Refined a posteriori error estimation for classical and pressure-robust Stokes finite element methods

P. L. LEDERER

Institute for Analysis and Scientific Computing, TU Wien

C. MERDON

Weierstrass Institute for Applied Analysis and Stochastics, Berlin

J. SCHÖBERL

Institute for Analysis and Scientific Computing, TU Wien

ABSTRACT. Recent works showed that pressure-robust modifications of mixed finite element methods for the Stokes equations outperform their standard versions in many cases. This is achieved by divergence-free reconstruction operators and results in pressure independent velocity error estimates which are robust with respect to small viscosities. In this paper we develop a posteriori error control which reflects this robustness.

The main difficulty lies in the volume contribution of the standard residual-based approach that includes the L^2-norm of the right-hand side. However, the velocity is only steered by the divergence-free part of this source term. An efficient error estimator must approximate this divergence-free part in a proper manner, otherwise it can be dominated by the pressure error.

To overcome this difficulty a novel approach is suggested that uses arguments from the stream function and vorticity formulation of the Navier–Stokes equations. The novel error estimators only take the curl of the right-hand side into account and so lead to provably reliable, efficient and pressure-independent upper bounds in case of a pressure-robust method in particular in pressure-dominant situations. This is also confirmed by some numerical examples with the novel pressure-robust modifications of the Taylor–Hood and mini finite element methods.

incompressible Navier–Stokes equations and mixed finite elements and pressure robustness and a posteriori error estimators and adaptive mesh refinement

1. INTRODUCTION

This paper studies a posteriori error estimators for the velocity of the Stokes equation with a special focus on pressure-robust finite element methods. Pressure-robustness is closely related to the L^2-orthogonality of divergence-free functions onto gradients of H^1-functions. In particular, the exact velocity $\boldsymbol{u}$ of the Stokes equations (with zero boundary data),

$$-\nu\Delta\boldsymbol{u} + \nabla p = \boldsymbol{f} \text{ in } \Omega \quad \text{and} \quad \boldsymbol{u} \in \boldsymbol{V}_0 := \{\boldsymbol{v} \in H_0^1(\Omega)^2 : \operatorname{div}\boldsymbol{v} = 0\},$$

is orthogonal onto any $q \in L^2(\Omega)$ in the sense that $\int_\Omega q\operatorname{div}(\boldsymbol{u}) \, dx = 0$. Consequently, $\boldsymbol{u}$ also solves the Stokes equations with $\boldsymbol{f}$ replaced by $\boldsymbol{f} + \nabla q$ for $q \in H^1(\Omega)$. This invariance property is in general not preserved for discretely divergence-free testfunctions of most classical finite element methods that

relax the divergence constraint to attain inf-sup stability. With an inf-sup-stable pair of a discrete velocity space $\boldsymbol{V}_h$ and some discrete pressure space Q_h and the discretely divergence-free functions $\boldsymbol{V}_{0,h} \subset \boldsymbol{V}_h$, the consistency error from the relaxed divergence constraint can be expressed by the discrete dual norm , for any $q \in L^2(\Omega)$,

$$(1) \qquad \|\nabla q\|_{\boldsymbol{V}_{0,h}^\star} := \sup_{\boldsymbol{v}_h \in V_{0,h} \setminus \{\boldsymbol{0}\}} \frac{\int_\Omega q \mathrm{div} \boldsymbol{v}_h \; dx}{\|\nabla \boldsymbol{v}_h\|_{L^2}} \leqslant \begin{cases} \min_{q_h \in Q_h} \|q - q_h\|_{L^2} & \text{if } \boldsymbol{V}_{0,h} \nsubseteq \boldsymbol{V}_0, \\ 0 & \text{if } \boldsymbol{V}_{0,h} \subseteq \boldsymbol{V}_0. \end{cases}$$

Besides some expensive or exotic divergence-free methods like the Scott-Vogelius finite element method [32, 37], most classical inf-sup stable mixed finite element methods, including the popular Taylor–Hood [18] and mini finite element families [5] have $V_{0,h} \nsubseteq \boldsymbol{V}_0$ and so the term from (1) appears in their a priori velocity gradient error estimate [8] scaled with $1/\nu$, i.e.

$$(2) \qquad \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2}^2 \leqslant \inf_{\substack{\boldsymbol{v}_h \in V_{0,h}, \\ \boldsymbol{u}_h = \boldsymbol{v}_h \text{ on } \partial\Omega}} \|\nabla(\boldsymbol{u} - \boldsymbol{v}_h)\|_{L^2}^2 + \frac{1}{\nu^2}\|\nabla p\|_{V_{0,h}^\star}^2.$$

This factor $1/\nu$ causes a locking phenomenon. Indeed, for $\nu \to 0$ or very complicated pressures, the pressure contribution may dominate and lead to a very bad solution for the discrete velocity $\boldsymbol{u}_h$ [21, 24, 19, 25].

By a trick of [21] one can introduce a reconstruction operator Π, that maps discretely divergence-free functions onto exactly divergence-free ones, into the right-hand side and so transform any classical finite element method into a pressure-robust one. This replaces the pressure-dependent term in (2) by a small consistency error of optimal order [21, 9, 22, 25, 20] and independent of $1/\nu$. Although this fixes the locking phenomenon and leads to huge gains in many numerical examples, efficient a posteriori error control for these methods is an open problem. Efficient error estimators for the velocity error not only have to cope with the variational crime but also, and more importantly, have to mimic the pressure-independence.

Standard residual-based a posteriori error estimators η usually have the form

$$\|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2} \lesssim \eta := \eta_{\mathrm{vol}} + \text{other terms}$$

with a volume contribution η_{vol} and some other terms, like norms of the normal jumps of $\boldsymbol{u}_h$, data oscillations or consistency errors. In the standard residual-based error estimator for classical finite element methods [17, 36, 34, 33] the volume contribution takes the form (for any $q \in H^1(\Omega)$ and piecewise Laplacian $\Delta_\mathcal{T}$)

$$(3) \qquad \eta_{\mathrm{vol}} = \nu^{-1}\|\nabla q\|_{V_{0,h}^\star} + \nu^{-1}\|h_\mathcal{T}(\boldsymbol{f} - \nabla q + \nu\Delta_\mathcal{T}\boldsymbol{u}_h)\|_{L^2}$$

$$\lesssim \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2} + \nu^{-1}\left(\|p - q\|_{L^2} + \min_{q_h \in Q_h}\|q - q_h\|_{L^2} + \mathrm{osc}_\mathrm{k}(\boldsymbol{f} - \nabla q, \mathcal{T})\right).$$

The inequality above states efficiency, i.e. beeing also a lower bound of the real error, and its dependence on the choice of q. Note, that for $q \in Q_h$ the terms $\|\nabla q\|_{V_{0,h}^\star} \leqslant \min_{q_h \in Q_h}\|q - q_h\|_{L^2} = 0$ vanish, but $\|p - q\|_{L^2}$ remains, whereas for $q = p$ the term $\|p - q\|_{L^2}$ vanishes, but the other two remain. If the velocity error is at best as good as the error in the pressure (scaled by $1/\nu$), as it is the case for classical pressure-inrobust methods, this estimate is fine (e.g. for q chosen as an H^1-interpolation of p_h). As a result classical a posteriori error estimates, see e.g. [17, 36, 34, 33], often perform the error analysis in a norm that combines the velocity error and the pressure error. A pressure-robust method, however, allows for a decoupled error analysis of velocity error and pressure error and so gives more control over both.

For a pressure-robust finite element method, the term (3) can be replaced by

$$(4) \qquad \eta_{\mathrm{vol}} = \nu^{-1}\|h_\mathcal{T}(\boldsymbol{f} - \nabla q + \nu\Delta_\mathcal{T}\boldsymbol{u}_h)\|_{L^2}$$

$$\lesssim \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2} + \nu^{-1}\left(\|p - q\|_{L^2} + \mathrm{osc}_\mathrm{k}(\boldsymbol{f} - \nabla q, \mathcal{T})\right).$$

Here, the choice $q = p$ leads to a pressure-independent efficient estimate. However, this cannot be considered a posteriori, since p is unknown. Hence, an efficient error estimator of this form for pressure-robust methods hinges on a good approximation of $q \approx p$ as already investigated in [15, 23].

The main result of this paper concerns a different approach to estimate the velocity error that yields an estimator with the volume contribution

$$(5) \qquad \eta_{\mathrm{curl}} = \nu^{-1} \| h_{\mathcal{T}}^2 \mathrm{curl}_{\mathcal{T}} (\boldsymbol{f} + \nu \Delta_{\mathcal{T}} \boldsymbol{u}_h) \|_{L^2}$$
$$\lesssim \| \nabla (\boldsymbol{u} - \boldsymbol{u}_h) \|_{L^2} + \nu^{-1} \mathrm{osc}_{\mathrm{k}} (h_{\mathcal{T}} \mathrm{curl}_{\mathcal{T}} (\boldsymbol{f} + \nu \Delta_{\mathcal{T}} \boldsymbol{u}_h), \mathcal{T}).$$

The advantage of η_{curl} over η_{vol} is that the curl operator automatically cancels any ∇q from the Helmholtz decomposition of $\boldsymbol{f} + \nu \Delta_{\mathcal{T}} \boldsymbol{u}_h$ and therefore no approximation of p as in (4) is needed. Also note, that η_{curl} is similar to the volume contribution of a residual-based error estimator for the Navier–Stokes equations in streamline and vorticity formulation [3]. However, the error estimator with this volume contribution is valid for any pressure-robust finite element method like the Scott–Vogelius finite element method [32, 37] or the novel family of pressure-robustly modified finite element methods of [21, 9, 22, 25, 20] that allow for an interesting interplay between the Fortin interpolator I and the reconstruction operator Π manifestated in the required assumption

$$(6) \qquad \int_{\Omega} (1 - \Pi I) \boldsymbol{v} \cdot \boldsymbol{\theta} \, dx \lesssim \| \nabla \boldsymbol{v} \|_{L^2} \| h_{\mathcal{T}}^2 \mathrm{curl} \boldsymbol{\theta} \|_{L^2} \quad \text{for all } \boldsymbol{\theta} \in H(\mathrm{curl}, \Omega) \text{ and } \boldsymbol{v} \in \boldsymbol{V}_0.$$

We prove this assumption for certain popular finite element methods including the Taylor–Hood and mini finite element methods, and some elements with discontinuous pressure approximations. However, we only focus on the two-dimensional case, since the proofs for the three-dimensional case are much more involved and therefore discussed in a future publication.

The rest of the paper is structured as follows. Section 2 introduces the Stokes equations and preliminaries as well as notation used throughout the paper. Section 3 focuses on classical finite element methods and their recently developed pressure-robust siblings that are based on a suitable reconstruction operator. Section 4 is concerned with standard residual-based error estimates for classical and pressure-robust finite element methods and the efficiency of its contributions, in particular (3) and (4), especially in the pressure-dominated regime. Section 5 derives some novel a posteriori error bounds with the volume contribution (5) that are efficient and easy to evaluate for the pressure-robust finite element methods that satisfy Assumption (6). In Section 6 this assumption is verified for many popular finite element methods and their pressure-robust siblings. Section 7 studies numerical examples and employs the local contributions of the a posteriori error estimates as refinement indicators for adaptive mesh refinement. The numerical examples verify the theory and show that the pressure-robust finite element methods converge with the optimal order also in non-smooth examples.

2. Model problem and preliminaries

This section states our model problem and the needed notation.

2.1. **Stokes equations and Helmholtz projector.** The Stokes model problem seeks a vector-valued velocity field $\boldsymbol{u}$ and a scalar-valued pressure field p on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$ with Dirichlet data $\boldsymbol{u} = \boldsymbol{u}_D$ along $\partial \Omega$ and

$$-\nu \Delta \boldsymbol{u} + \nabla p = \boldsymbol{f} \quad \text{and} \quad \mathrm{div} \boldsymbol{u} = 0 \quad \text{in } \Omega.$$

The weak formulation characterises $\boldsymbol{u} \in H^1(\Omega)^2$ by $\boldsymbol{u} = \boldsymbol{u}_D$ along $\partial \Omega$ and

$$\nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) - (p, \mathrm{div} \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) \qquad \text{for all } \boldsymbol{v} \in V := H_0^1(\Omega)^2,$$
$$(q, \mathrm{div} \boldsymbol{u}) = 0 \qquad \text{for all } q \in Q := L_0^2(\Omega).$$

In the set of divergence-free functions $\boldsymbol{V}_0 := \{ \boldsymbol{v} \in V \mid \mathrm{div} \boldsymbol{v} = 0 \}$, $\boldsymbol{u}$ satisfies

$$\nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{V}_0.$$

The Helmholtz decomposition decomposes every vector field into

$$\boldsymbol{f} = \nabla \alpha + \beta =: \nabla \alpha + \mathbb{P} \boldsymbol{f}$$

with $\alpha \in H^1(\Omega)/\mathbb{R}$ and $\beta =: \mathbb{P}\boldsymbol{f} \in L^2_\sigma(\Omega) := \{\boldsymbol{w} \in H(\mathrm{div}, \Omega) \,|\, \mathrm{div}\boldsymbol{w} = 0, \boldsymbol{w} \cdot \boldsymbol{n} = 0 \text{ along } \partial D\}$ [16]. Note in particular, that the continuous Helmholtz projector satisfies $\mathbb{P}(\nabla q) = 0$ for all $q \in H^1(\Omega)$ which implies

$$\nu(\nabla\boldsymbol{u}, \nabla\boldsymbol{v}) = (\mathbb{P}\boldsymbol{f}, \boldsymbol{v}) \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{V}_0,$$

hence $\boldsymbol{u}$ is steered only by the Helmholtz projector $\mathbb{P}\boldsymbol{f}$ of the right-hand side.

2.2. **Notation.** The set $\mathcal{T}$ denotes a regular triangulation of Ω into two dimensional simplices with edges $\mathcal{E}$ and nodes $\mathcal{N}$. The three edges of a simplex $T \in \mathcal{T}$ are denoted by $\mathcal{E}(T)$. Similarly, $\mathcal{N}(T)$ consists of the three nodes that belong to $T \in \mathcal{T}$, $\mathcal{N}(E)$ consists of the two nodes that belong to $E \in \mathcal{E}$ and $\mathcal{T}(z)$ for a vertex $z \in \mathcal{N}$ consists of all cells $T \in \mathcal{T}$ with $z \in \mathcal{N}(T)$. Finally we define $\mathcal{E}^\circ$ as the set of all inner.

As usual $L^2(\Omega)$, $H^1(\Omega)$, $H(\mathrm{div}, \Omega)$ and $H(\mathrm{curl}, \Omega)$ denote the Sobolev spaces and $L^2(\Omega)^2$, $H^1(\Omega)^2$ denote their vector-valued versions. Moreover, several discrete function spaces are used throughout the paper. The set $P_k(T)$ denotes scalar-valued polynomials up to order k that live on the simplex $T \in \mathcal{T}$ and generate the global piecewise polynomials of order k, i.e.

$$P_k(\mathcal{T}) := \{q_h \in L^2(\Omega) \,|\, \forall T \in \mathcal{T} : v_h|_T \in P_k(T)\}.$$

The function $\pi_{P_k(\omega)}$ denotes the L^2 best approximation into $P_k(\omega)$ for any subdomain $\omega \subset \Omega$. For approximation of functions in $H(\mathrm{div}, \Omega)$ we use the set of Brezzi-Douglas-Marini functions of order $k \geqslant 1$ denoted by $BDM_k(\mathcal{T}) := P_k(\mathcal{T})^2 \cap H(\mathrm{div}, \Omega)$ and the subset of Raviart-Thomas functions of order $k \geqslant 0$ denoted by $RT_k(\mathcal{T})$, see [28]. The functions I_{RT_k} and I_{BDM_k} denotes the standard interpolator into $RT_k(\mathcal{T})$ and $BDM_k(\mathcal{T})$, respectively, see e.g. [8]. We are also using lowest order Nédélec (type I) functions $\mathcal{N}_0(\mathcal{T})$ defined as the 90 degree rotated lowest order Raviart-Thomas functions with the corresponding interpolator $I_{\mathcal{N}_0}$, see [27].

The diameter of a simplex $T \in \mathcal{T}$ is denoted by h_T and $h_\mathcal{T} \in P_0(\mathcal{T})$ is the local mesh width function, i.e. $h_\mathcal{T}|_T := h_T$ for all $T \in \mathcal{T}$. Similarly, h_E denotes the diameter of the side $E \in \mathcal{E}$. At some point certain bubble functions are used. The cell bubble function on a cell $T \in \mathcal{T}$ is defined by $b_T = \prod_{z \in \mathcal{N}} \varphi_z$ where φ_z is the nodal basis function of the node $z \in \mathcal{N}$, i.e. $\varphi_z(z) = 1$ and $\varphi_z(y) = 0$ for $y \in \mathcal{N}\backslash\{z\}$. Similarly, the face bubble b_E for some side $E \in \mathcal{E}$ is defined by $b_E = \prod_{z \in \mathcal{E}} \varphi_z$. The vector $\boldsymbol{n}_E$ denotes the unit normal vector of the side $E \in \mathcal{E}$ with arbitrary but fixed orientation, such that the normal jump $[\boldsymbol{v} \cdot \boldsymbol{n}]$ of some function $\boldsymbol{v}$ has a well-defined sign. The vector $\boldsymbol{\tau}_E$ denotes a unit tangential vector of E.

3. Classical and pressure-robust finite element methods

This section recalls classical (usually not presssure-robust) inf-sup stable finite element methods and a pressure-robust modification of these methods.

3.1. **Classical inf-sup stable finite element methods.** Classical inf-sup stable finite element methods choose ansatz spaces $V_h \subseteq V = H^1_0(\Omega)^2$ and $Q_h \subseteq Q = L^2_0(\Omega)$ with the inf-sup property

$$(7) \qquad 0 < c_0 := \inf_{q_h \in Q_h\backslash\{0\}} \sup_{\boldsymbol{v}_h \in V_h\backslash\{\boldsymbol{0}\}} \frac{\int_\Omega q_h \mathrm{div}\boldsymbol{v}_h \; dx}{\|\nabla\boldsymbol{v}_h\|_0 \|q_h\|_{L^2}}.$$

This guarantees surjectivity of the discrete divergence operator

$$\mathrm{div}_h\boldsymbol{v}_h = \Pi_{Q_h}(\mathrm{div}\boldsymbol{v}_h) := \mathrm{argmin}_{q_h \in Q_h} \|\mathrm{div}\boldsymbol{v}_h - q_h\|_{L^2},$$

but also leads to the set of only discretely divergence-free testfunctions

$$V_{0,h} = \{\boldsymbol{v}_h \in V_h \,|\, \mathrm{div}_h\boldsymbol{v}_h = 0\},$$

that in general is not a subset of the really divergence-free functions $\boldsymbol{V}_0$. Table 1 lists some classical finite element methods that are inf-sup stable and are considered in this paper. Besides the Scott-Vogelius finite element method (on a barycentric refined mesh bary($\mathcal{T}$) to ensure the inf-sup stability [32, 37]), all of them are not divergence-free. The space $P_k^+(\mathcal{T})$ in case of the P2-bubble [13] or the mini finite element methods [5] indicates that the $P_k(\mathcal{T})$ space is enriched with the standard

TABLE 1. List of classical finite element methods that are considered in this paper and their expected velocity gradient error convergence order k.

FEM name & reference & order	abbr.	V_h	Q_h
Bernardi–Raugel FEM [7] ($k=1$)	BR	$P_1^{\mathrm{BR}}(\mathcal{T}) \cap V$	$P_0(\mathcal{T})$
Mini FEM [5] ($k=1$)	MINI	$P_1^+(\mathcal{T})^2 \cap V$	$P_1(\mathcal{T}) \cap H^1(\Omega)$
$P_{k+1} \times P_{k-1}$ FEM ($k \geqslant 1$)	P2P0,...	$P_{k+1}(\mathcal{T})^2 \cap V$	$P_{k-1}(\mathcal{T})$
P2-bubble FEM [13] ($k=2$)	P2B	$P_2^+(\mathcal{T})^2 \cap V$	$P_1(\mathcal{T})$
Taylor–Hood FEM [18] ($k \geqslant 2$)	TH_k	$P_k(\mathcal{T})^2 \cap V$	$P_{k-1}(\mathcal{T}) \cap H^1(\Omega)$
Scott-Vogelius FEM [32, 37] (k=2)	SV	$P_2(\mathrm{bary}(\mathcal{T}))^2 \cap V$	$P_1(\mathrm{bary}(\mathcal{T}))$

cell bubbles b_T for all $T \in \mathcal{T}$. For the Bernardi–Raugel finite element method normal-weighted face bubbles are added [7] defining the space $P_1^{\mathrm{BR}}(\mathcal{T}) := P_1(\mathcal{T})^2 \cup \{b_E \boldsymbol{n}_E : E \in \mathcal{E}\}$.

The relaxation of the divergence constraint leads to the usual best approximation error in the pressure ansatz space, i.e.

$$
\begin{aligned}
(8) \qquad \|\nabla p\|_{V_{0,h}^\star} &:= \sup_{\boldsymbol{v}_h \in V_{0,h} \setminus \{\boldsymbol{0}\}} \frac{\int_\Omega p \mathrm{div} \boldsymbol{v}_h \; dx}{\|\nabla \boldsymbol{v}_h\|_{L^2}} \\
&= \sup_{\boldsymbol{v}_h \in V_{0,h} \setminus \{\boldsymbol{0}\}} \frac{\int_\Omega (p - q_h) \mathrm{div} \boldsymbol{v}_h \; dx}{\|\nabla \boldsymbol{v}_h\|_{L^2}} \leqslant \min_{q_h \in Q_h} \|p - q_h\|_{L^2},
\end{aligned}
$$

and divergence-free methods are characterised by

$$
V_{0,h} \subseteq \boldsymbol{V}_0 \;\Leftrightarrow\; \|\nabla p\|_{V_{0,h}^\star} = 0 \text{ for all } p \in L^2(\Omega).
$$

For completeness, we shortly prove the classical a priori error estimate in the following theorem for the discrete solution $\boldsymbol{u}_h \in \boldsymbol{u}_{D,h} + V_h$ (where $\boldsymbol{u}_{D,h}$ is some suitable approximation of $\boldsymbol{u}_D$) and $p_h \in Q_h$ defined by

$$
\begin{aligned}
(9) \qquad \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) - (p_h, \mathrm{div} \boldsymbol{v}_h) &= (\boldsymbol{f}, \boldsymbol{v}_h) && \text{for all } \boldsymbol{v}_h \in V_h, \\
(q_h, \mathrm{div} \boldsymbol{u}_h) &= 0 && \text{for all } q_h \in Q_h,
\end{aligned}
$$

or, equivalently,

$$
\nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \qquad \text{for all } \boldsymbol{v}_h \in V_{0,h}.
$$

Theorem 3.1 (A priori estimate for classical finite element methods). *For the discrete velocity $\boldsymbol{u}_h$ of* (9)*, it holds*

$$
\|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2}^2 \leqslant \inf_{\substack{\boldsymbol{v}_h \in V_{0,h}, \\ \boldsymbol{u}_h = \boldsymbol{v}_h \; on \; \partial\Omega}} \|\nabla(\boldsymbol{u} - \boldsymbol{v}_h)\|_{L^2}^2 + \frac{1}{\nu^2} \|\nabla p\|_{V_{0,h}^\star}^2.
$$

Proof. The best approximation $\boldsymbol{w}_h \in V_{0,h}$ with boundary data $\boldsymbol{w}_h = \boldsymbol{u}_h$ along $\partial\Omega$ of $\boldsymbol{u}$ in the H^1-seminorm satisfies in particular the orthogonality $(\nabla(\boldsymbol{u} - \boldsymbol{w}_h), \nabla(\boldsymbol{u}_h - \boldsymbol{w}_h)) = 0$ and therefore allows for the Pythagoras theorem

$$
\begin{aligned}
(10) \qquad \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2}^2 &= \|\nabla(\boldsymbol{u} - \boldsymbol{w}_h)\|_{L^2}^2 + \|\nabla(\boldsymbol{u}_h - \boldsymbol{w}_h)\|_{L^2}^2 \\
&= \inf_{\substack{\boldsymbol{v}_h \in V_{0,h}, \\ \boldsymbol{u}_h = \boldsymbol{v}_h \; \text{on} \; \partial\Omega}} \|\nabla(\boldsymbol{u} - \boldsymbol{v}_h)\|_{L^2}^2 + \|\nabla(\boldsymbol{u}_h - \boldsymbol{w}_h)\|_{L^2}^2.
\end{aligned}
$$

The same orthogonality allows to estimate

$$
\begin{aligned}
\|\nabla(\boldsymbol{u}_h - \boldsymbol{w}_h)\|_{L^2}^2 &= (\nabla(\boldsymbol{u} - \boldsymbol{u}_h), \nabla(\boldsymbol{u}_h - \boldsymbol{w}_h)) \\
&= \nu^{-1}(p, \mathrm{div}(\boldsymbol{u}_h - \boldsymbol{w}_h)) \leqslant \nu^{-1} \|\nabla p\|_{V_{0,h}^\star} \|\nabla(\boldsymbol{u}_h - \boldsymbol{w}_h)\|_{L^2}.
\end{aligned}
$$

$\square$

The malicious influence of the pressure-dependent error and the factor $1/\nu$ in front of it for classical finite element methods that are not divergence-free was demonstrated and observed in many benchmark examples, see e.g. [24, 25, 19, 20].

3.2. **Pressure-robust finite element methods.** A method is called pressure-robust if its discrete velocity is pressure-independent, i.e. if the a priori error estimate for the velocity error is independent of the pressure.

The key feature behind pressure-robustness for the Stokes problem is that the testfunctions in the right-hand side are divergence-free. This can be achieved e.g. by fully divergence-free finite element methods (like the Scott-Vogelius finite element method) or, focused on in this paper, by the application of some reconstruction operator Π in the right-hand side of the equation (and in further terms in case of the stationary and transient Navier–Stokes equations [25, 1]).

Hence, the modified pressure-robust finite element method (of any classical pair of inf-sup stable spaces V_h and Q_h) searches $\boldsymbol{u}_h \in \boldsymbol{u}_{D,h} + V_h$ and $p_h \in Q_h$ with

$$(11) \qquad \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) - (p_h, \mathrm{div}\boldsymbol{v}_h) = (\boldsymbol{f}, \Pi \boldsymbol{v}_h) = (\mathbb{P}\boldsymbol{f}, \Pi \boldsymbol{v}_h) \qquad \text{for all } \boldsymbol{v}_h \in V_h,$$

$$(q_h, \mathrm{div}\boldsymbol{u}_h) = 0 \qquad \text{for all } q_h \in Q_h.$$

The operator Π maps discretely divergence-free functions onto exactly divergence-free ones, i.e.

$$(12) \qquad \Pi : V_h \to H(\mathrm{div}, \Omega) \quad \text{with} \quad \mathrm{div}(\Pi \boldsymbol{v}_h) = 0 \quad \text{for all } \boldsymbol{v}_h \in V_{0,h}.$$

This implicitly defines a modified discrete Helmholtz projector

$$\mathbb{P}_h^\star \boldsymbol{f} = \mathrm{argmin}_{\boldsymbol{v}_h \in V_{0,h}} \|\boldsymbol{f} - \Pi \boldsymbol{v}_h\|_{L^2}$$

with $\mathbb{P}_h^\star(\nabla q) = 0$ for any $q \in H^1(\Omega)$ or $\|\nabla q\|_{(\Pi V_{0,h})^\star}^2 = 0$ for all $q \in L^2(\Omega)$ and so allows for a pressure-independent and locking-free a priori velocity error estimate.

Theorem 3.2 (A priori estimate for pressure-robust finite element methods)**.** *For the solution $\boldsymbol{u}_h$ of* (11) *with a reconstruction operator Π that satisfies* (12), *it holds*

$$\|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2}^2 \leqslant \inf_{\substack{\boldsymbol{v}_h \in V_{0,h}, \\ \boldsymbol{u}_h = \boldsymbol{v}_h \ on \ \partial\Omega}} \|\nabla(\boldsymbol{u} - \boldsymbol{v}_h)\|_{L^2}^2 + \|\Delta\boldsymbol{u} \circ (1 - \Pi)\|_{V_{0,h}^\star}^2$$

with the consistency error

$$(13) \qquad \|\Delta\boldsymbol{u} \circ (1 - \Pi)\|_{V_{0,h}^\star}^2 := \sup_{\boldsymbol{v}_h \in V_{0,h} \setminus \{\boldsymbol{0}\}} \frac{\int_\Omega \Delta\boldsymbol{u} \cdot (1 - \Pi)\boldsymbol{v}_h \ dx}{\|\nabla \boldsymbol{v}_h\|_{L^2}}.$$

Note, that divergence-free methods (like the Scott-Vogelius finite element method) allow for $\Pi = 1$ and so attain the same estimate as Theorem 3.1.

Proof. Similar to the proof of Theorem 3.1, it remains to estimate the second term on the right-hand side of (10). Using the orthogonality $(\nabla(\boldsymbol{u} - \boldsymbol{w}_h), \nabla(\boldsymbol{u}_h - \boldsymbol{w}_h)) = 0$ we get $\|\nabla(\boldsymbol{u}_h - \boldsymbol{w}_h)\|_{L^2}^2 = (\nabla(\boldsymbol{u} - \boldsymbol{u}_h), \nabla(\boldsymbol{u}_h - \boldsymbol{w}_h))$. The insertion of $\boldsymbol{f} = -\nu\Delta\boldsymbol{u} + \nabla p$ and $\int_\Omega \nabla p \cdot \Pi(\boldsymbol{u}_h - \boldsymbol{w}_h) = 0$ (thanks to (12)) then further shows

$$(\nabla(\boldsymbol{u} - \boldsymbol{u}_h), \nabla(\boldsymbol{w}_h - \boldsymbol{u}_h)) = (\Delta\boldsymbol{u}, \boldsymbol{u}_h - \boldsymbol{w}_h) + \frac{1}{\nu}(\boldsymbol{f}, \Pi(\boldsymbol{u}_h - \boldsymbol{w}_h))$$

$$= (\Delta\boldsymbol{u}, \boldsymbol{u}_h - \boldsymbol{w}_h) + (\Delta\boldsymbol{u}, \Pi(\boldsymbol{u}_h - \boldsymbol{w}_h))$$

$$\leqslant \|\Delta\boldsymbol{u} \circ (1 - \Pi)\|_{V_{0,h}^\star} \|\nabla(\boldsymbol{u}_h - \boldsymbol{w}_h)\|_{L^2}.$$

This concludes the proof. $\qquad \qquad \square$

To gain optimal convergence behavior of (13), the reconstruction operator additionally has to satisfy another important property that concerns the consistency error of the modified method. For a finite element method with optimal H^1-velocity convergence order k and pressure L^2-convergence order q we require, for all $\boldsymbol{v}_h \in V_{0,h}$,

$$(14) \qquad (\boldsymbol{g}, (1 - \Pi)\boldsymbol{v}_h) \lesssim \|h_{\mathcal{T}}^{q+1} D^{q-1}\boldsymbol{g}\|_{L^2(\Omega)} \|\nabla \boldsymbol{v}_h\|_{L^2} \qquad \text{for any } \boldsymbol{g} \in H^{q-1}(\Omega)^2.$$

In particular, for $\Delta\boldsymbol{u} \in H^{q-1}(\Omega)^2$, this property directly implies

$$(15) \qquad \|\Delta\boldsymbol{u} \circ (1 - \Pi)\|_{V_{0,h}^\star} \lesssim \|h_{\mathcal{T}}^{q+1} D^{q-1}\Delta\boldsymbol{u}\|_{L^2(\Omega)}$$

and so ensures that the modified method still converges with the optimal order.

Table 2. Suitable reconstruction operators Π for the classical FEMs of Table 1.

FEM name	abbr.	Π & reference
Bernardi–Raugel FEM	BR	I_{BDM_1}, see [25]
Mini FEM	MINI	see [20]
$P_{k+1} \times P_{k-1}$ FEM $(k \geqslant 1)$	P2P0, P3P1, ...	I_{BDM_k}
P2-bubble FEM	P2B	I_{BDM_2}, see [22, 25]
Taylor–Hood FEM $(k \geqslant 1)$	TH_k	see [20]
Scott-Vogelius FEM	SV	1 (identity)

To be more precise, we require that the reconstruction operator satisfies some local splitting and orthogonality property that can be formulated by

$$(16) \qquad (1-\Pi)\boldsymbol{v}_h = \sum_{K \in \mathcal{K}} \boldsymbol{\sigma}_K|_K \text{ with } \|\boldsymbol{\sigma}_K\|_{L^2(K)} \lesssim h_K \|\nabla \boldsymbol{v}_h\|_{L^2(K)} \text{ and}$$

$$\int_K \boldsymbol{\sigma}_K \cdot \boldsymbol{g}_h \ dx = 0 \quad \text{for all } \boldsymbol{g}_h \in P_{q-1}(K),$$

with $h_K := \operatorname{diam}(K)$. Reconstruction operators Π with the properties (12)-(14) were already successfully designed for finite element methods with discontinuous pressure spaces, like the nonconforming Crouzeix-Raviart finite element method [21, 9], or the Bernardi–Raugel [25] and P^2-bubble finite element methods [22, 25]. In all these cases Π can be chosen as standard BDM interpolators with elementwise-orthogonality with resepect to $\mathcal{K} = \mathcal{T}$. Recently, also for Taylor–Hood and mini finite element methods (with $k = q$) of arbitrary order such an operator was found [20]. For these vertex-based constructions Property (16) holds with $\mathcal{K} = \{\omega_z : z \in \mathcal{N}\}$. Table 2 summarizes suitable reconstruction operators, that satisfy the needed properties, for the methods from Table 1.

4. (Limits of) Standard a posteriori residual-based error bounds

This section states and proves a posteriori error bounds for the classical and the pressure-robust finite element methods by classical means. The resulting bounds reflect the pressure-robustness but are, in case of a pressure-robust finite element method, rather unhandy as their efficiency relies on a good approximation of $\mathbb{P}\boldsymbol{f}$. To stress this observation, the analysis is performed in some detail.

First, we define the residual for the Stokes equations by

$$r(\boldsymbol{v}) := \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \ dx - \int_\Omega \nu \nabla \boldsymbol{u}_h : \nabla \boldsymbol{v} \ dx \quad \text{for all } \boldsymbol{v} \in \boldsymbol{V}_0.$$

The dual norm of the residual r with respect to $\boldsymbol{V}_0$ defined by

$$\|r\|_{\boldsymbol{V}_0^\star} := \sup_{\boldsymbol{v}_h \in V_0 \backslash \{\boldsymbol{0}\}} \frac{r(\boldsymbol{v})}{\|\nabla \boldsymbol{v}\|_{L^2}}$$

enters the generalised error bound as the central object of a posteriori error estimation. The error analysis also assumes the existence of a Fortin interpolation operator I that maps from $\boldsymbol{V}_0$ to $V_{0,h}$ and has first-order approximation properties and is H^1-stable, i.e, for all $\boldsymbol{v} \in \boldsymbol{V}_0$, it holds

$$(17) \qquad \|(1-I)\boldsymbol{v}\|_{L^2(T)} \lesssim h_T \|\nabla \boldsymbol{v}\|_{L^2(\omega_T)} \quad \text{for all } T \in \mathcal{T},$$

$$(18) \qquad \|\nabla I \boldsymbol{v}\|_{L^2} \lesssim \|\nabla \boldsymbol{v}\|_{L^2}.$$

For many classical finite element methods such an operator can be found in [8]. For its existence and design in the Taylor–Hood case we refer to [26, 12]. Some more details are given in Section 6 below.

The following theorem establishes a general estimate similar to [17, Theorem 5.1] and can be extended to nonconforming methods in a similar fashion. However, our focus is on the consistency errors (8) and (13) and the dependency on ν.

Theorem 4.1. *The following velocity error estimates hold:*

(a) In general, the L^2 gradient error can be estimated by

$$\|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2}^2 \leqslant \nu^{-2}\|r\|_{\boldsymbol{V}_0^\star}^2 + 1/c_0^2\|\operatorname{div}\boldsymbol{u}_h\|_{L^2}^2.$$

(b) *For the discrete solution $\boldsymbol{u}_h$ of the modified method* (11) *(or of the classical method* (9) *with $\Pi = 1$), the dual norm of the residual r can be bounded by*

$$\|r\|_{\boldsymbol{V}_0^\star} \lesssim \eta_{class}(\sigma, q) := \eta_{vol}(\sigma, q) + \eta_{avg}(\sigma) + \eta_{jump}(\sigma) + \eta_{cons,1}(\sigma) + \eta_{cons,2}(q)$$

for arbitrary $q \in H^1(\Omega)$ and $\sigma \in H^1(\mathcal{T})^{2\times2}$. The subterms read

$$\eta_{vol}(\sigma, q) := \|h_{\mathcal{T}}(\boldsymbol{f} - \nabla q + \nu\mathrm{div}_h(\sigma))\|_{L^2}$$
$$\eta_{avg}(\sigma) := \nu\|\nabla\boldsymbol{u}_h - \sigma\|_{L^2}$$
$$\eta_{jump}(\sigma) := \|h_{\mathcal{E}}^{1/2}[\nu\sigma\boldsymbol{n}_E]\|_{L^2(\bigcup\mathcal{E}^\circ)}$$
$$\eta_{cons,1}(\sigma) := \|\nu\mathrm{div}_h(\sigma) \circ (1 - \Pi)\|_{V_{0,h}^\star}$$
$$\eta_{cons,2}(q) := \|\nabla q\|_{(\Pi V_{0,h})^\star},$$

Note that q acts as a conforming approximation of the pressure p and σ acts as an approximation of $\nabla\boldsymbol{u}$ (in particular $\sigma = \nabla\boldsymbol{u}_h$ is allowed).

Proof. The proof of (a) can be found in [2, 11] and is based on the decomposition $\nu\nabla(\boldsymbol{u} - \boldsymbol{u}_h) = \nu\nabla\boldsymbol{z} + y$ into some $\boldsymbol{z} \in \boldsymbol{V}_0$ and some remainder

$$y \in Y := \left\{ y \in L^2(\Omega)^{d\times d} \mid \int_\Omega y : \nabla v \, dx = 0 \text{ for all } v \in \boldsymbol{V}_0 \right\}.$$

The orthogonality relations between $\boldsymbol{z}$ and y lead to

$$\|\nu^{1/2}\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2}^2 = \|\nu^{1/2}\nabla\boldsymbol{z}\|_{L^2}^2 + \|\nu^{-1/2}y\|_{L^2}^2.$$

Since

$$\|\nu^{1/2}\nabla\boldsymbol{z}\|_{L^2}^2 = \int_\Omega \nu\nabla(\boldsymbol{u} - \boldsymbol{u}_h) : \nabla\boldsymbol{z} \, dx = r(\boldsymbol{z}) \leqslant \nu^{-1/2}\|r\|_{\boldsymbol{V}_0^\star}\|\nu^{1/2}\nabla\boldsymbol{z}\|_{L^2},$$

one arrives at $\|\nu^{1/2}\nabla\boldsymbol{z}\|_{L^2} \leqslant \nu^{-1/2}\|r\|_{\boldsymbol{V}_0^\star}$. This is in fact an identity, since

$$r(\boldsymbol{v}) = \int_\Omega \nu\nabla\boldsymbol{z} : \nabla\boldsymbol{v} \, dx \leqslant \nu^{1/2}\|\nu^{1/2}\nabla\boldsymbol{z}\|_{L^2}\|\nabla\boldsymbol{v}\|_{L^2} \quad \text{for any } \boldsymbol{v} \in \boldsymbol{V}_0.$$

Furthermore, there exists some $w \in L^2(\Omega)$ such that (see [2])

$$\|\nu^{-1/2}y\|_{L^2}^2 = \int_\Omega \nabla(\boldsymbol{u} - \boldsymbol{u}_h) : y \, dx = \int_\Omega w\mathrm{div}(\boldsymbol{u} - \boldsymbol{u_h}) \, dx$$
$$\leqslant \|w\|_{L^2}\|\mathrm{div}(\boldsymbol{u} - \boldsymbol{u_h})\|_{L^2} \leqslant \nu^{1/2}/c_0\|\nu^{-1/2}y\|_{L^2}\|\mathrm{div}\boldsymbol{u}_h\|_{L^2}.$$

Hence, $\|\nu^{-1/2}y\|_{L^2} \leqslant \nu^{1/2}/c_0\|\mathrm{div}\boldsymbol{u}_h\|_{L^2}$. This concludes the proof of (a) and it remains to prove (b).

Given any $\boldsymbol{v} \in \boldsymbol{V}_0$, subtraction of its Fortin interpolation $I\boldsymbol{v} \in V_{0,h}$ and (11) lead to

$$r(\boldsymbol{v}) = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \, dx - \int_\Omega \nu\nabla\boldsymbol{u}_h : \nabla\boldsymbol{v} \, dx$$
$$= \int_\Omega \boldsymbol{f} \cdot (1 - \Pi I)\boldsymbol{v} \, dx - \int_\Omega \nu\nabla\boldsymbol{u}_h : \nabla(1 - I)\boldsymbol{v} \, dx$$
$$= \int_\Omega \boldsymbol{f} \cdot (1 - \Pi I)\boldsymbol{v} \, dx - \int_\Omega \nu\sigma : \nabla(1 - I)\boldsymbol{v} \, dx - \int_\Omega \nu(\nabla\boldsymbol{u}_h - \sigma) : \nabla(1 - I)\boldsymbol{v} \, dx$$
$$= \int_\Omega (\boldsymbol{f} - \nabla q + \nu\mathrm{div}_h\sigma) \cdot (1 - \Pi I)\boldsymbol{v} \, dx + \sum_T \int_{\partial T} (\nu\sigma\boldsymbol{n}) \cdot (1 - I)\boldsymbol{v} \, ds$$
$$- \int_\Omega \nu(\nabla\boldsymbol{u}_h - \sigma) : \nabla(1 - I)\boldsymbol{v} \, dx + \int_\Omega \nu\mathrm{div}_h\sigma \cdot (1 - \Pi)I\boldsymbol{v} \, dx + \int_\Omega \nabla q \cdot \Pi I\boldsymbol{v} \, dx \,.$$

In the last step it was used that $\int \nabla q \cdot \boldsymbol{v} \, dx = 0$ for any $q \in H^1(\Omega)$, since $\boldsymbol{v} \in \boldsymbol{V}_0$ is divergence-free. The third integral is estimated by a Cauchy inequality and the H^1-stability of I. The last two integrals

are estimated by discrete dual norms and the H^1-stability of I. Properties (17)-(18) of I and (16) of Π show

$$\|h_{\mathcal{T}}^{-1}(1-\Pi I)\boldsymbol{v}\|_{L^2(T)} \lesssim \|h_{\mathcal{T}}^{-1}(1-I)\boldsymbol{v}\|_{L^2(T)} + \|h_{\mathcal{T}}^{-1}(1-\Pi)I\boldsymbol{v}\|_{L^2(T)}$$
$$\lesssim \|\nabla\boldsymbol{v}\|_{L^2(\omega_T)} + \|h_{\mathcal{T}}\nabla(I\boldsymbol{v})\|_{L^2(\omega_T)} \lesssim \|\nabla\boldsymbol{v}\|_{L^2(\omega_T)}$$

and hence together with some Cauchy inequalities

$$\int_\Omega (\boldsymbol{f} - \nabla q + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h)\cdot(1-\Pi I)\boldsymbol{v}\ dx$$
$$\leqslant \sum_{T\in\mathcal{T}} \|h_{\mathcal{T}}(\boldsymbol{f} - \nabla q + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h)\|_{L^2(T)}\|h_{\mathcal{T}}^{-1}(1-\Pi I)\boldsymbol{v}\|_{L^2(T)}$$
$$\lesssim \eta_{\mathrm{vol}}(\sigma,q)\left(\sum_{T\in\mathcal{T}}\|\nabla\boldsymbol{v}\|_{L^2(\omega_T)}^2\right)^{1/2} \lesssim \eta_{\mathrm{vol}}(\sigma,q)\|\nabla\boldsymbol{v}\|_{L^2}.$$

Similar arguments hold for the edge-based integral using a trace inequality and Properties (17)-(18), i.e.

$$\sum_{E\in\mathcal{E}^\circ}\int_E [\nu\sigma\boldsymbol{n}]\cdot(\boldsymbol{v}-I\boldsymbol{v})\ ds \leqslant \sum_{E\in\mathcal{E}^\circ} \|[\nu\sigma\boldsymbol{n}]\|_{L^2(E)}\|\boldsymbol{v}-I\boldsymbol{v}\|_{L^2(E)}$$
$$\leqslant \sum_{E\in\mathcal{E}^\circ} h_E^{1/2}\|[\nu\sigma\boldsymbol{n}]\|_{L^2(E)}\|\nabla\boldsymbol{v}\|_{L^2(\omega_E)}$$
$$\leqslant \|h_{\mathcal{E}}^{1/2}[\nu\sigma\boldsymbol{n}]\|_{L^2(\mathcal{E}^\circ)}\|\nabla\boldsymbol{v}\|_{L^2} = \eta_{\mathrm{jump}}(\sigma)\|\nabla\boldsymbol{v}\|_{L^2}.$$

This concludes the proof of (b). $\qquad\square$

Remark 4.2. *Some remarks are in order:*

• *The existence of w in the last part of the proof of (a) needs $\boldsymbol{u}-\boldsymbol{u}_h \in H_0^1(\Omega)^2$. In case of inhomogeneous Dirichlet boundary data or nonconforming discretisations $\boldsymbol{u}_h \notin H^1(\Omega)^2$, one can introduce a function $\boldsymbol{w}\in H^1(\Omega)$ (e.g. the harmonic extension of the boundary data error $\boldsymbol{u}_D - \boldsymbol{u}_{D,h}$ [6] plus some H^1-conforming boundary-data preserving interpolation of $\boldsymbol{u}_h$ [2, 17, 11]) with $\boldsymbol{w}=\boldsymbol{u}_D$ along $\partial\Omega$ and attains $\boldsymbol{u}-\boldsymbol{w}\in H_0^1(\omega)$. Then, a modified estimation of the second term yields*

$$\|\nu^{-1/2}y\|_{L^2} \leqslant \nu^{1/2}/c_0\|\mathrm{div}\boldsymbol{w}\|_{L^2} + \nu^{1/2}\|\nabla_h(\boldsymbol{u}_h - \boldsymbol{w})\|_{L^2}.$$

• *The term $\eta_{cons,1}(\sigma) = \|\nu\Delta_{\mathcal{T}}(\mathrm{div}\sigma)\circ(1-\Pi)\|_{V_{0,h}^\star}$ only appears for $\Pi\neq 1$ as in the novel pressure-robust methods and equals the consistency error (13) for $\sigma = \nabla\boldsymbol{u}_h$.*

• *Recall that $\eta_{cons,2}(q) = 0$ if Π satisfies (12) or if $q\in Q_h$ and $\Pi = 1$.*

The following theorem studies the efficiency of the contributions of the standard residual error estimators from Theorem 4.1 for the explicit choice $\sigma = \nabla\boldsymbol{u}_h$.

Theorem 4.3 (Efficiency for $\sigma = \nabla\boldsymbol{u}_h$). *For $\sigma = \nabla\boldsymbol{u}_h$ all terms of the residual-based error estimator of Theorem 4.1 are efficient possibly up to data oscillations*

$$\mathrm{osc}_k(\bullet,\mathcal{T})^2 := \sum_{T\in\mathcal{T}} h_T^2\|(1-\pi_{P_k(T)})\bullet\|_{L^2(T)}^2$$

and up to pressure contributions (either from the lack of pressure-robustness or from the quality of the approximation of p by q) in the following sense.

(a) *For the divergence term it holds $\|\mathrm{div}\boldsymbol{u}_h\|_{L^2} \leqslant \|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2}$.*

(b) *For the volume term $\eta_{vol}(q,\nabla\boldsymbol{u}_h)$, it holds*

$$\nu^{-1}\|h_T(\boldsymbol{f}-\nabla q + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h)\|_{L^2} \lesssim \|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2}$$
$$+ \nu^{-1}\left(\|p-q\|_{L^2} + \mathrm{osc}_k(\boldsymbol{f}-\nabla q,\mathcal{T})\right).$$

(c) *For the jump term $\eta_{jump}(\nabla\boldsymbol{u}_h)$, it holds*

$$\nu^{-1}\|h_{\mathcal{E}}^{1/2}[\nu\nabla\boldsymbol{u}_h\boldsymbol{n}_E]\|_{L^2(\bigcup\mathcal{E}^\circ)} \lesssim \|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2} + \nu^{-1}\mathrm{osc}_k(\boldsymbol{f}-\nabla p,\mathcal{T}).$$

(d) *If Π satisfies (16), the consistency error $\eta_{cons,1}(\nabla\boldsymbol{u})$ is efficient in the sense*

$$\|\Delta_{\mathcal{T}}\boldsymbol{u}_h \circ (1-\Pi)\|_{V_{0,h}^\star} \lesssim \|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2}$$
$$+ \nu^{-1}\left(\mathrm{osc}_k(\boldsymbol{f}-\nabla p,\mathcal{T}) + \mathrm{osc}_{q-1}(\boldsymbol{f}-\nabla p,\mathcal{K})\right)$$

(e) *For the consistency error $\eta_{cons,2}(q)$, it holds*

$$\|\nabla q\|_{(\Pi V_{0,h})^\star} \leqslant \begin{cases} 0 & \text{if } \Pi \text{ satisfies (12)}, \\ \min_{q_h \in Q_h} \|q-q_h\|_{L^2} & \text{if } \Pi=1 \text{ without (12)}. \end{cases}$$

Proof. The proof of (a) simply uses $\mathrm{div}\boldsymbol{u}=0$ to estimate

$$\|\mathrm{div}\boldsymbol{u}_h\|_{L^2} = \|\mathrm{div}(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2} \leqslant \|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2}.$$

The last inequality follows from the identity $\|\nabla\boldsymbol{v}\|^2 = \|\mathrm{curl}\boldsymbol{v}\|^2 + \|\mathrm{div}\boldsymbol{v}\|^2$ for any $\boldsymbol{v} \in H_0^1(\Omega)^2$, see e.g. [4, Remark 2.6].

The proof of (b) and (c) is standard and employs the bubble-technique of Verfürth, see e.g. [34, 35] or into the proof of Theorem 5.4 below.

To show (d), observe that Property (16) leads to

$$\int_\Omega \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h \cdot (1-\Pi)\boldsymbol{v}_h \ dx = \sum_{K\in\mathcal{K}} \int_K \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h \cdot \sigma_K \ dx$$
$$= \sum_{K\in\mathcal{K}} \int_K (\boldsymbol{f}-\nabla p + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h)\cdot\sigma_K \ dx - \int_K (1-\pi_{P_{q-1}(K)})(\boldsymbol{f}-\nabla p)\cdot\sigma_K \ dx$$
$$\lesssim \sum_{K\in\mathcal{K}} h_K \left(\|\boldsymbol{f}-\nabla p + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h\|_{L^2(K)}\right.$$
$$\left. + \|(1-\pi_{P_{q-1}(K)})(\boldsymbol{f}-\nabla p)\|_{L^2(K)}\right)\|h_K^{-1}\sigma_K\|_{L^2(K)}$$
$$\lesssim \left(\sum_{K\in\mathcal{K}} \|h_K(\boldsymbol{f}-\nabla p + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h)\|_{L^2(K)}^2\right)^{1/2} \|\nabla\boldsymbol{v}_h\|_{L^2} + \mathrm{osc}_{q-1}(\boldsymbol{f}-\nabla p,\mathcal{K})\|\nabla\boldsymbol{v}_h\|_{L^2}$$
$$= (\eta_{\mathrm{vol}}(p,\nabla\boldsymbol{u}) + \mathrm{osc}_{q-1}(\boldsymbol{f}-\nabla p,\mathcal{K}))\|\nabla\boldsymbol{v}_h\|_{L^2}.$$

A division by $\|\nabla\boldsymbol{v}_h\|_{L^2}$ and the result from (b) conclude the proof of (d).

The proof of (e) is straight forward and employs integration by parts and the orthogonality of $\mathrm{div}(\boldsymbol{v}_h)$ onto all $q_h \in Q_h$ if $\Pi=1$ does not satisfy (12). Otherwise, if Π satisfies (12), the assertion follows from $\mathrm{div}(\Pi\boldsymbol{v}_h)=0$. □

Remark 4.4. *Theorem 4.3.(b) shows the pressure-dependence also in the efficiency estimate. The volume term $\eta_{vol}(q,\nabla\boldsymbol{u}_h)$ scales with the term $\nu^{-1}\|p-q\|_{L^2}$. Hence, a pressure-robust method is only efficient with a good approximation $q \approx p$. In the hydrostatic (worst) case with $\boldsymbol{u}_h=0$ and $\boldsymbol{f}=\nabla p$, $\eta_{vol}(q,\nabla\boldsymbol{u})$ is not zero (hence inefficient with efficiency index infinity) as long as $q \neq p$ is inserted. To compute the correct pressure is in general impossible or expensive. Some strategy to find an approximation that at least yields a higher-order term is discussed in [23].*

Note however, that $\eta_{vol}(q,\nabla\boldsymbol{u}_h)$ is efficient for a classical pressure-inrobust method with $q_h = p_h$ (or some suitable H^1-approximation), since then the discrete velocity error and its velocity error also depends on $\nu^{-1}\|p-p_h\|_{L^2}$, see e.g. our numerical examples in Section 7.

5. Refined residual-based error bounds

This section offers an alternative a posteriori error estimator and is related to the stream function and vorticity formulation of the Navier–Stokes equations. The analysis employs the two-dimensional curl operators for vector and scalar fields

$$\mathrm{curl}\boldsymbol{\phi} := (\partial\phi_2/\partial x - \partial\phi_1/\partial y) \quad \text{for } \boldsymbol{\phi}=(\phi_1,\phi_2) \in H^1(\Omega)^2,$$

$$\mathrm{curl}\phi := \begin{pmatrix} -\partial\phi/\partial y \\ \partial\phi/\partial x \end{pmatrix} \quad \text{for } \phi \in H^1(\Omega).$$

The outcome of this alternative approach is a different volume term that only takes $\mathrm{curl}(\boldsymbol{f})$ into account and so automatically cancels the gradient part of the Helmholtz decomposition. Hence, no knowledge or good approximation of $\mathbb{P}\boldsymbol{f}$ is needed. The resulting terms are related to the terms in [3] where error indicators for discretisations of the streamline and vorticity formulation were derived. However, our error estimator holds for pressure-robust finite element methods for the velocity and pressure formulation of the Navier–Stokes equations.

Given a Fortin interpolator I and a reconstruction operator Π with (12) (possibly $\Pi = 1$ for divergence-free finite element methods like the Scott-Vogelius finite element method), the novel approach exploits that $\Pi I\boldsymbol{v}$ for some divergence-free function $\boldsymbol{v} \in \boldsymbol{V}_0$ is again a divergence-free function in $L_\sigma^2(\Omega)$. Our analysis needs the following assumption on the two operators additional to (12) and (17)-(18).

Assumption 5.1. *For every $\boldsymbol{v} \in \boldsymbol{V}_0$, the Fortin interpolator I and the reconstruction operator Π satisfy*

$$\Pi I\boldsymbol{v} \in L_\sigma^2(\Omega) \quad \text{and hence} \quad \int_\Omega (1 - \Pi I)\boldsymbol{v} \cdot \nabla q \ dx = 0 \quad \text{for all } q \in H^1(\Omega),$$

and the estimate

$$\int_\Omega (1 - \Pi I)\boldsymbol{v} \cdot \boldsymbol{\theta} \ dx \lesssim \|\nabla\boldsymbol{v}\|_{L^2}\|h_\mathcal{T}^2 \mathrm{curl}\boldsymbol{\theta}\|_{L^2} \quad \text{for all } \boldsymbol{\theta} \in H(\mathrm{curl}, \Omega).$$

Theorem 5.2 (Novel error estimator for pressure-robust methods). *For $\boldsymbol{u}_h$ of (11) and any $\sigma \in H^1(\mathcal{T})^{2\times 2}$ (that approximates or equals $\nabla\boldsymbol{u}_h$), the error estimator*

$$\eta_{new}(\sigma) := \eta_{curl}(\sigma) + \eta_{jump}(\sigma) + \eta_{jump,2}(\sigma) + \eta_{avg}(\sigma) + \eta_{cons,1}(\sigma)$$

with the subterms

$$\eta_{curl}(\sigma) := \|h_\mathcal{T}^2 \mathrm{curl}_\mathcal{T}(\boldsymbol{f} + \nu\mathrm{div}_h\sigma)\|_{L^2}$$
$$\eta_{jump}(\sigma) := \|h_\mathcal{E}^{1/2}[\nu\sigma\boldsymbol{n}_E]\|_{L^2(\mathcal{E}^\circ)}$$
$$\eta_{jump,2}(\sigma) := \|h_\mathcal{E}^{3/2}[(\boldsymbol{f} + \nu\mathrm{div}_h\sigma) \cdot \boldsymbol{\tau_E}]\|_{L^2(\mathcal{E}^\circ)}$$
$$\eta_{avg}(\sigma) := \nu\|\nabla\boldsymbol{u}_h - \sigma\|_{L^2}$$
$$\eta_{cons,1}(\sigma) := \|\nu\mathrm{div}_h(\sigma) \circ (1 - \Pi)\|_{V_{0,h}^\star}$$

satisfies

$$\|r\|_{\boldsymbol{V}_0^\star} \lesssim \eta(q) \quad \text{and hence} \quad \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2}^2 \lesssim \nu^{-2}\eta(q)^2 + \frac{1}{c_0^2}\|\mathrm{div}\boldsymbol{u}_h\|_{L^2}^2.$$

Note in particular, that the volume contribution $\eta_{vol}(q, \sigma)$ from Theorem 4.1 has been replaced by the quantity $\eta_{curl}(\sigma)$ that is pressure-independent (or q-independent).

Proof. As in the estimation of $\|r\|_{\boldsymbol{V}_0^\star}$ in the proof of Theorem 4.1.(b), we subtract the Fortin interpolation $I\boldsymbol{v}$ of any testfunction $\boldsymbol{v}$ by employing (11), i.e.

$$r(v) = \int_\Omega \boldsymbol{f} \cdot (\boldsymbol{v} - \Pi I\boldsymbol{v}) \ dx - \nu\int_\Omega \nabla\boldsymbol{u}_h : \nabla(\boldsymbol{v} - I\boldsymbol{v}) \ dx.$$

Given any $\sigma \in H^1(\mathcal{T})^{2\times 2}$, an (element-wise) integration by parts shows

$$r(v) = \int_\Omega (\boldsymbol{f} + \nu\mathrm{div}_h\sigma) \cdot (\boldsymbol{v} - \Pi I\boldsymbol{v}) \ dx + \nu\int_\Omega (\sigma - \nabla\boldsymbol{u}_h) : \nabla(\boldsymbol{v} - I\boldsymbol{v}) \ dx$$
$$+ \nu\sum_{E\in\mathcal{E}^\circ}\int_E [\sigma\boldsymbol{n}] \cdot (\boldsymbol{v} - I\boldsymbol{v}) \ ds + \nu\int_\Omega (\mathrm{div}_h\sigma) \cdot (\Pi I\boldsymbol{v} - I\boldsymbol{v}) \ dx =: A + B + C + D.$$

The terms B, C and D are estimated as in Theorem 4.1.(b) by

$$B := \nu \sum_{T \in \mathcal{T}} \int_T (\sigma - \nabla \boldsymbol{u}_h) : \nabla(\boldsymbol{v} - I\boldsymbol{v}) \ dx \leqslant \nu \|\sigma - \nabla \boldsymbol{u}_h\|_{L^2} \|\nabla \boldsymbol{v}\|_{L^2}$$

$$C := \nu \sum_{E \in \mathcal{E}^\circ} \int_E [\sigma \boldsymbol{n}] \cdot (\boldsymbol{v} - I\boldsymbol{v}) \ ds \leqslant \nu \|h_{\mathcal{E}}^{1/2}[\sigma \boldsymbol{n}]\|_{L^2(\mathcal{E}^\circ)} \|\nabla \boldsymbol{v}\|_{L^2}$$

$$D := \nu \int_\Omega (\mathrm{div}_h \sigma) \cdot (\Pi I \boldsymbol{v} - I\boldsymbol{v}) \ dx \leqslant \nu \|(\mathrm{div}_h \sigma) \circ (1 - \Pi)\|_{V_{0,h}^\star} \|\nabla \boldsymbol{v}\|_{L^2}.$$

It remains to estimate term A. As $\boldsymbol{v} - \Pi I\boldsymbol{v}$ is exactly divergence free and has a zero normal trace we can apply Theorem 3.1, chapter 1 in [16] to find a scalar potential $\psi \in H_0^1(\Omega)$ with $\mathrm{curl}\psi = \boldsymbol{v} - \Pi I\boldsymbol{v}$. In the following we bound the weighted L^2 norm of ψ. Note that from $h_{\mathcal{T}}^{-4}\psi \in L^2(\Omega)$ follows $h_{\mathcal{T}}^{-2}\psi \in h_{\mathcal{T}}^2 \mathrm{curl}(H(\mathrm{curl}, \Omega))$, due to the surjectivity of the curl operator (de Rham complex) and so

$$\|h_{\mathcal{T}}^{-2}\psi\|_{L^2(\Omega)} = \frac{\int_\Omega h_{\mathcal{T}}^{-2}\psi h_{\mathcal{T}}^{-2}\psi \ dx}{\|h_{\mathcal{T}}^{-2}\psi\|_{L^2(\Omega)}}$$

$$\leqslant \sup_{\boldsymbol{\theta} \in H(\mathrm{curl}, \Omega)} \frac{\int_\Omega h_{\mathcal{T}}^{-2}\psi h_{\mathcal{T}}^2 \mathrm{curl}\boldsymbol{\theta} \ dx}{\|h_{\mathcal{T}}^2 \mathrm{curl}\boldsymbol{\theta}\|_{L^2(\Omega)}} = \sup_{\boldsymbol{\theta} \in H(\mathrm{curl}, \Omega)} \frac{\int_\Omega \psi \mathrm{curl}\boldsymbol{\theta} \ dx}{\|h_{\mathcal{T}}^2 \mathrm{curl}\boldsymbol{\theta}\|_{L^2(\Omega)}}.$$

On the other hand one can bound the supremum by $\|h_{\mathcal{T}}^{-2}\psi\|_{L^2(\Omega)}$ with a simple Cauchy Schwarz estimate. Using Assumption 5.1 it follows by an integration by parts and $\psi \in H_0^1(\Omega)$ that

$$(19) \qquad \|h_{\mathcal{T}}^{-2}\psi\|_{L^2(\Omega)} = \sup_{\boldsymbol{\theta} \in H(\mathrm{curl}, \Omega)} \frac{\int_\Omega \psi \mathrm{curl}\boldsymbol{\theta} \ dx}{\|h_{\mathcal{T}}^2 \mathrm{curl}\boldsymbol{\theta}\|_{L^2(\Omega)}}$$

$$= \sup_{\boldsymbol{\theta} \in H(\mathrm{curl}, \Omega)} \frac{\int_\Omega \mathrm{curl}\psi \cdot \boldsymbol{\theta} \ dx}{\|h_{\mathcal{T}}^2 \mathrm{curl}\boldsymbol{\theta}\|_{L^2(\Omega)}} \lesssim \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}.$$

With $\boldsymbol{\theta}_h := \boldsymbol{f} + \nu \mathrm{div}_h \sigma$ and $\psi = 0$ on $\partial\Omega$ a piecewise integration by parts yields

$$A := \int_\Omega \boldsymbol{\theta}_h \cdot (\boldsymbol{v} - \Pi I\boldsymbol{v}) \ dx = \int_\Omega \boldsymbol{\theta}_h \cdot \mathrm{curl}\psi \ dx$$

$$= \sum_{T \in \mathcal{T}} \int_T \mathrm{curl}\boldsymbol{\theta}_h \psi \ dx + \sum_{E \in \mathcal{E}^\circ} \int_E [\boldsymbol{\theta}_h \cdot \boldsymbol{\tau}_E]\psi \ ds$$

$$\lesssim \sum_{T \in \mathcal{T}} \|h_T^2 \mathrm{curl}\boldsymbol{\theta}_h\|_{L^2(T)} \|h_T^{-2}\psi\|_{L^2(T)} + \sum_{E \in \mathcal{E}^\circ} \|h_E^{3/2}[\boldsymbol{\theta}_h \cdot \boldsymbol{\tau}_E]\|_{L^2(E)} \|h_E^{-3/2}\psi\|_{L^2(E)}$$

$$\lesssim \left(\|h_{\mathcal{T}}^2 \mathrm{curl}_{\mathcal{T}} \boldsymbol{\theta}_h\|_{L^2(\Omega)} + \|h^{3/2}[\boldsymbol{\theta}_h \cdot \boldsymbol{\tau}_E]\|_{L^2(\mathcal{E}^\circ)} \right) \left(\|h_{\mathcal{T}}^{-2}\psi\|_{L^2(\Omega)} + \|h_{\mathcal{E}}^{-3/2}\psi\|_{L^2(\mathcal{E}^\circ)} \right).$$

Using a standard scaling argument we get, for each edge $E \in \mathcal{E}^\circ$,

$$\|h_E^{-3/2}\psi\|_{L^2(E)} \lesssim h_T^{-2}\|\psi\|_{L^2(T)} + h_T^{-1}\|\nabla\psi\|_{L^2(T)}.$$

For the second term in the previous estimate we have

$$h_T^{-1}\|\nabla\psi\|_{L^2(T)} = h_T^{-1}\|\mathrm{curl}\psi\|_{L^2(T)} = h_T^{-1}\|\boldsymbol{v} - \Pi I\boldsymbol{v}\|_{L^2(T)} \lesssim \|\nabla \boldsymbol{v}\|_{L^2(\omega_T)}.$$

Together with (19) and an overlap argument this leads to

$$\|h_{\mathcal{T}}^{-3/2}\psi\|_{L^2(\mathcal{E}^\circ)} \lesssim \|h_{\mathcal{T}}^{-2}\psi\|_{L^2(\Omega)} + \|h_{\mathcal{T}}^{-1}\nabla\psi\|_{L^2(\Omega)} \lesssim \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}.$$

This concludes the estimate for A, i.e.

$$A \lesssim (\eta_{\mathrm{curl}}(\sigma) + \eta_{\mathrm{jump},2}(\sigma)) \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}.$$

The collection of all separate estimates for A to D shows

$$r(v) \lesssim \eta(\sigma) \|\nabla \boldsymbol{v}\|_{L^2}$$

and a division by $\|\nabla \boldsymbol{v}\|_{L^2}$ concludes the proof. $\qquad\qquad \square$

The same techniques also a yield a novel error estimate for classical methods.

Proposition 5.3 (Novel error estimator for classical methods). *For $\boldsymbol{u}_h$ of* (9) *and any $\sigma \in H^1(\mathcal{T})^{2\times 2}$ (that approximates or equals $\nabla \boldsymbol{u}_h$), the error estimator*

$$\eta_{new}(\sigma) \quad := \quad \eta_{curl}(\sigma) \quad + \quad \eta_{jump}(\sigma) \quad + \quad \eta_{avg}(\sigma) \quad + \quad \|(\boldsymbol{f} \quad + \quad \nu \mathrm{div}_h \sigma) \quad \circ \quad (1 \quad - \quad \Pi)\|_{V^\star_{0,h}}$$

satisfies

$$\|r\|_{\boldsymbol{V}^\star_0} \lesssim \eta(q) \quad \text{and hence} \quad \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|^2_{L^2} \lesssim \nu^{-2}\eta(q)^2 + \frac{1}{c_0^2}\|\mathrm{div}\boldsymbol{u}_h\|^2_{L^2}.$$

Note, that Π is used only in the error estimator here, but not in the calculation of $\boldsymbol{u}_h$. It is not allowed to set $\Pi = 1$ if the classical method is not divergence-free, i.e. Π has to satisfy (12). *The difference to the previous theorem lies in the appeerence of $\boldsymbol{f}$ in the consistency error $\|(\boldsymbol{f} + \nu \mathrm{div}_h \sigma) \circ (1 - \Pi)\|_{V^\star_{0,h}}$.*

Proof. The proof follows the proof of Theorem 5.2 but one has to add the term $\int_\Omega \boldsymbol{f} \cdot (\boldsymbol{v} - \Pi \boldsymbol{v}) \, dx$ which can be added to the estimate of term C. $\qquad\square$

The next theorem establishes the efficiency of the novel terms $\eta_{\mathrm{curl}}(\sigma)$ and $\eta_{\mathrm{jump},2}(\sigma)$ for $\sigma = \nabla \boldsymbol{u}_h$. For the efficiency of the other terms see Theorem 4.3.

Theorem 5.4 (Efficiency for $\sigma = \nabla \boldsymbol{u}_h$). *It holds*

(a) $\nu^{-1}h_T^2\|\mathrm{curl}_{\mathcal{T}}(\boldsymbol{f} + \nu \Delta_{\mathcal{T}}\boldsymbol{u}_h)\|_{L^2(T)} \lesssim \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(T)}$
$\hspace{6cm} + \nu^{-1}h_T\mathrm{osc}_k(\mathrm{curl}(\boldsymbol{f} + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h), T),$

(b) $\nu^{-1}h_E^{3/2}\|[(\boldsymbol{f} + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h) \cdot \boldsymbol{\tau_E}]\|_{L^2(E)} \lesssim \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\omega_E)}$
$\hspace{6cm} + \nu^{-1}h_E\mathrm{osc}_k(\mathrm{curl}_{\mathcal{T}}(\boldsymbol{f} + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h), \mathcal{T}(E))$
$\hspace{6cm} + \nu^{-1}h_E\mathrm{osc}_k([f \cdot \boldsymbol{\tau_E}], E) + \mathrm{osc}_k(\boldsymbol{f} - \nabla p, \mathcal{T}(E)),$

for all $T \in \mathcal{T}$ and $E \in \mathcal{E}^\circ$.

Proof. The proof employs the standard Verfürth bubble-technique. To shorten the notion in the proof of (a), we define

$$Q_T := \mathrm{curl}(\boldsymbol{f} + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h)|_T \quad \text{for any } T \in \mathcal{T}.$$

Then, it holds (similarly to [35])

$$\|\pi_{P_k(T)}Q_T\|_{L^2(T)}$$
$$\lesssim \sup_{\boldsymbol{v}_T \in P_k(T)^2} \int_T \pi_{P_k(T)}Q_T \cdot (b_T^2 \boldsymbol{v}_T) \, dx \, / \|\boldsymbol{v}_T\|_{L^2(T)}$$
$$\leqslant \sup_{\boldsymbol{v}_T \in P_k(T)^2} \frac{\int_T Q_T b_T^2 \boldsymbol{v}_T \, dx}{\|\boldsymbol{v}_T\|_{L^2(T)}} + \sup_{\boldsymbol{v}_T \in P_k(T)^2} \frac{\|Q_T - \pi_{P_k(T)}Q_T\|_{L^2(T)}\|b_T^2\boldsymbol{v}_T\|_{L^2(T)}}{\|\boldsymbol{v}_T\|_{L^2(T)}}.$$

Testing the continuous system with the (divergence-free) testfunction $\mathrm{curl}(b_T^2\boldsymbol{v}_T) \in H^2(T)^2 \cap H_0^1(\Omega)^2$ and an integration by parts leads to

$$\int_T Q_T b_T^2 \boldsymbol{v}_T \, dx = \int_T (\boldsymbol{f} + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h) \cdot \mathrm{curl}(b_T^2\boldsymbol{v}_T) \, dx$$
$$= \int_T \nu\nabla(\boldsymbol{u} - \boldsymbol{u}_h) : \nabla\mathrm{curl}(b_T^2\boldsymbol{v}_T) \, dx$$
$$\leqslant \nu\|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(T)}\|\nabla\mathrm{curl}(b_T^2\boldsymbol{v}_T)\|_{L^2(T)}.$$

A discrete inverse inequality shows $\|\nabla\mathrm{curl}(b_T^2\boldsymbol{v}_T)\|_{L^2(T)} \lesssim h_T^{-2}\|b_T^2\boldsymbol{v}_T\|_{L^2(T)}$. This and the norm equivalence $\|b_T^2\boldsymbol{v}_T\|_{L^2(T)} \approx \|\boldsymbol{v}_T\|_{L^2(T)}$ lead to

$$h_T^2\|\pi_{P_k(T)}Q_T\|_{L^2(T)} \lesssim \nu\|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(T)} + h_T^2\|Q_T - \pi_{P_k(T)}Q_T\|_{L^2(T)}.$$

This concludes the proof of (a).

In the proof of (b), we use the notation

$$Q_E := [\boldsymbol{f} + \nu\Delta_{\mathcal{T}}\boldsymbol{u}_h] \cdot \boldsymbol{\tau_E} \quad \text{for any } E \in \mathcal{E}$$

and the face bubble b_E with support ω_E for every face $E \in \mathcal{E}$. Then,

$$\|\pi_{P_k(E)} Q_E\|_{L^2(E)} \lesssim \sup_{\boldsymbol{v}_E \in P_k(E)^2} \frac{\int_E Q_E \cdot (b_E^2 \boldsymbol{v}_E) \, ds}{\|\boldsymbol{v}_E\|_{L^2(E)}} + \|Q_E - \pi_{P_k(E)} Q_E\|_{L^2(E)}.$$

Testing the continuous equation with the divergence-free testfunction $\mathrm{curl}(b_E^2 \boldsymbol{v}_E) \in H_0^1(\Omega)$ (where $\boldsymbol{v}_E$ is reasonably extended to ω_E) and an integration by parts show

$$\int_E Q_E \cdot (b_E^2 \boldsymbol{v}_E) \, ds$$
$$= \int_E [(\boldsymbol{f} + \nu \Delta_{\mathcal{T}} \boldsymbol{u}_h) \cdot \boldsymbol{\tau}_E] \cdot (b_E^2 \boldsymbol{v}_E) \, ds$$
$$= \int_{\omega_E} (\boldsymbol{f} + \nu \Delta_{\mathcal{T}} \boldsymbol{u}_h) : \mathrm{curl}(b_E^2 \boldsymbol{v}_E) \, dx - \int_{\omega_E} \mathrm{curl}(\boldsymbol{f} + \nu \Delta_{\mathcal{T}} \boldsymbol{u}_h) : (b_E^2 \boldsymbol{v}_E) \, dx$$
$$\leqslant \|\boldsymbol{f} + \nu \Delta_{\mathcal{T}} \boldsymbol{u}_h\|_{L^2(\omega_E)} \|\mathrm{curl}(b_E^2 \boldsymbol{v}_E)\|_{L^2(\omega_E)} + \|Q_T\|_{L^2(\omega_E)} \|b_E^2 \boldsymbol{v}_E\|_{L^2(\omega_E)}.$$

A discrete inverse inequality $\|\mathrm{curl}(b_E^2 \boldsymbol{v}_E)\|_{L^2(\omega_E)} \lesssim h_T^{-1} \|b_E^2 \boldsymbol{v}_E\|_{L^2(\omega_E)}$ and a scaling argument (see [35]), that yields $\|b_E^2 \boldsymbol{v}_E\|_{L^2(\omega_E)} \lesssim h_T^{1/2} \|\boldsymbol{v}_E\|_{L^2(E)}$, show

$$h_E^{3/2} \|\pi_{P_k(E)} Q_E\|_{L^2(E)} \lesssim h_T \|\boldsymbol{f} + \nu \Delta_{\mathcal{T}} \boldsymbol{u}_h\|_{L^2(\omega_E)} + h_T^2 \|Q_T\|_{L^2(\omega_E)}$$
$$+ h_E^{3/2} \|Q_E - \pi_{P_k(E)} Q_E\|_{L^2(E)}.$$

The proof of Theorem 4.3.(c) yields

$$\|\boldsymbol{f} + \nu \Delta_{\mathcal{T}} \boldsymbol{u}_h\|_{L^2(\omega_E)} \lesssim \nu \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2} + \mathrm{osc}_k(\boldsymbol{f} - \nabla q, \mathcal{T}(E)).$$

This and the already proven result from (a) conclude the proof. $\qquad\square$

6. Proof of Assumption 5.1 for certain finite element methods

This section proves Assumption 5.1 for certain finite element methods. For the analysis several standard interpolation operators that are related to the de Rahm complex (see e.g. [31]) are employed. These are a (projection based) nodal interpolation operator $I_{\mathcal{L}}$, the lowest order Raviart-Thomas interpolation operator I_{RT_0} and the lowest-order Nédélec interpolation operator $I_{\mathcal{N}_0}$. These operators satisfy in particular the commuting diagram properties in two dimensions (see [14])

$$(20) \qquad\qquad \mathrm{curl}(I_{\mathcal{L}} \boldsymbol{v}) = I_{RT_0}(\mathrm{curl} \boldsymbol{v}) \quad \text{and} \quad \nabla(I_{\mathcal{L}} \boldsymbol{v}) = I_{\mathcal{N}_0}(\nabla \boldsymbol{v})$$

for arbitrary sufficiently smooth functions $\boldsymbol{v}$. Furthermore we need a refined Helmholtz decomposition.

Lemma 6.1 ([31]). *It exists an operator $\Pi_{\mathcal{N}_0} : H(\mathrm{curl}, \Omega) \to \mathcal{N}_0(\mathcal{T})$ with the property: for every $\boldsymbol{\theta} \in H(\mathrm{curl}, \Omega)$ exists a decomposition*

$$\boldsymbol{\theta} - \Pi_{\mathcal{N}_0} \boldsymbol{\theta} = \nabla \phi + \boldsymbol{y}$$

with $\phi \in H^1(\Omega)$, $\boldsymbol{y} \in H^1(\Omega)^2$, and

$$h_T^{-1} \|\boldsymbol{y}\|_{L^2(T)} + \|\nabla \boldsymbol{y}\|_{L^2(T)} \lesssim \|\mathrm{curl} \boldsymbol{\theta}\|_{L^2(T)} \quad \text{for all } T \in \mathcal{T}.$$

Proof. In [31] a proof for three dimensions is given. The two dimensional case follows similarly. $\qquad\square$

Lemma 6.2 (Regular decomposition). *For each $\boldsymbol{\theta} \in H(\mathrm{curl}, \omega)$ there exists a decomposition with $\alpha \in H^2(\omega)$ and $\boldsymbol{\beta} \in H^1(\omega)^2$ such that*

$$\boldsymbol{\theta} = \nabla \alpha + \boldsymbol{\beta},$$

with

$$\|\nabla \boldsymbol{\beta}\|_{L^2(\omega)} \lesssim \|\mathrm{curl} \boldsymbol{\theta}\|_{L^2(\omega)} \quad \text{and} \quad \int_\omega \boldsymbol{\beta} \, dx = 0.$$

Proof. Let $q := \mathrm{curl}\boldsymbol{\theta}$ and $\tilde{\omega}$ be a convex domain such that $\omega \subset \tilde{\omega}$. We define $\tilde{q}$ as a trivial extension of q by zero, i.e. $\tilde{q}|_{\omega} = q$ and $\tilde{q}|_{\tilde{\omega}\backslash\omega} = 0$. In the next step we seek the solution $w \in H^1(\tilde{\omega})$ of the Poisson problem $\Delta w = \mathrm{curlcurl}w = \tilde{q}$ on $\tilde{\omega}$. Using a regularity estimate for the Poisson problem on the convex domain $\tilde{\omega}$, it follows for $\tilde{\boldsymbol{\beta}} := \mathrm{curl}w$ and $\boldsymbol{\beta} := \tilde{\boldsymbol{\beta}}|_{\omega} - \int_{\omega} \tilde{\boldsymbol{\beta}}\ dx\ /|\omega|$ that

$$||\nabla\boldsymbol{\beta}||_{L^2(\omega)} \lesssim ||\nabla\tilde{\boldsymbol{\beta}}||_{L^2(\tilde{\omega})} \lesssim ||w||_{H^2(\tilde{\omega})} \lesssim ||\tilde{q}||_{L^2(\tilde{\omega})} = ||q||_{L^2(\omega)} = ||\mathrm{curl}\boldsymbol{\theta}||_{L^2(\omega)}.$$

Since $\mathrm{curl}(\boldsymbol{\theta} - \boldsymbol{\beta}) = 0$ in ω, its exists a vector potential $\alpha \in H^2(\omega)$ such that $\nabla\alpha = \boldsymbol{\theta} - \boldsymbol{\beta}$. This concludes the proof. $\qquad\square$

Theorem 6.3 (Proof of Assumption 5.1 for finite element methods with P_0 pressure space). *If the reconstruction operator Π and the Fortin operator I satisfy* (16) *and*

$$(21) \qquad \int_E (1 - I)\boldsymbol{v} \cdot \vec{n}_E\ ds = \int_E (1 - \Pi I)\boldsymbol{v} \cdot \vec{n}_E\ ds = 0 \quad \text{for all } E \in \mathcal{E},$$

also Assumption 5.1 is satisfied.

Remark 6.4. *Condition* (21) *is satisfied for the Forint interpolators for the $P_2 \times P_0$, $P_3 \times P_0$ and the Bernardi–Raugel finite element methods [8, Section 8.4.3]. For these methods the reconstruction operator Π is the standard interpolation into the space BDM_1 or RT_0 [25].*

Proof. Since every function $\boldsymbol{g} \in H^1(T)$ with $\int_E \boldsymbol{g} \cdot \boldsymbol{n}\ ds = 0$ along all edges $E \in \mathcal{E}(T)$ of T satisfies a discrete Friedrichs inequality $\|\boldsymbol{g}\|_{L^2(T)} \lesssim h_T\|\nabla\boldsymbol{g}\|_{L^2(T)}$, see e.g. [10], it follows together with (16)

$$\|(1 - \Pi I)\boldsymbol{v}\|_{L^2(T)} \leqslant \|(1 - I)\boldsymbol{v}\|_{L^2(T)} + \|(1 - \Pi)(I\boldsymbol{v})\|_{L^2(T)}$$
$$\lesssim h_T\|\nabla\boldsymbol{v}\|_{L^2(T)} + h_T\|\nabla I\boldsymbol{v}\|_{L^2(T)} \lesssim h_T\|\nabla\boldsymbol{v}\|_{L^2(T)}.$$

Since $(1 - \Pi I)\boldsymbol{v}$ is divergence-free, it holds $(1 - \Pi I)\boldsymbol{v} = \mathrm{curl}\psi$ for some $\psi \in H_0^1(\Omega) \cap H^2(\Omega)$, see Corollary 3.2 in [16]. Condition (21) implies that the standard interpolator into RT_0 vanishes, i.e. $I_{RT_0}\mathrm{curl}\psi = 0$. Moreover, by the commuting properties (20) of the de Rham complex, it also holds $\mathrm{curl}(I_{\mathcal{L}}\psi) = I_{RT_0}\mathrm{curl}\psi = 0$. An integration by parts and standard interpolation estimates yield

$$\int_{\Omega} \boldsymbol{\theta} \cdot (1 - \Pi I)\boldsymbol{v}\ dx = \int_{\Omega} \boldsymbol{\theta} \cdot \mathrm{curl}(\psi - I_{\mathcal{L}}\psi)\ dx = \int_{\Omega} \mathrm{curl}\boldsymbol{\theta} \cdot (\psi - I_{\mathcal{L}}\psi)\ dx$$
$$\leqslant \|h_{\mathcal{T}}^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2}\|h_{\mathcal{T}}^{-2}(\psi - I_{\mathcal{L}}\psi)\|_{L^2}$$
$$\leqslant \|h_{\mathcal{T}}^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2}\|h_{\mathcal{T}}^{-1}\nabla(\psi - I_{\mathcal{L}}\psi)\|_{L^2}$$
$$= \|h_{\mathcal{T}}^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2}\|h_{\mathcal{T}}^{-1}\mathrm{curl}(\psi)\|_{L^2}$$
$$= \|h_{\mathcal{T}}^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2}\|h_{\mathcal{T}}^{-1}(1 - \Pi I)\boldsymbol{v}\|_{L^2} \leqslant \|h_{\mathcal{T}}^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2}\|\nabla\boldsymbol{v}\|_{L^2},$$

where we used that the curl is just the rotated gradient in two dimensions. This concludes the proof. $\qquad\square$

Theorem 6.5 (Proof of Assumption 5.1 for finite element methods with discontinuous P_1 pressure space). *If the reconstruction operator Π and the Fortin operator I satisfy*

$$(22) \qquad \int_T (1 - I)\boldsymbol{v}\ dx = \int_T (1 - \Pi I)\boldsymbol{v}\ dx = 0 \quad \text{for all } T \in \mathcal{T},$$

also Assumption 5.1 is satisfied.

Remark 6.6. *Condition* (22) *is satisfied by the P_2-bubble finite element method and its Fortin interpolator [8, Section 8.6.2]. A suitable reconstruction operator Π is the standard interpolation into the space BDM_2 or RT_1 [22, 25]. Moreover, the result generalises to all $P_k \times P_{k-2}$ finite element methods with $k > 2$.*

Proof. A triangle inequality, interpolation properties of Π, a Poincaré inequality, and the H^1-stability of I show

$$\|(1 - \Pi I)\boldsymbol{v}\|_{L^2(T)} \leqslant \|(1 - \Pi)I\boldsymbol{v}\|_{L^2(T)} + \|(1 - I)\boldsymbol{v}\|_{L^2(T)} \lesssim h_T\|\nabla\boldsymbol{v}\|_{L^2(T)}.$$

To estimate the dual norm, Lemma 6.1 yields $\boldsymbol{\theta} - \Pi_{\mathcal{N}_0}\boldsymbol{\theta} = \nabla\phi + \boldsymbol{y}$ with

$$\|h_{\mathcal{T}}\boldsymbol{y}\|_{L^2} \lesssim \|h_{\mathcal{T}}^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2}.$$

Also note that due to $\Pi_{\mathcal{N}_0}\boldsymbol{\theta} \in H(\mathrm{curl}, T)$ we can use the regular decomposition from Lemma 6.2 to find

$$\Pi_{\mathcal{N}_0}\boldsymbol{\theta}|_T = \nabla\alpha_T + \boldsymbol{\beta}_T \quad \text{for all } T \in \mathcal{T}$$

with some $\alpha_T \in H^2(T)$ and $\boldsymbol{\beta}_T \in [H^1(T)]^2$ such that $\int_T \boldsymbol{\beta}_T \ dx = 0$ and

$$\|\nabla\boldsymbol{\beta}_T\|_{L^2(T)} \lesssim \|\mathrm{curl}(\Pi_{\mathcal{N}_0}\boldsymbol{\theta})\|_{L^2(T)} \lesssim \|\mathrm{curl}\boldsymbol{\theta}\|_{L^2(T)}.$$

Together with the projection property of $I_{\mathcal{N}_0}$, the commuting properties (20) of the de Rham complex and the continuity of the nodal interpolation $I_{\mathcal{L}}$ for H^2 functions, the Helmholtz decomposition can be cast into the discrete version

$$\Pi_{\mathcal{N}_0}\boldsymbol{\theta}|_T = I_{\mathcal{N}_0}(\nabla\alpha_T + \boldsymbol{\beta}_T) = \nabla(I_{\mathcal{L}}\alpha_T) + I_{\mathcal{N}_0}\boldsymbol{\beta}_T.$$

The combination of all decompositions defines some function $\alpha_{\mathcal{T}} \in P_1(\mathcal{T})$ and $\beta_{\mathcal{T}} \in P_1(\mathcal{T})^2$ with

$$\|h_{\mathcal{T}}^2\nabla_h\boldsymbol{\beta}_{\mathcal{T}}\|_{L^2} \lesssim \|h_{\mathcal{T}}^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2}.$$

Since $\boldsymbol{z} := (1-\Pi I)\boldsymbol{v}$ is orthogonal onto piecewise constants (by (22)), in particular the piecewise constant function $\nabla(I_{\mathcal{L}}\alpha)_{\mathcal{T}} \in P_0(\mathcal{T})^2$, and gradients (because $\boldsymbol{z}$ is divergence-free and has zero boundary data), it follows

$$\int_{\Omega} \boldsymbol{\theta} \cdot (1 - \Pi I)\boldsymbol{v} \ dx = \int_{\Omega} \boldsymbol{z} \cdot \boldsymbol{\theta} \ dx = \int_{\Omega} \boldsymbol{z} \cdot (\boldsymbol{\theta} - \Pi_{ND_0}\boldsymbol{\theta}) \ dx + \int_{\Omega} \boldsymbol{z} \cdot \Pi_{ND_0}\boldsymbol{\theta} \ dx$$

$$= \int_{\Omega} \boldsymbol{z} \cdot \boldsymbol{y} \ dx + \int_{\Omega} \boldsymbol{z} \cdot \boldsymbol{\beta}_{\mathcal{T}} \ dx$$

$$= \int_{\Omega} h_{\mathcal{T}}^{-1}\boldsymbol{z} \cdot h_{\mathcal{T}}\boldsymbol{y} \ dx + \int_{\Omega} h_{\mathcal{T}}^{-1}\boldsymbol{z} \cdot h_{\mathcal{T}}\boldsymbol{\beta}_{\mathcal{T}} \ dx$$

$$\lesssim \|h_{\mathcal{T}}^{-1}\boldsymbol{z}\|_{L^2}(\|h_{\mathcal{T}}\boldsymbol{y}\|_{L^2} + \|h_{\mathcal{T}}^2\nabla_h\boldsymbol{\beta}_{\mathcal{T}}\|_{L^2})$$

$$\lesssim \|h_{\mathcal{T}}^{-1}\boldsymbol{z}\|_{L^2}\|h_{\mathcal{T}}^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2} \lesssim \|\nabla\boldsymbol{v}\|_{L^2}\|h_{\mathcal{T}}^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2}.$$

Note, that we used an elementwise Poincaré inequality for $\boldsymbol{\beta}_{\mathcal{T}}$ (which has piecewise integral mean zero). This concludes the proof. □

Theorem 6.7 (Proof of Assumption 5.1 for the mini finite element method). *The mini finite element method family with the reconstruction operator from [20] and a Fortin operator I with the property (see e.g. [8, Section 8.4.2])*

(23)
$$\int_T (1 - I)\boldsymbol{v} \ ds = 0 \quad \text{for all } T \in \mathcal{T}$$

satisfies Assumption 5.1.

Proof. For the mini finite element method, the reconstruction operator is given in [20]. It in particular satisfies (16) in the sense

(24)
$$(1 - \Pi)I\boldsymbol{v} = \sum_{y \in \mathcal{N}} \boldsymbol{\sigma}_y$$

where $\boldsymbol{\sigma}_y \in BDM_2(\mathcal{T}(\omega_y))$ satisfies $\|\boldsymbol{\sigma}_y\|_{L^2(\omega_y)} \lesssim h_y\|\nabla I\boldsymbol{v}\|_{L^2(\omega_y)}$ on the nodal patch ω_y of the node $y \in \mathcal{N}$ and (at least) the local orthogonality

$$\int_{\omega_y} \boldsymbol{\sigma}_y \ dx = 0.$$

Furthermore we have $\boldsymbol{\sigma}_y \cdot n = 0$ on the boundary $\partial \omega_y$. This time, the operators I and Π do not share the same orthogonality on cell-wise constants as in Theorem 6.5, but one can split up the L^2-norm by a triangle inequality

$$\|(1 - \Pi I)\boldsymbol{v}\|_{L^2} \leqslant \|(1 - \Pi)I\boldsymbol{v}\|_{L^2} + \|(1 - I)\boldsymbol{v}\|_{L^2}.$$

Due to (23) the norm $\|(1 - I)\boldsymbol{v}\|_{L^2(T)}$ can be estimated as in Theorem 6.5 and it remains to estimate $\|(1 - \Pi)I\boldsymbol{v}\|_{L^2}$. For the first one, it holds

$$\begin{aligned}
\|(1 - \Pi)I\boldsymbol{v}\|_{L^2(T)}^2 &= \sum_{z \in \mathcal{N}(T)} \int_T \boldsymbol{\sigma}_z \cdot (1 - \Pi)I\boldsymbol{v} \; dx \\
&\leqslant \sum_{z \in \mathcal{N}(T)} \|\boldsymbol{\sigma}_z\|_{L^2(\omega_z)} \|(1 - \Pi)I\boldsymbol{v}\|_{L^2(T)} \\
&\leqslant h_T \|\nabla I\boldsymbol{v}\|_{L^2(\omega_T)} \|(1 - \Pi)I\boldsymbol{v}_h\|_{L^2(T)}
\end{aligned}$$

and hence

$$\|(1 - \Pi)I\boldsymbol{v}\|_{L^2(T)} \lesssim h_T \|\nabla I\boldsymbol{v}\|_{L^2(\omega_T)} \lesssim h_T \|\nabla \boldsymbol{v}\|_{L^2(\omega_T)}.$$

For the estimate of the dual norm, inserting the decomposition from Lemma 6.1 leads to

$$\int_\Omega \boldsymbol{\theta} \cdot (1 - \Pi I)\boldsymbol{v} \; dx = \int_\Omega \boldsymbol{z} \cdot \boldsymbol{\theta} \; dx = \int_\Omega \boldsymbol{z} \cdot \boldsymbol{y} \; dx + \int_\Omega \boldsymbol{z} \cdot \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx.$$

The first integral can be estimated as in Theorem 6.5 and it remains to estimate the second integral where we employ the decomposition (24) for $(1 - \Pi)I\boldsymbol{v} = \sum_{y \in \mathcal{N}} \boldsymbol{\sigma}_y$ and its orthogonality properties, i.e.

$$(25) \qquad \int_\Omega \boldsymbol{z} \cdot \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx = \int_\Omega (1 - I)\boldsymbol{v} \cdot \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx + \sum_{y \in \mathcal{N}} \int_{\omega_y} h_y^{-1}\boldsymbol{\sigma}_y \cdot h_y \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx$$

and we bound both integrals separately. The first integral of (25) can be estimated exactly as in Theorem 6.5 due to (23) by a element-wise Helmholtz decomposition such that

$$\int_\Omega (1 - I)\boldsymbol{v} \cdot \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx \lesssim \|\nabla \boldsymbol{v}\|_{L^2} \|h_{\mathcal{T}}^2 \mathrm{curl}\boldsymbol{\theta}\|_{L^2}.$$

For the second integral, first note that due to $\Pi_{\mathcal{N}_0}\boldsymbol{\theta} \in H(\mathrm{curl}, \omega_y)$ we can use the regular decomposition of Lemma 6.2 on each patch to get

$$\Pi_{\mathcal{N}_0}\boldsymbol{\theta}|_{\omega_y} = \nabla \alpha_y + \boldsymbol{\beta}_y \quad \text{for all } y \in \mathcal{N}.$$

with some $\alpha_y \in H^1(\omega_y)$ and $\boldsymbol{\beta}_y \in [H^1(\omega_y)]^2$ such that $\int_{\omega_y} \boldsymbol{\beta}_y \; dx = 0$ and

$$\|\boldsymbol{\beta}_y\|_{H^1(\omega_y)} \lesssim \|\mathrm{curl}(\Pi_{ND_0}\boldsymbol{\theta})\|_{L^2(\omega_y)} \lesssim \|\mathrm{curl}\boldsymbol{\theta}\|_{L^2(\omega_y)}.$$

Next note, that on each element $T \subset \omega_y$ we have $\Pi_{ND_0}\boldsymbol{\theta}|_T \in [H^1(T)]^2$ and thus

$$\nabla \alpha_y|_T = \Pi_{ND_0}\boldsymbol{\theta}|_T - \boldsymbol{\beta}_y|_T \in [H^1(T)]^2 \quad \Rightarrow \quad \alpha_y|_T \in H^2(T).$$

Together with the projection property of $I_{\mathcal{N}_0}$, the commuting properties (20) of the de Rham complex and the continuity of the nodal interpolation $I_{\mathcal{N}}$ for H^2 functions, the Helmholtz decomposition can be cast into the discrete version

$$\Pi_{ND_0}\boldsymbol{\theta}|_{\omega_y} = I_{\mathcal{N}_0}(\nabla \alpha_y + \boldsymbol{\beta}_y) = \nabla(I_{\mathcal{L}}\alpha_y) + I_{\mathcal{N}_0}\boldsymbol{\beta}_y.$$

Finally, a scaling argument and a Poincaré inequality shows

$$\|I_{\mathcal{N}_0}\boldsymbol{\beta}_y\|_{L^2(\omega_y)} \lesssim \|\boldsymbol{\beta}_y\|_{L^2(\omega_y)} + h_y \|\nabla \boldsymbol{\beta}_y\|_{L^2(\omega_y)} \lesssim h_y \|\nabla \boldsymbol{\beta}_y\|_{L^2(\omega_y)} \lesssim h_y \|\mathrm{curl}\boldsymbol{\theta}\|_{L^2(\omega_y)}.$$

Furthermore, note that the reconstruction operator is orthogonal on gradients of continuous P_1-functions like $\nabla(I_{\mathcal{L}}\alpha_y)$ due to [20, Proposition 16.ii], i.e. $\int_{\omega_y} \sigma_y \cdot \nabla(I_{\mathcal{L}}\alpha_y) \; dx = 0$. Now, the second integral of (25) is bounded by

$$\sum_{y \in \mathcal{N}} \int_{\omega_y} h_y^{-1}\sigma_y \cdot h_y\Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx \lesssim \sum_{y \in \mathcal{N}} \|h_y^{-1}\sigma_y\|_{L^2(\omega_y)}\|h_y^2\nabla\boldsymbol{\beta}_y\|_{L^2(\omega_y)}$$
$$\lesssim \sum_{y \in \mathcal{N}} \|\nabla\boldsymbol{v}\|_{L^2}\|h_y^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2(\omega_y)} \lesssim \|\nabla\boldsymbol{v}\|_{L^2}\|h_{\mathcal{T}}^2\mathrm{curl}\boldsymbol{\theta}\|_{L^2}.$$

The combination of all previous results concludes the proof. $\qquad\square$

Theorem 6.8 (Proof of Assumption 5.1 for the Taylor–Hood finite element method)**.** *The Taylor–Hood finite element method family with the reconstruction operator from* [20] *and the Fortin operator I from* [26, 12] *in two dimensions with the property*

$$(26) \qquad\qquad \int_\Omega (1 - I)\boldsymbol{v} \cdot \boldsymbol{w} \; ds = 0 \quad \text{for all } \boldsymbol{w} \in \widetilde{\mathcal{N}_0}(\mathcal{T}),$$

where $\widetilde{\mathcal{N}_0}(\mathcal{T})$ is a subset of $\mathcal{N}_0(\mathcal{T})$ as defined in [26, 12], *satisfy Assumption 5.1.*

Remark 6.9. *The proof requires some assumption on the mesh, i.e. we require that each interior face $E \in E^\circ$ has at most one node on the boundary $\partial\Omega$. This assumption was also needed in* [26] *for the construction of a stable the Fortin interpolator and was later removed in* [12]. *Maybe similar arguments can be used in our case.*

Proof. A triangle inequality, properties of Π, and the H^1-stability of I show

$$\|(1 - \Pi I)\boldsymbol{v}\|_{L^2(T)} \lesssim \|(1 - \Pi)I\boldsymbol{v}\|_{L^2(T)} + \|(1 - I)\boldsymbol{v}\|_{L^2(T)} \lesssim h_T\|\nabla\boldsymbol{v}\|_{L^2(T)}.$$

Again using the decomposition from Lemma 6.1 and the orthogonality between gradients and $(1 - \Pi I)\boldsymbol{v}$ leads to

$$\int_\Omega (1 - \Pi I)\boldsymbol{v} \cdot \boldsymbol{\theta} \; dx = \int_\Omega (1 - \Pi I)\boldsymbol{v} \cdot \boldsymbol{y} \; dx + \int_\Omega (1 - \Pi I)\boldsymbol{v} \cdot \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx$$

The first integral can be estimated similarly as in the proof of Theorem 6.7. For the second integral we use $(1 - \Pi)I\boldsymbol{v} = \sum_{y \in \mathcal{N}} \boldsymbol{\sigma}_y$ to get

$$\int_\Omega (1 - \Pi I)\boldsymbol{v} \cdot \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx = \int_\Omega (1 - I)\boldsymbol{v} \cdot \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx + \sum_{y \in \mathcal{N}} \int_{\omega_y} h_y^{-1}\boldsymbol{\sigma}_y \cdot h_y\Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx \, .$$

Similarly as in the proof of Theorem 6.7 we bound the first term. However the integral (using the orthogonality (26))

$$\int_\Omega (1 - I)\boldsymbol{v} \cdot \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx = \int_\Omega (1 - I)\boldsymbol{v} \cdot (1 - I_{\widetilde{\mathcal{N}_0}})\Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx$$

needs a different treatment. To estimate this integral we have to design a proper interpolation $I_{\widetilde{\mathcal{N}_0}}(\Pi_{\mathcal{N}_0}\boldsymbol{\theta})$ of $\Pi_{\mathcal{N}_0}\boldsymbol{\theta}$ into the space $\widetilde{\mathcal{N}_0}(\mathcal{T})$. To do so, we can write $\Pi_{\mathcal{N}_0}\boldsymbol{\theta}$ as a linear combination

$$\Pi_{\mathcal{N}_0}\boldsymbol{\theta} = \sum_{E \in \mathcal{E}} \alpha_E N_E \quad \text{with coefficients } \alpha_E := \int_E \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \cdot \boldsymbol{\tau}_E \; ds$$

and Nédélec basis functions N_E with $\int_F N_E\boldsymbol{\tau}_F \; ds = \delta_{EF}$ for $E, F \in \mathcal{E}$. Then, we choose $I_{\widetilde{\mathcal{N}_0}}(\Pi_{\mathcal{N}_0}\boldsymbol{\theta})$ as

$$I_{\widetilde{\mathcal{N}_0}}(\Pi_{\mathcal{N}_0}\boldsymbol{\theta}) := \sum_{E \in \mathcal{E}^0} \alpha_E \widetilde{N}_E$$

where $\mathcal{E}^0$ are the interior edges and $\widetilde{N}_E$ are the modified basis functions as in [12], i.e. $\widetilde{N}_E = N_E$ for all edges E with two interior endpoints and $\widetilde{N}_E = N_E \pm N_F$ for interior edges E with one boundary endpoint and F is a boundary edge with the same boundary endpoint and in the same triangle of E. The sign depends on the orientation of the tangent vectors. Assume a boundary triangle T_E with
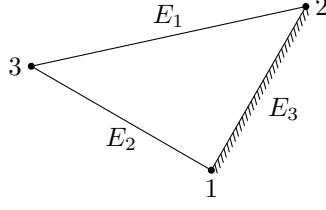
FIGURE 1. Enumeration of the vertices and edges in a boundary triangle with boundary edge E_3.

nodes $1, 2, 3$, boundary edge $E_3 = \text{conv}\{1, 2\}$ and two adjacent interior edges E_1 and E_2 as depicted in Figure 1. We further assume, that the tangential vectors are pointing from the lower to the larger node number. Then, according to [12], the modified basis functions read $\widetilde{N}_{E_2} = N_{E_2} + N_{E_3}$ and $\widetilde{N}_{E_1} = N_{E_1} - N_{E_3}$. Hence, locally on T, we have

$$\left((1 - I_{\widetilde{\mathcal{N}_0}})\Pi_{\mathcal{N}_0}\boldsymbol{\theta}\right)|_T = \alpha_{E_1}N_{E_1} + \alpha_{E_2}N_{E_2} + \alpha_{E_3}N_{E_3} - (\alpha_{E_1}\widetilde{N}_{E_1} + \alpha_{E_2}\widetilde{N}_{E_2})$$
$$= N_{E_3}(\alpha_{E_3} + \alpha_{E_1} - \alpha_{E_2}).$$

The definition of α_{E_j} and an easy calculation plus the Stokes theorem show

$$\alpha_{E_3} + \alpha_{E_1} - \alpha_{E_2} = \int_{\partial T} \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \cdot \boldsymbol{\tau} \; ds = \int_T \text{curl}(\Pi_{\mathcal{N}_0}\boldsymbol{\theta}) \; dx$$

and hence the estimate

$$\|(1 - I_{\widetilde{\mathcal{N}_0}})\Pi_{\mathcal{N}_0}\boldsymbol{\theta}\|_{L^2(T)} \leqslant \left|\int_T \text{curl}(\Pi_{\mathcal{N}_0}\boldsymbol{\theta}) \; dx\right| \|N_E\|_{L^2(T)} \lesssim h_T\|\text{curl}(\Pi_{\mathcal{N}_0}\boldsymbol{\theta})\|_{L^2(T)}.$$

On interior triangles, it holds $\Pi_{\mathcal{N}_0}\boldsymbol{\theta} - I_{\widetilde{\mathcal{N}_0}}(\Pi_{\mathcal{N}_0}\boldsymbol{\theta}) = 0$ and hence

$$\int_\Omega (1 - I)\boldsymbol{v} \cdot \Pi_{\mathcal{N}_0}\boldsymbol{\theta} \; dx = \int_\Omega (1 - I)\boldsymbol{v} \cdot (\Pi_{\mathcal{N}_0}\boldsymbol{\theta} - I_{\widetilde{\mathcal{N}_0}}(\Pi_{\mathcal{N}_0}\boldsymbol{\theta})) \; dx$$
$$\lesssim \sum_{T \in \mathcal{T}(\partial\Omega)} h_T^2\|\text{curl}(\Pi_{\mathcal{N}_0}\boldsymbol{\theta})\|_{L^2(T)}\|\nabla\boldsymbol{v}\|_{L^2(\omega_T)} \lesssim \|h_\mathcal{T}^2\text{curl}\boldsymbol{\theta}\|_{L^2}\|\nabla\boldsymbol{v}\|_{L^2}.$$

This concludes the proof.

$\square$

7. NUMERICAL EXPERIMENTS

In the following two numerical examples, the novel error estimator

$$\mu_{\text{new}}^2 := \nu^{-2}\eta_{\text{new}}(\nabla\boldsymbol{u}_h)^2 + \|\text{div}\boldsymbol{u}_h\|_{L^2}^2$$

from Theorem 5.2 (for pressure-robust methods) or Proposition 5.3 (for classical methods) is compared to the classical error estimator

$$\mu_{\text{class}}^2 := \nu^{-2}\eta_{\text{class}}(\nabla u_h, p_h)^2 + \|\text{div}\boldsymbol{u}_h\|_{L^2}^2$$

from Theorem 4.1, with respect to the H^1-seminorm $\text{err}_{H^1}(u_h) := \|\nabla u - \nabla u_h\|_{L^2}$. Our adaptive mesh refinement algorithm follows the loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE} \rightarrow \text{SOLVE} \rightarrow \ldots$$

and employs the local contributions to the error estimator as element-wise refinement indicators. In the marking step, an element $T \in \mathcal{T}$ is marked for refinement if $\mu(T) \geqslant \frac{1}{4}\max_{K \in \mathcal{T}}\mu(K)$. The refinement step refines all marked elements plus further elements in a closure step to guarantee a regular triangulation. The implementation and numerical examples where performed with NGSolve/Netgen [30], [29].

TABLE 3. The H^1 error and the old and new error estimators including the efficiency for the example of section 7.1 for varying ν using the classical Taylor Hood element TH$_2$ and its pressure robust modification.

ν	$\mathrm{err}_{H^1}(u_h)$	(classical) μ_{class}	$\frac{\mu_{\mathrm{class}}}{\mathrm{err}_{H^1}(u_h)}$	$\mathrm{err}_{H^1}(u_h)$	(p-robust) μ_{new}	$\frac{\mu_{\mathrm{new}}}{\mathrm{err}_{H^1}(u_h)}$
10^1	$1.27\cdot 10^{-3}$	$1.99\cdot 10^{-2}$	$1.58\cdot 10^1$	$1.30\cdot 10^{-3}$	$5.19\cdot 10^{-2}$	$3.98\cdot 10^1$
10^0	$1.30\cdot 10^{-3}$	$1.42\cdot 10^{-2}$	$1.09\cdot 10^1$	$1.30\cdot 10^{-3}$	$3.47\cdot 10^{-2}$	$2.66\cdot 10^1$
10^{-1}	$3.12\cdot 10^{-3}$	$1.12\cdot 10^{-1}$	$3.58\cdot 10^1$	$1.30\cdot 10^{-3}$	$3.29\cdot 10^{-2}$	$2.53\cdot 10^1$
10^{-2}	$2.85\cdot 10^{-2}$	1.11	$3.90\cdot 10^1$	$1.30\cdot 10^{-3}$	$3.28\cdot 10^{-2}$	$2.51\cdot 10^1$
10^{-3}	$2.85\cdot 10^{-1}$	$1.11\cdot 10^1$	$3.90\cdot 10^1$	$1.30\cdot 10^{-3}$	$3.27\cdot 10^{-2}$	$2.51\cdot 10^1$
10^{-4}	2.85	$1.11\cdot 10^2$	$3.90\cdot 10^1$	$1.30\cdot 10^{-3}$	$3.27\cdot 10^{-2}$	$2.51\cdot 10^1$
10^{-5}	$2.85\cdot 10^1$	$1.11\cdot 10^3$	$3.90\cdot 10^1$	$1.30\cdot 10^{-3}$	$3.27\cdot 10^{-2}$	$2.51\cdot 10^1$
10^{-6}	$2.85\cdot 10^2$	$1.11\cdot 10^4$	$3.90\cdot 10^1$	$1.30\cdot 10^{-3}$	$3.27\cdot 10^{-2}$	$2.51\cdot 10^1$

Remark 7.1. *For reducing the costs of the estimator, we estimated the consistency error $\eta_{cons,1}(\nabla u_h) = \|\nu\mathrm{div}_h(\nabla u_h) \circ (1-\Pi)\|_{V_{0,h}^\star}$ according to* (16) *by*

$$\eta_{cons,1}(\nabla u_h) \lesssim \nu \left(\sum_{K\in\mathcal{K}} h_K^2 \|(1-\pi_{P_{q-1}(K)})\Delta_h u_h\|_{L^2(K)}^2 \right)^{1/2}.$$

7.1. **Smooth example on unit square.** This example concerns the Stokes problem for

$$\boldsymbol{u}(x,y) := \mathrm{curl}\left(x^2(x-1)^2 y^2(y-1)^2\right) \quad \text{and} \quad p(x,y) := x^5 + y^5 - 1/3$$

on the unit square $\Omega := (0,1)^2$ with matching right-hand side $\boldsymbol{f} := -\nu\Delta\boldsymbol{u} + \nabla p$ for variable viscosity ν.

Table 3 lists the error of the classical and pressure-robust Taylor-Hood finite element methods with their error estimators μ_{class} and μ_{new} on a fixed mesh with 1139 degrees of freedom but varying viscosities $\nu \in (10^{-6}, 10)$. As expected by the a priori error estimates of Theorems 3.1 and Theorem 3.2, the error of the classical solution scales with ν^{-1}, while the error of the pressure-robust method is ν-invariant. Another observation is that both error estimators are efficient for their respective discrete solution.

Figure 2 compares the errors and error estimators of the Taylor–Hood finite element method of order 2 and the MINI finite element method with and without the pressure robust modification for uniform mesh refinement as in the case $\nu = 1$ and a pressure-dominant case with $\nu = 10^{-3}$.

In the pressure dominant case $\nu = 10^{-3}$ the right hand side $\boldsymbol{f}$ tends to have a large irrotational part. The left plot of Figure 2 confirms once again that the velocity error scales with $1/\nu$ and that pressure-robust methods result in much more accurate solutions. For the classical methods both estimators μ_{new} and μ_{class} are efficient, i.e. have comparable overestimation factors and the same optimal convergence order as the velocity error. In case of the MINI finite element method, all quantities even converge quadratically. This is due to the dominance of the pressure error and the higher approximation order of the pressure. In this sense, we are in a pre-asymptotic range and the error will convergence linearly as soon as the ν^{-3}-weighted pressure error is of same magnitude (as it is the case for $\nu = 1$ from the very beginning). Also for the classical MINI element μ_{new} and μ_{class} are efficient with a comparable overestimation factor.

For the pressure-robust methods we observe that for both elements the novel estimator μ_{new} is much smaller than μ_{class}. To be more precise, it scales with $\mu_{\mathrm{new}} \approx 1/\nu\, \mu_{\mathrm{class}}$ in case of the Taylor-Hood method as expected by the theory. This is again due to the discrete pressure that is used in μ_{class} (p_h replaced by some better approximation of p would reduce the gap between μ_{new} and μ_{class}). Hence, μ_{new} is efficient and μ_{class} is not efficient for the pressure-robust Taylor–Hood finite element method. In case of the pressure-robustly modified MINI method, the velocity error and the novel estimator μ_{new} now have the expected optimal linear order of the MINI finite element method. Otherwise, the conclusions are similar to the ones for the Taylor–Hood method.
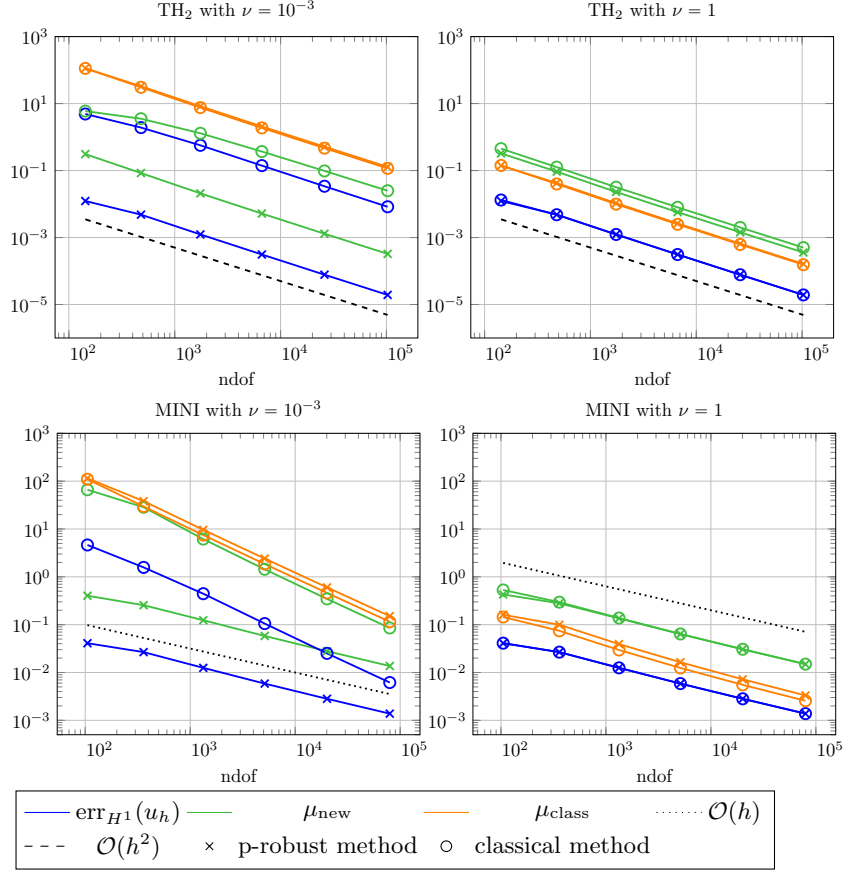
FIGURE 2. The H^1-error, μ_{class} and μ_{new} for the example of section 7.1 with $\nu = 1$ and $\nu = 10^{-3}$. At the top the TH$_2$ element, and at the bottom the MINI element.

In this case $\nu = 1$ the irrotational part and the rotational part of the right hand side $\boldsymbol{f}$ have the same magnitude, thus the pressure error has not such a big impact on the accuracy of the discrete velocity. Accordingly, there is only little to no improvement by the application of the pressure-robust modification. Thus, in the right plots of Figure 2 we can see that the velocity error of both methods, the pressure robust and the classical one, is of the same magnitude and order. Both estimators are efficient with slightly less overestimation by η_{class}.

7.2. **L-shape example.** This example studies a velocity $\boldsymbol{u}$ and a pressure p_0 on the L-shaped domain $\Omega := (-1, 1)^2 \setminus ((0, 1) \times (-1, 0))$ taken from [34] that satisfy $-\nu \Delta \boldsymbol{u} + \nabla p_0 = 0$. The fields are defined in polar coordinates and read

$$\boldsymbol{u}(r, \varphi) := r^\alpha \begin{pmatrix} (\alpha + 1) \sin(\varphi) \psi(\varphi) + \cos(\varphi) \psi'(\varphi) \\ -(\alpha + 1) \cos(\varphi) \psi(\varphi) + \sin(\varphi) \psi'(\varphi) \end{pmatrix}^T,$$
$$p_0 := \nu^{-1} r^{(\alpha-1)} ((1 + \alpha)^2 \psi'(\varphi) + \psi'''(\varphi))/(1 - \alpha)$$

where

$$\psi(\varphi) := 1/(\alpha + 1) \sin((\alpha + 1)\varphi) \cos(\alpha \omega) - \cos((\alpha + 1)\varphi)$$
$$- 1/(\alpha - 1) \sin((\alpha - 1)\varphi) \cos(\alpha \omega) + \cos((\alpha - 1)\varphi)$$

and $\alpha = 856399/1572864 \approx 0.54$, $\omega = 3\pi/2$. To have a nonzero right-hand side we add the pressure $p_+ := \sin(xy\pi)$, i.e. $p := p_0 + p_+$ and $\boldsymbol{f} := \nabla(p_+)$. We generate a pressure dominant case by using a small viscosity $\nu = 10^{-3}$. In Figure 3 and 4 the velocity error and the novel estimator η_{new} are
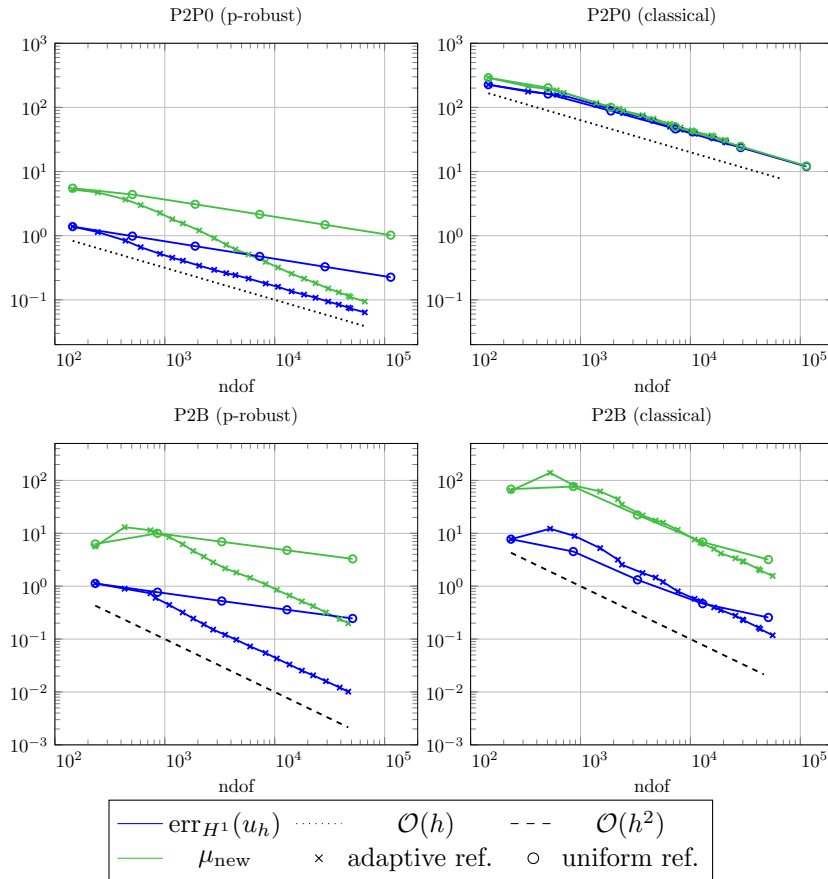
FIGURE 3. Error for L-shape example of section 7.2 using the discontinuous pressure elements P2P0 (top) and the P2B (bottom) with $\nu = 10^{-3}$

plotted for the classical and modified version of four different finite element methods and uniform and adaptive mesh refinement. For this example an adaptive refinement is expected to refine the generic singularity of the velocity in the corner $(0,0)$.

We first discuss the pressure-robust variants of the finite element methods. Looking at the left plots of Figure 3 and 4 we can see that there is a major difference between adaptive and uniform mesh refinement. The adaptive algorithm results in optimal orders of the velocity error and the estimator, while uniform refinement only leads to suboptimal orders as the singularity is not resolved well enough. The only exception is the MINI finite element method which pre-asymptotically converges with quadratic speed. This is again thanks to the better polynomial order in the pressure ansatz space and the smooth pressure p^+. Asymptotically also the MINI finite element method shows the suboptimal behaviour in case of uniform mesh refinement and first-order convergence in case of adaptive mesh refinement. In all cases, the new error estimator μ_{new} is efficient and gives reasonable refinement indicators.

In case of the classical variants of the finite element methods, totally different observations can be made. In the right pictures of Figure 3 and 4 we first note that the error is much larger compared to the pressure-robust method. Furthermore similar as before only adaptive mesh refinement leads to optimal orders. However, it is important to note that the gap between the velocity error of the classical method and the velocity error of the pressure-robust method stays as large as in the beginning also under adaptive mesh refinement. A possible explanation is given by Figure 5 which shows that the classical method refines the mesh almost uniformly. This is reasonable in the sense that the pressure
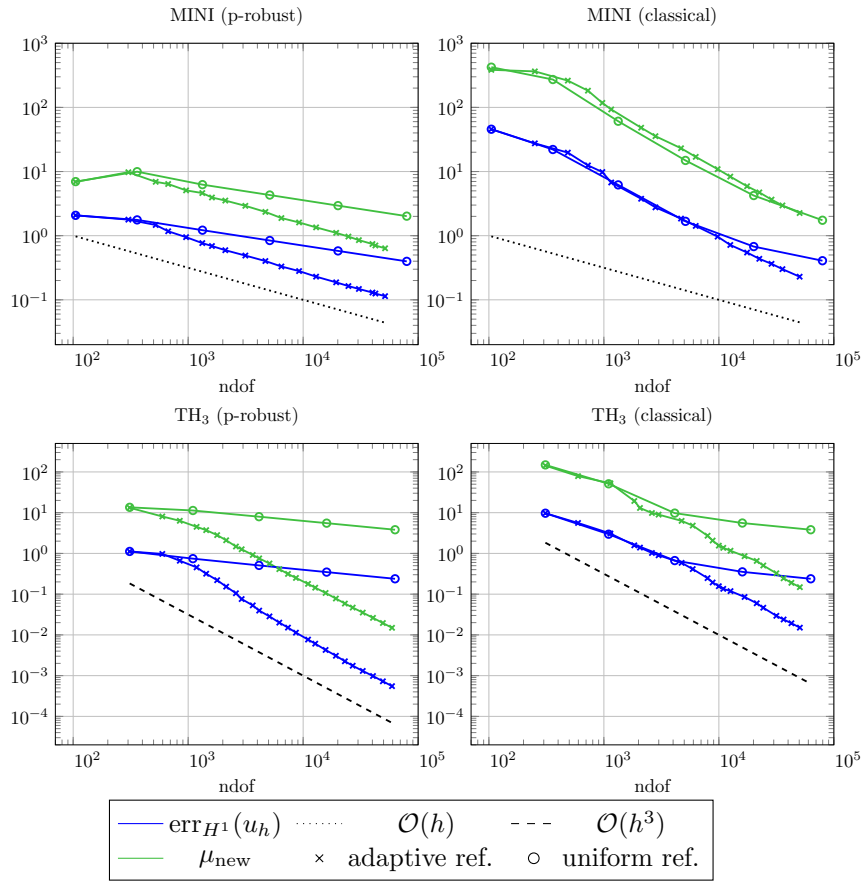
FIGURE 4. Error for L-shape example of section 7.2 using the continuous pressure elements MINI (top) and the TH$_3$ (bottom) with $\nu = 10^{-3}$
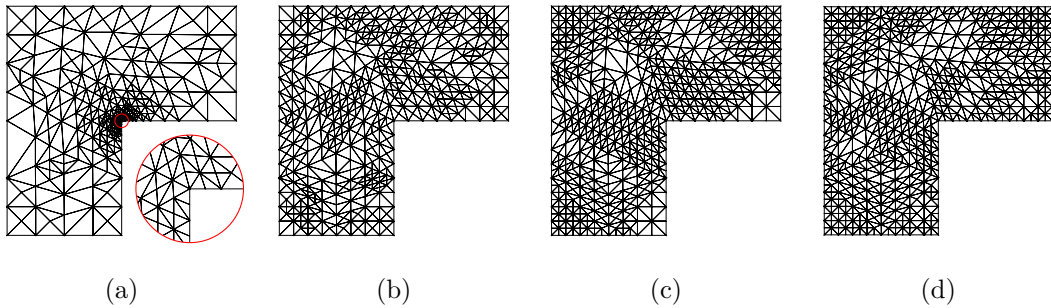

FIGURE 5. (a): according to μ_{new} adaptively refined mesh with 4584 degrees of freedom for the pressure-robust Taylor–Hood method; (b): according to μ_{new} adaptively refined mesh with 3971 degrees of freedom for the classical Taylor–Hood method; (c): according to μ_{class} adaptively refined mesh with 5119 degrees of freedom for the pressure-robust Taylor–Hood method; (d): according to μ_{class} adaptively refined mesh with 5320 degrees of freedom for the classical Taylor–Hood method

error of the smooth pressure p_+ dominates the (real and the estimated) discretisation error in the beginning. The pressure-robust method on the other hand is not polluted by this influence and can concentrate immediately on the corner singularity. However, it is important that the error estimator is also pressure-robust. If the refinement indicators are taken from μ_{class}, the corner singularity remains unrefined until the dominance of the pressure error in the error bound is removed. Hence, the main conclusion is that only a pressure-robust finite element method with a pressure-robust error estimator leads to optimal meshes with the smallest velocity error.

References

[1] N. Ahmed, A. Linke, and C. Merdon. On really locking-free mixed finite element methods for the transient incompressible stokes equations. *WIAS Preprint 2368 (accepted by SINUM)*, 2017.

[2] M. Ainsworth and W. Dörfler. Reliable a posteriori error control for nonconformal finite element approximation of Stokes flow. *Math. Comp.*, 74(252):1599–1619 (electronic), 2005.

[3] M. Amara, M. Ben Younes, and C. Bernardi. Error indicators for the Navier-Stokes equations in stream function and vorticity formulation. *Numer. Math.*, 80(2):181–206, 1998.

[4] C. Amrouche, C. Bernardi, M. Dauge, and V. Girault. Vector potentials in three-dimensional non-smooth domains. *Math. Methods Appl. Sci.*, 21(9):823–864, 1998.

[5] D. N. Arnold, F. Brezzi, and M. Fortin. A stable finite element for the Stokes equations. *Calcolo*, 21(4):337–344 (1985), 1984.

[6] S. Bartels, C. Carstensen, and G. Dolzmann. Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis. *Numer. Math.*, 99(1):1–24, 2004.

[7] Christine Bernardi and Geneviève Raugel. Analysis of some finite elements for the Stokes problem. *Math. Comp.*, 44(169):71–79, 1985.

[8] Daniele Boffi, Franco Brezzi, and Michel Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.

[9] C. Brennecke, A. Linke, C. Merdon, and J. Schöberl. Optimal and pressure-independent L^2 velocity error estimates for a modified Crouzeix-Raviart Stokes element with BDM reconstructions. *J. Comput. Math.*, 33(2):191–208, 2015.

[10] C. Carstensen, J. Gedicke, and D. Rim. Explicit error estimates for Courant, Crouzeix-Raviart and Raviart-Thomas finite element methods. *J. Comput. Math.*, 30(4):337–353, 2012.

[11] Carsten Carstensen and Christian Merdon. Computational Survey on A Posteriori Error Estimators for the Crouzeix–Raviart Nonconforming Finite Element Method for the Stokes Problem. *Comput. Methods Appl. Math.*, 14(1):35–54, 2014.

[12] Long Chen. A simple construction of a Fortin operator for the two dimensional Taylor-Hood element. *Comput. Math. Appl.*, 68(10):1368–1373, 2014.

[13] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 7(R-3):33–75, 1973.

[14] L. Demkowicz, P. Monk, L. Vardapetyan, and W. Rachowicz. de Rham diagram for hp finite element spaces. *Comput. Math. Appl.*, 39(7-8):29–38, 2000.

[15] J.-F. Gerbeau, C. Le Bris, and M. Bercovier. Spurious velocities in the steady flow of an incompressible fluid subjected to external forces. *Internat. J. Numer. Methods Fluids*, 25(6):679–695, 1997.

[16] Vivette Girault and Pierre-Arnaud Raviart. *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986. Theory and algorithms.

[17] Antti Hannukainen, Rolf Stenberg, and Martin Vohralík. A unified framework for a posteriori error estimation for the Stokes problem. *Numer. Math.*, 122(4):725–769, 2012.

[18] P. Hood and C. Taylor. Navier–Stokes equations using mixed interpolation. In J. T. Oden, R. H. Gallagher, O. C. Zienkiewicz, and C. Taylor, editors, *Finite Element Methods in Flow Problems*, pages 121–132. University of Alabama in Huntsville Press, 1974.

[19] Volker John, Alexander Linke, Christian Merdon, Michael Neilan, and Leo G. Rebholz. On the divergence constraint in mixed finite element methods for incompressible flows. *SIAM Review*, 59(3):492–544, 2017.

[20] P. L. Lederer, A. Linke, C. Merdon, and J. Schöberl. Divergence-free Reconstruction Operators for Pressure-Robust Stokes Discretizations With Continuous Pressure Finite Elements. *ArXiv e-prints*, September 2016.

[21] A. Linke. On the role of the Helmholtz decomposition in mixed methods for incompressible flows and a new variational crime. *Comput. Methods Appl. Mech. Engrg.*, 268:782–800, 2014.

[22] A. Linke, G. Matthies, and L. Tobiska. Robust arbitrary order mixed finite element methods for the incompressible stokes equations with pressure independent velocity errors. *ESAIM: M2AN*, 50(1):289–309, 2016.

[23] A. Linke and C. Merdon. Guaranteed energy error estimators for a modified robust Crouzeix-Raviart Stokes element. *J. Sci. Comput.*, 64(2):541–558, 2015.

[24] A. Linke and C. Merdon. On velocity errors due to irrotational forces in the Navier–Stokes momentum balance. *J. Comput. Phys.*, 313:654–661, 2016.

[25] A. Linke and C. Merdon. Pressure-robustness and discrete Helmholtz projectors in mixed finite element methods for the incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, 311:304–326, 2016.

[26] Kent-Andre Mardal, Joachim Schöberl, and Ragnar Winther. A uniformly stable fortin operator for the taylor–hood element. *Numerische Mathematik*, 123(3):537–551, 2013.

[27] J.-C. Nédélec. Mixed finite elements in **R**^{3}. *Numer. Math.*, 35(3):315–341, 1980.

[28] J. M. Thomas P. A. Raviart. Primal hybrid finite element methods for 2nd order elliptic equations. *Mathematics of Computation*, 31(138):391–413, 1977.

[29] J. Schöberl. NETGEN An advancing front 2D/3D-mesh generator based on abstract rules. *Computing and Visualization in Science*, 1(1):41–52, 1997.

[30] J. Schöberl. C++11 Implementation of Finite Elements in NGSolve. *Institute for Analysis and Scientific Computing, Vienna University of Technology*, 2014.

[31] Joachim Schöberl. A posteriori error estimates for Maxwell equations. *Math. Comp.*, 77(262):633–649, 2008.

[32] L. R. Scott and M. Vogelius. Conforming finite element methods for incompressible and nearly incompressible continua. In *Large-scale computations in fluid mechanics, Part 2 (La Jolla, Calif., 1983)*, volume 22 of *Lectures in Appl. Math.*, pages 221–244. Amer. Math. Soc., Providence, RI, 1985.

[33] Rolf Stenberg and Juha Videman. On the error analysis of stabilized finite element methods for the Stokes problem. *SIAM J. Numer. Anal.*, 53(6):2626–2633, 2015.

[34] R. Verfürth. A posteriori error estimators for the Stokes equations. *Numer. Math.*, 55(3):309–325, 1989.

[35] R. Verfürth. A posteriori error estimates for nonlinear problems. Finite element discretizations of elliptic equations. *Math. Comp.*, 62(206):445–475, 1994.

[36] Junping Wang, Yanqiu Wang, and Xiu Ye. Unified a posteriori error estimator for finite element methods for the Stokes equations. *Int. J. Numer. Anal. Model.*, 10(3):551–570, 2013.

[37] Shangyou Zhang. A new family of stable mixed finite elements for the 3D Stokes equations. *Math. Comp.*, 74(250):543–554, 2005.