma < \infty. \tag{2.2}$$

Note that this implies $h_T \leq \operatorname{diam}(T) \leq \sigma\, h_T$. Let $\mathcal{N}_\times$ (or $\mathcal{N}_\times^\Omega$) denote the set of all (or all interior) nodes. Let $\mathcal{F}_\times$ (or $\mathcal{F}_\times^\Omega$) denote the set of all (or all interior) facets. For $T \in \mathcal{T}_\times$, let $\mathcal{F}_T := \{F \in \mathcal{F}_\times : F \subseteq \partial T\}$ be the set of facets of T. Moreover,

$$\omega_\times(T) := \bigcup \{T' \in \mathcal{T}_\times : T \cap T' \neq \emptyset\} \subseteq \overline{\Omega} \tag{2.3}$$

denotes the element patch of T in $\mathcal{T}_\times$.



(a) $\mathcal{T}_\times$ (triangles) and $\mathcal{T}_\times^*$ (grey boxes).

(b) Refinement of a triangle by newest vertex bisection (NVB).

FIG. 1. Construction of the dual mesh $\mathcal{T}_\times^*$ (grey boxes) from the primal mesh $\mathcal{T}_\times$ (triangles) in two dimensions (left) and two-dimensional newest vertex bisection (NVB) (right). Each triangle has a reference edge (indicated by the double line). If edges are marked for refinement (indicated by dots), the resulting configurations are shown.

The associated *dual mesh* $\mathcal{T}_\times^*$ is obtained as follows: for $d = 2$, connect the center of gravity of an element $T \in \mathcal{T}_\times$ with the midpoint of an edge of ∂T. These lines define the nondegenerate closed polygons $V_i \in \mathcal{T}_\times^*$; see Fig. 1(a). For $d = 3$, we first connect the center of gravity of $T \in \mathcal{T}_\times$ with each center of gravity of the four faces of $F \in \mathcal{F}_T$ by straight lines. Then, as in the two-dimensional case, we connect each center of gravity of $F \in \mathcal{F}_T$ to the midpoints of the edges of the face F. Note that this forms polyhedrons $V_i \in \mathcal{T}_\times^*$. In two and three dimensions, each volume $V_i \in \mathcal{T}_\times^*$ is uniquely associated with a node a_i of $\mathcal{T}_\times$.

2.3 *Discrete spaces*

For a partition $\mathcal{M}$ of Ω and $p \in \mathbb{N}_0$, let

$$\mathcal{P}^p(\mathcal{M}) := \big\{v : \Omega \to \mathbb{R} \,:\, \forall M \in \mathcal{M}, \quad v|_M \text{ is a polynomial of degree } \le p\big\} \tag{2.4}$$

be the space of $\mathcal{M}$-piecewise polynomials of degree p. With this at hand, let

$$\mathcal{S}^1(\mathcal{T}_\times) := \mathcal{P}^1(\mathcal{T}_\times) \cap H^1(\Omega) = \big\{v_\times \in C(\Omega) \,:\, \forall T \in \mathcal{T}_\times, \quad v_\times|_T \text{ is affine}\big\}. \tag{2.5}$$

Then the discrete ansatz space

$$\mathcal{S}_0^1(\mathcal{T}_\times) := \mathcal{S}^1(\mathcal{T}_\times) \cap H_0^1(\Omega) = \big\{v_\times \in \mathcal{S}^1(\mathcal{T}_\times) \,:\, v_\times|_\Gamma = 0\big\} \tag{2.6}$$

consists of all $\mathcal{T}_\times$-piecewise affine and globally continuous functions that are zero on Γ. By convention, the discrete test space

$$\mathcal{P}_0^0\left(\mathcal{T}_\times^*\right) := \big\{v_\times^* \in \mathcal{P}^0\left(\mathcal{T}_\times^*\right) \,:\, v_\times^*|_\Gamma = 0\big\} \tag{2.7}$$

consists of all $\mathcal{T}_\times^*$-piecewise constant functions which are zero on all $V \in \mathcal{T}_\times^*$ with $\partial V \cap \Gamma \neq \emptyset$.

2.4 *Mesh refinements*

For local mesh refinement, we employ newest vertex bisection (NVB) (see, e.g., Stevenson, 2008, Karkulik *et al.*, 2013 and Fig. 1(b)). Below, we use the following notation: first, $\mathcal{T}' := \texttt{refine}(\mathcal{T}, \mathcal{M})$ denotes the coarsest conforming triangulation generated by NVB from a conforming triangulation $\mathcal{T}$ such that all marked elements $\mathcal{M} \subseteq \mathcal{T}$ have been refined, i.e., $\mathcal{M} \subseteq \mathcal{T} \setminus \mathcal{T}'$. Second, we simply write $\mathcal{T}' \in \texttt{refine}(\mathcal{T})$, if $\mathcal{T}'$ is an arbitrary refinement of $\mathcal{T}$, i.e., there exists a finite number of refinements steps $j = 1, \dots, n$ such that $\mathcal{T}' = \mathcal{T}_n'$ can be generated from $\mathcal{T} = \mathcal{T}_0'$ with marked elements $\mathcal{M}_j' \subseteq \mathcal{T}_j'$ and $\mathcal{T}_j' = \texttt{refine}(\mathcal{T}_{j-1}', \mathcal{M}_{j-1}')$. Note that NVB guarantees that there exist only finitely many shapes of triangles and patches in $\mathcal{T}' \in \texttt{refine}(\mathcal{T})$. These shapes are determined by $\mathcal{T}$. In particular, the meshes $\mathcal{T}' \in \texttt{refine}(\mathcal{T})$ are uniformly σ-shape regular (2.2), where σ depends only on $\mathcal{T}$.

2.5 *Vertex-centered FVM*

The FVM approximates the solution $u \in H^1_0(\Omega)$ of (1.5) by some $u_\times \in \mathscr{S}^1_0(\mathscr{T}_\times)$. The scheme is based on the balance equation over $\mathscr{T}^*_\times$ and reads in variational form as follows: find $u_\times \in \mathscr{S}^1_0(\mathscr{T}_\times)$ such that

$$\mathscr{A}_\times\left(u_\times, w^*_\times\right) = \left(f, w^*_\times\right)_\Omega = \sum_{a_i \in \mathscr{N}^\Omega_\times} w^*_\times|_{V_i} \int_{V_i} f \,\mathrm{d}x \quad \text{for all } w^*_\times \in \mathscr{P}^0_0\left(\mathscr{T}^*_\times\right). \tag{2.8}$$

For all $v_\times \in \mathscr{S}^1_0(\mathscr{T}_\times)$ and all $w^*_\times \in \mathscr{P}^0_0\left(\mathscr{T}^*_\times\right)$, the bilinear form reads

$$\mathscr{A}_\times\left(v_\times, w^*_\times\right) := \sum_{a_i \in \mathscr{N}^\Omega_\times} w^*_\times|_{V_i} \left(\int_{\partial V_i} (-\mathbf{A}\nabla v_\times + \mathbf{b} v_\times) \cdot \mathbf{n} \,\mathrm{d}s + \int_{V_i} c v_\times \,\mathrm{d}x \right).$$

To recall that the FVM is well posed on sufficiently fine triangulations $\mathscr{T}_\times$, we require the following interpolation operator (see, e.g., Erath, 2012 and Erath & Praetorius, 2016).

LEMMA 2.1 With $\chi^*_i \in \mathscr{P}^0\left(\mathscr{T}^*_\times\right)$ being the characteristic function of $V_i \in \mathscr{T}^*_\times$, define

$$\mathscr{I}^*_\times : \mathscr{C}(\overline{\Omega}) \to \mathscr{P}^0\left(\mathscr{T}^*_\times\right), \quad \mathscr{I}^*_\times v := \sum_{a_i \in \mathscr{N}_\times} v(a_i)\chi^*_i.$$

Then, for all $T \in \mathscr{T}_\times$, $F \in \mathscr{F}_T$ and $v_\times \in \mathscr{S}^1(\mathscr{T}_\times)$, it holds that

$$\int_T \left(v_\times - \mathscr{I}^*_\times v_\times\right) \mathrm{d}x = 0 = \int_F \left(v_\times - \mathscr{I}^*_\times v_\times\right) \mathrm{d}s, \tag{2.9}$$

$$\|v_\times - \mathscr{I}^*_\times v_\times\|_{L^2(T)} \le h_T \|\nabla v_\times\|_{L^2(T)}, \tag{2.10}$$

$$\|v_\times - \mathscr{I}^*_\times v_\times\|_{L^2(F)} \le C h_T^{1/2} \|\nabla v_\times\|_{L^2(T)}. \tag{2.11}$$

In particular, it holds that $\mathscr{I}^*_\times v_\times \in \mathscr{P}^0_0\left(\mathscr{T}^*_\times\right)$ for all $v_\times \in \mathscr{S}^1_0(\mathscr{T}_\times)$. The constant $C > 0$ depends only on the σ-shape regularity of $\mathscr{T}_\times$. □

The following lemma is a key observation for the FVM analysis. For Lipschitz continuous $\mathbf{A}$, the proof is found in Ewing *et al.* (2002) and Erath (2012). We note that the result transfers directly to the present situation (see Erath & Praetorius, 2016; 2017), where $\mathbf{A}$ satisfies (1.2)–(1.3), $\mathbf{b} \neq \mathbf{0}$ and $c \neq 0$.

LEMMA 2.2 There exists $C_{\mathrm{bil}} > 0$ such that for all $v_\times, w_\times \in \mathscr{S}^1_0(\mathscr{T}_\times)$,

$$|\mathscr{A}(v_\times, w_\times) - \mathscr{A}_\times\left(v_\times, \mathscr{I}^*_\times w_\times\right)| \le C_{\mathrm{bil}} \sum_{T \in \mathscr{T}_\times} h_T \, \|v_\times\|_{H^1(T)} \|w_\times\|_{H^1(T)}. \tag{2.12}$$

Moreover, let $\mathscr{T}_\times$ be sufficiently fine such that $C_{\mathrm{ell}} - C_{\mathrm{bil}} \|h_\times\|_{L^\infty(\Omega)} > 0$, where $C_{\mathrm{ell}} > 0$ is the ellipticity constant from (1.6). Then there exists $C_{\mathrm{stab}} > 0$ such that

$$\mathscr{A}_\times(v_\times, \mathscr{I}^*_\times v_\times) \ge C_{\mathrm{stab}} \, \|v_\times\|^2_{H^1(\Omega)} \quad \text{for all } v_\times \in \mathscr{S}^1_0(\mathscr{T}_\times). \tag{2.13}$$

In particular, the FVM system (2.8) admits a unique solution $u_\times \in \mathscr{S}^1_0(\mathscr{T}_\times)$. The constants $C_{\rm bil}$ and $C_{\rm stab}$ depend only on the data assumptions (1.2)–(1.4) and the σ-shape regularity of $\mathscr{T}_\times$ and Ω. □

2.6 *Weighted-residual a posteriori error estimator*

For all $v_\times \in \mathscr{S}^1_0(\mathscr{T}_\times)$, we define the volume residual $R_\times$ and the normal jump $J_\times$ by

$$R_\times(v_\times)|_T := f - \operatorname{div}_\times(-\mathbf{A}\nabla v_\times + \mathbf{b}v_\times) - cv_\times \qquad \text{for all } T \in \mathscr{T}_\times, \tag{2.14}$$

$$J_\times(v_\times)|_F := [\![\mathbf{A}\nabla v_\times]\!]_F \qquad \text{for all } F \in \mathscr{F}_\times^\Omega. \tag{2.15}$$

Here, $\operatorname{div}_\times$ denotes the $\mathscr{T}_\times$-piecewise divergence operator, and the normal jump reads $[\![\mathbf{g}]\!]|_F := (\mathbf{g}|_T - \mathbf{g}|_{T'}) \cdot \mathbf{n}$, where $\mathbf{g}|_T$ denotes the trace of $\mathbf{g}$ from T onto F and $\mathbf{n}$ points from T to T'. Let $\Pi_\times$ be the edgewise or elementwise integral mean operator, i.e.,

$$(\Pi_\times)|_\tau = \frac{1}{|\tau|}\int_\tau v\,\mathrm{d}x \qquad \text{for all } \tau \in \mathscr{T}_\times \cup \mathscr{F}_\times \text{ and all } v \in L^2(\tau).$$

For all $T \in \mathscr{T}_\times$, we define the local error indicators and oscillations by

$$\begin{aligned}\eta_\times(T, v_\times)^2 &:= h_T^2\,\|R_\times(v_\times)\|^2_{L^2(T)} + h_T\,\|J_\times(v_\times)\|^2_{L^2(\partial T\setminus\Gamma)},\\ \operatorname{osc}_\times(T, v_\times)^2 &:= h_T^2\,\|(1-\Pi_\times)R_\times(v_\times)\|^2_{L^2(T)} + h_T\,\|(1-\Pi_\times)J_\times(v_\times)\|^2_{L^2(\partial T\setminus\Gamma)}.\end{aligned} \tag{2.16}$$

Then the error estimator $\eta_\times$ and the oscillations $\operatorname{osc}_\times$ are defined by

$$\eta_\times(v_\times)^2 := \sum_{T\in\mathscr{T}_\times} \eta_\times(T, v_\times)^2 \qquad \text{and} \qquad \operatorname{osc}_\times^2(v_\times) := \sum_{T\in\mathscr{T}_\times} \operatorname{osc}_\times(T, v_\times)^2. \tag{2.17}$$

To abbreviate notation, we write $\eta_\times := \eta_\times(u_\times)$ and $\operatorname{osc}_\times := \operatorname{osc}_\times(u_\times)$. The following proposition is proved, e.g., in Carstensen *et al.* (2005) and Erath (2013).

Proposition 2.3 (Reliability and efficiency). The residual error estimator $\eta_\times$ satisfies

$$C_{\rm rel}^{-1}\,\|u - u_\times\|^2_{H^1(\Omega)} \le \eta_\times^2 \le C_{\rm eff}\left(\|u - u_\times\|^2_{H^1(\Omega)} + \operatorname{osc}_\times^2\right), \tag{2.18}$$

where $C_{\rm rel}$, $C_{\rm eff} > 0$ depend only on the σ-shape regularity of $\mathscr{T}_\times$, the data assumptions (1.2)–(1.4) and Ω. □

Note that a robust variant of this estimator with respect to an energy norm is found and analysed in Erath (2013, Theorems 4.9, 6.3 and Remark 6.1), where we additionally require the assumption $\|\operatorname{div}\mathbf{b} + c\|_{L^\infty(\Omega)} \le C\big(\frac{1}{2}\operatorname{div}\mathbf{b} + c\big)$ with $C > 0$. One of the key ingredients to prove Proposition 2.3 is (2.19) of the following lemma which will be employed below. The proof of the orthogonality relation (2.19) is well known and found, e.g., in Carstensen *et al.* (2005), Erath (2010; 2013). The discrete defect identity (2.20) is proved in Erath & Praetorius (2016, Lemma 16) for symmetric problems on arbitrary refinements of meshes. This result can easily be transferred to the present model problem (1.1).

LEMMA 2.4 Let $\mathcal{T}_\diamond \in \texttt{refine}(\mathcal{T}_0)$ and $\mathcal{T}_\times \in \texttt{refine}(\mathcal{T}_\diamond)$. Suppose that the discrete solutions $u_\times \in \mathcal{S}^1_0(\mathcal{T}_\times)$ or $u_\diamond \in \mathcal{S}^1_0(\mathcal{T}_\times)$ exist. Then there holds the L^2-orthogonality

$$\sum_{T\in\mathcal{T}_\diamond} \big(R_\diamond(u_\diamond), v^*_\diamond\big)_T - \sum_{F\in\mathcal{F}^\Omega_\diamond} \big(J_\diamond(u_\diamond), v^*_\diamond\big)_F = 0 \quad \text{for all } v^*_\diamond \in \mathcal{P}^0_0\big(\mathcal{T}^*_\diamond\big), \tag{2.19}$$

as well as the discrete defect identity

$$\sum_{T\in\mathcal{T}_\diamond} \big(R_\diamond(u_\diamond), v^*_\times\big)_T - \sum_{F\in\mathcal{F}^\Omega_\diamond} \big(J_\diamond(u_\diamond), v^*_\times\big)_F = \mathcal{A}_\times\big(u_\times - u_\diamond, v^*_\times\big) \quad \text{for all } v^*_\times \in \mathcal{P}^0_0\big(\mathcal{T}^*_\times\big). \quad \square \tag{2.20}$$

2.7 *Comparison result and a priori error estimate*

The following proposition states that the FVM error estimator is equivalent to the optimal *total error* (i.e., error plus oscillations) and so improves Proposition 2.3. The result is first proved in Erath & Praetorius (2016) for $\mathbf{b} = \mathbf{0}$ and $c = 0$ and generalized to the present model problem in Erath & Praetorius (2017).

PROPOSITION 2.5 Let $\mathcal{T}_\times$ be sufficiently fine such that $C_{\rm ell} - C_{\rm bil}\|h_\times\|_{L^\infty(\Omega)} > 0$ with $C_{\rm ell}$ and $C_{\rm bil}$ from (1.6) and (2.12), respectively. Then it holds that

$$C_1^{-1}\,\eta_\times \le \min_{v_\times\in\mathcal{S}^1_0(\mathcal{T}_\times)} \big(\|u - v_\times\|_{H^1(\Omega)} + \operatorname{osc}_\times(v_\times)\big) \le \|u - u_\times\|_{H^1(\Omega)} + \operatorname{osc}_\times \le C_1\,\eta_\times. \tag{2.21}$$

Moreover, if $u^{\rm FEM}_\times \in \mathcal{S}^1_0(\mathcal{T}_\times)$ denotes the FEM solution of $\mathcal{A}(u^{\rm FEM}_\times, w_\times) = (f, w_\times)_\Omega$ for all $w_\times \in \mathcal{S}^1_0(\mathcal{T}_\times)$, it holds that

$$\begin{aligned} C_2^{-1}\big(\|u - u_\times\|_{H^1(\Omega)} + \operatorname{osc}_\times\big) &\le \|u - u^{\rm FEM}_\times\|_{H^1(\Omega)} + \operatorname{osc}_\times\Big(u^{\rm FEM}_\times\Big) \\ &\le C_2\big(\|u - u_\times\|_{H^1(\Omega)} + \operatorname{osc}_\times\big). \end{aligned}$$

The constants C_1, $C_2 > 0$ depend only on Ω, the σ-shape regularity of $\mathcal{T}_\times$ and the data assumptions (1.2)–(1.4). $\square$

As a direct consequence of Proposition 2.5, one obtains the following *convergence* result and *a priori* estimate that confirms first-order convergence of FVM (see again Erath & Praetorius, 2016; 2017). Note that the statement even holds for $u \in H^1_0(\Omega)$, whereas in the literature standard FVM analysis usually requires, e.g., $u \in H^{1+\varepsilon}(\Omega)$ for some $\varepsilon > 0$.

COROLLARY 2.6 Let $\{\mathcal{T}_\times\}$ be a family of sufficiently fine and uniformly σ-shape-regular triangulations. Let $u \in H^1_0(\Omega)$ be the solution of (1.5). Then there holds convergence

$$\|u - u_\times\|_{H^1(\Omega)} + \operatorname{osc}_\times \to 0 \quad \text{as} \quad \|h_\times\|_{L^\infty(\Omega)} \to 0.$$

Moreover, additional regularity $u \in H^1_0(\Omega) \cap H^2(\Omega)$ implies first-order convergence

$$\|u - u_\times\|_{H^1(\Omega)} + \operatorname{osc}_\times = \mathcal{O}(\|h_\times\|_{L^\infty(\Omega)}). \qquad \square$$

3. Adaptive FVM

In this section, we apply an adaptive mesh-refining algorithm for FVM. We combine ideas from Mekchay & Nochetto (2005) and Erath & Praetorius (2016) to prove that adaptive FVM leads to linear convergence with optimal algebraic rates for the error estimator (and hence for the total error; see Proposition 2.5).

3.1 *Adaptive algorithm*

As in Erath & Praetorius (2016), we employ the following adaptive algorithm.

ALGORITHM 3.1. **Input:** Let $0 < \theta' \leq \theta \leq 1$ and $C_{\mathrm{mark}}, C'_{\mathrm{mark}} \geq 1$. Let $\mathcal{T}_0$ be a conforming triangulation of Ω that resolves possible discontinuities of $\mathbf{A}$.

Loop: For $\ell = 0, 1, 2, \ldots$, iterate the following steps (i)–(v):

(i) **Solve:** Compute the discrete solution $u_\ell \in \mathcal{S}^1_0(\mathcal{T}_\ell)$ from (2.8).

(ii) **Estimate:** Compute $\eta_\ell(T, u_\ell)$ and $\mathrm{osc}_\ell(T, u_\ell)$ from (2.16) for all $T \in \mathcal{T}_\ell$.

(iii) **Mark I:** Find $\mathcal{M}^\eta_\ell \subseteq \mathcal{T}_\ell$ of up to the multiplicative constant $C_{\mathrm{mark}} \geq 1$ minimal cardinality that satisfies the Dörfler marking criterion

$$\theta \sum_{T \in \mathcal{T}_\ell} \eta_\ell(T, u_\ell)^2 \leq \sum_{T \in \mathcal{M}^\eta_\ell} \eta_\ell(T, u_\ell)^2. \tag{3.1}$$

(iv) **Mark II:** Find $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of up to the multiplicative constant $C'_{\mathrm{mark}} \geq 1$ minimal cardinality that satisfies $\mathcal{M}^\eta_\ell \subseteq \mathcal{M}_\ell$ as well as the Dörfler marking criterion

$$\theta' \sum_{T \in \mathcal{T}_\ell} \mathrm{osc}_\ell(T, u_\ell)^2 \leq \sum_{T \in \mathcal{M}_\ell} \mathrm{osc}_\ell(T, u_\ell)^2. \tag{3.2}$$

(v) **Refine:** Generate a new triangulation $\mathcal{T}_{\ell+1} := \texttt{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$ by refinement of all marked elements.

Output: Adaptively refined triangulations $\mathcal{T}_\ell$, corresponding discrete solutions u_ℓ, estimators η_ℓ and data oscillations osc_ℓ for $\ell \geq 0$.

Due to the lack of standard Galerkin orthogonality (see Section 3.2), we additionally have to mark the oscillations (3.2). In practice, however, this marking is negligible, since θ' can be chosen arbitrarily small (see Erath & Praetorius, 2016, Remark 7 for more details).

3.2 *Quasi-Galerkin orthogonality*

Given $g \in L^2(\Omega)$, we consider the dual problem: find $\phi \in H^1_0(\Omega)$ such that

$$\mathcal{A}(v, \phi) = (g, v)_\Omega \quad \text{for all } v \in H^1_0(\Omega). \tag{3.3}$$

The Lax–Milgram theorem proves existence and uniqueness of $\phi \in H^1_0(\Omega)$. Let $0 < s \leq 1$. We suppose that the dual problem (3.3) is H^{1+s}-regular, i.e., there exists a constant $C_{\mathrm{dual}} > 0$ such that for all

$g \in L^2(\Omega)$, the solution of (3.3) satisfies

$$\phi \in H^1_0(\Omega) \cap H^{1+s}(\Omega) \quad \text{with} \quad \|\phi\|_{H^{1+s}(\Omega)} \leq C_{\text{dual}} \, \|g\|_{L^2(\Omega)}. \tag{3.4}$$

We refer to Grisvard (1985) for a discussion on this regularity assumption. The main result of this section is the following quasi-Galerkin orthogonality with respect to the operator-induced quasi-norm from (1.6). The proof is postponed to the end of this section.

PROPOSITION 3.2 Let $0 < s \leq 1$ and suppose that the dual problem (3.3) is H^{1+s}-regular (3.4). Let $\mathscr{T}_\diamond \in \texttt{refine}(\mathscr{T}_0)$ and $\mathscr{T}_\times \in \texttt{refine}(\mathscr{T}_\diamond)$. Then there exists $C_{\text{gal}} > 0$ such that

$$|||u - u_\times|||^2 \leq |||u - u_\diamond|||^2 - \tfrac{1}{2} \, |||u_\times - u_\diamond|||^2 + C_{\text{gal}} \, \|h_\times\|^{2s}_{L^\infty(\Omega)} \eta^2_\times + C_{\text{gal}} \, \text{osc}^2_\times. \tag{3.5}$$

The constant $C_{\text{gal}} > 0$ depends only on $C_{\text{dual}}, C_{\text{osc}}, C_{\text{rel}}, C_{\text{ell}}, C_{\text{cont}}$, $\text{diam}(\Omega)$ and $\|\mathbf{b}\|_{W^{1,\infty}(\Omega)}$ as well as on σ-shape regularity and all possible shapes of element patches in $\mathscr{T}_\times$.

For the FVM error, the classical Galerkin orthogonality fails, i.e., $\mathscr{A}(u - u_\times, v_\times) \neq 0$ for some $v_\times \in \mathscr{S}^1_0(\mathscr{T}_\times)$. However, there holds the following estimate (see, e.g. Erath & Praetorius, 2016).

LEMMA 3.3 The FVM error $u - u_\times$ satisfies

$$|\mathscr{A}(u - u_\times, v_\times)| \leq C_{\text{osc}} \, \|v_\times\|_{H^1(\Omega)} \, \text{osc}_\times \quad \text{for all } v_\times \in \mathscr{S}^1_0(\mathscr{T}_\times). \tag{3.6}$$

The constant $C_{\text{osc}} > 0$ depends only on the σ-shape regularity of $\mathscr{T}_\times$.

Proof. Standard calculations (see, e.g., Erath, 2013, Theorem 4.9) show

$$\mathscr{A}(u - u_\times, v_\times) = \sum_{T \in \mathscr{T}_\times} \int_T R_\times(u_\times) \, v_\times \, \mathrm{d}x + \sum_{F \in \mathscr{F}^\Omega_\times} \int_F J_\times(u_\times) \, v_\times \, \mathrm{d}s.$$

Together with (2.19) for $v^*_\times = \mathscr{I}^*_\times v_\times \in \mathscr{P}^0_0\left(\mathscr{T}^*_\times\right)$, this leads to

$$\mathscr{A}(u - u_\times, v_\times) = \sum_{T \in \mathscr{T}_\times} \int_T R_\times(u_\times) \, \left(v_\times - v^*_\times\right) \, \mathrm{d}x + \sum_{F \in \mathscr{F}^\Omega_\times} \int_F J_\times(u_\times) \, \left(v_\times - v^*_\times\right) \, \mathrm{d}s.$$

We apply (2.9) for the involved integrals and obtain

$$\begin{aligned} \mathscr{A}(u - u_\times, v_\times) = &\sum_{T \in \mathscr{T}_\times} \int_T (R_\times(u_\times) - \Pi_\times R_\times(u_\times)) \, \left(v_\times - v^*_\times\right) \, \mathrm{d}x \\ &+ \sum_{F \in \mathscr{F}^\Omega_\times} \int_F (J_\times(u_\times) - \Pi_\times J_\times(u_\times)) \, \left(v_\times - v^*_\times\right) \, \mathrm{d}s. \end{aligned}$$

The Cauchy–Schwarz inequality and (2.10)–(2.11) conclude the proof. □

LEMMA 3.4 Let $0 < s \leq 1$ and suppose that the dual problem (3.3) is H^{1+s}-regular (3.4). Then the FVM error satisfies

$$C_{\mathrm{aux}}^{-1} \|u - u_\times\|_{L^2(\Omega)}^2 \leq \|h_\times\|_{L^\infty(\Omega)}^{2s} \|u - u_\times\|_{H^1(\Omega)}^2 + \mathrm{osc}_\times^2. \tag{3.7}$$

The constant $C_{\mathrm{aux}} > 0$ depends only on the σ-shape regularity of $\mathcal{T}_\times$, $\mathrm{diam}(\Omega)$, C_{cont} and C_{dual} as well as on all possible shapes of element patches in $\mathcal{T}_\times$.

Proof. We split the proof into two steps.

Step 1. Let $\mathcal{I}_\times : H^1(\Omega) \to \mathcal{S}^1(\mathcal{T}_\times)$ be the Scott–Zhang projector (Scott & Zhang, 1990). Recall the following properties of $\mathcal{I}_\times$ for all $v \in H^1(\Omega)$ and $v_\times \in \mathcal{S}^1(\mathcal{T}_\times)$ and all $T \in \mathcal{T}_\times$:

- $\mathcal{I}_\times$ has a local projection property, i.e., $(\mathcal{I}_\times v)|_T = v_\times|_T$ if $v|_{\omega_\times(T)} = v_\times|_{\omega_\times(T)}$;
- $\mathcal{I}_\times$ preserves discrete boundary data, i.e., $v|_\Gamma = v_\times|_\Gamma$ implies that $(\mathcal{I}_\times v)|_\Gamma = v|_\Gamma$;
- $\mathcal{I}_\times$ is locally H^1-stable, i.e., $\|\nabla \mathcal{I}_\times v\|_{L^2(T)} \leq C_{\mathrm{sz}} \|\nabla v\|_{H^1(\omega_\times(T))}$;
- $\mathcal{I}_\times$ has a local approximation property, i.e., $\|v - \mathcal{I}_\times v\|_{L^2(T)} \leq C_{\mathrm{sz}}\, h_T \|\nabla v\|_{H^1(\omega_\times(T))}$.

The constant $C_{\mathrm{sz}} > 0$ depends only on the σ-shape regularity of $\mathcal{T}_\times$. In particular,

$$\|v - \mathcal{I}_\times v\|_{H^1(\Omega)} \lesssim \|v\|_{H^1(\Omega)} \quad \text{for all } v \in H^1(\Omega),$$

where the hidden constant depends only on C_{sz} and $\mathrm{diam}(\Omega)$. With the local projection property of $\mathcal{I}_\times$, we may apply the Bramble–Hilbert lemma. For $v \in H^2(\Omega)$, scaling arguments then prove that

$$\|v - \mathcal{I}_\times v\|_{H^1(T)} \lesssim \mathrm{diam}(\omega_\times(T)) \|v\|_{H^2(\omega_\times(T))} \quad \text{for all } T \in \mathcal{T}_\times, \tag{3.8}$$

where the hidden constant depends only on the shape of $\omega_\times(T)$ and on the operator norm of $A := 1 - \mathcal{I}_\times$ (and hence on $\mathrm{diam}(\Omega)$ and C_{sz}) Altogether, this proves the operator norm estimates

$$\|A := 1 - \mathcal{I}_\times : H^{1+t}(\Omega) \to H^1(\Omega)\| \leq C \|h_\times\|_{L^\infty(\Omega)}^t \quad \text{for } t \in \{0, 1\}, \tag{3.9}$$

where $C > 0$ depends only on C_{sz}, $\mathrm{diam}(\Omega)$ and all possible shapes of element patches in $\mathcal{T}_\times$. Interpolation arguments (Bergh & Löfström, 1976) conclude that (3.9) holds for all $0 \leq t \leq 1$. For $t = s$, this proves

$$\|v - \mathcal{I}_\times v\|_{H^1(\Omega)} \leq C \|h_\times\|_{L^\infty(\Omega)}^s \|v\|_{H^{1+s}(\Omega)} \quad \text{for all } v \in H^{1+s}(\Omega). \tag{3.10}$$

Step 2. With $g = v = u - u_\times$ in (3.3), it holds that

$$\|u - u_\times\|_{L^2(\Omega)}^2 = \mathcal{A}(u - u_\times, \phi) = \mathcal{A}(u - u_\times, \phi - \mathcal{I}_\times \phi) + \mathcal{A}(u - u_\times, \mathcal{I}_\times \phi).$$

Since we suppose $\phi \in H_0^1(\Omega) \cap H^{1+s}(\Omega)$, the first summand is bounded by (3.10). This yields

$$\begin{aligned} \mathcal{A}(u - u_\times, \phi - \mathcal{I}_\times \phi) &\lesssim \|u - u_\times\|_{H^1(\Omega)} \|\phi - \mathcal{I}_\times \phi\|_{H^1(\Omega)} \\ &\lesssim \|h_\times\|_{L^\infty(\Omega)}^s \|u - u_\times\|_{H^1(\Omega)} \|\phi\|_{H^{1+s}(\Omega)}, \end{aligned}$$

where the hidden constant depends only on $C_{\rm cont}$, $C_{\rm sz}$ and $\operatorname{diam}(\Omega)$. The second summand is bounded by (3.6) and H^1-stability of $\mathcal{I}_\times$. This yields that

$$\mathcal{A}(u-u_\times,\mathcal{I}_\times\phi)\lesssim \operatorname{osc}_\times\,\|\mathcal{I}_\times\phi\|_{H^1(\Omega)}\lesssim \operatorname{osc}_\times\,\|\phi\|_{H^1(\Omega)}\le \operatorname{osc}_\times\,\|\phi\|_{H^{1+s}(\Omega)},$$

where the hidden constant depends only on $C_{\rm osc}$, $C_{\rm sz}$ and $\operatorname{diam}(\Omega)$. Combining the latter three estimates with H^{1+s}-regularity (3.4), we prove

$$\begin{aligned}\|u-u_\times\|_{L^2(\Omega)}^2&\lesssim \big(\|h_\times\|_{L^\infty(\Omega)}^s\,\|u-u_\times\|_{H^1(\Omega)}+\operatorname{osc}_\times\big)\,\|\phi\|_{H^{1+s}(\Omega)}\\&\lesssim \big(\|h_\times\|_{L^\infty(\Omega)}^s\,\|u-u_\times\|_{H^1(\Omega)}+\operatorname{osc}_\times\big)\,\|u-u_\times\|_{L^2(\Omega)},\end{aligned}$$

where the hidden constant depends additionally on $C_{\rm dual}$. This concludes the proof. □

Proof of Proposition 3.2. Recall that $\mathcal{A}(v,w)=(\mathbf{A}\nabla v,\nabla w)_\Omega-(\mathbf{b}v,\nabla w)_\Omega+(cv,w)_\Omega$ and thus $\mathcal{A}(w,v)=(\mathbf{A}\nabla w,\nabla v)_\Omega-(\mathbf{b}w,\nabla v)_\Omega+(cw,v)_\Omega$. For $v,w\in H^1_0(\Omega)$, integration by parts proves

$$-(\mathbf{b}w,\nabla v)_\Omega=(\mathbf{b}\cdot\nabla w,v)_\Omega+(\operatorname{div}(\mathbf{b})\,w,v)_\Omega$$

and hence

$$\mathcal{A}(v,w)+\mathcal{A}(w,v)=2\mathcal{A}(v,w)+2(v,\mathbf{b}\cdot\nabla w)_\Omega+(\operatorname{div}(\mathbf{b})\,v,w)_\Omega.$$

By definition of $|||\cdot|||$, this proves

$$\begin{aligned}|||v+w|||^2&=|||v|||^2+|||w|||^2+\mathcal{A}(v,w)+\mathcal{A}(w,v)\\&=|||v|||^2+|||w|||^2+2\mathcal{A}(v,w)+2(v,\mathbf{b}\cdot\nabla w)_\Omega+(\operatorname{div}(\mathbf{b})\,v,w)_\Omega.\end{aligned}$$

This leads to

$$|||v|||^2=|||v+w|||^2-|||w|||^2-2\mathcal{A}(v,w)-2(v,\mathbf{b}\cdot\nabla w)_\Omega-(\operatorname{div}(\mathbf{b})\,v,w)_\Omega.$$

With $C_1:=C_{\rm ell}^{-1}\,(2\|\mathbf{b}\|_{L^\infty(\Omega)}+\|\operatorname{div}\mathbf{b}\|_{L^\infty(\Omega)})^2$, the Young inequality $ab\le\frac14\,a^2+b^2$ and norm equivalence (1.6) prove

$$\begin{aligned}-2(v,\mathbf{b}\cdot\nabla w)_\Omega-(\operatorname{div}(\mathbf{b})\,v,w)_\Omega&\le\|v\|_{L^2(\Omega)}\|w\|_{H^1(\Omega)}\,\big(2\|\mathbf{b}\|_{L^\infty(\Omega)}+\|\operatorname{div}\mathbf{b}\|_{L^\infty(\Omega)}\big)\\&\le\tfrac14\,|||w|||^2+C_1\,\|v\|_{L^2(\Omega)}^2.\end{aligned}$$

Choose $v=u-u_\times$ as well as $w=u_\times-u_\diamond$. So far, we have shown

$$|||u-u_\times|||^2\le|||u-u_\diamond|||^2-\tfrac34\,|||u_\times-u_\diamond|||^2-2\mathcal{A}(u-u_\times,u_\times-u_\diamond)+C_1\,\|u-u_\times\|_{L^2(\Omega)}^2.$$

We apply (3.6), norm equivalence (1.6) and the Young inequality $2ab \le \frac{1}{4}\,a^2 + 4b^2$ to see that

$$\begin{aligned} -2\mathscr{A}(u-u_\times, u_\times - u_\diamond) &\le 2\,C_{\rm osc}\,\|u_\times - u_\diamond\|_{H^1(\Omega)}\,{\rm osc}_\times \\ &\le 2\,C_{\rm osc} C_{\rm ell}^{-1/2}\,|||u_\times - u_\diamond|||\,{\rm osc}_\times \le \tfrac{1}{4}\,|||u_\times - u_\diamond|||^2 + 4\,C_{\rm osc}^2 C_{\rm ell}^{-1}\,{\rm osc}_\times^2. \end{aligned}$$

Next, Lemma 3.4 and reliability (2.18) lead to

$$C_{\rm aux}^{-1}\,\|u-u_\times\|_{L^2(\Omega)}^2 \le \|h_\times\|_{L^\infty(\Omega)}^{2s}\,\|u-u_\times\|_{H^1(\Omega)}^2 + {\rm osc}_\times^2 \le C_{\rm rel}\,\|h_\times\|_{L^\infty(\Omega)}^{2s}\,\eta_\times^2 + {\rm osc}_\times^2.$$

Combining the latter three estimates, we prove

$$\begin{aligned} |||u-u_\times|||^2 \le |||u-u_\diamond|||^2 - \tfrac{1}{2}\,|||u_\times - u_\diamond|||^2 \\ + C_1 C_{\rm aux} C_{\rm rel}\,\|h_\times\|_{L^\infty(\Omega)}^{2s}\,\eta_\times^2 + \Big(4\,C_{\rm osc}^2 C_{\rm ell}^{-1} + C_1 C_{\rm aux}\Big)\,{\rm osc}_\times^2. \end{aligned}$$

Choosing $C_{\rm gal} = \max\big\{C_1 C_{\rm aux} C_{\rm rel}\,,\, 4C_{\rm osc}^2 C_{\rm ell}^{-1} + C_1 C_{\rm aux}\big\}$, we conclude the proof. □

3.3 *Linear convergence and general quasi-orthogonality*

The following properties (A1)–(A2) of the estimator and (B1)–(B2) of the oscillations are some key observations to prove linear convergence of Algorithm 3.1. The proofs for a symmetric problem are based on scaling arguments and can be found in the literature, (see, e.g., Cascón *et al.*, 2008, Section 3.1 for (A1)–(A2) and Erath & Praetorius, 2016, Section 3.3 for (B1)–(B2)). These proofs apply almost verbatim to the present nonsymmetric problem with $\mathbf{b} \neq \mathbf{0}$. Therefore, the details are left to the reader.

LEMMA 3.5 There exist constants $0 < q < 1$ and $C > 0$ such that for all $\mathscr{T}_\diamond \in \texttt{refine}(\mathscr{T}_0)$, all $\mathscr{T}_\times \in \texttt{refine}(\mathscr{T}_\diamond)$ and all $v_\times \in \mathscr{S}_0^1(\mathscr{T}_\times)$, $v_\diamond \in \mathscr{S}_0^1(\mathscr{T}_\diamond)$, it holds that

(stability of estimator on nonrefined elements)

$$\left|\left(\sum_{T\in\mathscr{T}_\times\cap\mathscr{T}_\diamond}\eta_\times(T,v_\times)^2\right)^{1/2} - \left(\sum_{T\in\mathscr{T}_\times\cap\mathscr{T}_\diamond}\eta_\diamond(T,v_\diamond)^2\right)^{1/2}\right| \le C\,\|v_\times - v_\diamond\|_{H^1(\Omega)}, \tag{A1}$$

(reduction of estimator on refined elements)

$$\sum_{T\in\mathscr{T}_\times\setminus\mathscr{T}_\diamond}\eta_\times(T,v_\times)^2 \le q\sum_{T\in\mathscr{T}_\diamond\setminus\mathscr{T}_\times}\eta_\diamond(T,v_\diamond)^2 + C\,\|v_\times - v_\diamond\|_{H^1(\Omega)}^2, \tag{A2}$$

(stability of oscillations on nonrefined elements)

$$\left| \left(\sum_{T \in \mathcal{T}_\times \cap \mathcal{T}_\diamond} \operatorname{osc}_\times(T, v_\times)^2 \right)^{1/2} - \left(\sum_{T \in \mathcal{T}_\times \cap \mathcal{T}_\diamond} \operatorname{osc}_\diamond(T, v_\diamond)^2 \right)^{1/2} \right| \leq C \, \|h_\times\|_{L^\infty(\Omega)} \|v_\times - v_\diamond\|_{H^1(\Omega)}, \quad \text{(B1)}$$

(reduction of oscillations on refined elements)

$$\sum_{T \in \mathcal{T}_\times \setminus \mathcal{T}_\diamond} \operatorname{osc}_\times(T, v_\times)^2 \leq q \sum_{T \in \mathcal{T}_\diamond \setminus \mathcal{T}_\times} \operatorname{osc}_\diamond(T, v_\diamond)^2 + C \, \|h_\times\|^2_{L^\infty(\Omega)} \|v_\times - v_\diamond\|^2_{H^1(\Omega)}. \quad \text{(B2)}$$

The constants $0 < q < 1$ and $C > 0$ depend only on the σ-shape regularity (2.2) and on the data assumptions (1.2)–(1.4). □

THEOREM 3.6 (Linear convergence). Let $0 < \theta' \leq \theta \leq 1$. There exists $H > 0$ such that the following statement is valid provided that $\|h_0\|_{L^\infty(\Omega)} \leq H$ and that the dual problem (3.3) is H^{1+s}-regular (3.4) for some $0 < s \leq 1$: there exist $C_{\text{lin}} > 0$ and $0 < q_{\text{lin}} < 1$ such that Algorithm 3.1 guarantees linear convergence in the sense of

$$\eta^2_{\ell+n} \leq C_{\text{lin}} q^n_{\text{lin}} \, \eta^2_\ell \quad \text{for all } \ell, n \in \mathbb{N}_0. \quad (3.11)$$

The constant H depends only on the σ-shape regularity (2.2), on the data assumptions (1.2)–(1.4), C_{gal}, θ and θ', whereas C_{lin} and q_{lin} additionally depend on C_{cont} and C_{rel}.

Proof. We split the proof into three steps.

Step 1. There exist constants $C > 0$ and $0 < q < 1$ which depend only on $0 < \theta \leq 1$, C_{ell} and the constants in (A1)–(A2), such that

$$\eta^2_{\ell+1} \leq q \, \eta^2_\ell + C \, |||u_{\ell+1} - u_\ell|||^2 \quad \text{for all } \ell \in \mathbb{N}_0. \quad (3.12)$$

Furthermore, there exist constants $C > 0$ and $0 < q < 1$ which depend only on $0 < \theta' \leq 1$, C_{ell} and the constants in (B1)–(B2), such that

$$\operatorname{osc}^2_{\ell+1} \leq q \operatorname{osc}^2_\ell + C \, \|h_{\ell+1}\|^2_{L^\infty(\Omega)} \, |||u_{\ell+1} - u_\ell|||^2 \quad \text{for all } \ell \in \mathbb{N}_0. \quad (3.13)$$

The proofs of (3.12) and (3.13) rely only on (A1)–(A2) with the Dörfler marking (3.1) and (B1)–(B2) with marking (3.2), respectively. For details, we refer, e.g., to Erath & Praetorius (2016, Proposition 10, steps 1 and 2).

Step 2. Without loss of generality, we may assume that the constants $C > 0$ and $0 < q < 1$ in (3.12)–(3.13) are the same. With free parameters $\gamma, \mu > 0$, we define

$$\Delta_\times := |||u - u_\times|||^2 + \gamma \, \eta^2_\times + \mu \operatorname{osc}^2_\times.$$

We employ the quasi-Galerkin orthogonality (3.5) and obtain

$$\Delta_{\ell+1} \le |||u-u_\ell|||^2 + \Big[\gamma + C_{\rm gal}\,\|h_{\ell+1}\|^{2s}_{L^\infty(\Omega)}\Big]\,\eta^2_{\ell+1} + \big[\mu + C_{\rm gal}\big]\,{\rm osc}^2_{\ell+1} - \tfrac12\,|||u_{\ell+1}-u_\ell|||^2.$$

Using (3.12)–(3.13), we further derive

$$\begin{aligned}\Delta_{\ell+1} \le |||u-u_\ell|||^2 &+ \Big[\gamma + C_{\rm gal}\,\|h_{\ell+1}\|^{2s}_{L^\infty(\Omega)}\Big]\,q\,\eta^2_\ell + \big[\mu + C_{\rm gal}\big]\,q\,{\rm osc}^2_\ell\\ &- \Big(\tfrac12 - C\,\Big[\gamma + C_{\rm gal}\,\|h_{\ell+1}\|^{2s}_{L^\infty(\Omega)}\Big] - C\,\|h_{\ell+1}\|^2_{L^\infty(\Omega)}\big[\mu + C_{\rm gal}\big]\Big)\,|||u_{\ell+1}-u_\ell|||^2.\end{aligned}$$

Let $H > 0$ be a free parameter and suppose that $\|h_0\|_{L^\infty(\Omega)} \le H$. We estimate $\|h_{\ell+1}\|_{L^\infty(\Omega)} \le \|h_0\|_{L^\infty(\Omega)} \le H$. Norm equivalence (1.6) and reliability (2.18) prove

$$|||u-u_\ell|||^2 \le C_{\rm cont}\,\|u-u_\ell\|^2_{H^1(\Omega)} \le C_{\rm cont}C_{\rm rel}\,\eta^2_\ell.$$

Let $\varepsilon > 0$ be a free parameter. Combining the last two estimates, we see that

$$\begin{aligned}\Delta_{\ell+1} \le (1-\varepsilon)\,|||u-u_\ell|||^2 + \gamma\,&\Big[\Big(1+\gamma^{-1}C_{\rm gal}\,H^{2s}\Big)\,q + \gamma^{-1}\varepsilon\,C_{\rm cont}C_{\rm rel}\Big]\,\eta^2_\ell + \mu\,\Big[1+\mu^{-1}C_{\rm gal}\Big]\,q\,{\rm osc}^2_\ell\\ &- \Big(\tfrac12 - C\,\Big[\gamma + C_{\rm gal}\,H^{2s}\Big] - C\,H^2\big[\mu + C_{\rm gal}\big]\Big)\,|||u_{\ell+1}-u_\ell|||^2.\end{aligned}$$

Step 3. It only remains to fix the four free parameters γ, μ, ε and H.

- Choose $\gamma > 0$ sufficiently small such that $\gamma C < 1/2$.
- Choose $\mu > 0$ sufficiently large such that $q_{\rm osc} := \big[1+\mu^{-1}C_{\rm gal}\big]\,q < 1$.
- Choose H sufficiently small such that
  - $C\,\big[\gamma + C_{\rm gal}\,H^{2s}\big] + C\,H^2\big[\mu + C_{\rm gal}\big] < 1/2$,
  - $\big(1+\gamma^{-1}C_{\rm gal}\,H^{2s}\big)\,q < 1$.
- Choose $0 < \varepsilon < 1$ such that $q_{\rm est} := \big[\big(1+\gamma^{-1}C_{\rm gal}\,H^{2s}\big)\,q + \gamma^{-1}\varepsilon\,C_{\rm cont}C_{\rm rel}\big] < 1$.

With $q_{\rm lin} := \max\{\,1-\varepsilon\,,\,q_{\rm est}\,,\,q_{\rm osc}\,\}$, we then obtain that

$$\begin{aligned}\Delta_{\ell+1} &\le (1-\varepsilon)\,|||u-u_\ell|||^2 + \gamma\,\Big[\Big(1+\gamma^{-1}C_{\rm gal}\,H^{2s}\Big)\,q + \gamma^{-1}\varepsilon\,C_{\rm cont}C_{\rm rel}\Big]\,\eta^2_\ell + \mu\,\big[1+\mu^{-1}C_{\rm gal}\big]\,q\,{\rm osc}^2_\ell\\ &\le \max\{\,1-\varepsilon\,,\,q_{\rm est}\,,\,q_{\rm osc}\,\}\,\Delta_\ell = q_{\rm lin}\,\Delta_\ell.\end{aligned}$$

Induction on n, norm equivalence (1.6), reliability (2.18) and ${\rm osc}^2_\ell \le \eta^2_\ell$ prove

$$\gamma\,\eta^2_{\ell+n} \le \Delta_{\ell+n} \le q^n_{\rm lin}\Delta_\ell \le q^n_{\rm lin}\,(C_{\rm rel}C_{\rm cont} + \gamma + \mu)\,\eta^2_\ell \quad \text{for all } \ell, n \in \mathbb{N}_0.$$

This concludes linear convergence (3.11) with $C_{\rm lin} = (C_{\rm rel}C_{\rm cont} + \gamma + \mu)\gamma^{-1}$. □

REMARK 3.7 In the above proof, we could apply the relation $\operatorname{osc}_\ell^2 \leq \eta_\ell^2$. Hence, we could avoid using (3.13). Consequently, Algorithm 3.1 would not need marking (3.2) of oscillations. However, the expression $\big(1+\gamma^{-1}C_{\mathrm{gal}}\,H^{2s}\big)\,q$ in Step 3 of the foregoing proof would become $\big(1+\gamma^{-1}C_{\mathrm{gal}}\,\big(1+H^{2s}\big)\big)\,q$ which is not less than 1 as required for the analysis. Hence, the overall proof of linear convergence (3.11) would fail.

From the linear convergence (3.11), we immediately obtain the so-called *general quasi-orthogonality* (see, e.g., Carstensen *et al.*, 2014, Proposition 4.11 or Erath & Praetorius, 2016, Proposition 10, step 5).

COROLLARY 3.8 (General quasi-orthogonality). Let (u_k) be the sequence of solutions of Algorithm 3.1. Then there exists $C>0$ such that

$$\sum_{k=\ell}^{\infty} \|u_{k+1}-u_k\|_{H^1(\Omega)}^2 \leq C\,\eta_\ell^2 \quad \text{for all } \ell\in\mathbb{N}_0. \tag{A3}$$

The constant $C>0$ has the same dependencies as C_{lin} from (3.11).

3.4 *Optimal algebraic convergence rates*

In order to prove optimal convergence rates of Algorithm 3.1, we need one further property of the error estimator, namely the so-called *discrete reliability* (A4). The proof of the next lemma follows as for the symmetric case in Erath & Praetorius (2016, Proposition 15). While the proof is thus omitted, we note that the main difficulties over the well-known FEM proof (Cascón *et al.*, 2008) arise in the handling of the piecewise constant test spaces on $\mathscr{T}_\times^*$ and $\mathscr{T}_\diamond^*$, and the fact that these test spaces are not nested.

LEMMA 3.9 (Discrete reliability). There exists a constant $C>0$ such that for all $\mathscr{T}_\diamond \in \texttt{refine}(\mathscr{T}_0)$ and all $\mathscr{T}_\times \in \texttt{refine}(\mathscr{T}_\diamond)$, it holds that

$$\|u_\times-u_\diamond\|_{H^1(\Omega)}^2 \leq C\left(\sum_{T\in\mathscr{T}_\times} h_T^2\|u_\times-u_\diamond\|_{H^1(T)}^2 + \sum_{T\in\mathscr{R}_\diamond}\eta_\diamond(T,u_\diamond)^2\right), \tag{A4}$$

where $\mathscr{R}_\diamond := \big\{T\in\mathscr{T}_\diamond : \exists\, T'\in\mathscr{T}_\diamond\backslash\mathscr{T}_\times \text{ with } T\cap T'\neq\emptyset\big\}$ consists of all refined elements $\mathscr{T}_\diamond\backslash\mathscr{T}_\times$ plus one additional layer of neighboring elements. The constant $C>0$ depends only on the σ-shape regularity (2.2), the data assumptions (1.2)–(1.4) and Ω. Note that for a sufficiently fine initial mesh $\mathscr{T}_0$, e.g., $C\,\|h_0\|_{L^\infty(\Omega)}^2 \leq 1/2$, (A4) leads to discrete reliability as stated in Carstensen *et al.* (2014). □

Let $\mathbb{T} := \texttt{refine}(\mathscr{T}_0)$ be the set of all possible triangulations obtained by NVB. For $N\geq 0$, let $\mathbb{T}_N := \big\{\mathscr{T}_\times\in\mathbb{T} : \#\mathscr{T}_\times - \#\mathscr{T}_0 \leq N\big\}$. For $s>0$, define

$$\|u\|_{\mathbb{A}_s} := \sup_{N\in\mathbb{N}_0}\inf_{\mathscr{T}_\times\in\mathbb{T}_N}(N+1)^s\,\eta_\times. \tag{3.14}$$

Note that $\|u\|_{\mathbb{A}_s}<\infty$ implies an algebraic decay $\eta_\times = \mathscr{O}\big((\#\mathscr{T}_\times)^{-s}\big)$ along the optimal sequence of meshes (which minimize the error estimator). Optimal convergence of the adaptive algorithm thus means

that for all $s > 0$ with $\|u\|_{\mathbb{A}_s} < \infty$, the adaptive algorithm leads to $\eta_\ell = \mathcal{O}\big((\#\mathcal{T}_\ell)^{-s}\big)$. The work Carstensen *et al.* (2014, Theorem 4.1) proves in a general framework the following Theorem 3.10, if the adaptive algorithm applied to a numerical scheme and a corresponding estimator satisfies (A1)–(A4).

THEOREM 3.10 (Optimal algebraic convergence rates). Suppose that the dual problem (3.3) is H^{1+s}-regular (3.4) for some $0 < s \leq 1$. Let the initial mesh $\mathcal{T}_0$ be sufficiently fine, i.e, there exists a constant $H > 0$ such that $\|h_0\|_{L^\infty(\Omega)} \leq H$. Finally, suppose that there is a constant $C_{\mathrm{MNS}} \geq 1$ such that $\#\mathcal{M}_\ell \leq C_{\mathrm{MNS}}\#\mathcal{M}_\ell^\eta$ for all $\ell \in \mathbb{N}_0$. Then there exists a bound $0 < \theta_{\mathrm{opt}} \leq 1$ such that for all $0 < \theta < \theta_{\mathrm{opt}}$ and all $s > 0$ with $\|u\|_{\mathbb{A}_s} < \infty$, there exists a constant $C_{\mathrm{opt}} > 0$ such that Algorithm 3.1 guarantees

$$\eta_\ell \leq C_{\mathrm{opt}}(\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-s} \quad \text{for all } \ell \in \mathbb{N}. \tag{3.15}$$

The constant θ_{opt} depends only on Ω, H, uniform σ-shape regularity of the triangulations $\mathcal{T}_\times \in \texttt{refine}(\mathcal{T}_0)$ and the data assumptions (1.2)–(1.4). The constant C_{opt} additionally depends on s, the constant q_{lin} from (3.11), the use of NVB and on C_{MNS}. □

REMARK 3.11 A direct consequence of the assumption $\#\mathcal{M}_\ell \leq C_{\mathrm{MNS}}\#\mathcal{M}_\ell^\eta$ in Theorem 3.10 is that data oscillation marking (3.2) is negligible with respect to the overall number of marked elements (see also Erath & Praetorius, 2016, Remark 7). In practice, (3.1) already implies (3.2) since $\theta' > 0$ can be chosen arbitrarily small. Furthermore, efficiency (2.18) is not required to show (3.11) and (3.15) but guarantees (optimal) linear convergence also for the FVM error.

4. Numerical examples

In extension of our theory, we consider the model problem (1.1) with inhomogeneous Dirichlet boundary conditions. For all experiments in two dimensions, we run Algorithm 3.1 with $\theta = 1 = \theta'$ and $\theta = 0.5 = \theta'$ for uniform mesh refinement and adaptive mesh refinement, respectively.

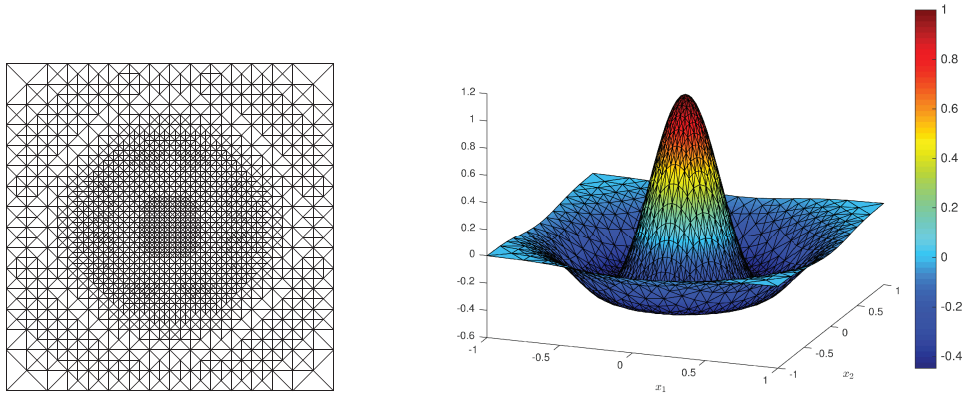
4.1 *Experiment with a smooth solution*

On the square $\Omega = (-1, 1)^2$, we prescribe the exact solution $u(x_1, x_2) = \big(1 - 10x_1^2 - 10x_2^2\big)\, e^{-5(x_1^2+x_2^2)}$ with $x = (x_1, x_2) \in \mathbb{R}^2$. We choose the diffusion matrix

$$\mathbf{A} = \begin{pmatrix} 10 + \cos x_1 & 9x_1x_2 \\ 9x_1x_2 & 10 + \sin x_2 \end{pmatrix},$$

the velocity $\mathbf{b} = (\sin x_1, \cos x_2)^{\mathrm{T}}$ and the reaction $c = 1$. Note that (1.2) holds with $\lambda_{\min} = 0.82293$ and $\lambda_{\max} = 10.84096$ and (1.4) with $\frac{1}{2}\mathrm{div}\,\mathbf{b} + c > 0$. The right-hand side f is calculated appropriately. The uniform initial mesh $\mathcal{T}_0$ consists of 16 triangles.

In Fig. 2(a) we see an adaptively generated mesh after 16 refinements. Figure 2(b) plots the smooth solution on the mesh $\mathcal{T}_{16}$. Both uniform and adaptive mesh refinements lead to the optimal convergence order $\mathcal{O}(N^{-1/2})$ with respect to the number N of elements since u is smooth; see Fig. 3. The oscillations are of higher order and decrease with $\mathcal{O}(N^{-1})$.



(a) $\mathscr{T}_{16}$ (3842 elements). (b) Solution ($\mathscr{T}_{16}$).

FIG. 2. Experiment with a smooth solution from Section 4.1: adaptively generated mesh $\mathscr{T}_{16}$ from a uniform initial triangulation $\mathscr{T}_0$ with 16 elements (left) and discrete FVM solution calculated on $\mathscr{T}_{16}$ (right).

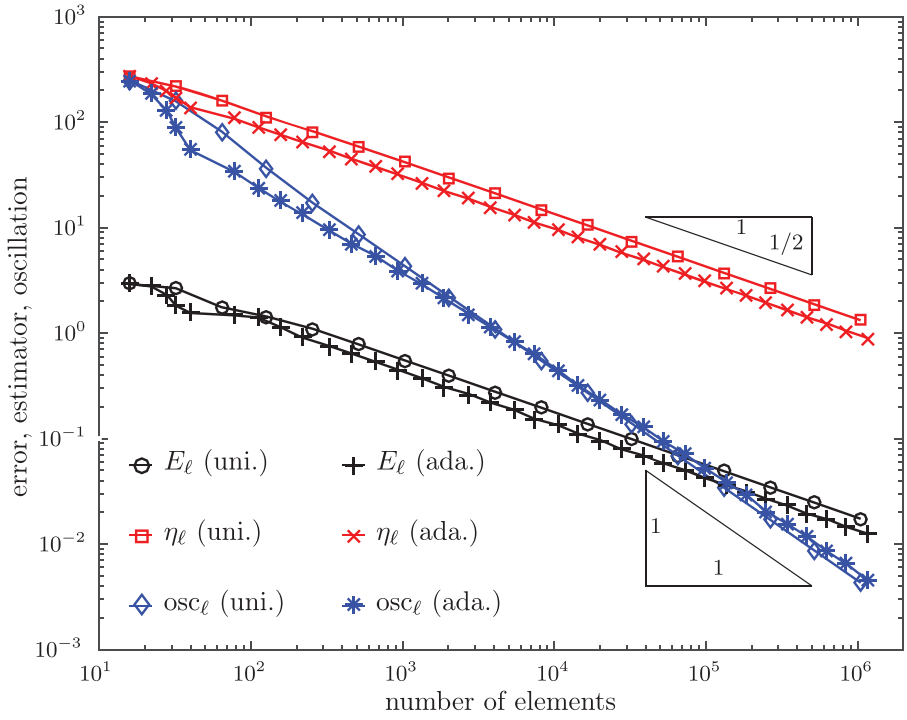


FIG. 3. Experiment with a smooth solution from Section 4.1: error $E_\ell = \|u - u_\ell\|_{H^1(\Omega)}$, weighted-residual error estimator η_ℓ and data oscillations osc_ℓ for uniform and adaptive mesh refinement.

Table 1 shows the experimental validation of the additional assumption in Theorem 3.10, i.e., marking for the data oscillations is negligible; see also Remark 3.11.

TABLE 1 *Experiment with a smooth solution from Section 4.1: we compute* $\widetilde{C}_{\mathrm{MNS}} := \#\mathcal{M}_\ell / \#\mathcal{M}_\ell^\eta \leq 1.3$. *Hence, the additional assumption in Theorem 3.10 is experimentally verified. Furthermore, we compute* $\widetilde{\theta}' := \mathrm{osc}_\ell\left(\mathcal{M}_\ell^\eta\right)^2 / \mathrm{osc}_\ell^2 \geq 0.2$ *with* $\mathrm{osc}_\ell\left(\mathcal{M}_\ell^\eta\right)^2 := \sum_{T\in\mathcal{M}_\ell^\eta} \mathrm{osc}_\ell\,(T,u_\ell)^2$, *i.e., the choice* $\theta = 0.5$, $\theta' = 0.2$ *would guarantee* $\mathcal{M}_\ell = \mathcal{M}_\ell^\eta$ *in Algorithm 3.1*

ℓ	$\#\mathcal{T}_\ell$	$\frac{\#\mathcal{M}_\ell}{\#\mathcal{M}_\ell^\eta}$	$\frac{\mathrm{osc}_\ell(\mathcal{M}_\ell^\eta)^2}{\mathrm{osc}_\ell^2}$
0	16	1.000	0.631
1	22	1.000	0.615
2	28	1.000	0.704
3	32	1.000	0.769
4	40	1.214	0.338
5	78	1.111	0.446
6	112	1.133	0.292
7	156	1.119	0.410
8	216	1.062	0.394
9	331	1.198	0.264
10	460	1.014	0.472
11	660	1.049	0.371
12	944	1.027	0.431
13	1,338	1.025	0.400
14	1,910	1.018	0.387
15	2,748	1.026	0.374
16	3,842	1.015	0.358
17	5,430	1.003	0.449
18	7,438	1.013	0.359
19	10,590	1.003	0.445
20	14,478	1.019	0.323
21	20,286	1.004	0.430
22	27,558	1.004	0.457
23	38,450	1.010	0.324
24	52,422	1.000	0.540
25	72,454	1.007	0.404
26	98,232	1.000	0.508
27	135,172	1.004	0.446
28	184,142	1.000	0.606
29	251,896	1.002	0.475
30	342,148	1.001	0.488
31	461,674	1.000	0.617
32	635,266	1.004	0.416
33	852,730	1.000	0.664
34	1,172,122	1.002	0.464

4.2 *Experiment with a generic singularity*

On the L-shaped domain $\Omega = (-1,1)^2 \backslash \big([0,1] \times [-1,0]\big)$ we consider the exact solution $u(x_1,x_2) = r^{2/3}\sin(2\varphi/3)$ in polar coordinates $r \in \mathbb{R}_0^+$, $\varphi \in [0, 2\pi[$ and $(x_1,x_2) = r(\cos\varphi, \sin\varphi)$. It is well known that u has a generic singularity at the reentrant corner (0, 0), which leads to $u \in H^{1+2/3-\varepsilon}(\Omega)$ for all $\varepsilon > 0$. We choose the diffusion matrix

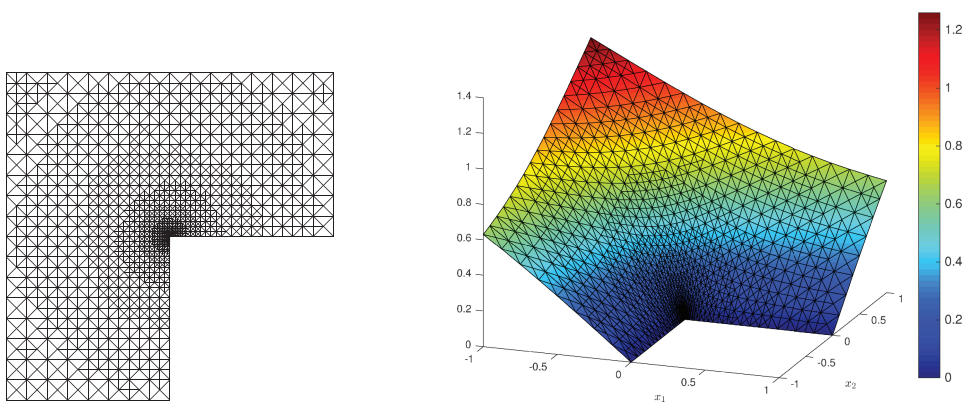
$$\mathbf{A} = \begin{pmatrix} 5 + \big(x_1^2 + x_2^2\big)\cos x_1 & \big(x_1^2 + x_2^2\big)^2 \\ \big(x_1^2 + x_2^2\big)^2 & 5 + \big(x_1^2 + x_2^2\big)\sin x_2 \end{pmatrix}$$

so that (1.2) holds with $\lambda_{\min} = 0.46689$ and $\lambda_{\max} = 5.14751$, $\mathbf{b} = (1,1)^{\mathrm{T}}$ and $c = 1$ so that (1.4) holds with $\frac{1}{2}\operatorname{div}\mathbf{b} + c = 1$. The right-hand side f is calculated appropriately. The uniform initial mesh $\mathscr{T}_0$ consists of 12 triangles. An adaptively generated mesh after 16 refinements and a plot of the discrete solution are shown in Fig. 4.

We observe the expected suboptimal convergence order of $\mathscr{O}(N^{-1/3})$ for uniform mesh refinement. We regain the optimal convergence order of $\mathscr{O}(N^{-1/2})$ for adaptive mesh refinement; see Fig. 5. As in the experiment of Section 4.1, the oscillations are of higher order $\mathscr{O}(N^{-1})$. We refer to Table 2 for the experimental validation of the additional assumption in Theorem 3.10 that marking for the data oscillations is negligible.

4.3 *Convection-dominated experiment*

The final example is taken from Mekchay & Nochetto (2005). On the square $\Omega = (0,1)^2$, we fix the diffusion $\mathbf{A} = 10^{-3}\mathbf{I}$ and the convection velocity $\mathbf{b} = (x_2, 1/2 - x_1)^{\mathrm{T}}$. The reaction and right-hand side are $c = f = 0$. Thus, (1.2) holds with $\lambda_{\min} = \lambda_{\max} = 10^{-3}$ and (1.4) with $\frac{1}{2}\operatorname{div}\mathbf{b} + c = 0$.



(a) $\mathscr{T}_{16}$ (2534 elements). (b) Solution ($\mathscr{T}_{16}$).

FIG. 4. Experiment with a generic singularity of the solution in the reentrant corner (0, 0) from Section 4.2: adaptively generated mesh $\mathscr{T}_{16}$ from a uniform initial triangulation $\mathscr{T}_0$ with 12 elements (left) and discrete FVM solution calculated on $\mathscr{T}_{16}$ (right).

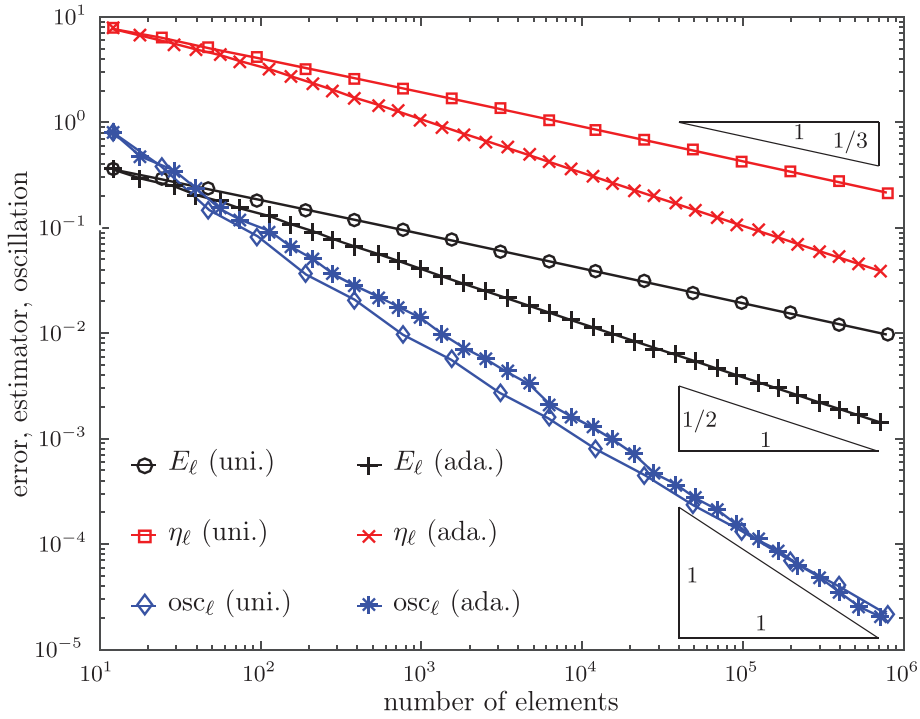


FIG. 5. Experiment with a generic singularity of the solution from Section 4.2: error $E_\ell = \|u-u_\ell\|_{H^1(\Omega)}$, weighted-residual error estimator η_ℓ and data oscillations osc_ℓ for uniform and adaptive mesh refinement.

On the Dirichlet boundary Γ, we prescribe a continuous piecewise linear function by

$$u(x_1,x_2)|_\Gamma = \begin{cases} 1 & \text{on } \{0.2005 \le x_1 \le 0.4995, x_2 = 0\}, \\ 0 & \text{on } \Gamma\backslash\{0.2 \le x_1 \le 0.5; x_2 = 0\}, \\ \text{linear} & \text{on } \{0.2 \le x_1 \le 0.2005 \text{ or } 0.4995 \le x_1 \le 0.5; x_2 = 0\}. \end{cases}$$

The model has a moderate convection dominance with respect to the diffusion and simulates the transport of a pulse from Γ to the interior and back to Γ. For this example, we do not know the analytical solution. The uniform initial mesh $\mathcal{T}_0$ consists of 32 triangles. In Fig. 6(a), we see the solution with strong oscillations on a uniformly generated mesh with 8,192 elements. The oscillations are due to the convection dominance. For the next refinement step (16,384 elements, not plotted), however, the oscillations disappear since the shock region at the boundary is refined enough. Our adaptive Algorithm 3.1, which *also* has a mandatory oscillation marking, provides a stable solution on a mesh with only 779 elements; see Fig. 6(b). In Fig. 7, we plot adaptively generated meshes after 14 and 20 mesh refinements. We see a strong refinement in the shock region. A similar observation can be found in Mekchay & Nochetto (2005). We remark that this strategy works only for this moderate convection-dominated problem. For $\mathbf{A} = 10^{-8}\mathbf{I}$, we cannot see any stabilization effects by Algorithm 3.1 (not displayed). Hence, only a stabilization of the numerical scheme, e.g., FVM with upwinding, would avoid these instabilities. However, the analysis of such schemes is beyond the scope of this work; see

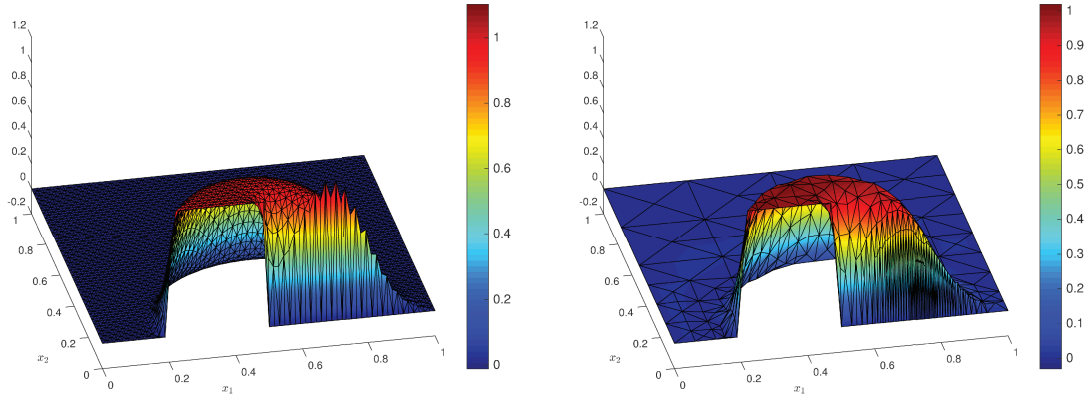
TABLE 2 *Experiment with a generic singularity of the solution from Section 4.2: we compute* $\widetilde{C}_{\mathrm{MNS}} := \#\mathcal{M}_\ell/\#\mathcal{M}_\ell^\eta \leq 1.8$. *Hence, the additional assumption in Theorem 3.10 is experimentally verified. Furthermore, we compute* $\widetilde{\theta}' := \mathrm{osc}_\ell(\mathcal{M}_\ell^\eta)^2/\mathrm{osc}_\ell^2 \geq 0.02$ *with* $\mathrm{osc}_\ell(\mathcal{M}_\ell^\eta)^2 := \sum_{T\in\mathcal{M}_\ell^\eta}\mathrm{osc}_\ell(T,u_\ell)^2$, *i.e., the choice* $\theta = 0.5$, $\theta' = 0.02$ *would guarantee* $\mathcal{M}_\ell = \mathcal{M}_\ell^\eta$ *in Algorithm 3.1*

ℓ	$\#\mathcal{T}_\ell$	$\frac{\#\mathcal{M}_\ell}{\#\mathcal{M}_\ell^\eta}$	$\frac{\mathrm{osc}_\ell(\mathcal{M}_\ell^\eta)^2}{\mathrm{osc}_\ell^2}$
0	12	1.667	0.135
1	18	1.750	0.086
2	29	1.600	0.027
3	40	1.375	0.057
4	56	1.400	0.252
5	74	1.667	0.079
6	114	1.286	0.148
7	153	1.188	0.243
8	212	1.111	0.256
9	284	1.065	0.390
10	380	1.194	0.168
11	539	1.068	0.328
12	721	1.050	0.346
13	991	1.007	0.466
14	1,356	1.003	0.482
15	1,852	1.020	0.386
16	2,534	1.000	0.630
17	3,413	1.009	0.443
18	4,684	1.000	0.597
19	6,341	1.003	0.443
20	8,568	1.002	0.490
21	11,564	1.000	0.640
22	15,590	1.000	0.539
23	21,071	1.000	0.569
24	28,304	1.017	0.437
25	38,350	1.000	0.670
26	51,122	1.016	0.414
27	69,135	1.000	0.563
28	92,367	1.000	0.528
29	123,666	1.008	0.463
30	166,532	1.000	0.703
31	221,144	1.020	0.378
32	298,213	1.000	0.549
33	397,086	1.000	0.597
34	532,432	1.017	0.409
35	712,738	1.000	0.666

also Section 5. We observe the above stabilization effects also in the convergence plot of the estimator; see Fig. 8. Note that the estimator for adaptive mesh refinement is faster in the asymptotic convergence than the estimator for uniform mesh refinement. Additionally, the convergence rate for the estimator

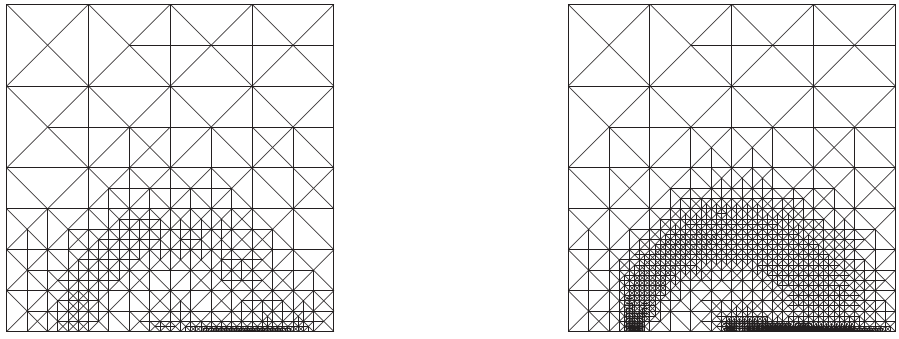
TABLE 3 *Experimental results on marking strategy for the convection-dominated experiment from Section 4.3: we compute* $\widetilde{C}_{\mathrm{MNS}} := \#\mathscr{M}_\ell/\#\mathscr{M}_\ell^\eta \leq 3$ *and see that the additional assumption in Theorem 3.10 is experimentally verified. In addition, we compute* $\widetilde{\theta}' := \operatorname{osc}_\ell(\mathscr{M}_\ell^\eta)^2/\operatorname{osc}_\ell^2 \geq 0.03$ *with* $\operatorname{osc}_\ell(\mathscr{M}_\ell^\eta)^2 := \sum_{T\in\mathscr{M}_\ell^\eta} \operatorname{osc}_\ell(T,u_\ell)^2$*, i.e., the choice* $\theta = 0.5$*,* $\theta' = 0.03$ *would guarantee* $\mathscr{M}_\ell = \mathscr{M}_\ell^\eta$ *in Algorithm 3.1*

ℓ	$\#\mathscr{T}_\ell$	$\frac{\#\mathscr{M}_\ell}{\#\mathscr{M}_\ell^\eta}$	$\frac{\operatorname{osc}_\ell(\mathscr{M}_\ell^\eta)^2}{\operatorname{osc}_\ell^2}$
0	32	1.125	0.434
1	48	1.400	0.201
2	59	1.500	0.266
3	72	1.667	0.196
4	90	2.500	0.177
5	110	1.333	0.266
6	154	1.583	0.085
7	187	1.500	0.124
8	238	1.786	0.055
9	280	1.296	0.234
10	332	1.371	0.154
11	405	1.412	0.124
12	511	1.537	0.083
13	628	1.521	0.146
14	779	1.559	0.077
15	1,100	1.600	0.064
16	1,428	1.605	0.063
17	1,837	1.643	0.037
18	2,416	1.594	0.058
19	3,195	1.437	0.060
20	4,336	1.583	0.048
21	5,664	1.402	0.072
22	7,666	1.445	0.047
23	10,186	1.351	0.067
24	13,919	1.258	0.078
25	19,041	1.230	0.112
26	26,248	1.182	0.106
27	36,592	1.142	0.135
28	50,806	1.112	0.180
29	70,367	1.082	0.196
30	97,946	1.058	0.227
31	135,122	1.057	0.236
32	186,959	1.028	0.311
33	255,994	1.021	0.311
34	351,880	1.022	0.289
35	484,157	1.015	0.328
36	662,325	1.006	0.381
37	902,659	1.005	0.384



(a) Solution $\mathscr{T}_8$ (uniform, 8192 elements). (b) Solution $\mathscr{T}_{14}$ (adaptive, 779 elements).

FIG. 6. Convection-dominated experiment from Section 4.3: the discrete FVM solution on a uniformly generated mesh $\mathscr{T}_8$ (left) and adaptively generated mesh $\mathscr{T}_{14}$ (right). The algorithm starts with a uniform initial triangulation $\mathscr{T}_0$ with 32 elements.



(a) $\mathscr{T}_{14}$ (779 elements). (b) $\mathscr{T}_{20}$ (4336 elements).

FIG. 7. Convection-dominated experiment from Section 4.3: adaptively generated meshes $\mathscr{T}_{14}$ (left) and $\mathscr{T}_{20}$ (right) from a uniform initial triangulation $\mathscr{T}_0$ with 32 elements.

is suboptimal for uniform mesh refinement. For adaptive mesh refinement, we regain the optimal convergence order of $\mathcal{O}(N^{-1/2})$; see Fig. 8. As in the previous experiments, the oscillations are of higher order. In Table 3, we also see that the oscillation marking for this convection-dominated problem is for more refinement steps dominant than for the previous problems; see also the discussion in Mekchay & Nochetto (2005).

5. Conclusions

In this work, we have proved linear convergence of an adaptive vertex-centered FVM with generically optimal algebraic rates to the solution of a general second-order linear elliptic PDE. Besides marking based on the local contributions of the *a posteriori* error estimator, we additionally had to mark

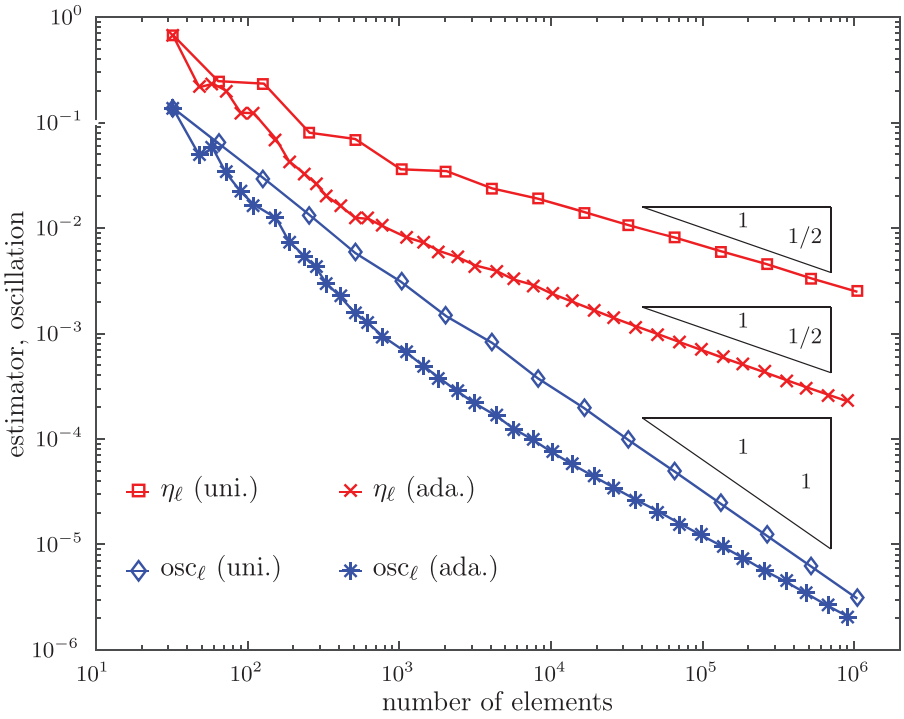


FIG. 8. Convection-dominated experiment from Section 4.3: weighted-residual error estimator η_ℓ and data oscillations osc_ℓ for uniform and adaptive mesh refinement.

the oscillations to overcome the lack of a classical Galerkin orthogonality property. In the case of dominating convection, FVMs provide a natural upwind stabilization. Although there exist estimators for these upwind discretizations also (see Erath, 2013), we were not able to provide a rigorous convergence result for the related adaptive mesh-refinement strategy. Note that the upwind direction and thus the corresponding error-indicator contributions are defined over the boundary of the control volumes of the dual mesh. As mentioned above, the dual meshes are not nested for a sequence of locally refined triangulations. This makes it difficult to show (A1)–(A2) and (B1)–(B2). We stress that the other error-indicator contributions are defined over the elements of the primal mesh and can hence be treated by the developed techniques.

Funding

Austrian Science Fund (FWF) research project *Optimal adaptivity for BEM and FEM-BEM coupling* (P27005 to D.P.), research program *Taming complexity in partial differential systems.* (F65 to D.P.).

References

Bergh, J. & Löfström, J. (1976) *Interpolation Spaces. An Introduction*. Grundlehren der Mathematischen Wissenschaften, vol. 223. Berlin-New York: Springer.

Bespalov, A., Haberl, A. & Praetorius, D. (2017) Adaptive fem with coarse initial mesh guarantees optimal convergence rates for compactly perturbed elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, **317**, 318–340.

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

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

Carstensen, C., Lazarov, R. D. & Tomov, S. Z. (2005) Explicit and averaging *a posteriori* error estimates for adaptive finite volume methods. *SIAM J. Numer. Anal.*, **42**, 2496–2521.

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

Cascón, J. M. & Nochetto, R. H. (2012) Quasioptimal cardinality of AFEM driven by nonresidual estimators. *IMA J. Numer. Anal.*, **32**, 1–29.

Dörfler, W. (1996) A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, **33**, 1106–1124.

Erath, C. (2010) Coupling of the finite volume method and the boundary element method-theory, analysis, and numerics. *Ph.D. Thesis*, University of Ulm, Ulm.

Erath, C. (2012) Coupling of the finite volume element method and the boundary element method: an *a priori* convergence result. *SIAM J. Numer. Anal.*, **50**, 574–594.

Erath, C. (2013) *A posteriori* error estimates and adaptive mesh refinement for the coupling of the finite volume method and the boundary element method. *SIAM J. Numer. Anal.*, **51**, 1777–1804.

Erath, C. & Praetorius, D. (2016) Adaptive vertex-centered finite volume methods with convergence rates. *SIAM J. Numer. Anal.*, **54**, 2228–2255.

Erath, C. & Praetorius, D. (2017) Céa-type quasi-optimality and convergence rates for (adaptive) vertex-centered FVM. *Finite Volumes for Complex Applications VIII-Methods and Theoretical Aspects* (C. Cances & P. Omnes eds), vol. 199. Berlin: Springer.

Ewing, R. E., Lin, T. & Lin, Y. (2002) On the accuracy of the finite volume element method based on piecewise linear polynomials. *SIAM J. Numer. Anal.*, **39**, 1865–1888.

Feischl, M., Führer, T. & Praetorius, D. (2014) Adaptive FEM with optimal convergence rates for a certain class of nonsymmetric and possibly nonlinear problems. *SIAM J. Numer. Anal.*, **52**, 601–625.

Grisvard, P. (1985) *Elliptic Problems in Nonsmooth Domains*. Boston: Pitman.

Karkulik, M., Pavlicek, D. & Praetorius, D. (2013) On 2D newest vertex bisection: optimality of mesh-closure and H^1-stability of L_2-projection. *Constr. Approx.*, **38**, 213–234.

Mekchay, K. & Nochetto, R. H. (2005) Convergence of adaptive finite element methods for general second order linear elliptic PDEs. *SIAM J. Numer. Anal.*, **43**, 1803–1827.

Morin, P., Nochetto, R. H. & Siebert, K. G. (2000) Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, **38**, 466–488.

Scott, L. R. & Zhang, S. (1990) Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, **54**, 483–493.

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

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