Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

An analogue of grad-div stabilization in nonconforming methods for incompressible flows

Mine Akbas[1], Alexander Linke[2], Leo G. Rebholz[3], Philipp W. Schroeder[4]

submitted: November 15, 2017

[1] Duzce University
Department of Mathematics
81620 Duzce
Turkey
E-Mail: mineakbas@duzce.edu.tr

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: alexander.linke@wias-berlin.de

[3] Clemson University
Department of Mathematical Sciences
Clemson, SC 29634
USA
E-Mail: rebholz@clemson.edu

[4] Georg-August-University Göttingen
Institute for Numerical and Applied Mathematics
37083 Göttingen
Germany
E-Mail: p.schroeder@math.uni-goettingen.de

No. 2448
Berlin 2017

2010 *Mathematics Subject Classification.* 35Q30, 65M15, 65M60, 76M10.

Key words and phrases. Incompressible Navier–Stokes equations, mixed finite element methods, grad-div stabilization, Discontinuous Galerkin method, nonconforming finite elements.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

An analogue of grad-div stabilization in nonconforming methods for incompressible flows

Mine Akbas, Alexander Linke, Leo G. Rebholz, Philipp W. Schroeder

Abstract

Grad-div stabilization is a classical remedy in conforming mixed finite element methods for incompressible flow problems, for mitigating velocity errors that are sometimes called poor mass conservation. Such errors arise due to the relaxation of the divergence constraint in classical mixed methods, and are excited whenever the spacial discretization has to deal with comparably large and complicated pressures. In this contribution, an analogue of grad-div stabilization is presented for nonconforming flow discretizations of Discontinuous Galerkin or nonconforming finite element type. Here the key is the penalization of the jumps of the normal velocities over facets of the triangulation, which controls the measure-valued part of the distributional divergence of the discrete velocity solution. Furthermore, we characterize the limit for arbitrarily large penalization parameters, which shows that the proposed nonconforming Discontinuous Galerkin methods remain robust and accurate in this limit. Several numerical examples illustrate the theory and show their relevance for the simulation of practical, nontrivial flows.

1 Introduction

Classical conforming and inf-sup stable mixed finite element methods for the incompressible (Navier–)Stokes equations, such as the mini [2], the Bernardi–Raugel [3] and the Taylor–Hood elements [43], relax the divergence constraint $\nabla \cdot \boldsymbol{u} = 0$, in order to construct optimally convergent spacial discretizations on regular unstructured triangulations [19, 4]. Indeed, the discrete velocities solutions $\boldsymbol{u}_h$ are not divergence-free, but only *discretely divergence-free*, i.e., it holds $\nabla_h \cdot \boldsymbol{u}_h \coloneqq \pi_0(\nabla \cdot \boldsymbol{u}_h) = 0$, where $\nabla_h \cdot$ denotes the discrete divergence operator and π_0 denotes the L^2 best approximation in the discrete pressure space.

While relaxing the divergence constraint facilitates the construction of inf-sup stable discretizations, it was soon realized that *discretely divergence-free* is sometimes not good enough. For example, already in 1989 D. Pelletier, A. Fortin and R. Camarero titled in their article [36] by the provocative question “Are FEM solution of incompressible flows really incompressible (or how simple flows can cause headaches!)” and pointed to the problem of *poor mass conservation*. Poor mass conservation describes velocity errors that are excited when the pressure is comparably large and complicated [23]. This happens in many flow problems, including Boussinesq flows [20, 36, 13, 18, 15, 16], potential and generalized Beltrami flows [30, 29, 23], quasi-geostrophic flows [42, 10, 29], electrophoresis [37], and two-phase flows with surface tension [17, 28]. Physical problems where poor mass conservation is not strong are rather limited, and include for example pressure-driven Stokes and pressure-driven, low-Reynolds number Navier–Stokes flow through a channel with zero exterior forcing, where for the momentum balance approximately holds $-\nu\Delta\boldsymbol{u} + \nabla p \approx \boldsymbol{0}$, i.e., where the pressure gradient is proportional to the friction force.

DOI 10.20347/WIAS.PREPRINT.2448 Berlin 2017

Phenomenologically, poor mass conservation is often accompanied by comparably large violations of the divergence error $\|\nabla \cdot \boldsymbol{u}_h\|_{L^2}$, see for example [16, Table 1]. To address this, in 1988 T. Hughes proposed to enhance the Navier–Stokes momentum balance of conforming mixed finite element methods by a consistent term [14]

$$\gamma(\nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{v}_h),$$

which penalizes large divergence errors, and is nowadays often called grad-div stabilization [33]; here $\gamma \geqslant 0$ denotes the grad-div stabilization parameter. Indeed, grad-div stabilization for conforming mixed finite element methods has recently been investigated in depth, both from a theoretical and computational point of view [32, 33, 35, 16, 34, 7, 31, 22, 1]. A better understanding of grad-div stabilization was achieved, when the limit behavior for arbitrarily large stabilization parameters $\gamma \to \infty$ was investigated [7, 31, 22]; it turned out that grad-div stabilization is not so much a stabilization, but instead a kind of penalization procedure. On a fixed grid, for $\gamma \to \infty$ the grad-div stabilized discrete velocity solution $\boldsymbol{u}_h^\gamma$ converges to a divergence-free velocity solution $\boldsymbol{u}_h^\infty$ which is the solution of a *divergence-free conforming mixed finite element method* with the same discrete velocity space, but with a richer discrete pressure space. Since divergence-free, conforming mixed finite element methods are *pressure-robust* [23, 30], i.e. their velocity error does not depend on the continuous pressure, grad-div stabilized discrete velocities behave in a more robust manner against large and complicated continuous pressures. However, this theoretical understanding also revealed limitations of grad-div stabilization in that large grad-div stabilization parameters can cause classical Poisson locking phenomena, whenever the limiting divergence-free mixed method is not inf-sup stable [22]. However, on certain mesh families and for certain conforming mixed finite elements, grad-div over-stabilization can be avoided [22, Corollary 1, Case 2].

Of course, nonconforming mixed methods like the Crouzeix–Raviart finite element method [11] are as much endangered by poor mass conservation as conforming ones, since they are not pressure-robust (their velocity error depends on the continuous pressure). In [27, 26] it was recognized that pressure-robustness of a mixed method for incompressible flows does not depend on the fact that the discrete *velocity trial functions* are divergence-free, but it only depends on the *discrete velocity test functions*. They have to be divergence-free in the *weak sense* of $\boldsymbol{H}(\mathrm{div})$, in order to be orthogonal to any gradient field in the $\boldsymbol{L}^2(\Omega)$ scalar product for vector fields [23, 27]. Recall that a vector field $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega)$ is said to be *weakly divergence-free*, if its *distributional divergence* [23]

$$\phi \mapsto -\int_\Omega \boldsymbol{v} \cdot \nabla \phi \,\mathrm{d}\boldsymbol{x}$$

vanishes for all $\phi \in C_0^\infty(\Omega)$, i.e. if it is orthogonal in $\boldsymbol{L}^2(\Omega)$ to all smooth gradient fields [23]. This shows that it is actually a very *strong* property for a vector field $\boldsymbol{v}$ to be *weakly divergence-free*, at least compared to being only *discretely divergence-free*. Indeed, applying the general definition of the distributional divergence to nonconforming finite element methods and Discontinuous Galerkin (DG) methods, it turns out that the distributional divergence of a velocity test function $\boldsymbol{v}_h$ is given by

$$\phi \mapsto \sum_{K \in \mathcal{T}_h} \int_K \phi \, \nabla \cdot \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} - \sum_{F \in \mathcal{F}_h^i} \oint_F \phi([\![\boldsymbol{v}_h]\!] \cdot \boldsymbol{n}_F) \,\mathrm{d}\boldsymbol{s},$$

since $\boldsymbol{v}_h$ is elementwise polynomial and an integration by parts can be applied. Therefore, the distributional divergence of $\boldsymbol{v}_h$ vanishes only if it holds $\nabla \cdot \boldsymbol{v}_h = 0$ elementwise for all $K \in \mathcal{T}_h$ and $[\![\boldsymbol{v}_h]\!] \cdot \boldsymbol{n}_F = 0$ for all $F \in \mathcal{F}_h^i$. Instead, if a vector field $\boldsymbol{v}_h$ is *discretely divergence-free*, this implies

for usual, i.e., not pressure-robust [23], discretizations that only

$$\phi_h \mapsto \sum_{K\in\mathcal{T}_h} \int_K \phi_h \nabla\cdot \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} - \sum_{F\in\mathcal{F}_h^i} \oint_F \phi_h([\![\boldsymbol{v}_h]\!]\cdot\boldsymbol{n}_F)\,\mathrm{d}\boldsymbol{s},$$

vanishes for all $\phi_h \in Q_h$ from a *finite-dimensional* space Q_h of pressure test functions. Hence, an analogue for grad-div stabilization for nonconforming methods has to penalize the elementwise defined broken divergence $\nabla_h\cdot\boldsymbol{v}_h$, *and* the facet jumps $[\![\boldsymbol{v}_h]\!]\cdot\boldsymbol{n}_F$ of the normal velocities (i.e. the mass flux). Therefore, in the following a penalization procedure with a penalization parameter γ is proposed, which enforces in the limit $\gamma\to\infty$ — on a fixed grid — that the discrete velocity solution $\boldsymbol{u}_h^\infty$ will be elementwise divergence-free and the jumps of the normal velocities over inner facets will vanish, likewise. Then, this discrete velocity solution $\boldsymbol{u}_h^\infty$ will be shown to be the discrete velocity solution of another DG method, where also the *velocity test functions* $\boldsymbol{v}_h$ are weakly divergence-free, and therefore this discrete velocity solution will be pressure-robust, i.e., the discrete velocity solution is pressure-independent [23].

In the context of the Crouzeix–Raviart finite element method, the importance of penalizing the jumps of the normal velocities was recognized by E. Burman and P. Hansbo in 2005 [6]. However, they did not investigate possible problems of over-stabilization in the case of small kinematic viscosities, if the jumps of the normal velocities are strongly penalized. In the context of DG methods, in [41] an elementwise divergence penalization has been used (seemingly) for the first time. It was shown that even this incomplete stabilization method can improve poor mass conservation for both inf-sup stable $\mathbb{Q}_k^{\mathrm{dc}}/\mathbb{Q}_{k-1}^{\mathrm{dc}}$ and equal-order DG methods $\mathbb{Q}_k^{\mathrm{dc}}/\mathbb{Q}_k^{\mathrm{dc}}$ on tensor-product meshes in practical applications of non-isothermal flows. Further, it was recognized independently in the recent works [25] and [24] that in DG methods, the mass balance across interior facets has to be accounted for, in addition to the classical (broken) grad-div stabilization. For laminar and turbulent flows, the authors of Ref. [25] — improving earlier results [24] — use a velocity-correction time integration with equal-order $\mathbb{Q}_k^{\mathrm{dc}}/\mathbb{Q}_k^{\mathrm{dc}}$ elements on tensor-product meshes and an implementation in `deal.ii`. The additional terms they favor are different from the stabilization that is proposed in this work, see [25, 4.2.3 and 4.2.4]. Further, the authors do not rigorously justify their approach by numerical analysis and do not consider a possible lack of robustness against over-stabilization.

A major contribution of this work is to propose an analogue of grad-div stabilization for DG methods as motivated above. We will prove error estimates which show that the same positive results for grad-div stabilization in conforming methods also hold in this setting with these penalizations. Moreover, for DG methods using the pair $\mathbb{P}_k^{\mathrm{dc}}/\mathbb{P}_{k-1}^{\mathrm{dc}}$ on simplices, we are able to characterize the limit behavior of the penalization procedure for arbitrarily large parameters. It turns out that the method is perfectly robust against over-stabilization since in the limit, one obtains the weakly divergence-free inf-sup stable DG methods proposed by B. Cockburn, G. Kanschat, and D. Schötzau [8, 9]. Further, some remarks are made concerning DG methods on tensor product meshes, and the nonconforming Crouzeix–Raviart element. Our theory is restricted to the incompressible Stokes problem, however, an extension to the Navier–Stokes problem would be straightforward, although naturally with more technical details. Several numerical examples for the Stokes problem, the Navier–Stokes problem and some coupled Boussinesq flows show the practical relevance of the proposed penalization procedure.

Organization of the article This article is arranged as follows. In Section 2 we provide some notation and mathematical preliminaries to allow for a cleaner analysis to follow. Section 3 considers the

case of the proposed stabilization for DG on simplices, Section 4 considers tensor product meshes, and Section 5 considers the enhancement in Crouzeix–Raviart elements. Several numerical tests of concept are given in Sections 3 and 4, and in Section 6, we consider applications of the proposed penalization outside of the Stokes setting, to Navier–Stokes equations and to Boussinesq. In all numerical tests the (sometimes dramatic) improvement offered by the proposed penalization is clear. Finally, conclusions are drawn in Section 7, and future research directions are discussed.

2 Stokes problem and DG setting

We consider a domain $\Omega \subset \mathbb{R}^d$, d=2,3, to be a simply connected set with smooth boundary, or a convex polygon. For $K \subseteq \Omega$ we use the standard Sobolev spaces $W^{m,p}(K)$ for scalar-valued functions with associated norms $\|\cdot\|_{W^{m,p}(K)}$ and seminorms $|\cdot|_{W^{m,p}(K)}$ for $m \geqslant 0$ and $p \geqslant 1$. Spaces and norms for vector- and tensor-valued functions are indicated with bold letters. We use the Lebesgue space $L^p(K) = W^{0,p}(K)$ and the Hilbert space $H^m(K) = W^{m,2}(K)$. Additionally, the closed subspaces $H^1_0(K)$ consisting of $H^1(K)$-functions with vanishing trace on ∂K and the set $L^2_0(K)$ of $L^2(K)$-functions with zero mean in K play an important role. The $L^2(K)$-inner product is denoted by $(\cdot,\cdot)_K$ and, if $K = \Omega$, the domain is omitted completely when no confusion can arise.

2.1 Continuous Stokes problem

We consider the stationary Stokes problem with no-slip boundary conditions:

$$\begin{cases} -\nu\Delta\boldsymbol{u} + \nabla p = \boldsymbol{f} & \text{in } \Omega, \\ \nabla\cdot\boldsymbol{u} = 0 & \text{in } \Omega, \\ \boldsymbol{u} = \boldsymbol{0} & \text{on } \partial\Omega. \end{cases} \tag{1}$$

With $\boldsymbol{V} = \boldsymbol{H}^1_0(\Omega)$ and $Q = L^2_0(\Omega)$, the weak formulation of (1) reads: Find $(\boldsymbol{u}, p) \in \boldsymbol{V} \times Q$ s.t., $\forall\,(\boldsymbol{v}, q) \in \boldsymbol{V} \times Q$,

$$\nu a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) - b(\boldsymbol{u}, q) = (\boldsymbol{f}, \boldsymbol{v}). \tag{2}$$

The bilinear forms are given by

$$a(\boldsymbol{w}, \boldsymbol{v}) = \int_\Omega \nabla\boldsymbol{w} : \nabla\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} \quad \text{and} \quad b(\boldsymbol{w}, q) = -\int_\Omega q(\nabla\cdot\boldsymbol{w})\,\mathrm{d}\boldsymbol{x}. \tag{3}$$

Weakly divergence-free velocities belong to

$$\boldsymbol{V}^{\mathrm{div}} = \{\boldsymbol{v} \in \boldsymbol{V} \colon b(\boldsymbol{v}, q) = 0,\ \forall\, q \in Q\}. \tag{4}$$

2.2 Discontinuous Galerkin setting

Let $\mathcal{T}_h$ be a shape-regular FE partition of Ω without hanging nodes and mesh size $h = \max_{K\in\mathcal{T}_h} h_K$, where h_K denotes the diameter of the particular element $K \in \mathcal{T}_h$. Since the subsequent velocity approximation will not be $\boldsymbol{H}^1$-conforming, the broken Sobolev space is introduced as follows:

$$\boldsymbol{H}^m(\mathcal{T}_h) = \left\{\boldsymbol{w} \in \boldsymbol{L}^2(\Omega) \colon \boldsymbol{w}\big|_K \in \boldsymbol{H}^m(K),\ \forall\, K \in \mathcal{T}_h\right\}. \tag{5}$$

Define the broken gradient $\nabla_h\colon \boldsymbol{H}^1(\mathcal{T}_h) \to \boldsymbol{L}^2(\Omega)$ by

$$(\nabla_h \boldsymbol{w})\big|_K \coloneqq \nabla\big(\boldsymbol{w}\big|_K\big),$$

and similarly define the broken divergence. We additionally introduce the space

$$\boldsymbol{H}(\mathrm{div};\Omega) = \big\{\boldsymbol{w} \in \boldsymbol{L}^2(\Omega)\colon\ \nabla\cdot\boldsymbol{w} \in L^2(\Omega),\ \boldsymbol{w}\cdot\boldsymbol{n}\big|_{\partial\Omega} = 0\big\}. \tag{6}$$

In our context it is worth to remind the reader that the expression $\nabla\cdot\boldsymbol{w} \in L^2(\Omega)$ has the meaning that the *distributional divergence* of $\boldsymbol{w}$ can be expressed as a $L^2(\Omega)$ function, i.e., there exists $s \in L^2(\Omega)$ (called the weak divergence of $\boldsymbol{w}$) such that it holds for all $\phi \in C_0^\infty(\Omega)$ $-\int_\Omega \boldsymbol{v}\cdot\nabla\phi\,\mathrm{d}\boldsymbol{x} = \int_\Omega s\,\phi\,\mathrm{d}\boldsymbol{x}$, see [23].

The skeleton $\mathcal{F}_h$ denotes the set of all facets with $\mathcal{F}_K = \{F \in \mathcal{F}_h\colon F \subset \partial K\}$ and h_F represents the diameter of each facet $F \in \mathcal{F}_h$. Note that $h_F \leqslant h_K$ holds true for all $F \in \mathcal{F}_K$ and additionally, we define $N_\partial = \max_{K\in\mathcal{T}_h} \mathrm{card}(\mathcal{F}_K)$. Moreover, $\mathcal{F}_h = \mathcal{F}_h^i \cup \mathcal{F}_h^\partial$ where $\mathcal{F}_h^i$ is the subset of interior facets and $\mathcal{F}_h^\partial$ collects all boundary facets $F \subset \partial\Omega$. To any $F \in \mathcal{F}_h$ we assign a unit normal vector $\boldsymbol{n}_F$ where, for $F \in \mathcal{F}_h^\partial$, this is the outer unit normal vector $\boldsymbol{n}$. If $F \in \mathcal{F}_h^i$, there are two adjacent elements K^+ and K^- sharing the facet $F = \overline{\partial K^+} \cap \overline{\partial K^-}$ and $\boldsymbol{n}_F$ points in an arbitrary but fixed direction. Let ϕ be any piecewise smooth (scalar-, vector- or tensor-valued) function with traces from within the interior of $K^\pm$ denoted by $\phi^\pm$, respectively. Then, we define the jump $[\![\cdot]\!]_F$ and average $\big\{\!\!\big\{\cdot\big\}\!\!\big\}_F$ operator across interior facets $F \in \mathcal{F}_h^i$ by

$$[\![\phi]\!]_F = \phi^+ - \phi^- \quad \text{and} \quad \big\{\!\!\big\{\phi\big\}\!\!\big\}_F = \frac{1}{2}\big(\phi^+ + \phi^-\big). \tag{7}$$

For boundary facets $F \in \mathcal{F}_h^\partial$ we set $[\![\phi]\!]_F = \big\{\!\!\big\{\phi\big\}\!\!\big\}_F = \phi$. These operators act componentwise for vector- and tensor-valued functions. Frequently, the subscript indicating the facet is omitted. Note that if $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{div};\Omega)$, then $[\![\boldsymbol{w}]\!]\cdot\boldsymbol{n}_F = 0$ for all $F \in \mathcal{F}_h^i$; cf. [12, Lemma 1.24]. This is why the proposed stabilization of DG methods in this work is also sometimes called '$\boldsymbol{H}(\mathrm{div})$-stabilization'.

3 Mass flux penalization applied to inf-sup stable DG methods on simplicial meshes

In the following, $\mathbb{P}_k(K)$ (vector-valued: $\mathbb{P}_k(K)$) denotes the space of all polynomials on K with degree less or equal to k. Restricting ourselves to simplicial meshes, the global discrete spaces are

$$\boldsymbol{V}_h = \big\{\boldsymbol{v}_h \in \boldsymbol{L}^2(\Omega)\colon \boldsymbol{v}_h\big|_K \in \mathbb{P}_k(K),\ \forall\, K \in \mathcal{T}_h\big\}, \tag{8a}$$

$$Q_h = \big\{q_h \in L^2_0(\Omega)\colon q_h\big|_K \in \mathbb{P}_{k-1}(K),\ \forall\, K \in \mathcal{T}_h\big\}. \tag{8b}$$

This finite element (FE) pair is also called $\mathbb{P}_k^{\mathrm{dc}}/\mathbb{P}_{k-1}^{\mathrm{dc}}$ and it forms an inf-sup stable velocity/pressure pair; cf., for example, [38, Section 6.4]. Moreover, on simplicial meshes, we have the important property $\nabla_h\cdot\mathbb{P}_k^{\mathrm{dc}} \subset \mathbb{P}_{k-1}^{\mathrm{dc}}$.

Remark 3.1: On tensor-product meshes, which we consider in the next section, the inf-sup stable pressure/velocity pair $\mathbb{Q}_k^{\mathrm{dc}}/\mathbb{Q}_{k-1}^{\mathrm{dc}}$ is the common element choice. A key departure for such elements from the simplicial mesh framework is that $\nabla_h\cdot\mathbb{Q}_k^{\mathrm{dc}} \not\subset \mathbb{Q}_{k-1}^{\mathrm{dc}}$, which creates an additional obstacle to handle in the analysis since here the additional penalization of the broken divergence is needed.

On $\boldsymbol{V}_h$ the following discrete trace inequality is valid; cf. [12, Remark 1.47]:

$$\forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h\colon \quad \|\boldsymbol{v}_h\|^2_{\boldsymbol{L}^2(\partial K)} \leqslant C_{\mathrm{tr}} N_\partial h_K^{-1} \|\boldsymbol{v}_h\|^2_{\boldsymbol{L}^2(K)}\,, \quad \forall\, K \in \mathcal{T}_h. \tag{9}$$

A similar trace inequality holds true for the pressure space Q_h. The appearance below of certain traces of velocity facet values, and their normal derivatives, leads to the technical assumption for the proofs below that the involved velocities belong (at least) to $\boldsymbol{H}^{\frac{3}{2}+\varepsilon}(\mathcal{T}_h)$ for some $\varepsilon > 0$; cf. [38, Section 2.1.3]. Relaxing this assumption is possible [21], of course, but beyond the scope of this contribution. We thus define the compound space

$$\boldsymbol{V}(h) = \boldsymbol{V}_h \oplus \Big[\boldsymbol{V} \cap \boldsymbol{H}^{\frac{3}{2}+\varepsilon}(\mathcal{T}_h)\Big]. \tag{10}$$

We consider the symmetric interior penalty (SIP) method with σ sufficiently large to guarantee the coercivity estimates below, and so define the bilinear form

$$\begin{aligned} a_h(\boldsymbol{w}_h, \boldsymbol{v}_h) = &\int_\Omega \nabla_h \boldsymbol{w}_h : \nabla_h \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} + \sum_{F\in\mathcal{F}_h} \frac{\sigma}{h_F} \oint_F [\![\boldsymbol{w}_h]\!] \cdot [\![\boldsymbol{v}_h]\!] \,\mathrm{d}\boldsymbol{s} \\ &- \sum_{F\in\mathcal{F}_h} \oint_F \{\!\{\nabla \boldsymbol{w}_h\}\!\} \boldsymbol{n}_F \cdot [\![\boldsymbol{v}_h]\!] \,\mathrm{d}\boldsymbol{s} - \sum_{F\in\mathcal{F}_h} \oint_F [\![\boldsymbol{w}_h]\!] \cdot \{\!\{\nabla \boldsymbol{v}_h\}\!\} \boldsymbol{n}_F \,\mathrm{d}\boldsymbol{s}. \end{aligned} \tag{11}$$

The natural discrete energy norm corresponding to the SIP bilinear form for $\boldsymbol{w} \in \boldsymbol{V}(h)$ is given by

$$|||\boldsymbol{w}|||_e^2 = \|\nabla_h \boldsymbol{w}\|^2_{\boldsymbol{L}^2(\Omega)} + \sum_{F\in\mathcal{F}_h} \frac{\sigma}{h_F} \|[\![\boldsymbol{w}]\!]\|^2_{\boldsymbol{L}^2(F)}\,. \tag{12}$$

The discrete bilinear form for the pressure-velocity coupling is defined by

$$b_h(\boldsymbol{w}_h, q_h) = -\int_\Omega q_h (\nabla_h \cdot \boldsymbol{w}_h) \,\mathrm{d}\boldsymbol{x} + \sum_{F\in\mathcal{F}_h} \oint_F \{\!\{q_h\}\!\} ([\![\boldsymbol{w}_h]\!] \cdot \boldsymbol{n}_F) \,\mathrm{d}\boldsymbol{s}. \tag{13}$$

As mentioned above, the FE pair $\boldsymbol{V}_h/Q_h$ is discretely inf-sup stable. More precisely, there exists an even smaller discrete velocity space $\boldsymbol{W}_h \subset \boldsymbol{V}_h$, with $\boldsymbol{W}_h \subset \boldsymbol{H}(\mathrm{div};\Omega)$, such that $\boldsymbol{W}_h/Q_h$ is also inf-sup stable; cf., for example, [38, Section 6.4]. This ensures the existence of a $\beta^* > 0$, independent of h, such that

$$\beta^* \|q_h\|_{L^2(\Omega)} \leqslant \sup_{\boldsymbol{w}_h\in\boldsymbol{W}_h\setminus\{\boldsymbol{0}\}} \frac{b_h(\boldsymbol{w}_h, q_h)}{|||\boldsymbol{w}_h|||_e} \leqslant \sup_{\boldsymbol{v}_h\in\boldsymbol{V}_h\setminus\{\boldsymbol{0}\}} \frac{b_h(\boldsymbol{v}_h, q_h)}{|||\boldsymbol{v}_h|||_e}, \quad \forall\, q_h \in Q_h. \tag{14}$$

We will also utilize a stronger energy norm on $\boldsymbol{V}(h)$:

$$\forall\, \boldsymbol{w} \in \boldsymbol{V}(h)\colon \quad |||\boldsymbol{w}|||^2_{e,\sharp} := |||\boldsymbol{w}|||_e^2 + \sum_{K\in\mathcal{T}_h} h_K \|\nabla \boldsymbol{w} \cdot \boldsymbol{n}_K\|^2_{\boldsymbol{L}^2(\partial K)}\,. \tag{15}$$

Then, there exists a $M > 0$, independent of h, such that

$$\forall\, (\boldsymbol{w}, \boldsymbol{v}_h) \in \boldsymbol{V}(h) \times \boldsymbol{V}_h\colon \quad a_h(\boldsymbol{w}, \boldsymbol{v}_h) \leqslant M |||\boldsymbol{w}|||_{e,\sharp} |||\boldsymbol{v}_h|||_e. \tag{16}$$

Concerning a proof, see for example [12, Section 4.2.3] for a scalar-valued analogue. Moreover, note that the $|||\cdot|||_e$ and $|||\cdot|||_{e,\sharp}$ norms are uniformly equivalent on $\boldsymbol{V}_h$. That is, there exists a $C > 0$ such that $C|||\boldsymbol{v}_h|||_{e,\sharp} \leqslant |||\boldsymbol{v}_h|||_e \leqslant |||\boldsymbol{v}_h|||_{e,\sharp}$ for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$; cf. [12, Lemma 4.20] (scalar-valued).

The key idea of this work is to add the following mass flux penalization term to the DG formulation:

$$j_h(\boldsymbol{w}_h, \boldsymbol{v}_h) = \sum_{F\in\mathcal{F}_h} \frac{1}{h_F} \oint_F (\llbracket \boldsymbol{w}_h \rrbracket \cdot \boldsymbol{n}_F)(\llbracket \boldsymbol{v}_h \rrbracket \cdot \boldsymbol{n}_F) \,\mathrm{d}\boldsymbol{s}. \tag{17}$$

We will show the remarkable positive impact this term can have, as it improves mass conservation as well as the pressure-robustness of the solution. In fact, the analytical and numerical results we obtain are similar to what is found with using grad-div stabilization in conforming methods for Stokes problems, which is why we characterize this penalization as an analogue to grad-div stabilization for nonconforming methods. This term penalizes normal jumps and therefore, roughly speaking, the difference between a fully discontinuous DG velocity and a normal-continuous $\boldsymbol{H}(\mathrm{div})$ velocity. However, we emphasize an important advantage over grad-div stabilization in conforming methods: in the proposed DG setting, over-stabilization is never possible, as it will be shown. For a discussion on the issue of over-stabilization in conforming methods with grad-div stabilization, see [22, Corollary 1].

For all $\boldsymbol{v}_h \in \boldsymbol{V}_h$ we introduce the notation

$$|\boldsymbol{v}_h|_{\mathrm{nj}}^2 = \sum_{F\in\mathcal{F}_h} \frac{1}{h_F} \oint_F (\llbracket \boldsymbol{v}_h \rrbracket \cdot \boldsymbol{n}_F)^2 \,\mathrm{d}\boldsymbol{s} = \sum_{F\in\mathcal{F}_h} \frac{1}{h_F} \left\| \llbracket \boldsymbol{v}_h \rrbracket \cdot \boldsymbol{n}_F \right\|_{L^2(F)}^2,$$

and note that $|\boldsymbol{v}_h|_{\mathrm{nj}} \leqslant |||\boldsymbol{v}_h|||_e$ for all $\boldsymbol{v}_h \in \boldsymbol{V}(h)$.

For approximating (2), the proposed DG method with mass flux penalization is given by: Find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ satisfying $\forall\, (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h \times Q_h$,

$$\nu a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) + b_h(\boldsymbol{v}_h, p_h) + \gamma j_h(\boldsymbol{u}_h, \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h), \tag{18}$$

$$-b_h(\boldsymbol{u}_h, q_h) = 0, \tag{19}$$

where $\gamma \geqslant 0$ is the penalization parameter.

Discretely divergence-free DG velocities belong to

$$\boldsymbol{V}_h^{\mathrm{div}} = \{\boldsymbol{v}_h \in \boldsymbol{V}_h \colon b_h(\boldsymbol{v}_h, q_h) = 0,\ \forall\, q_h \in Q_h\},$$

and thanks to the inf-sup condition, an equivalent and pressure-free formulation of (18)–(19) can be expressed as:

$$\boldsymbol{u}_h \in \boldsymbol{V}_h^{\mathrm{div}} \text{ s.t. } \nu a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) + \gamma j_h(\boldsymbol{u}_h, \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h)\ \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h^{\mathrm{div}}.$$

3.1 Energy estimate

Making use of the discrete coercivity property of the SIP bilinear form a_h, cf., for example, [38, Lemma 6.6] or [12, Section 6.1.2.1], we also easily obtain discrete coercivity with a constant $C_\sigma > 0$, independent of h, in the following sense:

$$\forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h \colon \quad a_h(\boldsymbol{v}_h, \boldsymbol{v}_h) + \gamma j_h(\boldsymbol{v}_h, \boldsymbol{v}_h) \geqslant C_\sigma |||\boldsymbol{v}_h|||_e^2 + \gamma |\boldsymbol{v}_h|_{\mathrm{nj}}^2 \geqslant C_\sigma |||\boldsymbol{v}_h|||_e^2. \tag{20}$$

The well-posedness of the formulation thus follows using this discrete coercivity property and discrete inf-sup stability. The following energy estimate follows immediately as well.

LEMMA 3.2 (*Energy estimate*)

Let $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and assume that $\sigma > 0$ is sufficiently large to guarantee discrete coercivity. Then, with a constant $C > 0$, one obtains the following estimate for the FEM solution $(\boldsymbol{u}_h, p_h)$ to (18)–(19)*:*

$$\frac{\nu C_\sigma}{2} |||\boldsymbol{u}_h|||_e^2 + \gamma |\boldsymbol{u}_h|_{\mathrm{nj}}^2 \leqslant C\nu^{-1} \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}^2 , \tag{21a}$$

$$\|\nabla_h \cdot \boldsymbol{u}_h\|_{L^2(\Omega)}^2 \leqslant C\gamma^{-1}\nu^{-1} \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}^2 , \tag{21b}$$

$$\|p_h\|_{L^2(\Omega)}^2 \leqslant C \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}^2 . \tag{21c}$$

PROOF: Testing with $(\boldsymbol{v}_h, q_h) = (\boldsymbol{u}_h, p_h)$ in (18)–(19), together with coercivity from (20) and Cauchy–Schwarz leads to

$$\nu C_\sigma |||\boldsymbol{u}_h|||_e^2 + \gamma |\boldsymbol{u}_h|_{\mathrm{nj}}^2 \leqslant \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)} \|\boldsymbol{u}_h\|_{\boldsymbol{L}^2(\Omega)} . \tag{22}$$

Further estimating the right-hand side requires a DG analogue of the Poincaré–Friedrichs (PF) inequality; cf., for example, [12, Corollary 5.4]. Then, Young's inequality can be invoked to obtain

$$\|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)} \|\boldsymbol{u}_h\|_{\boldsymbol{L}^2(\Omega)} \leqslant \frac{1}{2} \frac{C_{\mathrm{PF}}}{\nu C_\sigma} \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\nu C_\sigma}{2} |||\boldsymbol{u}_h|||_e^2 . \tag{23}$$

Reordering shows the first bound. For an estimate of the divergence of the DG solution, choose $q_h = \nabla_h \cdot \boldsymbol{u}_h$ in (19). After an additional application of Cauchy–Schwarz and inserting $\frac{h_F}{\gamma} \frac{\gamma}{h_F} = 1$, we have

$$\begin{aligned} \|\nabla_h \cdot \boldsymbol{u}_h\|_{L^2(\Omega)}^2 &= \sum_{F \in \mathcal{F}_h} \oint_F \left\{\!\!\left\{ \nabla \cdot \boldsymbol{u}_h \right\}\!\!\right\} (\llbracket \boldsymbol{u}_h \rrbracket \cdot \boldsymbol{n}_F) \, \mathrm{d}\boldsymbol{s} \\ &\leqslant \left(\sum_{F \in \mathcal{F}_h} \frac{h_F}{\gamma} \left\| \left\{\!\!\left\{ \nabla \cdot \boldsymbol{u}_h \right\}\!\!\right\} \right\|_{L^2(F)}^2 \right)^{1/2} \left(\sum_{F \in \mathcal{F}_h} \frac{\gamma}{h_F} \|\llbracket \boldsymbol{u}_h \rrbracket \cdot \boldsymbol{n}_F\|_{L^2(F)}^2 \right)^{1/2} . \end{aligned} \tag{24}$$

For any $q_h \in Q_h$, the discrete trace inequality (9) yields

$$\begin{aligned} \sum_{F \in \mathcal{F}_h} \oint_F \left| \left\{\!\!\left\{ q_h \right\}\!\!\right\} \right|^2 \mathrm{d}\boldsymbol{s} &\leqslant \sum_{F \in \mathcal{F}_h} \left[\left\| q_h^+ \right\|_{L^2(F)}^2 + \left\| q_h^- \right\|_{L^2(F)}^2 \right] \\ &\leqslant \sum_{K \in \mathcal{T}_h} \|q_h\|_{L^2(\partial K)}^2 \\ &\leqslant \sum_{K \in \mathcal{T}_h} C_{\mathrm{tr}} N_\partial h_K^{-1} \|q_h\|_{L^2(K)}^2 . \end{aligned}$$

Thus, with a generic constant $C > 0$, using $h_F \leqslant h_K$ for all $F \in \mathcal{F}_K$, we obtain

$$\left(\sum_{F \in \mathcal{F}_h} \frac{h_F}{\gamma} \left\| \left\{\!\!\left\{ \nabla \cdot \boldsymbol{u}_h \right\}\!\!\right\} \right\|_{L^2(F)}^2 \right)^{1/2} \leqslant C\gamma^{-1/2} \|\nabla_h \cdot \boldsymbol{u}_h\|_{L^2(\Omega)} . \tag{25}$$

For the second term, we use the energy estimate (21a) for the velocity; that is,

$$\left(\sum_{F\in\mathcal{F}_h}\frac{\gamma}{h_F}\left\|[\![\boldsymbol{u}_h]\!]\cdot\boldsymbol{n}_F\right\|^2_{L^2(F)}\right)^{1/2}=\gamma^{1/2}|\boldsymbol{u}_h|_{\mathrm{nj}}\leqslant C\nu^{-1/2}\left\|\boldsymbol{f}\right\|_{\boldsymbol{L}^2(\Omega)}. \tag{26}$$

Using (26) and (25) in (24) shows the second bound of the lemma.

For the pressure bound, we invoke the discrete inf-sup condition (14) in $\boldsymbol{W}_h/Q_h$ and obtain

$$\beta^*\left\|p_h\right\|_{L^2(\Omega)}\leqslant\sup_{\boldsymbol{w}_h\in\boldsymbol{W}_h\setminus\{\boldsymbol{0}\}}\frac{b_h(\boldsymbol{w}_h,p_h)}{|||\boldsymbol{w}_h|||_e}=\sup_{\boldsymbol{w}_h\in\boldsymbol{W}_h\setminus\{\boldsymbol{0}\}}\left[\frac{(\boldsymbol{f},\boldsymbol{w}_h)}{|||\boldsymbol{w}_h|||_e}-\nu\frac{a_h(\boldsymbol{u}_h,\boldsymbol{w}_h)}{|||\boldsymbol{w}_h|||_e}\right]. \tag{27}$$

Here, the fact that $j_h(\boldsymbol{u}_h,\boldsymbol{w}_h)=0$ since $\boldsymbol{w}_h\in\boldsymbol{W}_h\subset\boldsymbol{H}(\mathrm{div};\Omega)$ has been used. A further estimation of the right-hand side uses Cauchy–Schwarz and Poincaré–Friedrichs for the term involving $\boldsymbol{f}$ together with boundedness of a_h (16) and uniform equivalence of $|||\cdot|||_e$ and $|||\cdot|||_{e,\sharp}$ on $\boldsymbol{V}_h$ for the viscous term. Thus, the supremum vanishes thereby leading to

$$\beta^*\left\|p_h\right\|_{L^2(\Omega)}\leqslant C[\left\|\boldsymbol{f}\right\|_{\boldsymbol{L}^2}+\nu|||\boldsymbol{u}_h|||_e]. \tag{28}$$

Inserting the energy estimate (21a) for the velocity concludes the proof. ■

3.2 Error estimate

For the error analysis that follows, it is important that the following property holds. A proof is straightforward and based on the consistency of both the DG method and $\boldsymbol{H}(\mathrm{div})$-stabilization.

COROLLARY 3.3 (*Galerkin orthogonality*)

Let $(\boldsymbol{u},p)\in\boldsymbol{V}\times Q$ solve (2) *and $(\boldsymbol{u}_h,p_h)\in\boldsymbol{V}_h\times Q_h$ solve* (18)–(19)*. If additionally $\boldsymbol{u}\in\boldsymbol{H}^{\frac{3}{2}+\varepsilon}(\mathcal{T}_h)$ and $p\in\boldsymbol{H}^{\frac{1}{2}+\varepsilon}(\mathcal{T}_h)$ for $\varepsilon>0$, then, for all $(\boldsymbol{v}_h,q_h)\in\boldsymbol{V}_h\times Q_h$:*

$$\nu a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{v}_h)+b_h(\boldsymbol{v}_h,p-p_h)-b_h(\boldsymbol{u}-\boldsymbol{u}_h,q_h)+\gamma j_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{v}_h)=0. \tag{29}$$

Under the assumptions of the previous corollary we decompose the error as

$$\begin{aligned}\boldsymbol{u}-\boldsymbol{u}_h&=(\boldsymbol{u}-\boldsymbol{\pi}_h\boldsymbol{u})-(\boldsymbol{u}_h-\boldsymbol{\pi}_h\boldsymbol{u})=\boldsymbol{\eta}^{\boldsymbol{u}}-\boldsymbol{e}_h^{\boldsymbol{u}},\\ p-p_h&=(p-\pi_0 p)-(p_h-\pi_0 p)=\eta^p-e_h^p,\end{aligned}$$

where $(\boldsymbol{\pi}_h,\pi_0)\colon\boldsymbol{V}\times Q\to\boldsymbol{V}_h\times Q_h$ represent appropriate approximation operators and we refer to $(\boldsymbol{\eta}^{\boldsymbol{u}},\eta^p)$ and $(\boldsymbol{e}_h^{\boldsymbol{u}},e_h^p)$ as approximation and discretization error, respectively. For the pressure, π_0 simply denotes the local L^2-projection onto Q_h. For the approximation operator for the velocity, we require $\boldsymbol{\pi}_h\colon\boldsymbol{V}\to\boldsymbol{W}_h\subset\boldsymbol{V}_h$, and recall from above that $\boldsymbol{W}_h\subset\boldsymbol{H}(\mathrm{div};\Omega)$. It is well-known that functions belonging to $\boldsymbol{H}(\mathrm{div};\Omega)$ have a continuous normal component across interior facets. Specifically, we define the operator $\boldsymbol{\pi}_h$ to be the Brezzi–Douglas–Marini (BDM) interpolation operator of order k; cf. [4, Sections 2.3 and 2.5]. A very important property of the BDM$_k$ interpolator $\boldsymbol{\pi}_h$ is the following commuting diagram property:

$$\forall\,\boldsymbol{w}\in\boldsymbol{H}(\mathrm{div};\Omega)\colon\qquad\nabla\cdot(\boldsymbol{\pi}_h\boldsymbol{w})=\pi_0(\nabla\cdot\boldsymbol{w}). \tag{30}$$

Note that for the Stokes velocity $\boldsymbol{u}$, we have that $\nabla\cdot\boldsymbol{u}=0$ holds weakly, which implies that $\nabla\cdot(\boldsymbol{\pi}_h\boldsymbol{u})=0$ also holds weakly. This property is important for the following theorem.

THEOREM 3.4 (*Error estimate*)

Let $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and assume that $\sigma > 0$ is sufficiently large to guarantee discrete coercivity. Under the assumptions of the previous corollary, the following error estimate holds true:

$$|||\boldsymbol{u} - \boldsymbol{u}_h|||_e \leqslant C\Big[|||\boldsymbol{\eta}^{\boldsymbol{u}}|||_{e,\sharp} + \gamma^{-1/2}\nu^{-1/2}\,\|\eta^p\|_{L^2(\Omega)}\Big] \tag{31a}$$

$$\|\pi_0 p - p_h\|_{L^2(\Omega)} \leqslant C\nu|||\boldsymbol{u} - \boldsymbol{u}_h|||_{e,\sharp} \tag{31b}$$

$$\|p - p_h\|_{L^2(\Omega)} \leqslant C\nu|||\boldsymbol{\eta}^{\boldsymbol{u}}|||_{e,\sharp} + \left[C\sqrt{\frac{\nu}{\gamma}} + 1\right]\|\eta^p\|_{L^2(\Omega)} \tag{31c}$$

REMARK 3.5: Due to the fact that the BDM interpolator has optimal approximation properties, one obtains the standard convergence rate of h^k whenever the exact solution $(\boldsymbol{u}, p)$ is smooth enough. Thus, the overall convergence rate of the proposed stabilized method remains the same as the unstabilized method, see e.g. [38]. Moreover, the accuracy of the discretization is robust and optimal with respect to the limit $\gamma \to \infty$, i.e., over-stabilization is not possible.

REMARK 3.6: Although the mass flux penalization does not alter the spacial convergence rate, it can dramatically lower the velocity error for certain flows. In particular, in the unstabilized DG method, the coefficient of the pressure error term in the velocity estimate is ν^{-1} [38], whereas we obtain $\gamma^{-1/2}\nu^{-1/2}$ due to the penalization. This is analogous to the effect of using grad-div stabilization in conforming methods, and as we show below, can dramatically reduce the velocity error when ν is small and/or the pressure is large. Further, due to (31b) in the limit $\gamma \to \infty$ also the discrete pressure is pressure-robust in some sense, see [23, Remark 4.5].

PROOF: Testing with $(\boldsymbol{v}_h, q_h) = (\boldsymbol{e}_h^{\boldsymbol{u}}, e_h^p) \in \boldsymbol{V}_h^{\mathrm{div}} \times Q_h$ in Corollary 3.3, inserting the error splitting and reordering leads to

$$\begin{aligned}\nu a_h(\boldsymbol{e}_h^{\boldsymbol{u}}, \boldsymbol{e}_h^{\boldsymbol{u}}) + \gamma j_h(\boldsymbol{e}_h^{\boldsymbol{u}}, \boldsymbol{e}_h^{\boldsymbol{u}}) + b_h(\boldsymbol{\eta}^{\boldsymbol{u}}, e_h^p) - b_h(\boldsymbol{e}_h^{\boldsymbol{u}}, e_h^p) + b_h(\boldsymbol{e}_h^{\boldsymbol{u}}, e_h^p)\\ = \nu a_h(\boldsymbol{\eta}^{\boldsymbol{u}}, \boldsymbol{e}_h^{\boldsymbol{u}}) + \gamma j_h(\boldsymbol{\eta}^{\boldsymbol{u}}, \boldsymbol{e}_h^{\boldsymbol{u}}) + b_h(\boldsymbol{e}_h^{\boldsymbol{u}}, \eta^p).\end{aligned} \tag{32}$$

On the left-hand side, due to (30), $\nabla\cdot\boldsymbol{\eta}^{\boldsymbol{u}} = 0$ holds weakly, and thus $b_h(\boldsymbol{\eta}^{\boldsymbol{u}}, e_h^p) = 0$ since also $\boldsymbol{\eta}^{\boldsymbol{u}} \in \boldsymbol{H}(\mathrm{div};\Omega)$. The other two mixed terms on the left-hand side cancel, as they are negative of each other. On the right-hand side, we observe that $\boldsymbol{\eta}^{\boldsymbol{u}} \in \boldsymbol{H}(\mathrm{div};\Omega)$, and thus $[\![\boldsymbol{\eta}^{\boldsymbol{u}}]\!]\cdot\boldsymbol{n}_F = 0$ for all facets $F \in \mathcal{F}_h$, and so $j_h(\boldsymbol{\eta}^{\boldsymbol{u}}, \boldsymbol{e}_h^{\boldsymbol{u}}) = 0$. For the remaining mixed term on the right-hand side, note that $\nabla\cdot\boldsymbol{e}_h^{\boldsymbol{u}}\big|_K \in \mathbb{P}_{k-1}(K)$ and since η^p is orthogonal to $\mathbb{P}_{k-1}$, we obtain

$$\begin{aligned}b_h(\boldsymbol{e}_h^{\boldsymbol{u}}, \eta^p) &= -\int_\Omega \eta^p(\nabla_h\cdot\boldsymbol{e}_h^{\boldsymbol{u}})\,\mathrm{d}\boldsymbol{x} + \sum_{F\in\mathcal{F}_h}\oint_F \{\!\{\eta^p\}\!\}([\![\boldsymbol{e}_h^{\boldsymbol{u}}]\!]\cdot\boldsymbol{n}_F)\,\mathrm{d}\boldsymbol{s}\\ &= \sum_{F\in\mathcal{F}_h}\oint_F \{\!\{\eta^p\}\!\}([\![\boldsymbol{e}_h^{\boldsymbol{u}}]\!]\cdot\boldsymbol{n}_F)\,\mathrm{d}\boldsymbol{s}.\end{aligned}$$

In (32), applying discrete coercivity (20) on the left-hand side and boundedness (16) plus Cauchy–Schwarz on the right-hand side now results in

$$\begin{aligned}\nu C_\sigma|||\boldsymbol{e}_h^{\boldsymbol{u}}|||_e^2 + \gamma|\boldsymbol{e}_h^{\boldsymbol{u}}|_{\mathrm{nj}}^2 &\leqslant \sqrt{\nu}M|||\boldsymbol{\eta}^{\boldsymbol{u}}|||_{e,\sharp}\sqrt{\nu}|||\boldsymbol{e}_h^{\boldsymbol{u}}|||_e\\ &+ \left(\sum_{F\in\mathcal{F}_h}\frac{h_F}{\gamma}\left\|\{\!\{\eta^p\}\!\}\right\|_{L^2(F)}^2\right)^{1/2}\left(\sum_{F\in\mathcal{F}_h}\frac{\gamma}{h_F}\left\|[\![\boldsymbol{e}_h^{\boldsymbol{u}}]\!]\cdot\boldsymbol{n}_F\right\|_{L^2(F)}^2\right)^{1/2} = \mathfrak{T}_1 + \mathfrak{T}_2.\end{aligned}$$

We further estimate these right-hand side using Young's inequality:

$$\mathfrak{T}_1 \leqslant \frac{\nu M^2}{2C_\sigma} |||\boldsymbol{\eta}^{\boldsymbol{u}}|||_{e,\sharp}^2 + \frac{\nu C_\sigma}{2} |||\boldsymbol{e}_h^{\boldsymbol{u}}|||_e^2,$$
$$\mathfrak{T}_2 \leqslant \frac{1}{2}\left(\sum_{F\in\mathcal{F}_h} \frac{h_F}{\gamma} \left\| \left\{\!\!\left\{ \eta^p \right\}\!\!\right\} \right\|_{L^2(F)}^2 \right) + \frac{\gamma}{2} |\boldsymbol{e}_h^{\boldsymbol{u}}|_{\mathrm{nj}}^2.$$

Thus, the terms involving the velocity discretization error $\boldsymbol{e}_h^{\boldsymbol{u}}$ can be absorbed in the left-hand side. For the term involving the average of η^p, we use the discrete trace inequality, analogously as for the energy estimate, and obtain

$$\sum_{F\in\mathcal{F}_h} \frac{h_F}{\gamma} \left\| \left\{\!\!\left\{ \eta^p \right\}\!\!\right\} \right\|_{L^2(F)}^2 \leqslant C\gamma^{-1} \left\| \eta^p \right\|_{L^2(\Omega)}^2. \tag{34}$$

Combining all above estimates leads to

$$\nu C_\sigma |||\boldsymbol{e}_h^{\boldsymbol{u}}|||_e^2 + \gamma |\boldsymbol{e}_h^{\boldsymbol{u}}|_{\mathrm{nj}}^2 \leqslant C\nu |||\boldsymbol{\eta}^{\boldsymbol{u}}|||_{e,\sharp}^2 + C\gamma^{-1} \left\| \eta^p \right\|_{L^2(\Omega)}^2. \tag{35}$$

Reordering, dropping the positive $\boldsymbol{H}(\mathrm{div};\Omega)$-stabilization term on the left-hand side, and taking the square root reveals

$$|||\boldsymbol{e}_h^{\boldsymbol{u}}|||_e \leqslant C\Big[|||\boldsymbol{\eta}^{\boldsymbol{u}}|||_{e,\sharp} + \gamma^{-1/2}\nu^{-1/2} \left\| \eta^p \right\|_{L^2(\Omega)} \Big]. \tag{36}$$

Application of the triangle inequality and the fact that $|||\boldsymbol{\eta}^{\boldsymbol{u}}|||_e \leqslant |||\boldsymbol{\eta}^{\boldsymbol{u}}|||_{e,\sharp}$ gives the claim for the velocity error estimate. For the pressure estimate, we again invoke the discrete inf-sup condition (14) in $\boldsymbol{W}_h/Q_h$ and the error splitting:

$$\beta^* \left\| e_h^p \right\|_{L^2(\Omega)} \leqslant \sup_{\boldsymbol{w}_h\in\boldsymbol{W}_h\setminus\{\boldsymbol{0}\}} \frac{b_h(\boldsymbol{w}_h, e_h^p)}{|||\boldsymbol{w}_h|||_e} = \sup_{\boldsymbol{w}_h\in\boldsymbol{W}_h\setminus\{\boldsymbol{0}\}} \left[\frac{b_h(\boldsymbol{w}_h, \eta^p) - b_h(\boldsymbol{w}_h, p-p_h)}{|||\boldsymbol{w}_h|||_e} \right]. \tag{37}$$

It remains to further estimate the numerator in the last term. Arguing similarly as above, $\nabla\cdot\boldsymbol{w}_h\big|_K \in \mathbb{P}_{k-1}(K)$ as $\boldsymbol{W}_h \subset \boldsymbol{V}_h$, and using the fact that η^p is orthogonal to $\mathbb{P}_{k-1}$ yields

$$b_h(\boldsymbol{w}_h, \eta^p) = \sum_{F\in\mathcal{F}_h} \oint_F \left\{\!\!\left\{ \eta^p \right\}\!\!\right\} (\llbracket \boldsymbol{w}_h \rrbracket \cdot \boldsymbol{n}_F)\, \mathrm{d}\boldsymbol{s} = 0, \tag{38}$$

where the last equality makes use of $\llbracket \boldsymbol{w}_h \rrbracket \cdot \boldsymbol{n}_F = 0$ for all $F \in \mathcal{F}_h$. Then, Corollary 3.3 and (16) together with $\boldsymbol{w}_h$ being normal-continuous leads to

$$-b_h(\boldsymbol{w}_h, p-p_h) = -b_h(\boldsymbol{w}_h, \pi_0 p - p_h) = \nu a_h(\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{w}_h) \leqslant \nu M |||\boldsymbol{u}-\boldsymbol{u}_h|||_{e,\sharp} |||\boldsymbol{w}_h|||_e. \tag{39}$$

This proves (31b). Equivalence of the $|||\cdot|||_e$- and $|||\cdot|||_{e,\sharp}$-norm, the velocity error estimate (31a) and the triangle inequality conclude the proof. ■

Table 1: Errors for the no-flow problem with the mass flux penalized (controlled by γ) and broken divergence penalized (controlled by γ_{gd}) DG method $\mathbb{P}_3^{\mathrm{dc}}/\mathbb{P}_2^{\mathrm{dc}}$, for $\nu = 10^{-3}$ and using a structured triangular mesh with $h = \frac{1}{32}$.

γ	γ_{gd}	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{L}^2}$	$\|\nabla_h(\boldsymbol{u} - \boldsymbol{u}_h)\|_{\boldsymbol{L}^2}$	$\|p - p_h\|_{L^2}$	$\|\nabla_h \cdot \boldsymbol{u}_h\|_{L^2}$
0	0	$3.94 \cdot 10^{-6}$	$1.29 \cdot 10^{-3}$	$1.31 \cdot 10^{-5}$	$9.91 \cdot 10^{-4}$
1	0	$3.19 \cdot 10^{-7}$	$1.01 \cdot 10^{-4}$	$1.27 \cdot 10^{-5}$	$4.52 \cdot 10^{-5}$
10	0	$3.48 \cdot 10^{-8}$	$1.1 \cdot 10^{-5}$	$1.27 \cdot 10^{-5}$	$4.77 \cdot 10^{-6}$
10^2	0	$3.52 \cdot 10^{-9}$	$1.11 \cdot 10^{-6}$	$1.27 \cdot 10^{-5}$	$4.79 \cdot 10^{-7}$
10^3	0	$3.72 \cdot 10^{-10}$	$1.12 \cdot 10^{-7}$	$1.27 \cdot 10^{-5}$	$4.8 \cdot 10^{-8}$
0	1	$3.47 \cdot 10^{-6}$	$1.12 \cdot 10^{-3}$	$1.38 \cdot 10^{-5}$	$5.11 \cdot 10^{-6}$
0	10	$3.48 \cdot 10^{-6}$	$1.12 \cdot 10^{-3}$	$1.38 \cdot 10^{-5}$	$5.15 \cdot 10^{-7}$
0	10^2	$3.48 \cdot 10^{-6}$	$1.12 \cdot 10^{-3}$	$1.38 \cdot 10^{-5}$	$5.16 \cdot 10^{-8}$
0	10^3	$3.48 \cdot 10^{-6}$	$1.12 \cdot 10^{-3}$	$1.38 \cdot 10^{-5}$	$5.16 \cdot 10^{-9}$

3.3 Numerical illustration of the error estimate

We now present results of a numerical experiment, in order to illustrate Theorem 3.4. Take as the domain the unit square $\Omega = (0,1)^2$, viscosity $\nu = 10^{-3}$, constant interior penalty parameter $\sigma = 4k^2$, and a third order ($k = 3$) DG method with $\mathbb{P}_3^{\mathrm{dc}}/\mathbb{P}_2^{\mathrm{dc}}$ elements on a structured triangular mesh with $h = \frac{1}{32}$. In order to show that the straight-forward addition of the well-known (broken) grad-div stabilization alone is not sufficient for DG methods, the term

$$\gamma_{\mathrm{gd}} \int_\Omega (\nabla_h \cdot \boldsymbol{u}_h)(\nabla_h \cdot \boldsymbol{v}_h)\,\mathrm{d}\boldsymbol{x} \tag{40}$$

is added to the left-hand side of (18)–(19). Therefore, the parameter γ now controls the amount of mass flux penalization, whereas γ_{gd} controls the amount of (broken) divergence penalization. We make use of the high-order finite element library `NGSolve` [39].

The problem we consider is a version of the no-flow problem where the exact solution is chosen to be $\boldsymbol{u} = \boldsymbol{0}$ and $p = \sin(2\pi x + 2\pi y)$. Note that $\int_\Omega p\,\mathrm{d}\boldsymbol{x} = 0$ and the corresponding forcing vector is the gradient field $\boldsymbol{f} = \nabla p$. The results of our experiment are shown in Table 1.

As the mass flux stabilization parameter γ increases (with fixed $\gamma_{\mathrm{gd}} = 0$), we observe convergence of the velocity to the no-flow solution. However, for fixed $\gamma = 0$ and increasing γ_{gd}, we do not observe an improvement in the velocity error, only in the (broken) divergence error. The discrete pressure is not significantly influenced by either stabilization.

In contrast to Theorem 3.4 which predicts an error reduction with rate $\gamma^{-1/2}$, Table 1 indicates a better (linear) reduction behaving like γ^{-1}. The purpose of the next section is to consider the limiting behavior of the method as $\gamma \to \infty$, and resolves this issue of scaling with γ.

3.4 Convergence as $\gamma \to \infty$ to the BDM solution

We now prove a limiting result for the mass flux penalized DG method as $\gamma \to \infty$. In particular, we will prove that the method converges to a weakly divergence-free BDM solution, with rate $O(\gamma^{-1})$. Since BDM optimally approximates Stokes, this result explains the linear convergence with γ^{-1} to the true solution, in the numerical test of the previous subsection.

To begin, we precisely define the BDM$_k$ space $\boldsymbol{W}_h$ by

$$\boldsymbol{W}_h = \big\{\boldsymbol{v}_h \in \boldsymbol{H}(\mathrm{div};\Omega)\colon \boldsymbol{v}_h\big|_K \in \mathbb{P}_k(K)\big\}.$$

The corresponding inf-sup stable FE pair is given by $\boldsymbol{W}_h/Q_h$, also denoted $\mathbb{BDM}_k/\mathbb{P}^{\mathrm{dc}}_{k-1}$. For more information on $\boldsymbol{H}(\mathrm{div})$-FEM, the reader is referred to [40]. Note that the pressure spaces for DG and $\boldsymbol{H}(\mathrm{div})$ method coincide. A sketch of the local degrees of freedom for both DG and $\boldsymbol{H}(\mathrm{div})$ methods in the 2D case with $k = 3$ is shown in Figure 1.

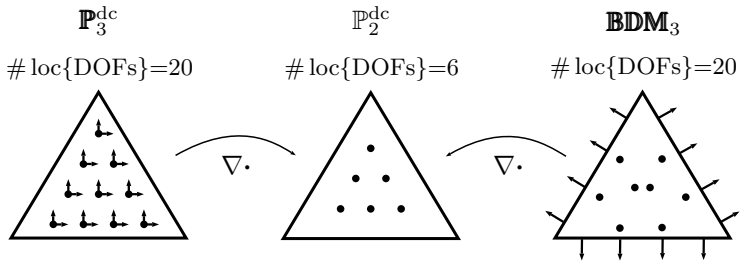

Figure 1: Shown above is a sketch of degrees of freedom in 2D for DG and $\boldsymbol{H}(\mathrm{div})$ methods.

We then introduce the following weakly divergence-free $\boldsymbol{H}(\mathrm{div})$-DG method: Find $(\widehat{\boldsymbol{u}}_h, \widehat{p}_h) \in \boldsymbol{W}_h \times Q_h$ such that for all $(\boldsymbol{w}_h, q_h) \in \boldsymbol{W}_h \times Q_h$,

$$\nu a_h(\widehat{\boldsymbol{u}}_h, \boldsymbol{w}_h) + b(\boldsymbol{w}_h, \widehat{p}_h) - b(\widehat{\boldsymbol{u}}_h, q_h) = (\boldsymbol{f}, \boldsymbol{w}_h). \tag{41}$$

Note that due to the $\boldsymbol{H}(\mathrm{div})$-conformity and in contrast to (18)–(19), the pressure-velocity coupling in (41) is the same as in the continuous weak formulation (2) of the Stokes problem. Since velocities in $\boldsymbol{W}_h$ are normal-continuous, the mass flux penalization naturally vanishes and the SIP bilinear form a_h in (41) acts only on the tangential component of the involved velocities.

The discretely divergence-free subspace of $\boldsymbol{W}_h$ is defined by

$$\boldsymbol{W}_h^{\mathrm{div}} = \{\boldsymbol{w}_h \in \boldsymbol{W}_h\colon b(\boldsymbol{w}_h, q_h) = 0,\ \forall\, q_h \in Q_h\},$$

and note that these discretely divergence-free $\boldsymbol{H}(\mathrm{div})$ velocities are even weakly divergence-free, i.e. $\boldsymbol{W}_h^{\mathrm{div}} = \{\boldsymbol{w}_h \in \boldsymbol{H}(\mathrm{div};\Omega)\colon\ \nabla\cdot\boldsymbol{w}_h = 0\}$.

The aim of this section is to show that the solution $\boldsymbol{u}_h$ of the stabilized DG method (18)–(19) converges to the weakly divergence-free $\boldsymbol{H}(\mathrm{div})$ solution $\widehat{\boldsymbol{u}}_h$ of (41) as $\gamma \to \infty$. To begin, notice that a_h defines a symmetric bilinear form on $\boldsymbol{V}_h$. Due to the inclusion $\boldsymbol{W}_h^{\mathrm{div}} \subset \boldsymbol{V}_h^{\mathrm{div}}$, and since both spaces are finite-dimensional, the orthogonal complement

$$\boldsymbol{R}_h^{\mathrm{div}} = \big(\boldsymbol{W}_h^{\mathrm{div}}\big)^{\perp} = \big\{\boldsymbol{v}_h \in \boldsymbol{V}_h^{\mathrm{div}}\colon a_h(\boldsymbol{v}_h, \boldsymbol{w}_h) = 0,\ \forall\, \boldsymbol{w}_h \in \boldsymbol{W}_h^{\mathrm{div}}\big\} \tag{42}$$

makes it possible to obtain the following inner direct sum decomposition:

$$\boldsymbol{V}_h^{\mathrm{div}} = \boldsymbol{W}_h^{\mathrm{div}} \oplus \boldsymbol{R}_h^{\mathrm{div}}, \quad \boldsymbol{W}_h^{\mathrm{div}} \cap \boldsymbol{R}_h^{\mathrm{div}} = \{\boldsymbol{0}\}, \quad (\perp \text{ w.r.t. } a_h \text{ inner product}). \tag{43}$$

Thus, $\boldsymbol{W}_h^{\mathrm{div}}$ contains weakly divergence-free, normal-continuous velocities whereas a velocity $\boldsymbol{v}_h' \in \boldsymbol{R}_h^{\mathrm{div}}$ is either not in $\boldsymbol{H}(\mathrm{div};\Omega)$, or $\boldsymbol{v}_h' \in \boldsymbol{H}(\mathrm{div};\Omega)$ but $\nabla\cdot\boldsymbol{v}_h' \neq 0$. The following lemma is the key property for showing the convergence of $\boldsymbol{u}_h \to \widehat{\boldsymbol{u}}_h$ as $\gamma \to \infty$.

LEMMA 3.7

The mapping $|\cdot|_{\mathrm{nj}} \colon \boldsymbol{V}_h \to \boldsymbol{R}$ defines a norm on $\boldsymbol{R}_h^{\mathrm{div}}$.

PROOF: We show that if $|\boldsymbol{v}_h'|_{\mathrm{nj}} = 0$, then $\boldsymbol{v}_h' \equiv \boldsymbol{0}$ for all $\boldsymbol{v}_h' \in \boldsymbol{R}_h^{\mathrm{div}}$, which is the only non-trivial property. Let $\boldsymbol{v}_h' \in \boldsymbol{R}_h^{\mathrm{div}}$ with $|\boldsymbol{v}_h'|_{\mathrm{nj}} = 0$. We know that $\boldsymbol{v}_h'$ is discretely divergence-free; that is, $b_h(\boldsymbol{v}_h', q_h) = 0$ for all $q_h \in Q_h$. Inserting $q_h = \nabla_h\cdot\boldsymbol{v}_h'$, multiplying by $h_F h_F^{-1} = 1$, using Cauchy–Schwarz and applying the discrete trace inequality leads to

$$\begin{aligned}\|\nabla_h\cdot\boldsymbol{v}_h'\|_{L^2(\Omega)}^2 &= \sum_{F\in\mathcal{F}_h} \oint_F \{\!\{ \nabla\cdot\boldsymbol{v}_h' \}\!\} (\llbracket \boldsymbol{v}_h' \rrbracket \cdot \boldsymbol{n}_F)\,\mathrm{d}\boldsymbol{s} \\ &\leqslant C\,\|\nabla_h\cdot\boldsymbol{v}_h'\|_{L^2(\Omega)}\,|\boldsymbol{v}_h'|_{\mathrm{nj}} = 0.\end{aligned}$$

Thus, $\nabla_h\cdot\boldsymbol{v}_h' = 0$. The fact that $|\boldsymbol{v}_h'|_{\mathrm{nj}} = 0$ also means that $\boldsymbol{v}_h'$ is normal-continuous and therefore in $\boldsymbol{H}(\mathrm{div};\Omega)$. Hence, we infer that $\boldsymbol{v}_h' \in \boldsymbol{W}_h^{\mathrm{div}}$ and since $\boldsymbol{W}_h^{\mathrm{div}} \cap \boldsymbol{R}_h^{\mathrm{div}} = \{\boldsymbol{0}\}$, $\boldsymbol{v}_h' \equiv \boldsymbol{0}$ is proven. ■

Next, decompose both the DG solution $\boldsymbol{u}_h = \boldsymbol{u}_h^0 + \boldsymbol{u}_h'$ with $(\boldsymbol{u}_h^0, \boldsymbol{u}_h') \in \boldsymbol{W}_h^{\mathrm{div}} \times \boldsymbol{R}_h^{\mathrm{div}}$ and the DG test functions $\boldsymbol{v}_h = \boldsymbol{v}_h^0 + \boldsymbol{v}_h'$ with $(\boldsymbol{v}_h^0, \boldsymbol{v}_h') \in \boldsymbol{W}_h^{\mathrm{div}} \times \boldsymbol{R}_h^{\mathrm{div}}$. Inserting this decomposition and using the properties of the spaces $\boldsymbol{W}_h^{\mathrm{div}}$ and $\boldsymbol{R}_h^{\mathrm{div}}$ leads to the following decoupled system:

$$\nu a_h\big(\boldsymbol{u}_h^0, \boldsymbol{v}_h^0\big) = \big(\boldsymbol{f}, \boldsymbol{v}_h^0\big), \quad \forall\, \boldsymbol{v}_h^0 \in \boldsymbol{W}_h^{\mathrm{div}}, \tag{44a}$$

$$\nu a_h(\boldsymbol{u}_h', \boldsymbol{v}_h') + \gamma j_h(\boldsymbol{u}_h', \boldsymbol{v}_h') = (\boldsymbol{f}, \boldsymbol{v}_h'), \quad \forall\, \boldsymbol{v}_h' \in \boldsymbol{R}_h^{\mathrm{div}}. \tag{44b}$$

Since the solution $\widehat{\boldsymbol{u}}_h$ to the weakly divergence-free $\boldsymbol{H}(\mathrm{div})$-FEM is unique, we infer from (44a) that $\boldsymbol{u}_h^0 = \widehat{\boldsymbol{u}}_h$. For the other part, testing with $\boldsymbol{v}_h' = \boldsymbol{u}_h'$ in (44b) and using discrete coercivity (20) leads to

$$\nu C_\sigma |\!|\!| \boldsymbol{u}_h' |\!|\!|_e^2 + \gamma |\boldsymbol{u}_h'|_{\mathrm{nj}}^2 \leqslant \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}\,\|\boldsymbol{u}_h'\|_{\boldsymbol{L}^2(\Omega)}\,. \tag{45}$$

From this inequality, we can directly obtain the limit behavior as $\gamma \to \infty$.

THEOREM 3.8 (*Convergence to divergence-free $\boldsymbol{H}(\mathrm{div})$ solution*)

Let $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and assume that $\sigma > 0$ is sufficiently large to guarantee discrete coercivity. Let $\boldsymbol{u}_h$ be the solution of the stabilized DG method (18)–(19) *with $\gamma \geqslant 0$, and $\widehat{\boldsymbol{u}}_h$ be the weakly divergence-free $\boldsymbol{H}(\mathrm{div})$ solution of* (41)*. Then there exists a constant C independent of γ such that*

$$|\!|\!| \boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h |\!|\!|_e \leqslant C\gamma^{-1}\,\|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)} \tag{46a}$$

$$\|p_h - \widehat{p}_h\|_{L^2(\Omega)} \leqslant C\gamma^{-1}\nu\,\|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)} \tag{46b}$$

REMARK 3.9: Although this result is for a fixed mesh, we note that the constant C includes constants arising from norm equivalences, and could potentially depend on h.

PROOF: First note that

$$\left|\!\left|\!\left|\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h\right|\!\right|\!\right|_e = \left|\!\left|\!\left|\boldsymbol{u}_h - \boldsymbol{u}_h^0\right|\!\right|\!\right|_e = \left|\!\left|\!\left|\boldsymbol{u}_h'\right|\!\right|\!\right|_e. \tag{47}$$

Due to the fact that in finite-dimensional spaces all norms are equivalent, an application of Lemma 3.7 on the left-hand side of (45) (drop the viscous energy norm multiplied by ν) and Poincaré–Friedrichs on the right-hand side leads to

$$\begin{aligned} C\gamma \left|\!\left|\!\left|\boldsymbol{u}_h'\right|\!\right|\!\right|_e^2 &\leqslant \nu C_\sigma \left|\!\left|\!\left|\boldsymbol{u}_h'\right|\!\right|\!\right|_e^2 + \gamma |\boldsymbol{u}_h'|_{\mathrm{nj}}^2 \\ &\leqslant \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)} \|\boldsymbol{u}_h'\|_{\boldsymbol{L}^2(\Omega)} \\ &\leqslant C_{\mathrm{PF}} \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)} \left|\!\left|\!\left|\boldsymbol{u}_h'\right|\!\right|\!\right|_e. \end{aligned}$$

Reordering shows the first claim. For the pressure convergence, we use the discrete inf-sup condition (14) for the FE pair $\boldsymbol{W}_h/Q_h$. Since both methods use the same pressure space, we can consider $q_h = p_h - \widehat{p}_h \in Q_h$:

$$\begin{aligned} \beta^* \|p_h - \widehat{p}_h\|_{L^2(\Omega)} &\leqslant \sup_{\boldsymbol{w}_h \in \boldsymbol{W}_h \setminus \{\boldsymbol{0}\}} \frac{b_h(\boldsymbol{w}_h, p_h - \widehat{p}_h)}{\left|\!\left|\!\left|\boldsymbol{w}_h\right|\!\right|\!\right|_e} \\ &= \sup_{\boldsymbol{w}_h \in \boldsymbol{W}_h \setminus \{\boldsymbol{0}\}} \frac{\nu a_h(\widehat{\boldsymbol{u}}_h - \boldsymbol{u}_h, \boldsymbol{w}_h)}{\left|\!\left|\!\left|\boldsymbol{w}_h\right|\!\right|\!\right|_e} \\ &\leqslant \nu M \left|\!\left|\!\left|\widehat{\boldsymbol{u}}_h - \boldsymbol{u}_h\right|\!\right|\!\right|_{e,\sharp}. \end{aligned}$$

Here, we used the boundedness of a_h (16). The final step is to acknowledge that $\left|\!\left|\!\left|\cdot\right|\!\right|\!\right|_e$- and $\left|\!\left|\!\left|\cdot\right|\!\right|\!\right|_{e,\sharp}$-norm are uniformly equivalent on $\boldsymbol{V}_h$. Reordering and (46a) shows the convergence rate for the pressure. ∎

We now illustrate Theorem 3.8 with a numerical experiment. We take $\Omega = (0,1)^2$, $\nu = 10^{-3}$, $\sigma = 4k^2$ and the exact solution as

$$\boldsymbol{u} = \begin{pmatrix} \pi \sin^2(\pi x)\sin(2\pi y) \\ -\pi \sin(2\pi x)\sin^2(\pi y) \end{pmatrix}, \quad p = \cos(\pi x)\sin(\pi y).$$

The corresponding right-hand side is

$$\boldsymbol{f} = \begin{pmatrix} -\nu 2\pi^3 (2\cos(2\pi x) - 1)\sin(2\pi y) - \pi \sin(\pi x)\sin(\pi y) \\ \nu 2\pi^3 \sin(2\pi x)(2\cos(2\pi y) - 1) + \pi \cos(\pi x)\cos(\pi y) \end{pmatrix}.$$

We use a structured triangular mesh with $h = \frac{1}{20}$, $k = 3$, and $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$. Results are shown in Table 2, and we observe the $O(\gamma^{-1})$ convergence in both pressure and velocity, as $\gamma \to 0$.

Table 2: Convergence behavior of $\boldsymbol{H}(\mathrm{div})$-stabilized DG solution $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ to the weakly divergence-free $\boldsymbol{H}(\mathrm{div})$-DG solution $(\widehat{\boldsymbol{u}}_h, \widehat{p}_h) \in \boldsymbol{W}_h \times Q_h$ as $\gamma \to \infty$.

γ	$\|\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h\|_{\boldsymbol{L}^2}$	$\|\nabla_h(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\|_{\boldsymbol{L}^2}$	$\|p_h - \widehat{p}_h\|_{L^2}$
0	$1.63 \cdot 10^{-5}$	$3.01 \cdot 10^{-3}$	$4.73 \cdot 10^{-6}$
1	$1.05 \cdot 10^{-6}$	$2.05 \cdot 10^{-4}$	$3.97 \cdot 10^{-7}$
10	$1.14 \cdot 10^{-7}$	$2.23 \cdot 10^{-5}$	$4.31 \cdot 10^{-8}$
10^2	$1.17 \cdot 10^{-8}$	$2.28 \cdot 10^{-6}$	$4.2 \cdot 10^{-9}$
10^3	$1.53 \cdot 10^{-9}$	$2.8 \cdot 10^{-7}$	$5.53 \cdot 10^{-10}$

Table 3: Errors for the no-flow problem with the mass flux stabilization (controlled by γ) and broken grad-div (controlled by γ_{gd}) stabilized DG method $\mathbb{Q}_3^{\mathrm{dc}}/\mathbb{Q}_2^{\mathrm{dc}}$ with $\nu = 10^{-3}$ on a structured quadratic mesh with $h = \frac{1}{32}$.

γ	γ_{gd}	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{L}^2}$	$\|\nabla_h(\boldsymbol{u} - \boldsymbol{u}_h)\|_{\boldsymbol{L}^2}$	$\|p - p_h\|_{L^2}$	$\|\nabla_h \cdot \boldsymbol{u}_h\|_{L^2}$
0	0	$4.28 \cdot 10^{-6}$	$1.87 \cdot 10^{-3}$	$3.14 \cdot 10^{-7}$	$1.82 \cdot 10^{-3}$
10^2	0	$1.58 \cdot 10^{-6}$	$4.6 \cdot 10^{-4}$	$3.01 \cdot 10^{-7}$	$4.92 \cdot 10^{-7}$
10^3	0	$1.58 \cdot 10^{-6}$	$4.6 \cdot 10^{-4}$	$3.01 \cdot 10^{-7}$	$4.92 \cdot 10^{-8}$
10^4	0	$1.58 \cdot 10^{-6}$	$4.6 \cdot 10^{-4}$	$3.01 \cdot 10^{-7}$	$4.91 \cdot 10^{-9}$
10^5	0	$1.58 \cdot 10^{-6}$	$4.6 \cdot 10^{-4}$	$3.01 \cdot 10^{-7}$	$4.86 \cdot 10^{-10}$
0	10^2	$1.61 \cdot 10^{-6}$	$5.25 \cdot 10^{-4}$	$5.5 \cdot 10^{-6}$	$5.57 \cdot 10^{-8}$
0	10^3	$1.61 \cdot 10^{-6}$	$5.25 \cdot 10^{-4}$	$5.5 \cdot 10^{-6}$	$5.57 \cdot 10^{-9}$
0	10^4	$1.61 \cdot 10^{-6}$	$5.25 \cdot 10^{-4}$	$5.51 \cdot 10^{-6}$	$5.6 \cdot 10^{-10}$
0	10^5	$1.61 \cdot 10^{-6}$	$5.25 \cdot 10^{-4}$	$7.94 \cdot 10^{-6}$	$8.1 \cdot 10^{-11}$
10^2	10^2	$1.19 \cdot 10^{-7}$	$3.45 \cdot 10^{-5}$	$4.8 \cdot 10^{-6}$	$4.87 \cdot 10^{-8}$
10^3	10^3	$1.38 \cdot 10^{-8}$	$4.02 \cdot 10^{-6}$	$4.79 \cdot 10^{-6}$	$4.87 \cdot 10^{-9}$
10^4	10^4	$1.41 \cdot 10^{-9}$	$4.09 \cdot 10^{-7}$	$4.8 \cdot 10^{-6}$	$4.89 \cdot 10^{-10}$
10^5	10^5	$1.42 \cdot 10^{-10}$	$4.11 \cdot 10^{-8}$	$7.37 \cdot 10^{-6}$	$7.53 \cdot 10^{-11}$

4 Stabilization of inf-sup stable DG methods on tensor-product meshes

As mentioned in the previous section, on tensor-product elements (quads and hexes) using $\mathbb{Q}_k^{\mathrm{dc}}/\mathbb{Q}_{k-1}^{\mathrm{dc}}$, the situation is slightly more involved since $\nabla_h \cdot \mathbb{Q}_k^{\mathrm{dc}} \not\subset \mathbb{Q}_{k-1}^{\mathrm{dc}}$. In order to demonstrate the difference, we repeat the no-flow test from the previous section with $\mathbb{Q}_k^{\mathrm{dc}}/\mathbb{Q}_{k-1}^{\mathrm{dc}}$ elements on a structured mesh consisting of squares. The results are shown in Table 3.

Interestingly, neither mass flux penalization alone ($\gamma > 0$, $\gamma_{\mathrm{gd}} = 0$), nor broken grad-div stabilization alone ($\gamma_{\mathrm{gd}} > 0$, $\gamma = 0$), is able to improve the pressure-robustness of the considered method, although each of them independently improve the (broken) divergence error. However, when they are added simultaneously ($\gamma > 0$, $\gamma_{\mathrm{gd}} > 0$), Table 3 clearly indicates that the resulting stabilized DG method yields better results. In the following, we briefly sketch how this can be shown by the numerical analysis.

The global discrete spaces in this setting are

$$\boldsymbol{V}_h = \big\{\boldsymbol{v}_h \in \boldsymbol{L}^2(\Omega)\colon \boldsymbol{v}_h\big|_K \in \mathbb{Q}_k(K),\ \forall\, K \in \mathcal{T}_h\big\}, \tag{48a}$$

$$Q_h = \big\{q_h \in L^2_0(\Omega)\colon q_h\big|_K \in \mathbb{Q}_{k-1}(K),\ \forall\, K \in \mathcal{T}_h\big\}. \tag{48b}$$

Again, we need an $\boldsymbol{H}(\mathrm{div})$-conforming space $\boldsymbol{W}_h \subset \boldsymbol{H}(\mathrm{div};\Omega)$ and we assume that $\boldsymbol{W}_h = \boldsymbol{V}_h \cap \boldsymbol{H}(\mathrm{div};\Omega)$. There also has to be an $\boldsymbol{H}(\mathrm{div})$ interpolator $\boldsymbol{\pi}_h\colon \boldsymbol{V} \to \boldsymbol{W}_h \subset \boldsymbol{V}_h$ which fulfills the commuting diagram property (30). Finally, the pair $\boldsymbol{W}_h/Q_h$ also has to be inf-sup stable.

For the sake of brevity in the analysis, we only allow one stabilization parameter γ for both mass flux and broken grad-div stabilization and redefine the stabilization bilinear form as

$$j_h(\boldsymbol{w}_h,\boldsymbol{v}_h) = \int_\Omega (\nabla_h\cdot\boldsymbol{w}_h)(\nabla_h\cdot\boldsymbol{v}_h)\,\mathrm{d}\boldsymbol{x} + \sum_{F\in\mathcal{F}_h}\frac{1}{h_F}\oint_F (\llbracket\boldsymbol{w}_h\rrbracket\cdot\boldsymbol{n}_F)(\llbracket\boldsymbol{v}_h\rrbracket\cdot\boldsymbol{n}_F)\,\mathrm{d}\boldsymbol{s}.$$

THEOREM 4.1 (*Error estimate on tensor-product meshes*)

Under the assumptions of Theorem 3.4, the following error estimate holds true for the stabilized DG method on tensor-product elements:

$$|||\boldsymbol{u}-\boldsymbol{u}_h|||_e \leqslant C\Big[|||\boldsymbol{\eta}^{\boldsymbol{u}}|||_{e,\sharp} + \gamma^{-1/2}\nu^{-1/2}\,\|\eta^p\|_{L^2(\Omega)}\Big], \tag{49a}$$

$$\|p-p_h\| \leqslant C\nu|||\boldsymbol{\eta}^{\boldsymbol{u}}|||_{e,\sharp} + \left[C\sqrt{\frac{\nu}{\gamma}}+2\right]\|\eta^p\|_{L^2(\Omega)}\,. \tag{49b}$$

PROOF: The proof is very similar to that of Theorem 3.4. We only comment on the parts where differences occur between simplicial and tensor-product elements.

First, we observe that also for the new stabilization term, it holds $j_h(\boldsymbol{\eta}^{\boldsymbol{u}},\boldsymbol{e}_h^{\boldsymbol{u}}) = 0$ since $\boldsymbol{\eta}^{\boldsymbol{u}} \in \boldsymbol{H}(\mathrm{div};\Omega)$ and $\nabla\cdot\boldsymbol{\eta}^{\boldsymbol{u}} = 0$ hold exactly. The main difference occurs in the treatment of the mixed term of b_h, where, after Cauchy–Schwarz and Young, we obtain

$$\begin{aligned} b_h(\boldsymbol{e}_h^{\boldsymbol{u}},\eta^p) &= -\int_\Omega \eta^p(\nabla_h\cdot\boldsymbol{e}_h^{\boldsymbol{u}})\,\mathrm{d}\boldsymbol{x} + \sum_{F\in\mathcal{F}_h}\oint_F \{\!\{\eta^p\}\!\}(\llbracket\boldsymbol{e}_h^{\boldsymbol{u}}\rrbracket\cdot\boldsymbol{n}_F)\,\mathrm{d}\boldsymbol{s} \\ &\leqslant C\gamma^{-1}\,\|\eta^p\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\,\|\nabla_h\cdot\boldsymbol{e}_h^{\boldsymbol{u}}\|^2_{L^2(\Omega)} + \frac{\gamma}{2}|\boldsymbol{e}_h^{\boldsymbol{u}}|^2_{\mathrm{nj}}. \end{aligned}$$

Applying the same ideas as in the proof for simplices, we arrive at

$$\nu C_\sigma|||\boldsymbol{e}_h^{\boldsymbol{u}}|||_e^2 + \gamma|\boldsymbol{e}_h^{\boldsymbol{u}}|^2_{\mathrm{nj}} + \gamma\,\|\nabla_h\cdot\boldsymbol{e}_h^{\boldsymbol{u}}\|^2_{L^2(\Omega)} \leqslant C\nu|||\boldsymbol{\eta}^{\boldsymbol{u}}|||^2_{e,\sharp} + C\gamma^{-1}\,\|\eta^p\|^2_{L^2(\Omega)}\,,$$

which yields the velocity error estimate. It is critical to note here that without the (broken) grad-div stabilization, the coefficient of $\leqslant C\nu|||\boldsymbol{\eta}^{\boldsymbol{u}}|||^2_{e,\sharp} + C\gamma^{-1}$ would be ν^{-1} instead of γ^{-1}.

For the pressure estimate, the discrete inf-sup condition (14) in $\boldsymbol{W}_h/Q_h$ is again essential. However, since $\nabla_h\cdot\mathbb{Q}_k^{\mathrm{dc}} \not\subset \mathbb{Q}_{k-1}^{\mathrm{dc}}$, for $\boldsymbol{w}_h \in \boldsymbol{W}_h$ we have to estimate

$$b_h(\boldsymbol{w}_h,\eta^p) = -\int_\Omega \eta^p(\nabla\cdot\boldsymbol{w}_h)\,\mathrm{d}\boldsymbol{x} \leqslant \|\eta^p\|_{L^2(\Omega)}\,\|\nabla\cdot\boldsymbol{w}_h\|_{L^2(\Omega)}\,. \tag{50}$$

From here, $\|\nabla \cdot \boldsymbol{w}_h\|_{L^2(\Omega)} \leqslant |||\boldsymbol{w}_h|||_e$ leads immediately to the same situation as in the proof of Theorem 3.4. ■

REMARK 4.2: Concerning the convergence of the stabilized $\boldsymbol{V}_h/Q_h$ DG solution to the $\boldsymbol{W}_h/Q_h$ $\boldsymbol{H}(\mathrm{div})$ solution as $\gamma \to 0$, basically the same arguments as in Section 3.2 can be applied. Unfortunately, to the best of the authors' knowledge, the $\boldsymbol{H}(\mathrm{div})$-FE space $\boldsymbol{W}_h$ is not known in the literature, and so we do not present the analysis here. We mention, however, that $\mathbb{RT}_{[k-1]} \subset \boldsymbol{W}_h \subset \mathbb{RT}_{[k]}$ where $\mathbb{RT}_{[k]}$ denotes the Raviart–Thomas space of order k on tensor-product elements; cf. [4].

5 Improving pressure-robustness in Crouzeix-Raviart approximations

The mass flux penalization can also be applied, with similar results, to the Crouzeix–Raviart (CR) element, and we include some results for completeness. In a sense, it is somewhat easier to analyze than the DG case considered above, and we find a similar fundamental result: penalization of the mass flux reduces the effect of the pressure error on the velocity error. Such a result is proven implicitly for CR elements in a recent work of Burman and Hansbo [6] for the Darcy–Stokes problem, where multiple stabilizations were used. Interestingly, the motivation for using the stabilization in that work was 'to control the nonconformity emanating from the pressure term', and in effect they proved something similar to what we prove above for DG: the scaling of the pressure term in the error estimate is improved from ν^{-1} to $\nu^{-1/2}\gamma^{-1/2}$.

However, there is (seemingly) a potential negative consequence for CR that is not an issue with DG: the use of the stabilization seemingly increases the scaling of the velocity error in the error estimate. Hence it is unclear whether large stabilization parameters can be used without negative consequence, as can be done in the DG case (at least, up to difficulties in linear solvers due to matrix conditioning).

The nonconforming Crouzeix–Raviart finite element velocity and pressure spaces are defined by:

$$\begin{aligned}\boldsymbol{V}_h &= \left\{\boldsymbol{v}_h \in \boldsymbol{L}^2(\Omega)\colon \boldsymbol{v}_h\big|_K \in \mathbb{P}_1(K),\ \forall\, K \in \mathcal{T}_h,\ \oint_F [\![\boldsymbol{v}_h]\!] = \boldsymbol{0},\ \forall\, F \in \mathcal{F}_h\right\},\\ Q_h &= \left\{q_h \in L^2_0(\Omega)\colon q_h\big|_K \in \mathbb{P}_0(K),\ \forall\, K \in \mathcal{T}_h\right\}.\end{aligned}$$

The discrete bilinear forms for CR elements are defined as

$$\begin{aligned}a_h(\boldsymbol{w}_h, \boldsymbol{v}_h) &= \int_\Omega \nabla_h \boldsymbol{w}_h : \nabla_h \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x},\\ b_h(\boldsymbol{w}_h, q_h) &= -\int_\Omega q_h (\nabla_h \cdot \boldsymbol{w}_h)\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

$\boldsymbol{V}_h$ is equipped with the norm

$$\|\boldsymbol{v}_h\|_{1,h} := \left(\int_\Omega \nabla_h \boldsymbol{v}_h : \nabla_h \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x}\right)^{1/2} = \left(\sum_{K\in\mathcal{T}_h} \|\nabla \boldsymbol{v}_h\|^2_{\boldsymbol{L}^2(K)}\right)^{1/2}.$$

We propose to consider CR together with the mass flux penalization: find $(\boldsymbol{u}_h, \boldsymbol{v}_h) \in \boldsymbol{V}_h \times Q_h$ such that for all $(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h \times Q_h$

$$\gamma j_h(\boldsymbol{u}_h, \boldsymbol{v}_h) + \nu a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) + b_h(\boldsymbol{v}_h, p_h) = (\boldsymbol{f}, v_h) \tag{51}$$

$$-b_h(\boldsymbol{u}_h, q_h) = 0. \tag{52}$$

Since $j_h(\boldsymbol{u}_h, \boldsymbol{u}_h) \geqslant 0$, the classical well-posedness results for CR (see e.g. [5]) will hold also for (51)–(52) with any fixed $\gamma \geqslant 0$.

We now present an error estimate, which follows from the recent work of Burman and Hansbo for a Darcy–Stokes problem with multiple stabilizations [6], and making the appropriate simplifications (hence we omit the proof).

THEOREM 5.1 (*Error estimate for Crouzeix–Raviart elements*)

Let $(\boldsymbol{u}, p) \in (\boldsymbol{V} \cap \boldsymbol{H}^2(\Omega)) \times (Q \cap H^1(\Omega))$ *be the Stokes solution, and* $(\boldsymbol{u}_h, p_h)$ *the solution to* (51)–(52) *with parameter* $\gamma > 0$. *Then it holds that*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,h} \leqslant Ch\Big[\big(1 + \gamma^{1/2}\nu^{-1/2}\big)|\boldsymbol{u}|_{\boldsymbol{H}^2(\Omega)} + \gamma^{-1/2}\nu^{-1/2}|p|_{H^1(\Omega)}\Big].$$

REMARK 5.2: This error estimates reveal, just as in the DG case, that the facet jump stabilization reduces the negative effect of the pressure on the velocity error by changing the coefficient of the pressure from ν^{-1} to $\nu^{-1/2}\gamma^{-1/2}$. However, the dependence on $\nu^{-1/2}$ in the term $\big(1 + \gamma^{1/2}\nu^{-1/2}\big)|\boldsymbol{u}|_{\boldsymbol{H}^2(\Omega)}$ seems to be unavoidable for the nonconforming CR-element, indicating the danger of over-stabilization, which is nothing more than a kind of Poisson locking for the divergence-free limit as $\gamma \to \infty$.

6 Application to more complex flows

We show here that the pressure-robust approach above can have a considerable impact on much more complicated problems than the steady incompressible Stokes equations. We consider first a numerical test for steady incompressible Navier–Stokes equations (NSE), followed by a numerical test for non-isothermal flow.

6.1 Steady Navier–Stokes equations

Consider now the steady incompressible Navier–Stokes equations

$$\begin{cases} -\nu\Delta\boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \boldsymbol{f} & \text{in } \Omega, \\ \nabla\cdot\boldsymbol{u} = 0 & \text{in } \Omega, \end{cases} \tag{53}$$

with inhomogeneous Dirichlet velocity boundary condition $\boldsymbol{g}_{\mathrm{D}}$. The new pressure-robust space discretization approach will be superior, whenever a difficult pressure spoils the velocity error.

In order to show this, we construct a simple polynomial potential flow $\boldsymbol{u} = \nabla\chi$ on the unit square $\Omega = (-1, 1)^2$, where χ is defined by the real part of the complex polynomial z^5, i.e., $\chi = \mathrm{Re}(z^5) =$

$\mathrm{Re}(x+iy)^5 = x^5 - 10x^3y^2 + 5xy^4$. The resulting flow consists of ten colliding jets, meeting at the stagnation point (0,0), see Figure 2. Indeed, $(\boldsymbol{u}, p) = (\nabla\chi, \frac{1}{2}|\nabla\chi|^2)$ solves the steady incompressible NSE (53) with appropriate inhomogeneous Dirichlet velocity boundary conditions, for all $\nu > 0$. Since p is a polynomial of order eight, a pressure-robust low-order DG space discretization should be comparably accurate — at least at non-negligible Reynolds numbers — as a high-order DG approximation which is not pressure-robust.

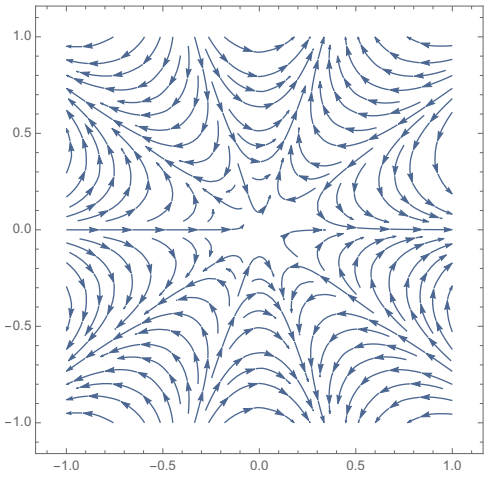

Figure 2: Streamlines for the steady incompressible Navier–Stokes benchmark with ten jets and a stagnation point at $(0,0)$.

The stabilized $\mathbb{P}_k^{\mathrm{dc}}/\mathbb{P}_{k-1}^{\mathrm{dc}}$ DG method for the stationary Navier–Stokes equations (53) is the following: Find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ satisfying $\forall\, (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h \times Q_h$,

$$\begin{aligned} \nu a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) + c_h(\boldsymbol{u}_h; \boldsymbol{u}_h, \boldsymbol{v}_h) + b_h(\boldsymbol{v}_h, p_h) + \gamma j_h(\boldsymbol{u}_h, \boldsymbol{v}_h) \\ = (\boldsymbol{f}, \boldsymbol{v}_h) + \nu a_h^\partial(\boldsymbol{g}_{\mathrm{D}}; \boldsymbol{v}_h) + \gamma j_h^\partial(\boldsymbol{g}_{\mathrm{D}}; \boldsymbol{v}_h), \\ b_h(\boldsymbol{u}_h, q_h) = b_h^\partial(\boldsymbol{g}_{\mathrm{D}}; q_h). \end{aligned}$$

Due to the weak imposition of a non-zero Dirichlet boundary condition $\boldsymbol{g}_{\mathrm{D}} = \boldsymbol{u}\big|_{\partial\Omega}$, additional boundary facet terms arise from the SIP bilinear form, the pressure-velocity coupling and the mass flux stabilization term:

$$\begin{aligned} a_h^\partial(\boldsymbol{g}_{\mathrm{D}}; \boldsymbol{v}_h) &= \sum_{F \in \mathcal{F}_h^\partial} \frac{\sigma}{h_F} \oint_F \boldsymbol{g}_{\mathrm{D}} \cdot \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{s} - \sum_{F \in \mathcal{F}_h^\partial} \oint_F \boldsymbol{g}_{\mathrm{D}} \cdot \nabla \boldsymbol{v}_h \boldsymbol{n} \,\mathrm{d}\boldsymbol{s}, \\ b_h^\partial(\boldsymbol{g}_{\mathrm{D}}; q_h) &= \sum_{F \in \mathcal{F}_h^\partial} \oint_F q_h (\boldsymbol{g}_{\mathrm{D}} \cdot \boldsymbol{n}) \,\mathrm{d}\boldsymbol{s}, \\ j_h^\partial(\boldsymbol{g}_{\mathrm{D}}; \boldsymbol{v}_h) &= \sum_{F \in \mathcal{F}_h^\partial} \frac{1}{h_F} \oint_F (\boldsymbol{g}_{\mathrm{D}} \cdot \boldsymbol{n})(\boldsymbol{v}_h \cdot \boldsymbol{n}) \,\mathrm{d}\boldsymbol{s}. \end{aligned}$$

In contrast to the Stokes problem, we also additionally have to deal with nonlinear inertia effects. For

Table 4: Errors for the Navier–Stokes potential flow problem with the mass flux penalized (controlled by γ) DG method $\mathbb{P}_4^{\mathrm{dc}}/\mathbb{P}_3^{\mathrm{dc}}$ with $\nu = 10^{-2}$ on an unstructured triangular mesh with $h = 0.1$.

γ	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\boldsymbol{L}^2}$	$\|\nabla_h(\boldsymbol{u}-\boldsymbol{u}_h)\|_{\boldsymbol{L}^2}$	$\|p-p_h\|_{L^2}$	$\|\nabla_h\cdot\boldsymbol{u}_h\|_{L^2}$
0	$1.96\cdot10^{-4}$	$1.11\cdot10^{-2}$	$7.79\cdot10^{-4}$	$9.15\cdot10^{-3}$
1	$3.96\cdot10^{-5}$	$3.22\cdot10^{-3}$	$2.64\cdot10^{-4}$	$2.44\cdot10^{-3}$
10	$7.79\cdot10^{-6}$	$6.28\cdot10^{-4}$	$1.15\cdot10^{-4}$	$4.19\cdot10^{-4}$
10^2	$8.86\cdot10^{-7}$	$7.12\cdot10^{-5}$	$1.01\cdot10^{-4}$	$4.57\cdot10^{-5}$
10^3	$9\cdot10^{-8}$	$7.22\cdot10^{-6}$	$1\cdot10^{-4}$	$4.61\cdot10^{-6}$

treating this, we choose the following standard convection term with upwinding:

$$c_h(\boldsymbol{w}_h;\boldsymbol{u}_h,\boldsymbol{v}_h) = \int_\Omega (\boldsymbol{w}_h\cdot\nabla_h)\boldsymbol{u}_h\cdot\boldsymbol{v}_h\,\mathrm{d}\boldsymbol{x} - \sum_{F\in\mathcal{F}_h^i}\oint_F (\{\!\{\boldsymbol{w}_h\}\!\}\cdot\boldsymbol{n}_F)[\![\boldsymbol{u}_h]\!]\cdot\{\!\{\boldsymbol{v}_h\}\!\}\,\mathrm{d}\boldsymbol{s} + \sum_{F\in\mathcal{F}_h^i}\oint_F \frac{1}{2}|\{\!\{\boldsymbol{w}_h\}\!\}\cdot\boldsymbol{n}_F|[\![\boldsymbol{u}_h]\!]\cdot[\![\boldsymbol{v}_h]\!]\,\mathrm{d}\boldsymbol{s}.$$

For the numerical test, we use $\nu = 10^{-2}$ and compute with the FE pair $\mathbb{P}_4^{\mathrm{dc}}/\mathbb{P}_3^{\mathrm{dc}}$ on an unstructured triangular mesh with $h = 0.1$. The results can be seen in Table 4, and show a clear and remarkable effect on the velocity error. Since the true velocity solution is a fourth degree polynomial, and the method becomes pressure-robust as $\gamma\to\infty$, we observe very small velocity errors for the larger penalization parameters.

6.2 Non-isothermal flows

We consider for our final test Rayleigh–Bénard with infinite Prandtl number (which corresponds physically to silicon oil) and Rayleigh number $Ra = 10^6$. Here we test the CR approximation, with and without the mass flux penalization. The problem setup is taken from [22], and we approximate the problem

$$\begin{aligned}\nabla p - \nu\Delta\boldsymbol{u} &= Ra T\mathbf{e}_2,\\ \nabla\cdot\boldsymbol{u} &= 0,\\ \boldsymbol{u}\cdot\nabla T - \Delta T &= 0,\end{aligned}$$

on $\Omega = (0,1)^2$, with no-slip boundary conditions on all walls, insulated boundary conditions on the top and bottom: $\nabla T\cdot\boldsymbol{n}\big|_{\Gamma_T\cup\Gamma_B} = 0$, and Dirichlet temperature conditions on the left and right walls, $T\big|_{\Gamma_L} = 1$, $T\big|_{\Gamma_R} = 0$. The system is approximated using CR elements for the velocity and pressure, and $S_h = \mathbb{P}_1\cap H^1(\Omega)$ for the temperature approximation. The scheme takes the form

$$\gamma j_h(\boldsymbol{u}_h,\boldsymbol{v}_h) + \nu a_h(\boldsymbol{u}_h,\boldsymbol{v}_h) + b_h(\boldsymbol{v}_h,p_h) = Ra(T_h\mathbf{e}_2,\boldsymbol{v}_h), \tag{54}$$

$$b_h(\boldsymbol{u}_h,q_h) = 0, \tag{55}$$

$$(\boldsymbol{u}_h\cdot\nabla T_h, s_h) + (\nabla T_h,\nabla s_h) = 0, \tag{56}$$

together with appropriate boundary conditions, and we resolve the nonlinear problem using Newton's method.

Simulations were run using (54)–(56), for $\gamma \in \{0, 0.1, 10, 10^3\}$ on a 48×48 uniform mesh that was additional refined along the boundary. Results are shown in Figure 3, and additionally we show a reference solution found using $(\mathbb{P}_2, \mathbb{P}_1, \mathbb{P}_2)$ elements on a 64×64 mesh (we note this solution matches that found in [22]). From the plots, we observe that with no stabilization, the CR solution is very poor: the velocity streamlines are very under-resolved, and the predicted temperature contours are visibly inaccurate. Clear improvement is seen in the solution as the stabilization parameter is increased, and with $\gamma = 10^3$, there is no visual difference between the stabilized CR solution and the reference solution.

7 Conclusions

We have proposed (in effect), analyzed and tested a penalization(s) for nonconforming methods that has an effect on solutions analogous to that of grad-div stabilization for conforming methods. The penalization is a mass flux penalization, and also a broken divergence stabilization in elements where the divergence of the velocity space is not contained in the pressure space.

Our theoretical results include error estimates that reduce the scaling of the velocity error by the pressure from ν^{-1} to $\gamma^{-1/2}\nu^{-1/2}$, analogous to the effect of grad-div for conforming methods [22]. Additionally, we prove limiting behavior results for the penalized DG method on simplex meshes, and in particular that the limit as $\gamma \to \infty$ is the associated (optimal) BDM solution. Therefore, over-stabilization is not possible in this DG context, surprisingly. We also consider the limiting behavior on tensor-product meshes, and find a limit that is seemingly not in the literature, but is in between successive RT element spaces (and is thus optimal).

Since there is now an overwhelming body of work done for grad-div stabilization, we expect that many similar results will also hold with our proposed penalization for DG methods. This includes further error analyses on more complex problems, linear algebra considerations, turbulence models, and so on.

Acknowledgments

The authors would especially like to thank Christoph Lehrenfeld for several fruitful discussions and the invaluable help he provided in using the finite element library `NGSolve` in the context of this work.

References

[1] D. ARNDT, H. DALLMANN, AND G. LUBE, *Local projection FEM stabilization for the time-dependent incompressible Navier–Stokes problem*, Numer. Methods Partial Differential Equations, 31 (2015), pp. 1224–1250.

[2] D. N. ARNOLD, F. BREZZI, AND M. FORTIN, *A stable finite element for the Stokes equations*, Calcolo, 21 (1984), pp. 337–344.

[3] C. BERNARDI AND G. RAUGEL, *Analysis of some finite elements for the Stokes problem*, Math. Comp., 44 (1985), pp. 71–79.

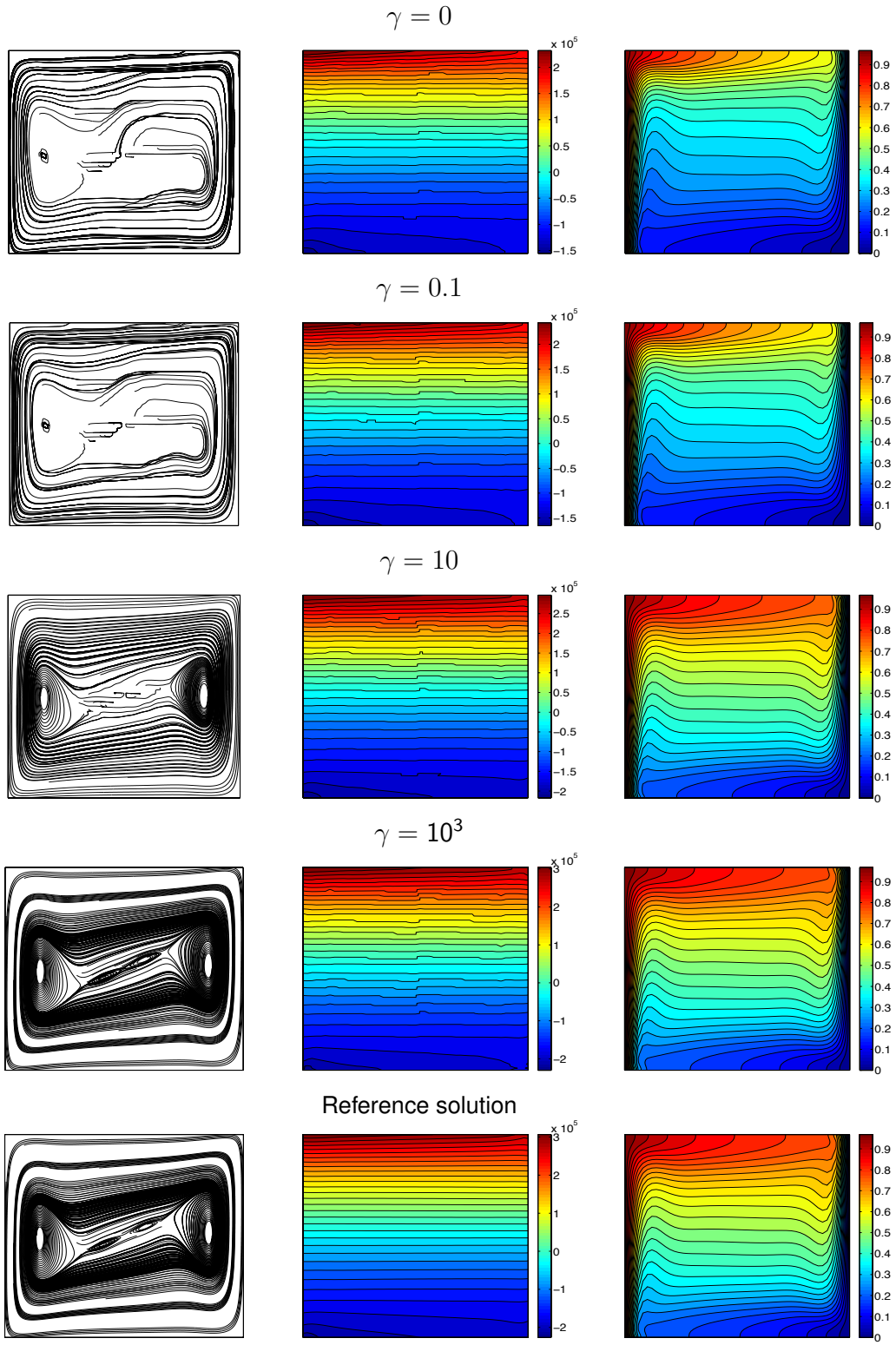

Figure 3: Shown above are plots of CR solutions' velocity streamlines (left), pressure contours (middle), and temperature contours(right), for varying γ, along with a reference solution at the bottom.

[4] D. BOFFI, F. BREZZI, AND M. FORTIN, *Mixed Finite Element Methods and Applications*, Springer-Verlag Berlin Heidelberg, 2013.

[5] S. BRENNER, *Forty years of the Crouzeix–Raviart element*, Numer. Methods Partial Differential Equations, 31 (2015), pp. 367–396.

[6] E. BURMAN AND P. HANSBO, *Stabilized Crouzeix–Raviart element for the Darcy–Stokes problem*, Numer. Methods Partial Differential Equations, 21 (2005), pp. 986–997.

[7] M. A. CASE, V. J. ERVIN, A. LINKE, AND L. G. REBHOLZ, *A connection between Scott–Vogelius and grad-div stabilized Taylor–Hood FE approximations of the Navier–Stokes equations*, SIAM J. Numer. Anal., 49 (2011), pp. 1461–1481.

[8] B. COCKBURN, G. KANSCHAT, AND D. SCHÖTZAU, *A locally conservative LDG method for the incompressible Navier–Stokes equations*, Math. Comp., 74 (2005), pp. 1067–1095.

[9] B. COCKBURN, G. KANSCHAT, AND D. SCHÖTZAU, *A note on discontinuous Galerkin divergence-free solutions of the Navier–Stokes equations*, J. Sci. Comput., 31 (2007), pp. 61–73.

[10] C. J. COTTER AND J. THUBURN, *A finite element exterior calculus framework for the rotating shallow-water equations*, J. Comput. Phys., 257 (2014), pp. 1506–1526.

[11] 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 (1973), pp. 33–75.

[12] D. A. DI PIETRO AND A. ERN, *Mathematical Aspects of Discontinuous Galerkin Methods*, Springer-Verlag Berlin, 2012.

[13] O. DOROK, W. GRAMBOW, AND L. TOBISKA, *Aspects of finite element discretizations for solving the Boussinesq approximation of the Navier–Stokes Equations*, Notes on Numerical Fluid Mechanics: Numerical Methods for the Navier-Stokes Equations., 47 (1994), pp. 50–61.

[14] L. FRANCA AND T. HUGHES, *Two classes of mixed finite element methods*, Comput. Methods Appl. Mech. Engrg., 69 (1988), pp. 89–129.

[15] P. FROLKOVIC, *Consistent velocity approximation for density driven flow and transport*, in Advanced Computational Methods in Engineering, Part 2, 1998, pp. 603–611.

[16] K. GALVIN, A. LINKE, L. G. REBHOLZ, AND N. WILSON, *Stabilizing poor mass conservation in incompressible flow problems with large irrotational forcing and application to thermal convection*, Comput. Methods Appl. Mech. Engrg., 237–240 (2012), pp. 166–176.

[17] S. GANESAN, G. MATTHIES, AND L. TOBISKA, *On spurious velocities in incompressible flow problems with interfaces*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 1193–1202.

[18] J.-F. GERBEAU, C. LE BRIS, AND M. BERCOVIER, *Spurious velocities in the steady flow of an incompressible fluid subjected to external forces*, Int. J. Numer. Meth. Fluids, 25 (1997), pp. 679–695.

[19] V. GIRAULT AND P.-A. RAVIART, *Finite Element Methods for Navier–Stokes Equations*, vol. 5 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1986.

[20] P. M. GRESHO, R. L. LEE, S. T. CHAN, AND J. M. LEONE, *A new finite element for incompressible or Boussinesq fluids*, in Proceedings of the Third International Conference on Finite Elements in Flow Problems, Wiley, 1980.

[21] T. GUDI, *A new error analysis for discontinuous finite element methods for linear elliptic problems*, Math. Comp., 79 (2010), pp. 2169–2189.

[22] E. W. JENKINS, V. JOHN, A. LINKE, AND L. G. REBHOLZ, *On the parameter choice in grad-div stabilization for the Stokes equations*, Adv. Comput. Math., 40 (2014), pp. 491–516.

[23] V. JOHN, A. LINKE, C. MERDON, M. NEILAN, AND L. G. REBHOLZ, *On the divergence constraint in mixed finite element methods for incompressible flows*, SIAM Rev., 59 (2017), pp. 492–544.

[24] S. M. JOSHI, P. J. DIAMESSIS, D. T. STEINMOELLER, M. STASTNA, AND G. N. THOMSEN, *A post-processing technique for stabilizing the discontinuous pressure projection operator in marginally-resolved incompressible inviscid flow*, Comput. & Fluids, 139 (2016), pp. 120–129.

[25] B. KRANK, N. FEHN, W. A. WALL, AND M. KRONBICHLER, *A high-order semi-explicit discontinuous Galerkin solver for 3D incompressible flow with application to DNS and LES of turbulent channel flow*, J. Comput. Phys., 348 (2017), pp. 634–659.

[26] A. LINKE, *A divergence-free velocity reconstruction for incompressible flows*, C. R. Math. Acad. Sci. Paris, 350 (2012), pp. 837–840.

[27] 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 (2014), pp. 782–800.

[28] 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 (2016), pp. 289–309.

[29] A. LINKE AND C. MERDON, *On velocity errors due to irrotational forces in the Navier–Stokes momentum balance*, J. Comput. Phys., 313 (2016), pp. 654–661.

[30] 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 (2016), pp. 304–326.

[31] A. LINKE, L. G. REBHOLZ, AND N. WILSON, *On the convergence rate of grad-div stabilized Taylor–Hood to Scott–Vogelius solutions for incompressible flow problems*, J. Math. Anal. Appl., 381 (2011), pp. 612–626.

[32] G. LUBE AND M. A. OLSHANSKII, *Stable finite-element calculation of incompressible flows using the rotation form of convection*, IMA J. Numer. Anal., 22 (2002), pp. 437–461.

[33] M. OLSHANSKII AND A. REUSKEN, *Grad-div stabilization for Stokes equations*, Math. Comp., 73 (2004), pp. 1699–1718.

[34] M. A. OLSHANSKII, *A low order Galerkin finite element method for the Navier–Stokes equations of steady incompressible flow: a stabilization issue and iterative methods*, Comput. Methods Appl. Mech. Eng., 191 (2002), pp. 5515–5536.

[35] M. A. OLSHANSKII, G. LUBE, T. HEISTER, AND J. LÖWE, *Grad-div stabilization and subgrid pressure models for the incompressible Navier–Stokes equations*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 3975–3988.

[36] D. PELLETIER, A. FORTIN, AND R. CAMARERO, *Are FEM solutions of incompressible flows really incompressible? (or how simple flows can cause headaches!)*, Int. J. Numer. Meth. Fluids, 9 (1989), pp. 99–112.

[37] G. REMPFER, G. B. DAVIES, C. HOLM, AND J. DE GRAAF, *Reducing spurious flow in simulations of electrokinetic phenomena*, J. Chem. Phys., 145 (2016), p. 044901.

[38] B. RIVIÈRE, *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations*, SIAM, 2008.

[39] J. SCHÖBERL, *C++11 Implementation of Finite Elements in NGSolve*, ASC Report 30/2014, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2014, https://www.asc.tuwien.ac.at/~schoeberl/wiki/publications/ngs-cpp11.pdf.

[40] P. W. SCHROEDER AND G. LUBE, *Divergence-free H(div)-FEM for time-dependent incompressible flows with applications to high Reynolds number vortex dynamics*, J. Sci. Comput., (2017), https://doi.org/10.1007/s10915-017-0561-1.

[41] P. W. SCHROEDER AND G. LUBE, *Stabilised dG-FEM for incompressible natural convection flows with boundary and moving interior layers on non-adapted meshes*, J. Comput. Phys., 335 (2017), pp. 760–779.

[42] J. THUBURN AND C. J. COTTER, *A framework for mimetic discretization of the rotating shallow-water equations on arbitrary polygonal grids*, SIAM J. Sci. Comput., 34 (2012), pp. B203–B225.

[43] R. VERFÜRTH, *Error estimates for a mixed finite element approximation of the Stokes equation*, RAIRO Anal. Numer., 18 (1984), pp. 175–182.