LOCAL MASS-CONSERVATION FOR STABILIZED LINEAR FINITE ELEMENT INTERPOLATIONS OF INCOMPRESSIBLE FLOW

B. GMEINER, C. WALUGA, AND B. WOHLMUTH

Abstract. In this work, we discuss a class of stabilized finite element discretizations for the incompressible Stokes problem using equal-order linear finite elements on simplicial meshes. By employing a two-level strategy, we obtain unconditional stability of the discrete problem and optimal error estimates. Moreover, we demonstrate that the proposed scheme preserves local mass-conservation on dual cells. This allows the direct coupling to vertex-centered finite volume discretizations of transport equations. Further, we can postprocess the fluxes independently for each dual box to obtain an element-wise conservative velocity approximation that can be used in cell-centered finite volume or discontinuous Galerkin schemes. Numerical examples are given to compare the quality of the discrete solution to that of solutions obtained by alternative stabilized schemes. Moreover, we demonstrate the coupling to cell- and vertex-centered finite volume methods for advective transport.

Key words. Stokes equations, equal-order interpolation, local mass conservation.

AMS subject classifications. 76D05, 76D07, 65F10, 65N30

1. Introduction

In this work, we consider the iso-viscous Stokes problem given by

\begin{align}
-\Delta u + \nabla p &= f, & \text{in } \Omega, \\
\text{div } u &= 0, & \text{in } \Omega,
\end{align}

as a prototype for incompressible flow in an open bounded domain \( \Omega \subset \mathbb{R}^d \), \( d = 2, 3 \). Here, \( u = [u_1, \ldots, u_d]^\top \) denotes the velocity field, \( p \) the pressure, and \( f = [f_1, \ldots, f_d]^\top \) denotes a forcing term. Below we agree on the convention, that vectorial quantities are printed in bold. For simplicity, we consider homogeneous Dirichlet boundary conditions, i.e., we set \( u = 0 \) on \( \partial \Omega \), and to make the pressure well-defined, we further set \( \int_{\Omega} p \, dx = 0 \). More general boundary conditions are considered in our numerical tests. These may be incorporated by the usual techniques; cf. [34] for an overview.

The Stokes problem written in the form of (1) is well-studied in the mathematical literature and there exists a number of widely-used discretization schemes. A common problem which is faced here is a compatibility condition on the discrete spaces in which the velocity and pressure are interpolated, the so-called LBB condition named after Ladyzhenskaya, Babuška and Brezzi. A seemingly natural placement of the degrees of freedom is employed in staggered discretizations, e.g. the MAC scheme [27]. Such staggered schemes often result in lightweight stencils that preserve desirable physical properties like local mass-conservation, which may be important when we want to avoid spurious sources and sinks if the computed flow field
is employed for transporting physical quantities \cite{13, 20}. However, these methods often have the drawback that they are not easily generalizable to unstructured meshes and/or higher orders. While finite element discretizations are usually geometrically more flexible, the most-appealing choices are known to be notoriously unstable \cite{8, 5}. For example it is well-known that equal-order conforming interpolations produce spurious pressure modes and violate the physical principle of local mass-conservation. The former problem can be controlled by adding stabilization terms on the discrete level. The latter problem is often addressed in the literature by either penalizing the divergence of the discrete velocity field (grad-div-stabilization; cf. \cite{33, 12}) or by enlarging the velocity space, thereby allowing piecewise discontinuous pressure approximations, which in turn allow for enhanced local mass-conservation. Among the most popular methods of this class are higher-order conforming finite elements \cite{19, 10}, and nonconforming schemes, such as the Crouzeix–Raviart element \cite{19} or the rapidly developing class of discontinuous Galerkin methods \cite{40, 35, 18, 16, 17, 14, 22}. However, while some of these schemes can be efficiently implemented on modern architectures, they often impose a considerable infrastructural demand on the underlying framework. In particular, if the required data-structures are not or incompletely implemented in a given framework, it is practically impossible to realize any of these methods efficiently without making massive extensions to the underlying codebase first. The situation gets even worse when we consider implementations involving highly optimized communication structures for parallel processing on extreme scales; cf., e.g., the discussion in \cite{25}. It is partly because of the facts listed above, that nowadays, still many successful academic and commercial computational fluid dynamics codes are based on a specific choice of finite differences, finite volumes or low-order continuous finite elements.

In this work, we aim to overcome the commonly criticized lack of local mass conservation in equal-order $P_1$ finite element discretizations. This is obtained by firstly introducing a two-level projection-based stabilization in Section 2. We discuss the relation to other stabilizations and give a stability and convergence analysis of the scheme. Secondly, using the conservation properties of the scheme on the two nested mesh-levels, we demonstrate in Section 3 that the discrete solution satisfies a local conservation property and discuss possibilities for locally postprocessing the discrete velocities for mass-conservation on dual and primal meshes. Section 4 gives some numerical examples in two dimensions and compares the proposed approach to some well-established stabilized and stable methods. Furthermore, we demonstrate the coupling of the finite-element solver with both cell- and vertex-centered finite volume methods and show examples that highlight the advantages of the proposed stabilization.

2. Two-level stabilized finite element discretization

Suppose a given polyhedral domain $\Omega \subset \mathbb{R}^d$ which is subdivided into simplices to obtain a quasi-uniform triangulation $T_{2h}$ satisfying the usual shape-regularity assumptions. Based on the macro triangulation $T_{2h}$, we define a finer triangulation $T_h$ which results from a uniform refinement of the original cells by introducing new vertices at all edges. The velocities and pressure are approximated in piecewise-linear function spaces associated with the fine-scale, i.e.,

$$V_h := \{ v \in [H_0^1(\Omega)]^d \cap C(\overline{\Omega})^d : v|_T \in [P_1(T)]^d, \forall T \in T_h \},$$
$$Q_h := \{ q \in C(\overline{\Omega}) : q|_T \in P_1(T), \forall T \in T_h \},$$
respectively, where $P_1(T)$ denotes the space of linear polynomials on a physical simplex $T$ of our triangulation. Moreover, we denote by $Q_{2h}$ the pressure space with respect to the macro-grid $T_{2h}$. By definition, the boundary conditions on the velocity are included in the respective discrete function spaces. To incorporate the mean-value condition of the pressure, we further define $Q_0^h := Q_h \cap L^2_0(\Omega)$ and $Q_2^h := Q_{2h} \cap L^2_0(\Omega)$.

We consider the following weak formulation for discretizing the Stokes problem with stabilization: find $(u_h, p_h) \in V_h \times Q_0^h$ such that

\begin{align}
(2a) & \quad a(u_h, v_h) + b(v_h, p_h) = f(v_h), \quad v_h \in V_h, \\
(2b) & \quad b(u_h, q_h) - c_{2h}(p_h, q_h) = 0, \quad q_h \in Q_h,
\end{align}

where we define the bilinear forms as

\begin{align*}
a(u, v) := \int_{\Omega} \nabla u : \nabla v \, dx, \quad \text{and} \quad b(u, q) := -\int_{\Omega} \text{div} u \cdot q \, dx,
\end{align*}

and the linear form as $f(v) := \int_{\Omega} f \cdot v \, dx$. The bilinear form $c_{2h}(\cdot, \cdot)$ is added to stabilize the equal-order finite element space $V_h \times Q_h$ in terms of the LBB condition. Given an operator $P_{2h} : Q_h \to Q_{2h}$, we set

\begin{align*}
(3) & \quad c_{2h}(p, q) := \alpha_1 \int_{\Omega} (p - P_{2h}p)(q - P_{2h}q) \, dx.
\end{align*}

To select $P_{2h}$, there are different suitable possibilities, e.g., $P_{2h}$ can be the locally defined nodal interpolation associated with $T_{2h}$, or a Scott–Zhang type operator [36].

For the stability analysis, we only require that the following abstract assumptions hold:

\begin{align}
(4a) & \quad P_{2h}q_{2h} = q_{2h}, \quad q_{2h} \in Q_{2h}, \\
(4b) & \quad \|P_{2h}q_h\|_0 \leq \gamma_s\|q_h\|_0, \quad q_h \in Q_h.
\end{align}

The motivation for this special choice of stabilization can be easily explained. Since, due to (4a), the kernel of the discrete operator associated with $c_{2h}$ is exactly $Q_{2h}$, we only penalize the checkerboard modes which are the cause of pressure instabilities in equal order methods; see Figure 1 for some illustration.

![Figure 1. Illustration of the action of the stabilization operator in 1D.](image-url)
Relation to other stabilized methods. There exists a variety of well-established stabilization techniques, which are typically based on Petrov-Galerkin formulations, artificial compressibility, interior-penalty or additional projection-terms; cf. e.g. [11, 6, 4, 23, 30, 8, 7] for an overview. Two popular choices, which we compare in the numerical examples are the pressure-stabilized Petrov-Galerkin (PSPG) scheme [11, 28, 39] and a projection-based stabilization due to Bochev et. al. [4], resulting from the choices

\[c_{\text{pspg}}(p, q) := \sum_{T \in \mathcal{T}_h} \delta h_T^2 \int_T \nabla p \cdot \nabla q \, dx,\]

\[c_{0,h}(p, q) := \sum_{T \in \mathcal{T}_h} \alpha_0 \int_T (p - \Pi_0 p)(q - \Pi_0 q) \, dx,\]

respectively, where \(h_T := \text{diam}(T)\) denotes the local element diameter, and \(\Pi_0\) is a (locally defined) projection onto piecewise constants; cf. [4] for details. For the residual-based PSPG scheme a consistency term may be added on the right hand side if \(f \neq 0\); see [28]. Moreover, since the stabilization acts as a Neumann condition on the pressure, Droux and Hughes proposed the inclusion of a non-symmetric interface consistency term for \(p\); cf. [21]. These modifications neither influence the stability nor the asymptotic rate of convergence of the method. However they can have a significant influence on the quality of the pressure solution. We note that the different scaling between (5a) and (5b) is related to the different orders of the operators; more precisely

\[\|p - \Pi_0 p\|_0^2 \leq c h_T^2 \|\nabla p\|_0^2.\]

While the stabilization parameters for the different approaches outlined above only need to be greater than zero for stability, their choice can have a significant effect on the quality of the discrete solution. We compare the sensitivity of the different methods on the stabilization parameter numerically in Section 4.

2.2. Uniform stability and convergence. As an important ingredient for the following analysis, we first recall the fact that \(V_h \times Q_{2h}^0\) forms a uniformly stable pairing [3, 24]. Let \(\beta_h > 0\) be the uniform inf-sup constant of the pairing \(V_h \times Q_{2h}^0\), and \(\beta_c > 0\) be the uniform coercivity constant of \(a(\cdot, \cdot)\) with respect to the \(H^1\)-Norm, i.e.,

\[\inf_{q_{2h} \in Q_{2h}^0} \sup_{v_h \in V_h} \frac{b(v_h, q_{2h})}{a(v_h, v_h)^{1/2} \|q_{2h}\|_0} \geq \beta_h,\]

\[a(v_h, v_h)^{1/2} \geq \beta_c \|v_h\|_1, \quad v_h \in V_h.\]

To simplify notation, we introduce the bilinear form

\[B_h(u, p; v, q) := a(u, v) + b(v, p) + b(u, q) - c_{2h}(p, q)\]
which allows us to write the discrete problem as: find \((u_h, p_h) \in V_h \times Q_h^0\) such that
\[
B_h(u_h, p_h; v_h, q_h) = f(v_h), \quad (v_h, p_h) \in V_h \times Q_h.
\]
Moreover, we set \(\|v; q\| := (\|v\|^2 + \|q\|^2)^{1/2}\). We recall that the choice of homogeneous Dirichlet boundary conditions automatically guarantees that \(\int_{\Gamma} \text{div} u_h \, dx = \int_{\partial \Gamma} u_h \cdot n \, ds = 0\). We are now prepared to show uniform stability of the discretization.

**Lemma 2.2.** For \(\alpha > 0\) the discrete Stokes formulation (7) is uniformly stable, i.e., there exists a stability constant \(\gamma_c > 0\) independent of \(h\) such that
\[
\sup_{v_h \in V_h} \sup_{q_h \in Q_h^0} \frac{B_h(w_h, r_h; v_h, q_h)}{\|v_h, q_h\|} \geq \gamma_c \|w_h, r_h\|, \quad w_h \in V_h, \ r_h \in Q_h^0.
\]
Moreover the stability constant \(\gamma_c\) depends only on \(\gamma_s, \beta_s, \beta_c\) and \(\alpha\).

**Proof.** The proof employs the decomposition \(r_h = r_h - (\text{Id} - \Pi_0)P_{2h}r_h + (\text{Id} - \Pi_0)P_{2h}r_h\) and the uniform inf-sup stability (6a) over \(V_h \times Q_{2h}\). For \(r_h \in Q_h^0\), we find \(r_{2h} := (\text{Id} - \Pi_0)P_{2h}r_h \in Q_{2h}^0\), and thus \(\|r_h - r_{2h}\|_0 = ((\text{Id} - \Pi_0)(r_h - P_{2h}r_h))_0 \leq \|r_h - P_{2h}r_h\|_0\). Moreover due to (6a), there exists a \(z_h \in V_h\) such that \(r > 0\) fixed there holds
\[
a(z_h, z_h) = \tau^2 \|r_{2h}\|_0^2, \quad \text{and} \quad b(z_h, r_{2h}) = \beta_0 \|r_{2h}\|_0^2.
\]
Setting \(v_h = w_h + z_h\) and \(q_h = -r_h\) and using Young’s inequality as well as the definition (3), we get in terms of (4b) and (6b) that
\[
\|v_h\|_1 \leq \|w_h\|_1 + \|z_h\|_1 \leq \|w_h\|_1 + \frac{\tau}{\beta_0} \|r_{2h}\|_0 \leq \|w_h\|_1 + \gamma_s \frac{\tau}{\beta_0} \|r_h\|_0.
\]
Moreover, we find
\[
B_h(w_h, r_h; v_h, q_h) = a(w_h, v_h) + b(v_h, r_h) + b(w_h, q_h) - c_{2h}(r_h, q_h) = a(w_h, w_h) + a(w_h, z_h) + b(w_h, r_h) + b(z_h, r_h) - b(w_h, r_h) + c_{2h}(r_h, r_h) \geq \frac{\beta_0}{2} \|r_{2h}\|_0^2 + \frac{\alpha_1}{2} \|r_h - P_{2h}r_h\|_0^2 \geq \frac{\beta_0}{2} \|w_h, w_h\| + \|z_h\|_0 \|r_{2h}\|_0^2 + \frac{\alpha_1}{2} \|r_h - P_{2h}r_h\|_0^2 \geq \frac{\beta_0}{2} \|w_h, w_h\| + \alpha_1 \|r_h - P_{2h}r_h\|_0^2.
\]
Setting \(\tau > 0\) small enough and using (6b), the uniform stability follows. \(\square\)

As it is common in projection-stabilized methods, the Galerkin orthogonality is not fulfilled exactly, hence we need to carefully analyze the consistency error. The convergence proof then follows by standard arguments, provided that \(P_{2h}\) satisfies an approximation property of the form
\[
\|p - P_{2hp}\| \leq Ch\|p\|_1, \quad p \in H^1(\Omega).
\]
We obtain the following convergence bound.

**Theorem 2.3.** Let \((u, p)\) be the solution of the Stokes problem (1) and let \((u_h, p_h) \in V_h \times Q_h\) be the solution of the discrete problem (2). Moreover assume that the assumptions (4) and (8) hold. Then
\[
\|u - u_h\|_1 + \|p - p_h\|_0 \leq Ch(\|u\|_2 + \|p\|_1)
\]
**Proof.** The proof follows the lines of [4, Sect. 5] and is thus not repeated. \(\square\)
Below we tacitly assume that the mean-value condition on the pressure is suitably treated in the discrete problem. We can realize this, for example, by a Lagrange multiplier or by suitable projection steps included in the iterative solver.

3. Local mass conservation

Although the discrete solution $u_h \in V_h$ satisfies (2b), it is not element-wise mass conservative. However, a local post-processing can recover the mass conservation on dual meshes. We work out the details only for the 2D case and comment on possible generalizations; the 3D case is technical and requires more notation but follows by essentially the same arguments.

3.1. Preliminary considerations. To start with, we associate with any conforming triangulation $\mathcal{T}_h$ a barycentric dual grid $B_h$. The dual boxes are linked to the vertices $p \in \mathcal{P}_h$ of the original mesh $\mathcal{T}_h$ in the following way: Let $p \in \mathcal{P}_h$ and denote by $\mathcal{E}_p$ the edges sharing the node $p$. Then, if $e \in \mathcal{E}_p$ is a boundary edge, we connect the midpoint $m_e$ of $e$ with $p$ and to the barycenter $m_T$ of the adjacent element $T$. Otherwise, if $e$ is an interior edge, we connect $m_e$ with the barycenters of the two adjacent elements. This construction results in a mesh $B_h$ of polygonal dual cells $B_p^h$, see Figure 2. Denoting by $T^p_h$ the set of all elements in $\mathcal{T}_h$ sharing the vertex $p$, it is easy to verify that

$$\frac{1}{3}|T| = |B_p^h \cap T|, \quad T \in T^p_h.$$ 

In 3D a similar construction can be done, which gives rise to polyhedrons. Then, the factor $1/3$ has to be replaced by $1/4$, or generally we find $1/(d+1)$.

![Figure 2](image_url)  

**Figure 2.** Example of an unstructured mesh $\mathcal{T}_{2h}$ (left) and the uniform refinement $\mathcal{T}_h$ (center) with the associated dual meshes $B_{2h}$ and $B_h$, respectively, printed in bold. We also show the overlap of these dual meshes (right).

We start with the following reformulation of the linear form $b(\cdot, \cdot)$ and note that $\text{div} \ V_h \subset \prod_{T \in \mathcal{T}_h} P_0(T)$. Using the nodal basis function $\phi_h^p \in Q_h$, we get

$$b(u_h, \phi_h^p) = \frac{1}{3} \int_T \sum_{T^p \in T_h^p} \text{div} u_h \, dx = \sum_{T^p \in T_h^p} \int_{T^p \cap B_h^p} \text{div} u_h \, dx = \int_{\partial B_h^p} u_h \cdot n \, ds.$$  

Here $n$ stands for the outer unit normal of $B_h^p$. To show that the two-level stabilization does not entirely destroy the mass conservation on the dual box, we use the hierarchical decomposition of $\phi_{2h}^p \in Q_{2h} \subset Q_h$ into the nodal basis functions associated with $\mathcal{T}_h$, i.e.,

$$\phi_{2h}^p = \phi_h^p + \frac{1}{2} \sum_{e \in \mathcal{E}_p} \phi_h^{m_e}.$$
Here $\mathcal{E}_{2h}^p$ is the set of all edges $e$ of the triangulation $\mathcal{T}_{2h}$ having the vertex $p$ as endpoint. Since $B_{2h}^p \in \mathcal{B}_{2h}$ cannot be written as the union of dual boxes in $\mathcal{B}_h$ (see the right picture in Figure 2), we do not directly obtain mass conservation. However observing that
\begin{equation}
\tag{11}
c_{2h}(\phi_h^p, \phi_{2h}^p) = 0, \quad p \in \mathcal{P}_{2h},
\end{equation}
where $\mathcal{P}_{2h}$ is the set of all vertices of the mesh $\mathcal{T}_{2h}$, we find in terms of \eqref{2b}, \eqref{10} and \eqref{11}
\begin{equation}
\tag{12}
b(\mathbf{u}_h, \phi_{2h}^p) = \frac{1}{2} \sum_{e \in \mathcal{E}_{2h}^p} b(\mathbf{u}_h, \phi_{h}^{m_e}) + b(\mathbf{u}_h, \phi_{h}^{p})
= \frac{1}{2} \sum_{e \in \mathcal{E}_{2h}^p} \int_{\partial B_{h}^{m_e}} \mathbf{u}_h \cdot \mathbf{n} \, ds + \int_{\partial B_{h}^{p}} \mathbf{u}_h \cdot \mathbf{n} \, ds = 0.
\end{equation}
This observation is crucial for the following construction.

3.2. Conservation on the macro dual-mesh. Based on the observation \eqref{12}, we first describe a method to construct conservative fluxes on each macro grid dual cell $B_{2h}^p \in \mathcal{B}_{2h}$ and subsequently describe a way to obtain element-wise conservative fluxes on $\mathcal{T}_h$ by solving small systems on each of these dual cells.

Let us first introduce the skeleton $\Gamma := \bigcup \partial B_{2h}^p$. $B_{2h}^p \in \mathcal{B}_{2h}^p$, with a uniquely defined normal vector $\mathbf{n}$ on each segment $\Gamma_i$ of $\Gamma$. We define a conservative flux by
\begin{equation}
\tag{13}
j(\mathbf{u}_h) := \mathbf{u}_h \cdot \mathbf{n} + \kappa(\mathbf{u}_h) \quad \text{on} \quad \Gamma,
\end{equation}
where $\kappa(\mathbf{u}_h)|_{\Gamma_i} \in P_0(\Gamma_i)$ denotes an oriented flux-correction of $\mathbf{u}_h \cdot \mathbf{n}$ which ensures local mass conservation. At the boundary $\partial \Omega$, we set $\kappa(\mathbf{u}_h) = 0$.

For each interior segment $\Gamma_i \in \Omega$, we define the flux correction as follows: Let $T_{2h}^i \in \mathcal{T}_{2h}$ denote the macro mesh element such that $\Gamma_i \subset T_{2h}^i$. Then $\Gamma_i$ is described by the barycenter $m_i$ of $T_{2h}^i$ and the midpoint $m_e$ of some edge $e \in \mathcal{E}_{2h}$, where by $\mathcal{E}_{2h}$ we shall denote the set of macro grid edges. Then, we take the fine dual box $B_{m_e} \in \mathcal{B}_h$ and consider only the part $B_i := B_{m_e} \cap T_{2h}^i$ overlapping with $T_{2h}^i$. We naturally divide $B_i$ by $\Gamma_i$ into the non-overlapping polygons $B_i^+$ and $B_i^-$ such that $\mathbf{n}|_{\Gamma_i}$ points from $B_i^-$ to $B_i^+$; cf. Figure 3. The correction is then defined by
\begin{equation}
\tag{14}
\kappa(\mathbf{u}_h)|_{\Gamma_i} := \frac{1}{2} |\Gamma_i|^{-1} \left( \int_{\partial B_i^+} \mathbf{u}_h \cdot \mathbf{n} \, ds - \int_{\partial B_i^-} \mathbf{u}_h \cdot \mathbf{n} \, ds \right) \in P_0(\Gamma_i).
\end{equation}

Remark 3.1. In case of non-homogeneous Dirichlet boundaries, we can either make sure that the boundary conditions in our discrete problem are applied in such a way that the preservation of piecewise polynomial boundary conditions is ensured (e.g., by a Scott–Zhang type operator), or we need to choose the correction $\kappa(\mathbf{u}_h)|_{\partial \Omega}$ accordingly.

As a preparation for the proof of our main result, we recall that for each dual box $B_{2h}^p \in \mathcal{B}_{2h}$ associated with a macro mesh node $p \in \mathcal{P}_{2h}$, we find
\[ B_h^p \subset B_{2h}^p \subset B_h^p \cup \bigcup_{e \in \mathcal{E}_{2h}^p} B_e^{m_e}, \]
as illustrated in Figure 2 (right). We can now prove the following local conservation property.
Figure 3. Sketch of the polygons \( B^+_{pi} \) and \( B^-_{pi} \) required for the reconstruction of the locally conservative fluxes. The corresponding macro grid element and its uniform refinement is depicted for illustration.

**Theorem 3.2.** Let \( u_h \in V_h \) denote the discrete velocity solution of (2). Moreover, let \( j(u_h) \) denote the reconstructed flux (13) as described above. Then

\[
\int_{\partial B^p_{2h}} j(u_h) \, ds = b(u_h, \phi^p_{2h}) = 0, \quad B^p_{2h} \in B_{2h}.
\]

**Proof.** We observe that \( \partial B^p_{2h} \) is composed of \( n_p \) segments \( \Gamma_i \), \( i = 1, \ldots, n_p \). Thus, by the piecewise definition of \( j(u_h) \) we can rewrite the integral over \( \partial B^p_{2h} \) as

\[
\sum_{i=1}^{n_p} \int_{\Gamma_i} j(u_h) \, ds = \sum_{i=1}^{n_p} \int_{\Gamma_i} (u_h \cdot n + \kappa(u_h)) \, ds
\]

\[
= \sum_{i=1}^{n_p} \int_{\Gamma_i} u_h \cdot n \, ds + \frac{1}{2} \sum_{e \in \mathcal{E}^p_{2h}} \left( \int_{\partial B^+_{eh} \cap \Omega} u_h \cdot n \, ds - \int_{\partial B^-_{eh} \cap \Omega} u_h \cdot n \, ds \right)
\]

\[
= \frac{1}{2} \sum_{e \in \mathcal{E}^p_{2h}} \left( \int_{\partial B^+_{eh} \cap \Omega} u_h \cdot n \, ds - \int_{\partial B^-_{eh} \cap \Omega} u_h \cdot n \, ds \right)
\]

\[
= \frac{1}{2} \sum_{e \in \mathcal{E}^p_{2h}} \left( \int_{\partial B^m_{eh} \cap \Omega} u_h \cdot n \, ds - \int_{\partial B^m_{eh} \cap B^p_{2h}} u_h \cdot n \, ds \right)
\]

\[
= \sum_{e \in \mathcal{E}^p_{2h}} \frac{1}{2} \int_{\partial B^m_{eh}} u_h \cdot n \, ds - \int_{\partial B^m_{eh} \cap B^p_{2h}} u_h \cdot n \, ds
\]

\[
= \frac{1}{2} \sum_{e \in \mathcal{E}^p_{2h}} \int_{\partial B^m_{eh}} u_h \cdot n \, ds + \int_{\partial B^p_{eh}} u_h \cdot n \, ds.
\]
Remark 3.3. Using a suitably chosen sub-mesh and a $H(\text{div})$-conforming discrete space (e.g., the linear Brezzi–Douglas–Marini finite element space [9]), we can construct a postprocessed solution $\tilde{u}_h$ which is strongly divergence-free. This reconstruction is inexpensive since it can be computed in a local fashion. Such a postprocessed velocity $\tilde{u}_h$ may be particularly interesting for a-posteriori error estimation. For details, we refer to the finite element literature on local flux-equilibration, cf. e.g., [31, 26] and the references cited therein.

Remark 3.4. The formal definition of $j(u_h)$ is convenient for proving our results, but not useful for implementation, since it involves dealing with the intersection of dual boxes and elements on the two mesh-levels. However, by some simple rearrangements, we can show that the correction is in fact only the weighted difference of the divergence between two elements: We consider again the notation used above. The dual half-box $B_i$ intersects exactly three triangles of $T_i$, $T_+$, and $T_-$. Moreover, $|B_i^+ \cap T_+| = \frac{1}{2}|T_+|$, $|B_i^- \cap T_-| = \frac{1}{3}|T_-|$, and $|B_i^+ \cap T_\pm| = \frac{1}{6}|T_\pm|$, $|B_i^- \cap T_\pm| = \frac{1}{6}|T_\pm|$. Using the definition (14), we can now integrate by parts and use the fact that $|\text{div } u_h|_T \in P_0(T)$ for all $T \in T_i$. This immediately yields

$$\kappa(u_h)|_{\Gamma_i} = (|\text{div } u_h|_{T_+} - |\text{div } u_h|_{T_-})/6,$$

since the contributions of $T_\pm$ cancel.

Based on the previous equivalent definition of the flux-correction $\kappa(u_h)$, we can make another interesting observation which is based on the fact that we have $|T_+| = |T_-|$ by construction.

Remark 3.5. If $u_h|_{T_{2h}} \in (P_1(T_{2h}))^2$ for some $T_{2h} \in T_{2h}$, then the flux-correction vanishes, i.e., $\kappa(u_h)|_{\Gamma_{2h}} = 0$, and thus $j(u_h)|_{\Gamma_{2h}} = u_h \cdot n|_{\Gamma_{2h}}$. Hence, discrete velocities living only on the macro-scale of a patch of elements around a vertex $p \in P_{2h}$ are already locally conservative with respect to the corresponding dual cell $B_{2h} \in B_{2h}$.

The conservative fluxes on the dual boxes can be directly used to couple an incompressible flow solver based on the two-level stabilization to a finite volume method for solving transport equations. However, if local conservation is desired with respect to the elements of the fine mesh $T_i$, we need to solve local problems in a postprocessing step. For this procedure, the previous considerations provide important preliminaries.

3.3. Element-wise conservative fluxes. As we can see in Figure 4, a macro dual-box $B_{2h}^p \in B_{2h}$ is geometrically composed of the patch of elements $T_i^p$ sharing all the same macro mesh node $p \in P_{2h}$ and additional triangles $T_i$, which are defined by the edges $e_i \in E_h$ opposing the vertex $p$ and the barycenter of the element neighboring at this edge. Here, $E_h$ denotes the set of edges of $T_i$. Since we already established a conservative flux $j(u_h)$ on the skeleton $\Gamma$ described by the boundaries of the macro dual cells, we can proceed locally for the edges contained in $B_{2h}^p$. Similar to (13), we define

$$j(u_h)|_e := u_h \cdot n|_e + \kappa(u_h)|_e \quad \text{on} \quad e \in E_h.$$

Our first observation is that we can locally equilibrate the fluxes on the elements $T_i$. Since we
are missing one flux-correction $\kappa(u_h)|_{e_i} \in P_0(e_i)$ on each $e_i$, we form the local mass balance
\[ \int_{\partial T_i} j(u_h) \, ds = 0, \]
and explicitly solve for $\kappa(u_h)|_{e_i}$; cf. Figure 4 (left). This is easy, since we already have conservative fluxes defined on two of three edges of $T_i$ by the construction of the previous section. Hence, the problem reduces to determining flux-corrections on the interior edges of the patch of elements $T_{p_h}$. Let us outline the procedure, which is similar to the construction in [1, Sect. 6.4.5]. For each interior node $p \in P_{2h}$, we have $n = |T_{p_h}| = |E_{p_h}|$, which generates a linear problem of the form
\[ L\kappa = r, \]
with $n$ unknowns and $n$ equations. Here, the rows $L_k$ and $r_k$ are associated with the mass-balance over the boundary of one element in $T_{p_h}$, and $\kappa$ is the vector containing the values $\kappa(u_h)|_{e|e}, e \in E_{p_h}$. For each $T_k \in T_{p_h}$ the mass-balance can be written as
\[ 0 \overset{!}{=} \int_{\partial T_k} j(u_h) \, ds \iff \sum_{e \in \partial T_k} \kappa(u_h)|_{e|e} = - \text{div } u_h|_{T_k}|T_k|, \iff L_k\kappa = r_k. \]
Assuming a counter-clockwise ordering of degrees of freedom inside the element patch as depicted in Figure 4, this leaves us with solving a system of the following form:
\[ \begin{pmatrix}
-1 & 1 & 0 & \ldots & 0 & 0 \\
0 & -1 & 1 & \ldots & 0 & 0 \\
0 & 0 & -1 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & -1 & 1 \\
1 & 0 & 0 & \ldots & 0 & -1 \\
\end{pmatrix} \begin{pmatrix}
\kappa_1 \\
\kappa_2 \\
\kappa_3 \\
\vdots \\
\kappa_{n-1} \\
\kappa_n \\
\end{pmatrix} = \begin{pmatrix}
\tau_1 \\
\tau_2 \\
\tau_3 \\
\vdots \\
\tau_{n-1} \\
\tau_n \\
\end{pmatrix}. \]
The operator $L$ has a non-trivial kernel $\ker L = \text{span}\{1\}$, where $1 := [1, 1, \ldots, 1]^\top$, since we can add any constant to each of the corrections in the interior of the patch and still obtain
locally conservative fluxes. This implies that a solution to (16) exists if and only if \( 1^T \cdot r = 0 \), which can be easily verified by substituting the definition of \( r \) and using the conservation properties derived above. For reconstructing unique flux-corrections inside each patch of elements, we thus solve for the solution producing the minimal Euclidean norm, i.e., we find

\[
\min_{\kappa \in \mathbb{R}^n} |\kappa|_2, \quad \text{such that} \quad L\kappa = r.
\]

This procedure is reasonable since the element-wise divergence of \( u_h \) converges to zero as \( h \to 0 \) and so should the correction. Thus only the solution with a minimal Euclidean norm yields an asymptotically correct approximation of the fluxes. For the dual cells located at the boundary we have no circular dependencies. Hence, the local system is uniquely solvable as long as we have one edge at a Dirichlet boundary. For cells located at a pure Neumann boundary, a similar procedure as for the interior cells needs to be carried out.

**Remark 3.6.** Given a locally conservative velocity on the elements in \( T_h \), we can equilibrate on each element to obtain conservation also on the fine scale dual mesh \( B_h \).

**Remark 3.7.** A similar but technically more involved reconstruction can be done for the second order Taylor-Hood element [38], which is not presented in this work for brevity.

4. **Numerical examples**

4.1. **Colliding flow benchmark.** The first test problem is concerned with the colliding-flow benchmark on the unit square \( \Omega = (-1,1)^2 \) which is used in e.g. [32, 2] for comparing low-order methods. Boundary conditions are chosen, such that the exact solution is given by

\[
\begin{align*}
\mathbf{u}(\mathbf{x}) &= (20xy^2, 5x^4 - 5y^4), \quad \text{and} \quad p(\mathbf{x}) = 60x^2y - 20y^3, \\
\mathbf{x} &:= (x,y).
\end{align*}
\]

We then first conduct a series of experiments in which we vary the stabilization parameters of the different methods to study their influence on the solution quality. By this we try to obtain suitable stabilization parameters for the different schemes under consideration to make a fair comparison possible. Our initial triangulation is a four element criss-cross mesh. To avoid pre-asymptotic effects we first refine this mesh five times. Then, we solve the problem for parameters \( \delta \in [0.001, 2] \) for the PSPG-stabilization, and \( \alpha_i \in [0.01, 20], \ i = 0,1, \) for both the projection-based stabilizations. We plot the results in Figure 5, where, for better interpretation, we normalize all errors by the minimum error obtained for the PSPG-method.

We can observe that for the PSPG-stabilization, we obtain a good balance of errors at around \( \delta \approx 1/12 \), which is in line with common recommendations for the practical choice of this parameter. After this point, all observed error-norms begin to increase, which is a clear sign of over-stabilization, i.e., unnecessarily strong constraints on degrees of freedom which would otherwise contribute to a better approximation of the solution. This behavior is to some extent shared by the \( P_0 \)-stabilized method which is not surprising, since also the limit \( \alpha_0 \to \infty \) would force the discrete pressure to be zero and reduce the problem to an unconstrained energy minimization. However, for the \( P_0 \)-stabilization the effect of over-stabilization can be considered much less dramatic, especially if we consider the fact that there is typically no reason to choose anything else than \( \alpha_0 = 1 \) in the iso-viscous case. For the proposed two-level-stabilization, the errors do not change significantly for \( \alpha_1 > 1 \). In fact, the solution is not much different to that obtained by a \( P_1 \)-iso-\( P_2 - P_1 \) element for all choices \( \alpha_1 \in [1, \infty) \). As mentioned before, this kind of (asymptotic) consistency comes from the fact that this stable pairing is recovered in the limit \( \alpha_1 \to \infty \) (if we assume exact arithmetics).
Taking the parameters identified by the above experiments, we continue by studying the qualitative characteristics of the different methods based on the benchmark shown above. To quantify the local mass-conservation characteristics of a discrete solution, we define

\[ |u_h|_{\text{div}, T} := \max_{T \in T_h} \int_{\partial T} u_h \cdot n \, ds, \]

and

\[ |j(u_h)|_{\text{div}, B_{2h}} := \max_{B \in B_{2h}} \int_{\partial B} j(u_h) \, ds, \]

where the postprocessed mass flux \( j(u_h) \) is defined as in the previous section. We then solve the problem on a series of uniformly refined meshes with levels \( l = 1, 2, \ldots \) using the PSPG stabilization, the \( P_0 \)-stabilized scheme, the \( P_1 \)-stabilized method investigated in this paper, and the classical \( P_1 \)-iso-\( P_2 \)-\( P_1 \) mixed finite element method. The results are listed in Table 1.

We can observe that all methods perform comparably in the standard norms. Also the observed local mass-defects do not substantially differ between the different stabilizations. The situation is different on the dual mesh \( B_{2h} \). Here, both, the PSPG-stabilization and
the $P_0$-stabilization converge asymptotically with an almost cubic rate. In contrast, the $P_1$-stabilization and the $P_1$-iso-$P_2 - P_1$ element yield exact conservation.

Table 1. Comparison of different stable $P_1 - P_1$ methods for the colliding flow benchmark. Rates for the mass-conservative schemes are omitted since they essentially result from effects of floating point arithmetic, hence contain no information about the rate of convergence with respect to the mesh-size.

### 4.2. Coupling with advective transport.

In this example, we couple the steady-state Stokes equation (1) to a simple transport equation. The change of mass in a control volume $V_i \subseteq \Omega$ is given by the integral conservation law

$$\partial_t \int_{V_i} c \, dx = - \int_{\partial V_i} c \mathbf{u} \cdot \mathbf{n} \, ds,$$

which equates the change of mass in the control volume with the advective transport through the boundary. Let us assume that $\Omega$ is partitioned into non-overlapping polygonal control-volumes $V_i$. Discretizing spatially in terms of finite volumes (FV) with first-order upwinding (sometimes called donor-cell splitting) and temporally by an forward Euler method on the uniform subdivision $t^k = k \Delta t$, we obtain the explicit marching scheme

$$c_i^{k+1} = c_i^k - \frac{\Delta t}{|V_i|} \sum_{j>i} \left\{ (c_i^k + c_j^k)(\mathbf{u}, \mathbf{n})_{ij} + (c_i^k - c_j^k)(\mathbf{u}, \mathbf{n})_{ij} \right\}, \quad k = 0, 1, \ldots,$$
where the degrees of freedom $c_k^i := |V_i|^{-1} \int_{V_i} c(x, t^k) \, dx$ denote the cell-averages in the control volume $V_i$ at $t = t^k$ and $(\mathbf{u}, \mathbf{n})_{ij} := \int_{\partial V_i \cap \partial V_j} \mathbf{u} \cdot \mathbf{n} \, ds$ denotes the net flux through an interior interface between two control volumes $V_i$ and $V_j$. Different types of finite volume methods can now be generated by specifying the partitioning of the domain $\Omega$. Below, we shall consider the vertex-centered (VC) scheme with respect to the macro-mesh, i.e., $V_i \in B_{2h}$ and the cell-centered (CC) scheme with respect to the fine mesh, i.e., $V_i \in T_h$. The time-step size $\Delta t > 0$ is always chosen small enough, such that the Courant–Friedrichs–Lewy condition is satisfied. For further details on the employed schemes we refer to [29].

For the sake of simplicity, we are not interested in a fully coupled simulation here. Hence, we first generate an exact flow field $\mathbf{u}$ as follows: Given the function $\psi_1(x) := xy(1 - x)(1 - y)$ and a yet unspecified but smooth enough function $\psi_2(x)$, we set

$$u(x) := 10 \left( \frac{\partial_y(\psi_1(x) \cdot \psi_2(x))}{-\partial_x(\psi_1(x) \cdot \psi_2(x))} \right),$$

which is by construction solenoidal and has free-slip boundary values. This flow field, together with $p(x) = 10(2x - 1)(2y - 1)$ is approximately reproduced using the differently stabilized finite element schemes discussed above, where we set boundary conditions given by the exact solution and we compute $\mathbf{f} = -\Delta \mathbf{u} + \nabla p$ to ensure that $(\mathbf{u}, p)$ solves (1). The resulting fluxes $\mathbf{u}_h \cdot \mathbf{n}$ and the postprocessed conservative fluxes $j(\mathbf{u}_h)$ on the boundaries of the control volumes $V_i$ are then used to convect the mass in the discrete sense.

4.2.1. Zeroth order accuracy. In this experiment, we follow the discussion of zeroth-order accuracy in [20], i.e., we investigate the capability of the schemes to reproduce constant initial concentrations $c(x, t = 0) = 1$ in $\Omega$. In this case, the solution is trivial for an incompressible flow field, since any convected mass is always instantly replaced and thus $c(x, t) = 1$ for all $t \geq 0$. For this, we solve for the interval $t \in (0, 1]$. The exact velocity field is generated by plugging the generator function $\psi_2(x) = (x - 0.5)^2 + (y - 0.5)^2$ into the construction given above, producing a complicated flow field with several recirculation zones in the interior of the domain as depicted in Figure 6 (left).

![Figure 6](image-url)  
**Figure 6.** Different velocity fields generated by (19) for different choices of the generator function $\psi_2$. left: $\psi_2(x) = (x - 0.5)^2 + (y - 0.5)^2$; right: $\psi_2(x) = \psi_1(x)/2$. Arrows are colored by magnitude (red: fast; blue: slow).
We conduct a series of experiments for an unstructured macro mesh $\mathcal{T}_{2h}$ with the cell-centered method defined on the refinement $\mathcal{T}_h$ and the vertex-centered method based on the dual mesh $\mathcal{B}_{2h}$. The results for $t = 1$ are depicted in Figures 7 and 8, respectively. We observe mass redistribution in the circular area around the big center vertex. Here, the non-conservative fluxes produce errors of up to 8% for the cell-centered approaches and $0.3 - 2.5\%$ for the vertex-centered scheme. The method using the newly proposed postprocessing strategies are in contrast capable of reproducing constant solutions (up to perturbations in the order of numerical roundoff). In terms of zeroth-order accuracy, it is thus of advantage to use the described post-process to recover local mass-conservation. If the proposed stabilization method is coupled to vertex-centered finite volumes, we even see a qualitative improvement already for the velocities without further postprocessing.

![Figure 7](image-url)

**Figure 7.** Results for $t = 1$ using a CC-FV method: we set the initial concentration to $c(x, t = 0) = 1$ and compare the velocity fields computed by the PSPG method ($\delta = 1/12$, top left), the $P_0$-stabilized method ($\alpha_0 = 1$, top right), the $P_1$-stabilized method ($\alpha_1 = 1$, bottom left) and the $P_1$-stabilized method with postprocessing ($\alpha_1 = 1$, bottom right).

4.2.2. **Mixing of two concentrations.** The next example takes a slightly more complicated initial profile which is transported by a single convection cell that can be generated by the
Figure 8. Results for $t = 1$ using a VC-FV method: we set the initial concentration to $c(x, t = 0) = 1$ and compare the velocity fields computed by the PSPG method ($\delta = 1/12$, top left), the $P_0$-stabilized method ($\alpha_0 = 1$, top right), the $P_1$-stabilized method ($\alpha_1 = 1$, bottom left) and the $P_1$-stabilized method with postprocessing ($\alpha_1 = 1$, bottom right).

choice $\psi_2(x) = \psi_1(x)/2$; cf. Figure 6 (right). We set the initial concentration to

$$c(x, t = 0) = \begin{cases} 2, & x < 0.5, \\ 1, & x \geq 0.5. \end{cases}$$

and solve for the interval $t \in (0, 10]$. The results for $t = 10$ are depicted in Figures 9 and 10, respectively. We observe that over-stabilization in the PSPG-scheme results in unphysical effects and that all stabilized methods display spurious sources and sinks. However, the effect is again less severe for the proposed $P_1$-stabilized method and can in fact be eliminated by postprocessing the velocity solution.

5. Conclusion

In this work, we proposed a projection-based stabilization for an equal-order discretization of incompressible flow problems. We discussed a way to obtain locally conservative velocities on dual meshes which can either be further postprocessed in an inherently parallel fashion or
directly coupled to finite-volume discretizations of transport equations. This provides a way to overcome the commonly criticized lack of local mass conservation in equal-order interpolations of incompressible flow problems. The numerical accuracy was compared to well-established equal-order schemes and the coupling to both cell- and vertex-centered finite volumes was demonstrated. Here we make the observation that the newly proposed stabilized method already delivers better results without further postprocessing, which indicates the importance of a sufficiently large kernel of the stabilization operator. By considering stabilization terms that have no effect on the macro grid pressures, we achieve better conservation properties. Finally, the postprocessing makes the resulting coupled schemes zeroth-order consistent in the way that exact mass-conservation is guaranteed.

Let us finally remark that the generalization towards other types of incompressible flow problems such as the Navier–Stokes equations is straightforward. In particular, the way in which we obtain conservative mass-fluxes is independent of the exact form of the momentum equation since it only relies on a nested mesh construction and a stabilizing operator which has the macro-level pressure space in its kernel. Using a suitable transformation of the momentum equation in terms of the mass flow rate per area $w := \rho u$, where $\rho$ denotes a spatially varying fluid density, the exact mass-conservation $\text{div } w = 0$ can even be achieved in this way for compressible Stokes models, which are widely used e.g. in mantle convection.

References


Figure 9. Results for $t = 10$ using a CC-FV method: the initial concentration is set to $c(x, t = 0) = 2$ for $x < 0.5$ and $c(x, t = 0) = 1$ otherwise. We compare the exact velocity field (top left) to an overstabilized PSPG method ($\delta = 1$, top right), a suitably stabilized PSPG method ($\delta = 1/12$, center left), the $P_0$-stabilized method ($\alpha_0 = 1$, center right), the $P_1$-stabilized method ($\alpha_1 = 1$, bottom left) and the $P_1$-stabilized method with postprocessing ($\alpha_1 = 1$, bottom right).
Figure 10. Results for $t = 10$ using a VC-FV method: the initial concentration is set to $c(x, t = 0) = 2$ for $x < 0.5$ and $c(x, t = 0) = 1$ otherwise. We compare the exact velocity field (top left) to an overstabilized PSPG method ($\delta = 1$, top right), a suitably stabilized PSPG method ($\delta = 1/12$, center left), the $P_0$-stabilized method ($\alpha_0 = 1$, center right), the $P_1$-stabilized method ($\alpha_1 = 1$, bottom left) and the $P_1$-stabilized method with postprocessing ($\alpha_1 = 1$, bottom right).