IMA Journal of Numerical Analysis (2019) **00**, 1–31
doi:10.1093/imanum/drz044

Quasi-optimal and pressure-robust discretizations of the Stokes equations by new augmented Lagrangian formulations

Christian Kreuzer and Pietro Zanotti*
TU Dortmund, Fakultät für Mathematik, D-44221 Dortmund, Germany
christian.kreuzer@tu-dortmund.de *Corresponding author: zanottipie@gmail.com

[Received on 8 February 2019; revised on 6 August 2019]

We approximate the solution of the stationary Stokes equations with various conforming and nonconforming inf-sup stable pairs of finite element spaces on simplicial meshes. Based on each pair, we design a discretization that is quasi-optimal and pressure-robust, in the sense that the velocity H^1-error is proportional to the best velocity H^1-error. This shows that such a property can be achieved without using conforming and divergence-free pairs. We also bound the pressure L^2-error, only in terms of the best velocity H^1-error and the best pressure L^2-error. Our construction can be summarized as follows. First, a linear operator acts on discrete velocity test functions, before the application of the load functional, and maps the discrete kernel into the analytical one. Second, in order to enforce consistency, we employ a new augmented Lagrangian formulation, inspired by discontinuous Galerkin methods.

Keywords: Stokes equations, finite elements, quasi-optimality, pressure-robustness, augmented Lagrangian formulations.

1. Introduction

We consider the discretization of the stationary Stokes equations

$$-\mu\Delta u + \nabla p = f \quad \text{and} \quad \operatorname{div} u = 0 \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega \tag{1.1}$$

with viscosity $\mu > 0$, in a bounded domain $\Omega \subseteq \mathbb{R}^d$, $d \in \{2, 3\}$. According to the classical approach of Brezzi (1974), we approximate the analytical velocity u and the analytical pressure p by means of discrete spaces V_h and Q_h, which are required to fulfil the so-called inf-sup condition. We additionally assume that V_h and Q_h are finite element spaces on a simplicial mesh of Ω.

To motivate our work, let us focus on the velocity H^1-error, i.e. the error between u and the discrete velocity u_h, measured in the H^1-norm. We refer to Boffi *et al.* (2013, Chapter 5) for the proof of the results listed hereafter. The Céa-type quasi-optimal estimate

$$\|\nabla(u - u_h)\|_{L^2(\Omega)} \leq c \inf_{w_h \in V_h} \|\nabla(u - w_h)\|_{L^2(\Omega)} \tag{1.2}$$

is well known for standard discretizations (see (2.2) and (2.11) below) with conforming and divergence-free pairs, i.e. under the assumptions $V_h \subseteq H^1_0(\Omega)^d$ and $\operatorname{div} V_h = Q_h$. Such pairs have attracted growing interest in recent years; see Scott & Vogelius (1985), Zhang (2007), Guzmán & Neilan (2014, 2018) and the references therein. Owing to (1.2), this class of discretizations seems particularly attractive, because it fully exploits, up to a constant, the approximation properties of the space V_h in the H^1-norm. This prevents, in particular, the following issues.

For standard discretizations with general conforming pairs (see (2.2) and (2.5) below) one typically has

$$\|\nabla(u-u_h)\|_{L^2(\Omega)} \le c\left(\inf_{w_h\in V_h}\|\nabla(u-w_h)\|_{L^2(\Omega)} + \frac{1}{\mu}\inf_{q_h\in Q_h}\|p-q_h\|_{L^2(\Omega)}\right). \tag{1.3}$$

Thus, if $\operatorname{div} V_h \neq Q_h$, the right-hand side suggests that the velocity H^1-error may not be robust with respect to the pressure. This is indeed the case and such an effect is known in the literature as poor mass conservation. It becomes extreme for purely irrotational loads or for small values of the viscosity; see, for instance, Linke (2014). Poor mass conservation discourages, in particular, the use of unbalanced pairs, i.e. pairs V_h/Q_h, so that the approximation power of V_h in the H^1-norm is higher than that of Q_h in the L^2-norm; cf. Remark 3.1.

Recall also that, in the nonconforming case $V_h \nsubseteq H_0^1(\Omega)^d$, estimates in the form

$$\|u-u_h\|_h \le c\left(\inf_{w_h\in V_h}\|u-w_h\|_h + \frac{1}{\mu}\inf_{q_h\in Q_h}\|p-q_h\|_{L^2(\Omega)} + |||(u,p)|||_h\right) \tag{1.4}$$

are often derived. Here $\|\cdot\|_h$ is an extension of the H^1-norm to $H_0^1(\Omega)^d + V_h$ and the seminorm $|||\cdot|||_h$ is defined on (a subspace of) $H_0^1(\Omega)^d \times L^2(\Omega)$. Since the lack of smoothness in V_h is commonly compensated by additional regularity of the load beyond $H^{-1}(\Omega)^d$, the seminorm $|||\cdot|||_h$ cannot be extended to $H_0^1(\Omega)^d \times L_0^2(\Omega)$ and potentially dominates the right-hand side of (1.4) for rough solutions. Therefore, an estimate like (1.3) cannot be expected to hold; cf. Remark 2.3.

There exists a vast literature concerned with the above-mentioned difficulties and several stabilization and correction techniques have been proposed over the years to deal with them. The discretization of Badia *et al.* (2014, Section 6) and the general framework in Veeser & Zanotti (2018a) indicate how the lack of quasi-optimality (1.4) can be avoided for nonconforming pairs . The over-penalized augmented Lagrangian formulation of Boffi & Lovadina (1997) or the grad-div stabilization of Olshanskii & Reusken (2004) may serve to mitigate the impact of poor mass conservation. More recently, Linke (2014), Linke *et al.* (2016) and Lederer *et al.* (2017) proposed a class of discretizations, which differ from standard ones only in the treatment of the load and enjoy the pressure-robust upper bound

$$\|u-u_h\|_h \le c\left(\inf_{w_h\in V_h}\|u-w_h\|_h + |||(u,0)|||_h\right) \tag{1.5}$$

for several conforming and nonconforming pairs. Remarkably, both (1.2) and this bound entail that u_h is insensitive to irrotational perturbations of the load, thus reproducing an important invariance property of u; see e.g. Linke (2014, Section 2). We refer to Gauger *et al.* (2018) and to the references therein for an extensive discussion on the importance of pressure-robustness in the discretization of the Navier–Stokes equations.

In this paper, we show that the quasi-optimal and pressure-robust estimate (1.2) is not a prerogative of conforming and divergence-free pairs, but can be achieved also by (carefully designed) discretizations, based on general inf-sup stable pairs. In this way, we combine the advantages of the various techniques listed above. We also bound the pressure L^2-error only in terms of the best velocity H^1-error and of the best pressure L^2-error. To the best of our knowledge, similar error bounds have previously only been obtained in Verfürth & Zanotti (2019) in the rather specific case of the lowest-order nonconforming Crouzeix–Raviart pair (Crouzeix & Raviart, 1973). In particular, our results make unbalanced pairs a valuable option if one is more interested in the analytical velocity rather than in the analytical pressure.

Our approach is guided by a few simple necessary conditions and builds on two main ingredients. First, we discretize the load with the help of an operator that maps V_h into $H^1_0(\Omega)^d$ and discretely divergence-free into exactly divergence-free functions. The importance of the latter property was first devised in Linke (2014). For this purpose, we solve local discrete Stokes-like problems with inf-sup stable Scott–Vogelius elements (Qin, 1994; Zhang, 2005; Guzmán & Neilan, 2018) on the Alfeld refinement of the mesh; see Fig. 1. Second, we discretize the weak form of the Laplace operator in a way inspired by discontinuous Galerkin (DG) methods in order to enforce the necessary consistency. The resulting discretization can be interpreted as a new augmented Lagrangian formulation; cf. Remark 3.8. The overall cost for assembling the proposed discretization is moderately higher than the one for assembling a standard discretization; cf. Remark 3.6.

The rest of the paper is organized as follows. In Section 2 we set up the abstract framework. In Section 3 we illustrate our construction by means of a model example. Various generalizations are then discussed in Section 4. Finally, in Section 5 we complement our theoretical findings through some numerical experiments.

2. Abstract framework

This section introduces an abstract discretization of (1.1) and the properties in which we are interested. Two basic results are also proved. We use standard notation for Lebesgue and Sobolev spaces.

2.1 *Quasi-optimal discretizations*

Let $\Omega \subseteq \mathbb{R}^d$, $d \in \{2,3\}$ be an open and bounded polytopic domain with Lipschitz-continuous boundary. The weak formulation of the stationary Stokes equations in Ω, with viscosity $\mu > 0$ and load $f \in H^{-1}(\Omega)^d$, looks for $u \in H^1_0(\Omega)^d$ and $p \in L^2_0(\Omega)$ such that

$$\begin{aligned}
&\forall\, v \in H^1_0(\Omega)^d, \qquad && \mu \int_\Omega \nabla u : \nabla v - \int_\Omega p \operatorname{div} v = \langle f, v\rangle, \\
&\forall\, q \in L^2_0(\Omega), && \int_\Omega q \operatorname{div} u = 0.
\end{aligned} \tag{2.1}$$

Here : denotes the Euclidean scalar product of $d \times d$ tensors and $\langle \cdot,\cdot\rangle$ is the dual pairing of $H^{-1}(\Omega)^d$ and $H^1_0(\Omega)^d$. Due to the boundary condition on the analytical velocity u, the analytical pressure p belongs to $L^2_0(\Omega) := \{q \in L^2(\Omega) \mid \int_\Omega q = 0\}$. Problem (2.1) is uniquely solvable, according to Boffi *et al.* (2013, Theorem 8.2.1).

Remark 2.1 (Alternative formulation). Most of our subsequent results remain unchanged in the case that the gradient is replaced by the symmetric gradient in the first equation of (2.1) and the homogeneous Neumann condition is imposed on (a portion of) $\partial\Omega$. The only notable difference is that a piecewise Korn inequality may fail to hold for some of the nonconforming pairs mentioned in Section 4.1; see Arnold (1993) and Brenner (2004). This problem, however, can be overcome, e.g. by an additional jump penalization in the spirit of Veeser & Zanotti (2018b, Section 3.3).

We consider discretizations that mimic the variational structure of problem (2.1). More precisely, we approximate u and p in finite-dimensional linear spaces V_h and Q_h. We require $Q_h \subseteq L^2_0(\Omega)$ and measure the pressure error in the L^2-norm $\|\cdot\|_{L^2(\Omega)}$. Instead, we allow for nonconforming discrete velocity spaces $V_h \nsubseteq H^1_0(\Omega)^d$. In order to measure the velocity error, we assume that an extension $\|\cdot\|_h$ of the H^1-norm $\|\nabla\cdot\|_{L^2(\Omega)}$ to $H^1_0(\Omega)^d + V_h$ is at our disposal. We replace the bilinear forms in (2.1) with discrete surrogates $a_h : V_h \times V_h \to \mathbb{R}$ and $b_h : V_h \times Q_h \to \mathbb{R}$. Moreover, we let $E_h : V_h \to H^1_0(\Omega)^d$ be a linear operator. Hence, we look for a discrete velocity $u_h \in V_h$ and a discrete pressure $p_h \in Q_h$ such that

$$\begin{aligned} \forall\, v_h \in V_h, \qquad & \mu\, a_h(u_h, v_h) + b_h(v_h, p_h) = \langle f, E_h v_h \rangle, \\ \forall\, q_h \in Q_h, \qquad & b_h(u_h, q_h) = 0. \end{aligned} \tag{2.2}$$

To ensure that this problem is uniquely solvable, we assume hereafter that a_h is coercive on V_h and that the pair V_h/Q_h is inf-sup stable, i.e.

$$\forall\, q_h \in Q_h, \qquad \beta \|q_h\|_{L^2(\Omega)} \le \sup_{v_h \in V_h} \frac{b_h(v_h, q_h)}{\|v_h\|_h} \tag{2.3}$$

for some constant $\beta > 0$; see Boffi *et al.* (2013, Corollary 4.2.1). Note, in particular, that the duality $\langle f, E_h v_h \rangle$ is well defined for all $f \in H^{-1}(\Omega)^d$ and $v_h \in V_h$, also in the nonconforming case.

We shall pay special attention to the following property, which guarantees that (u_h, p_h) is a near-best approximation of (u, p) in $V_h \times Q_h$.

Definition 2.2 (Quasi-optimality). Denote by (u, p) and (u_h, p_h) the solutions of (2.1) and (2.2), respectively, with load f and viscosity μ. We say that (2.2) is a quasi-optimal discretization of (2.1) when there is a constant $C \ge 1$ such that

$$\mu \|u - u_h\|_h + \|p - p_h\|_{L^2(\Omega)} \le C \left(\mu \inf_{w_h \in V_h} \|u - w_h\|_h + \inf_{q_h \in Q_h} \|p - q_h\|_{L^2(\Omega)} \right) \tag{2.4}$$

for all $f \in H^{-1}(\Omega)^d$ and $\mu > 0$. We denote by C_{qo} the smallest such constant.

According to Boffi *et al.* (2013, Theorem 5.2.5), discretization (2.2) is quasi-optimal if

$$\begin{aligned} V_h \subseteq H^1_0(\Omega)^d, \qquad & E_h = \mathrm{Id}_{V_h}, \\ a_h(w_h, v_h) = \int_\Omega \nabla w_h : \nabla v_h, \qquad & b_h(v_h, q_h) = -\int_\Omega q_h \operatorname{div} v_h; \end{aligned} \tag{2.5}$$

i.e. if V_h/Q_h is a conforming pair and a_h, b_h and E_h are simple restrictions of their continuous counterparts in (2.1). In Sections 3 and 4 we show that quasi-optimality can also be achieved with nonconforming pairs and/or for different choices of a_h and E_h.

Remark 2.3 (Smoothing by E_h). Since V_h is finite-dimensional, the operator E_h is bounded and the solution of (2.2) depends continuously on the H^{-1}-norm of f. This property, in turn, prevents the issue pointed out in the introduction concerning the seminorm $|\!|\!|\cdot|\!|\!|_h$ in (1.4). Of course, such an observation is of practical interest only if the norm of E_h is of moderate size, so that it does not affect the stability constant of (2.2) too much. We call E_h a 'smoothing' operator, because it increases the smoothness of the elements of V_h whenever $V_h \not\subseteq H_0^1(\Omega)^d$. For conforming pairs, one can let E_h be the identity as in (2.5). This choice is compatible with quasi-optimality but it is possibly not pressure-robust; compare with Section 2.2 below.

REMARK 2.4 (Computational feasibility). It is highly desirable that there are bases $\{\varphi_1, \dots, \varphi_N\}$ and $\{\psi_1, \dots, \psi_M\}$ of V_h and Q_h, respectively, such that the scalars

$$a_h(\varphi_i, \varphi_j), \qquad b(\varphi_i, \psi_k), \qquad \langle f, E_h\varphi_i \rangle$$

can be computed or approximated, up to a prescribed tolerance, with $\mathcal{O}(1)$ operations, for all $i, j = 1, \dots, N$ and $k = 1, \dots, M$. This 'computational feasibility' is not necessary for quasi-optimality but guarantees that the solution of (2.2) can be computed with optimal complexity.

2.2 *Quasi-optimal and pressure-robust discretizations*

The notion of quasi-optimality, in the sense of Definition 2.2, is not fully satisfactory if the approximations of u by u_h and of p by p_h are concerned separately. Indeed, each term in the left-hand side of (2.4) can be guaranteed to be smaller than a prescribed tolerance only if both summands on the right-hand side are sufficiently small. This observation motivates our interest in discretizations of (2.1) such that u_h is a near-best approximation of u in V_h. Ensuring that p_h is a near-best approximation of p in Q_h would also be of interest but goes beyond the purpose of this paper.

The analytical velocity u solving (2.1) can be equivalently characterized as the solution of an elliptic problem. In fact, the second equation imposes that u is divergence-free or, in other words, that it is an element of the kernel

$$Z := \left\{ z \in H_0^1(\Omega)^d \mid \operatorname{div} z = 0 \right\}.$$

Then, testing the first equation with an arbitrary element of Z, we obtain the reduced problem

$$\forall z \in Z, \qquad \mu \int_\Omega \nabla u : \nabla z = \langle f, z \rangle \tag{2.6}$$

which is uniquely solvable, according to the Lax–Milgram lemma and the Friedrichs inequality.

The same structure can be observed at the discrete level. To see this, we first introduce the discrete divergence $\underline{\operatorname{div}}_h : V_h \to Q_h$:

$$\forall q_h \in Q_h, \qquad \int_\Omega q_h \, \underline{\operatorname{div}}_h \, v_h = -b_h(v_h, q_h) \tag{2.7}$$

for all $v_h \in V_h$. The second equation of (2.2) imposes that u_h is discretely divergence-free, i.e. it is an element of the discrete kernel

$$Z_h := \{ z_h \in V_h \mid \underline{\operatorname{div}}_h \, z_h = 0 \}.$$

Then, testing the first equation with an arbitrary element of Z_h, we derive the discrete reduced problem

$$\forall z_h \in Z_h, \qquad \mu \, a_h(u_h, z_h) = \langle f, E_h z_h \rangle \tag{2.8}$$

which is uniquely solvable, since a_h is coercive on V_h. In the same vein as Brezzi (1974, Remark 2.1), it is worth recalling that this is a (possibly) nonconforming discretization of (2.6) because Z_h may fail to be a subspace of Z, even if $V_h \subseteq H_0^1(\Omega)^d$.

Similarly to Definition 2.2, we will be interested in the question of whether u_h is a near-best approximation of u in Z_h. In most cases, this actually amounts to asking whether u_h is near-best in V_h. In fact, if we consider, for instance, the conforming discretization (2.5), the inf-sup condition (2.3) implies

$$\inf_{z_h \in Z_h} \|u - z_h\|_h \le C \inf_{w_h \in V_h} \|u - w_h\|_h \tag{2.9}$$

with $C \le (1 + \beta^{-1})$ according to Boffi *et al.* (2013, Proposition 5.1.3) and Nochetto & Pyo (2004, Lemma 2.1). A sharper version of this bound is pointed out, e.g. in John *et al.* (2017, Remark 4.1), where the inverse of the inf-sup constant β is replaced by the operator norm of a Fortin operator. This is particularly convenient on certain 'extreme' domains. Inequality (2.9) actually holds for more general discretizations under the mild assumption $\|\underline{\mathrm{div}}_h\, z_h\|_{L^2(\Omega)} \le \widetilde{C}\|z - z_h\|_h$ for some $\widetilde{C} > 0$ and for all $z_h \in Z_h$ and $z \in Z$.

Definition 2.5 (Quasi-optimality and pressure-robustness). Denote by u and u_h the solutions of (2.6) and (2.8), respectively, with load f and viscosity μ. We say that (2.2) is a quasi-optimal and pressure-robust discretization of (2.1) when there is a constant $C \ge 1$ such that

$$\|u - u_h\|_h \le C \inf_{w_h \in V_h} \|u - w_h\|_h \tag{2.10}$$

for all $f \in H^{-1}(\Omega)^d$ and $\mu > 0$. We denote by $\mathrm{C}_{\mathrm{qopr}}$ the smallest such constant.

Problem (2.6) reveals that the analytical velocity u is independent of the pressure p. Equivalently, we observe that u depends only on the restriction of the load f to the kernel Z. For L^2-loads, this notable invariance property can be rephrased in terms of the Helmholtz–Hodge decomposition of f; see Linke (2014). The near-best estimate (2.10) prescribes that u_h depends only on u (and not on p). This readily implies that the above invariance property is preserved at the discrete level and justifies the designation 'pressure-robust'.

Discretization (2.2) is known to be quasi-optimal and pressure-robust if

$$\begin{gathered} V_h \subseteq H_0^1(\Omega)^d, \qquad \mathrm{div}\, V_h = Q_h, \qquad E_h = \mathrm{Id}_{V_h}, \\ a_h(w_h, v_h) = \int_\Omega \nabla w_h : \nabla v_h, \qquad b_h(v_h, q_h) = -\int_\Omega q_h \mathrm{div}\, v_h, \end{gathered} \tag{2.11}$$

i.e. if V_h/Q_h is a conforming and divergence-free pair and a_h, b_h and E_h are simple restrictions of their continuous counterparts in (2.1). In fact, in this case, we have $Z_h \subseteq Z$ and (2.8) is a conforming Galerkin discretization of (2.6). Therefore, Céa's lemma and (2.9) imply $\mathrm{C}_{\mathrm{qopr}} \le (1 + \beta^{-1})$. It is our purpose to show that quasi-optimality and pressure-robustness can also be achieved by discretizations other than (2.11).

2.3 *Necessary consistency conditions*

The left- and right-hand sides of (2.4) are seminorms on $Z \times L_0^2(\Omega)$ and the kernel of the latter is $(Z \cap Z_h) \times Q_h$, as a consequence of (2.9). Quasi-optimality actually prescribes that such seminorms are equivalent, because the converse of (2.4) immediately follows from the inclusion $(u_h, p_h) \in Z_h \times Q_h$. Hence, a simple necessary condition is that the kernels of the two seminorms coincide. In other words, whenever the solution (u, p) of (2.1) is in $Z_h \times Q_h$, it must also solve (2.2). This is an algebraic consistency condition, which can be rephrased in terms of the forms a_h and b_h and of the operator E_h, in the spirit of Veeser & Zanotti (2018a, Definition 2.7).

Lemma 2.6 (Consistency for quasi-optimality). Assume that (2.2) is a quasi-optimal discretization of (2.1). Then necessarily we have

$$\forall\, v_h \in V_h,\ p \in Q_h, \qquad \int_\Omega p(\underline{\operatorname{div}}_h\, v_h - \operatorname{div} E_h v_h) = 0 \tag{2.12a}$$

and

$$\forall\, u \in Z \cap Z_h,\ v_h \in V_h, \qquad a_h(u, v_h) = \int_\Omega \nabla u : \nabla E_h v_h. \tag{2.12b}$$

Proof. Denote by (u, p) the solution of (2.1) and assume first $u = 0$ and $p \in Q_h$. Quasi-optimality implies that the solution (u_h, p_h) of (2.2) satisfies $u_h = 0$ and $p_h = p$. Comparing the first equations of (2.1) and (2.2), we derive the identity $b_h(v_h, p) = -\int_\Omega p \operatorname{div} E_h v_h$ for all $v_h \in V_h$. Condition (2.12a) then follows from the definition of $\underline{\operatorname{div}}_h$ in (2.7). Next assume $u \in Z \cap Z_h$ and $p = 0$. Since quasi-optimality implies $u_h = u$ and $p_h = 0$, condition (2.12b) can be derived comparing the first equations of (2.1) and (2.2) as before. □

The conforming discretization (2.5) is a simple option to fulfil (2.12), but not the only one possible. Examples with nonconforming discrete velocity space can be found in Badia *et al.* (2014, Section 6) and Verfürth & Zanotti (2019). Standard nonconforming discretizations, like the one of Crouzeix & Raviart (1973), do not fulfil (2.12) because they do not employ a smoothing operator. It is also worth noticing that (2.12) involves the interplay of a_h and b_h with E_h. This indicates that the discretization of the differential operator in (1.1) and the one of the corresponding load should not be regarded as independent tasks.

Proceeding similarly to Lemma 2.6, we derive necessary conditions for quasi-optimality and pressure-robustness.

Lemma 2.7 (Consistency for quasi-optimality and pressure-robustness). Assume that (2.2) is a quasi-optimal and pressure-robust discretization of (2.1). Then necessarily we have

$$E_h(Z_h) \subseteq Z \tag{2.13a}$$

and

$$\forall\, u \in Z \cap Z_h,\ z_h \in Z_h, \qquad a_h(u, z_h) = \int_\Omega \nabla u : \nabla E_h z_h. \tag{2.13b}$$

Proof. Let $z_h \in Z_h$ be such that $\operatorname{div} E_h z_h \neq 0$. Assuming that $(u, p) = (0, \operatorname{div} E_h z_h)$ solves (2.1), we infer $\langle f, E_h z_h \rangle = -\|\operatorname{div} E_h z_h\|^2_{L^2(\Omega)} \neq 0$. Inserting this information in (2.8), we obtain $u_h \neq 0$. Therefore, we have $\|u - u_h\|_h > \inf_{v_h \in V_h} \|u - v_h\|_h = 0$, which contradicts quasi-optimality and pressure-robustness. This proves (2.13a). Assertion (2.13b) may be checked similarly to (2.12b) in Lemma 2.6. □

Condition (2.13b) is clearly necessary for (2.12b), while (2.13a) is neither necessary nor sufficient for (2.12a). We mention also that (2.13a) differs from the condition exploited in Linke *et al.* (2016) to achieve pressure-robustness, in that here E_h is required to map into $H^1_0(\Omega)^d$ and not only into $H_{\mathrm{div}}(\Omega)$; cf. Remark 2.3.

REMARK 2.8 (Failure of $E_h = \mathrm{Id}_{V_h}$). If V_h/Q_h is a conforming and divergence-free pair, the abstract discretization (2.2) with (2.11) verifies the first necessary condition in Lemma 2.7. If, instead, the pair is conforming but not divergence-free, we have $Z_h \nsubseteq Z$. In this case, the operator E_h cannot coincide with the identity on Z_h.

In the next sections we design some new discretizations, proceeding as follows. Given an inf-sup stable pair V_h/Q_h, together with the corresponding bilinear form b_h, we construct a_h and E_h so that the necessary conditions in Lemmas 2.6 and 2.7 hold true. Then, we use standard techniques from the analysis of saddle point problems to verify (2.4) and (2.10) and to bound the constants C_{qo} and $\mathrm{C}_{\mathrm{qopr}}$. Alternatively, one could exploit Veeser & Zanotti (2018a, Theorem 4.14), which guarantees that (2.13) is a sufficient condition for quasi-optimality and pressure-robustness. Such a result also provides a formula for $\mathrm{C}_{\mathrm{qopr}}$. Analogously, generalizing the framework of Veeser & Zanotti (2018a), one could also show that (2.12) is a sufficient condition for quasi-optimality and derive a formula for C_{qo}. We prefer to proceed as indicated, to make sure this paper can be read independently of Veeser & Zanotti (2018a).

3. A paradigmatic discretization

Assume that we are given an inf-sup stable pair V_h/Q_h, together with the corresponding bilinear form b_h. A possible strategy to fulfil the necessary conditions (2.12a) and (2.13a) is to employ a 'divergence-preserving' smoothing operator, i.e.

$$\forall\, v_h \in V_h, \qquad \operatorname{div} E_h v_h = \underline{\mathrm{div}}_h\, v_h. \tag{3.1}$$

Once such operator is given, conditions (2.12b) and (2.13b) prescribe the restriction of a_h on $(Z \cap Z_h) \times V_h$. Then, inspired by Arnold (1982) and Veeser & Zanotti (2018b), we extend the resulting form to $V_h \times V_h$ in a way that additionally ensures symmetry and coercivity. In order to keep the exposition as clear as possible, we first exemplify this idea in a model setting. We postpone various generalizations to the next section.

3.1 *The unbalanced $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair*

We consider hereafter pairs of finite element spaces on a face-to-face simplicial mesh $\mathcal{M}$ of Ω in the sense of Di Pietro & Ern (2012, Definition 1.36). We write c for a nondecreasing and non-negative function of the shape parameter of $\mathcal{M}$, which possibly depends also on different quantities (e.g. the space dimension), but not on other properties of $\mathcal{M}$ nor on the viscosity μ. Such a constant may change at different occurrences. We occasionally abbreviate $a \leq cb$ as $a \lesssim b$ and $c^{-1}b \leq a \leq cb$ as $a \eqsim b$.

For all integers $\ell \geq 0$, we denote by $\mathbb{P}_\ell(S)$ the space of polynomials with total degree $\leq \ell$ on a simplex $S \subseteq \mathbb{R}^d$. The space of H^k-conforming elementwise polynomials on $\mathcal{M}$ then reads

$$S^k_\ell := \{v \in H^k(\Omega) \mid \forall\, K \in \mathcal{M},\;\; v_{|K} \in \mathbb{P}_\ell(K)\} \tag{3.2}$$

with $k \in \{0,1\}$ and the convention $H^0(\Omega) := L^2(\Omega)$. Motivated by the homogeneous boundary condition in (1.1), we consider the subspaces

$$\mathring{S}^1_\ell := S^1_\ell \cap H^1_0(\Omega) \qquad \text{and} \qquad \widehat{S}^k_\ell := S^k_\ell \cap L^2_0(\Omega). \tag{3.3}$$

To exemplify our construction, we assume $d = 2$ for the remaining part of this section. We consider the conforming $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair, which is given by

$$V_h = (\mathring{S}^1_\ell)^2 \qquad \text{and} \qquad Q_h = \widehat{S}^0_{\ell-2}, \qquad b_h(v_h, q_h) = -\int_\Omega q_h \operatorname{div} v_h \tag{3.4}$$

with $\ell \geq 2$. The inf-sup condition (2.3) holds with $\beta^{-1} \leq c$; see Boffi *et al.* (2013, Remark 8.6.2).

REMARK 3.1 (Unbalanced pairs). The $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair is unbalanced in the sense that the approximation power $\ell - 1$ of the discrete pressure space in the L^2-norm is strictly less than the approximation power ℓ of the discrete velocity space in the H^1-norm. Other examples can be obtained, enriching the velocity space of any inf-sup stable pair. The use of conforming unbalanced pairs, in combination with the standard discretization (2.5), is discouraged by the error estimate (1.3) and Remark 2.8; see also Boffi *et al.* (2013, Remark 8.6.2). Still, quasi-optimal and pressure-robust discretizations based on such pairs would be a valuable option if one is more interested in the analytical velocity than in the analytical pressure.

The discrete divergence $\underline{\operatorname{div}}_h$ in the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair coincides with the L^2-orthogonal projection of the analytical divergence onto $\widehat{S}^0_{\ell-2}$. Since (2.7) actually holds for all discrete pressures in $S^0_{\ell-2}$, we can compute $\underline{\operatorname{div}}_h$ elementwise as follows:

$$\underline{\operatorname{div}}_h\, v_h = \Pi^K_{\ell-2} \operatorname{div} v_h \quad \text{in } K \tag{3.5}$$

for all $v_h \in (\mathring{S}^1_\ell)^2$ and $K \in \mathcal{M}$, where $\Pi^K_{\ell-2}$ is the L^2-orthogonal projection onto $\mathbb{P}_{\ell-2}(K)$. Therefore, denoting by Z^{ub}_h the discrete kernel, we conclude $Z^{\mathrm{ub}}_h \not\subseteq Z$.[1] This confirms that the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair is conforming but not divergence-free.

The abstract discretization (2.2) with (2.5), based on the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair, states $u_h \in (\mathring{S}^1_\ell)^2$ and $p_h \in \widehat{S}^0_{\ell-2}$ such that

$$\begin{aligned} \forall\, v_h \in (\mathring{S}^1_\ell)^2, \qquad & \mu \int_\Omega \nabla u_h : \nabla v_h - \int_\Omega p_h \operatorname{div} v_h = \langle f, v_h \rangle, \\ \forall\, q_h \in \widehat{S}^0_{\ell-2}, \qquad & \int_\Omega q_h \operatorname{div} u_h = 0. \end{aligned} \tag{3.6}$$

3.2 *Local inversion of the divergence*

Proceeding as in Verfürth & Zanotti (2019), we enforce (3.1) with the help of local right inverses of the divergence. Such operators can be defined solving discrete Stokes-like problems on the Alfeld

[1] The superscript '*ub*' stands for 'unbalanced'. Throughout this section, we use it to label spaces, forms and operators related to the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair.

FIG. 1. Generic triangle $K \in \mathcal{M}$ (left) and Alfeld refinement $\mathcal{M}_K$ (right).

refinement $\mathcal{M}_K$ of each triangle $K \in \mathcal{M}$. The refinement $\mathcal{M}_K$ is obtained connecting each vertex of K with the barycentre; cf. Fig. 1. We call $\mathcal{M}_K$ the 'Alfeld refinement', according to Fu *et al.* (2018), although it is also called the 'barycentric refinement' in other references (Guzmán & Neilan, 2018). The latter terminology has the disadvantage that it sometimes refers to different refinements; see e.g. Buffa & Christiansen (2007).

For $\ell \in \mathbb{N}$, we define the local spaces

$$\mathring{S}^1_\ell(\mathcal{M}_K) \qquad \text{and} \qquad \widehat{S}^0_{\ell-1}(\mathcal{M}_K) \tag{3.7}$$

on $\mathcal{M}_K$ similarly to the global spaces $\mathring{S}^1_\ell$ and $\widehat{S}^0_{\ell-1}$ in (3.3). In particular, all $v_k \in \mathring{S}^1_\ell(\mathcal{M}_K)$ vanish on ∂K and all $q_K \in \widehat{S}^0_{\ell-1}(\mathcal{M}_K)$ are such that $\int_K q_K = 0$. The pair $\mathring{S}^1_\ell(\mathcal{M}_K)^2/\widehat{S}^0_{\ell-1}(\mathcal{M}_K)$ is conforming and divergence-free in K.

According to Guzmán & Neilan (2018, Theorem 3.1) we have the local inf-sup stability

$$\forall\, q_K \in \widehat{S}^0_{\ell-1}(\mathcal{M}_K), \qquad \|q_K\|_{L^2(K)} \le c \sup_{v_K \in \mathring{S}^1_\ell(\mathcal{M}_K)^2} \frac{\int_K q_K \operatorname{div} v_K}{\|\nabla v_K\|_{L^2(K)}}. \tag{3.8}$$

This entails that we can define a linear operator $R^K_\ell : L^2(\Omega) \to H^1_0(\Omega)^2$ as follows. Given $q \in L^2(\Omega)$, let $u_K = u_K(q) \in \mathring{S}^1_\ell(\mathcal{M}_K)^2$ and $p_K = p_K(q) \in \widehat{S}^0_{\ell-1}(\mathcal{M}_K)$ solve

$$\begin{aligned} &\forall\, v_K \in \mathring{S}^1_\ell(\mathcal{M}_K)^2, && \int_K \nabla u_K : \nabla v_K - \int_K p_K \operatorname{div} v_K = 0, \\ &\forall\, q_K \in \widehat{S}^0_{\ell-1}(\mathcal{M}_K), && \int_K q_K \operatorname{div} u_K = \int_K q_K q. \end{aligned} \tag{3.9}$$

Hence, we set

$$R^K_\ell q := u_K \quad \text{in } K \qquad \text{and} \qquad R^K_\ell q := 0 \quad \text{in } \Omega \setminus K.$$

PROPOSITION 3.2 (Local right inverses). Let $K \in \mathcal{M}$ be a mesh element and $\ell \in \mathbb{N}$. The operator R^K_ℓ is well defined and, for all $q \in L^2(\Omega)$, we have

$$\|\nabla R^K_\ell q\|_{L^2(\Omega)} \le c\|q\|_{L^2(K)} \tag{3.10a}$$

and

$$q_{|K} \in \widehat{S}^0_{\ell-1}(\mathcal{M}_K) \quad \Longrightarrow \quad \operatorname{div} R^K_\ell q = q \quad \text{in } K. \tag{3.10b}$$

Proof. The operator R^K_ℓ is well defined and satisfies (3.10a) in view of the local inf-sup (3.8) and Boffi *et al.* (2013, Corollary 4.2.1). The property in (3.10b) directly follows from the second equation of problem (3.9), because $\operatorname{div} u_K \in \widehat{S}^0_{\ell-1}(\mathcal{M}_K)$. □

Remark 3.3 (Computation of the local right inverses). In what follows we shall need to compute $R^K_\ell q$ for all $K \in \mathcal{M}$ and various $q \in S^0_{\ell-1}$. An efficient strategy for this purpose is to precompute the solution of (3.9) on a reference triangle K_{ref} for all possible loads q_{ref} in a basis of $\mathbb{P}_{\ell-1}(K_{\mathrm{ref}})$. The computational complexity of this task depends only on ℓ. Then the solution of (3.9) in K can be obtained in terms of the corresponding solution in K_{ref} by means of the contravariant Piola transformation; see Boffi *et al.* (2013, Section 2.1.3).

Here we have considered the two-dimensional case only in order to be consistent with the simplification introduced in Section 3.1. The same construction is actually possible in any space dimension $d \geq 2$.

3.3 *A new augmented Lagrangian formulation*

We now propose a new discretization of the Stokes equations, based on the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair. The first ingredient of our construction is a linear operator $E^{\mathrm{ub}}_h : (\mathring{S}^1_\ell)^2 \to H^1_0(\Omega)^2$ fulfilling (3.1). In view of $Z^{\mathrm{ub}}_h \nsubseteq Z$ and Remark 2.8, the identity on $(\mathring{S}^1_\ell)^2$ cannot accommodate this property. Therefore, we introduce a 'divergence correction' $R^{\mathrm{ub}}_h : (\mathring{S}^1_\ell)^2 \to H^1_0(\Omega)^2$,

$$R^{\mathrm{ub}}_h v_h := \sum_{K \in \mathcal{M}} R^K_\ell (\underline{\operatorname{div}}_h\, v_h - \operatorname{div} v_h).$$

Proposition 3.4 (Divergence-preserving smoothing operator). The operator $E^{\mathrm{ub}}_h : (\mathring{S}^1_\ell)^2 \to H^1_0(\Omega)^2$ given by

$$E^{\mathrm{ub}}_h v_h := v_h + R^{\mathrm{ub}}_h v_h \tag{3.11}$$

fulfils (3.1) and is such that, for all $v_h \in (\mathring{S}^1_\ell)^2$,

$$\|\nabla(v_h - E^{\mathrm{ub}}_h v_h)\|_{L^2(\Omega)} \eqsim \|\underline{\operatorname{div}}_h\, v_h - \operatorname{div} v_h\|_{L^2(\Omega)}. \tag{3.12}$$

Proof. For all $v_h \in (\mathring{S}^1_\ell)^2$ and $K \in \mathcal{M}$,

$$\operatorname{div} E^{\mathrm{ub}}_h\, v_h = \operatorname{div} v_h + \operatorname{div} R^K_\ell (\underline{\operatorname{div}}_h\, v_h - \operatorname{div} v_h) \quad \text{in } K.$$

In view of (3.5) we have $\int_K (\underline{\operatorname{div}}_h\, v_h - \operatorname{div} v_h) = 0$. Since the inclusion $v_h \in (\mathring{S}^1_\ell)^2$ also implies $(\underline{\operatorname{div}}_h\, v_h - \operatorname{div} v_h)_{|K} \in \mathbb{P}_{\ell-1}(K)$, Proposition 3.2 and the identity above ensure that E^{ub}_h fulfils (3.1). This, in turn, easily implies the lower bound '$\gtrsim$' in (3.12). The corresponding upper bound '$\lesssim$' is a consequence of the identity $\|\nabla(v_h - E^{\mathrm{ub}}_h v_h)\|_{L^2(K)} = \|\nabla R^K_\ell v_h\|_{L^2(K)}$, $K \in \mathcal{M}$, combined with (3.10a). □

The second ingredient of our construction is a suitable bilinear form a_h. Accounting for the definition of E_h^{ub} in (3.11), the necessary conditions (2.12b) and (2.13b) prescribe

$$a_h(u, v_h) = \int_\Omega \nabla u : \nabla v_h + \int_\Omega \nabla u : \nabla R_h^{\mathrm{ub}} v_h \tag{3.13}$$

for all $u \in Z \cap Z_h^{\mathrm{ub}}$ and $v_h \in (\mathring{S}_\ell^1)^2$. A simple option would be to let the right-hand side define a_h on $(\mathring{S}_\ell^1)^2 \times (\mathring{S}_\ell^1)^2$. Still, it has to be noticed that the second summand $\int_\Omega \nabla u : \nabla R_h^{\mathrm{ub}} v_h = -\sum_{K\in\mathcal{M}} \int_K \Delta u \cdot R_h^{\mathrm{ub}} v_h$ cannot be expected to vanish. Therefore, it obstructs the symmetry and possibly also the nondegeneracy of a_h. To overcome this problem, we observe that R_h^{ub} vanishes on $Z \cap Z_h^{\mathrm{ub}}$, according to (3.12). This suggests reestablishing symmetry and nondegeneracy mimicking the construction of the symmetric interior penalty (DG-SIP) discretization of second-order problems; see Arnold (1982) or Di Pietro & Ern (2012, Section 4.2.1). Thus, we set $a_h = a_h^{\mathrm{ub}}$, where

$$\begin{aligned} a_h^{\mathrm{ub}}(w_h, v_h) := &\int_\Omega \nabla w_h : \nabla v_h + \int_\Omega \nabla w_h : \nabla R_h^{\mathrm{ub}} v_h \\ &+ \int_\Omega \nabla R_h^{\mathrm{ub}} w_h : \nabla v_h + \eta \int_\Omega \nabla R_h^{\mathrm{ub}} w_h : \nabla R_h^{\mathrm{ub}} v_h, \end{aligned} \tag{3.14}$$

where $\eta > 0$ is a penalty parameter. Note that a_h^{ub} fulfils the above-mentioned necessary conditions.

The abstract discretization (2.2) with the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair, $a_h = a_h^{\mathrm{ub}}$ and $E_h = E_h^{\mathrm{ub}}$ reads as follows. Find $u_h \in (\mathring{S}_\ell^1)^2$ and $p_h \in \widehat{S}_{\ell-2}^0$ such that

$$\begin{aligned} &\forall\, v_h \in (\mathring{S}_\ell^1)^2, \qquad \mu\, a_h^{\mathrm{ub}}(u_h, v_h) - \int_\Omega p_h \operatorname{div} v_h = \langle f, E_h^{\mathrm{ub}} v_h\rangle, \\ &\forall\, q_h \in \widehat{S}_{\ell-2}^0, \qquad \int_\Omega q_h \operatorname{div} u_h = 0. \end{aligned} \tag{3.15}$$

We begin our discussion on the new discretization by checking that a solution (u_h, p_h) exists and is unique. In view of the above-mentioned inf-sup stability of the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair, it suffices to prove that a_h^{ub} is coercive on $(\mathring{S}_\ell^1)^2$. We proceed similarly to Di Pietro & Ern (2012, Lemma 4.1.2).

Lemma 3.5 (Coercivity of a_h^{ub}). The bilinear form a_h^{ub} is coercive on $(\mathring{S}_\ell^1)^2$ for all $\eta > 1$ and we have

$$a_h^{\mathrm{ub}}(v_h, v_h) \geq \left(1 - \frac{1}{\eta}\right) \|\nabla v_h\|_{L^2(\Omega)}^2$$

for all $v_h \in (\mathring{S}_\ell^1)^2$.

Proof. Let $v_h \in (\mathring{S}_\ell^1)^2$. Setting $w_h = v_h$ in (3.14), we obtain

$$a_h^{\mathrm{ub}}(v_h, v_h) = \|\nabla v_h\|_{L^2(\Omega)}^2 + \eta \|\nabla R_h^{\mathrm{ub}} v_h\|_{L^2(\Omega)}^2 + 2\int_\Omega \nabla v_h : \nabla R_h^{\mathrm{ub}} v_h.$$

The Cauchy–Schwarz and the weighted Young's inequality further provide the upper bound $2\left|\int_\Omega \nabla v_h : \nabla R_h^{\mathrm{ub}} v_h\right| \leq \eta^{-1}\|\nabla v_h\|^2_{L^2(\Omega)} + \eta\|\nabla R_h^{\mathrm{ub}} v_h\|^2_{L^2(\Omega)}$. Inserting this inequality into the previous identity concludes the proof. □

Two remarks on discretization (3.15) are in order.

Remark 3.6 (Feasibility of the new discretization). Assume that $\{\varphi_1, \dots, \varphi_N\}$ and $\{\psi_1, \dots, \psi_M\}$ are nodal bases of $(\mathring{S}^1_\ell)^2$ and $\widehat{S}^0_{\ell-2}$, respectively. All functions φ_i and ψ_k, with $i = 1, \dots, N$ and $k = 1, \dots, M$, are locally supported. Hence, the construction of $E_h^{\mathrm{ub}}\varphi_i$ involves the solution of a limited number of local problems (3.9) and we have $\operatorname{supp}(E_h^{\mathrm{ub}}\varphi_i) \subseteq \operatorname{supp}(\varphi_i)$. Moreover, thanks to the local characterization of the discrete divergence (3.5), the entire computation of $E_h^{\mathrm{ub}}\varphi_i$ requires $\mathcal{O}(1)$ operations. This entails that the bilinear forms $a_h^{\mathrm{ub}}(\varphi_i, \varphi_j)$ and $\int_\Omega \psi_k \operatorname{div}\varphi_i$ and the linear form $\langle f, E_h^{\mathrm{ub}}\varphi_i\rangle$ can be evaluated with $\mathcal{O}(1)$ operations for all $i,j = 1, \dots, N$ and $k = 1, \dots, M$. Thus, the discretization (3.15) is computationally feasible, in the sense of Remark 2.4. Let us mention also that the stiffness matrices associated with a_h^{ub} and its counterpart in (3.6) are of course different but, for all $\eta > 1$, their condition numbers differ, at most, by the ratio of the continuity and the coercivity constants of a_h^{ub}. This ratio is bounded by $c\eta^2(\eta - 1)^{-1}$, as a consequence of Proposition 3.4 and Lemma 3.5.

Remark 3.7 (Use of the Scott–Vogelius pair). As mentioned in the introduction, this paper primarily aims to show that quasi-optimality and pressure-robustness, in the sense of Definition 2.5, is not a prerogative of the discretizations built with conforming and divergence-free pairs. Therefore, the fact that the Scott–Vogelius pair is involved through E_h in the construction of (3.7) may appear a contradiction. This dilemma can be clarified by noticing that we employ the Scott–Vogelius pair only in the solution of local problems and not of global ones. In particular, the system matrix associated with (3.15) is spectrally equivalent to and has the same dimension as the system matrix of the standard discretization (3.6) with the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair.

The following remarks connect (3.15) with other existing discretizations.

Remark 3.8 (Connection with augmented Lagrangian formulations). In view of (3.12), the last summand $\eta\int_\Omega \nabla R_h^{\mathrm{ub}} w_h : \nabla R_h^{\mathrm{ub}} v_h$ in the definition of a_h^{ub} penalizes the functions that are in the discrete kernel Z_h^{ub} and not in Z. More precisely, the penalization is equivalent to $\eta\int_\Omega \operatorname{div} w_h \operatorname{div} v_h$ on Z_h^{ub}. This indicates that (3.15) can be interpreted as a new augmented Lagrangian formulation for the Stokes problem; see Boffi *et al.* (2013, Section 6.1). The additional terms enforcing consistency and symmetry distinguish our formulation from previous ones.

Remark 3.9 (Connection with DG discretizations). The DG-SIP bilinear form in Arnold (1982) consists of four terms. The first two terms serve to accommodate consistency; see Di Pietro & Ern (2012, Section 4.2) or Veeser & Zanotti (2018b). In particular, the second one arises due to the use of possibly nonconforming, i.e. discontinuous, functions. The two remaining terms are designed to further enforce symmetry and coercivity, respectively, still preserving consistency. The same structure can be observed in the form a_h^{ub}. Here nonconformity has to be intended in the sense that $Z_h^{\mathrm{ub}} \nsubseteq Z$, i.e. discretely divergence-free functions are possibly not divergence-free. A notable difference from the DG-SIP bilinear form is that the coercivity of a_h^{ub} can be guaranteed for all $\eta > 1$ and not only for sufficiently large η.

Remark 3.10 (Connection with recovered finite element method discretizations). Rearranging terms in definition (3.14), in the same vein as Di Pietro & Ern (2012, Section 4.3.3), we see that the form a_h^{ub} can be rewritten as

$$a_h^{\mathrm{ub}}(w_h, v_h) = \int_\Omega \nabla E_h^{\mathrm{ub}} w_h : \nabla E_h^{\mathrm{ub}} v_h + (\eta - 1) \int_\Omega \nabla R_h^{\mathrm{ub}} w_h : \nabla R_h^{\mathrm{ub}} v_h. \tag{3.16}$$

This sheds additional light on the condition $\eta > 1$ in Lemma 3.5 and provides an interesting connection with the recovered finite element method of Georgoulis & Pryer (2018).

3.4 *Error estimates*

We now aim to show that, unlike (3.6), (3.15) is a quasi-optimal and pressure-robust discretization of (2.1). As a preliminary step we bound the consistency error generated by the last two terms in the definition of a_h^{ub}. Such terms can be expected to generate a consistency error, as they were artificially added to the right-hand side of (3.13).

LEMMA 3.11 (Consistency error). Let $\eta > 1$ be given. We have

$$\left| \int_\Omega \nabla z_h : \nabla E_h^{\mathrm{ub}} v_h - a_h^{\mathrm{ub}}(z_h, v_h) \right| \lesssim \eta \inf_{z \in Z} \|\nabla(z - z_h)\|_{L^2(\Omega)} \|\nabla v_h\|_{L^2(\Omega)} \tag{3.17}$$

for all $z_h \in Z_h^{\mathrm{ub}}$ and $v_h \in (\mathring{S}_\ell^1)^2$.

Proof. The definitions of a_h^{ub} and E_h^{ub} imply

$$\int_\Omega \nabla z_h : \nabla E_h^{\mathrm{ub}} v_h - a_h^{\mathrm{ub}}(z_h, v_h) = - \int_\Omega \nabla R_h^{\mathrm{ub}} z_h : \nabla (v_h + \eta R_h^{\mathrm{ub}} v_h).$$

The equivalence (3.12) reveals, in particular, $\|\nabla R_h^{\mathrm{ub}} z_h\|_{L^2(\Omega)} \lesssim \|\nabla(z - z_h)\|_{L^2(\Omega)}$ for all $z \in Z$. The characterization (3.5) of the discrete divergence $\underline{\mathrm{div}}_h$ and (3.12) entail also $\|\nabla(v_h + \eta R_h^{\mathrm{ub}} v_h)\|_{L^2(\Omega)} \lesssim \eta \|\nabla v_h\|_{L^2(\Omega)}$. Inserting these bounds into the identity above concludes the proof. □

Recall from Section 2.2 that the discrete velocity u_h solving (3.15) is in the discrete kernel Z_h^{ub} and can be equivalently characterized through the reduced problem

$$\forall\, z_h \in Z_h^{\mathrm{ub}}, \qquad \mu\, a_h^{\mathrm{ub}}(u_h, z_h) = \langle f, E_h^{\mathrm{ub}} z_h \rangle. \tag{3.18}$$

THEOREM 3.12 (Quasi-optimality and pressure-robustness). For all $\eta > 1$, problem (3.15) is a quasi-optimal and pressure-robust discretization of (2.1) with constant $\mathrm{C}_{\mathrm{qopr}} \leq c\eta^2(\eta - 1)^{-1}$.

Proof. Denote by $u \in Z$ and $u_h \in Z_h^{\mathrm{ub}}$ the solutions of problems (2.6) and (3.18), respectively, with load $f \in H^{-1}(\Omega)^2$ and viscosity $\mu > 0$. Let $z_h \in Z_h^{\mathrm{ub}}$ be arbitrary and define $v_h := u_h - z_h$. Lemma 3.5 and problem (3.18) reveal

$$\left(1 - \frac{1}{\eta}\right) \|\nabla(u_h - z_h)\|^2_{L^2(\Omega)} \leq \frac{1}{\mu} \langle f, E_h^{\mathrm{ub}} v_h \rangle - a_h^{\mathrm{ub}}(z_h, v_h).$$

Since $v_h \in Z_h^{\mathrm{ub}}$, we have $E_h^{\mathrm{ub}} v_h \in Z$ as a consequence of Proposition 3.4. Hence, problem (2.6) yields $\mu^{-1}\langle f, E_h^{\mathrm{ub}} v_h \rangle = \int_\Omega \nabla u : \nabla E_h^{\mathrm{ub}} v_h$. We insert this identity into the previous inequality and invoke

Proposition 3.4 and Lemma 3.11. Owing to the inclusion $u \in Z$, it results in

$$\|\nabla(u_h - z_h)\|_{L^2(\Omega)} \le c\eta^2(\eta - 1)^{-1}\|\nabla(u - z_h)\|_{L^2(\Omega)}.$$

We conclude by taking the infimum over all $z_h \in Z_h$ and recalling (2.9). □

Let us mention that a better bound of the constant $\mathrm{C_{qopr}}$ in terms of η, namely $\mathrm{C_{qopr}} \le c\eta(\eta-1)^{-1/2}$, could be obtained with the help of Veeser & Zanotti (2018a, Theorem 4.14). Both this estimate and the one in Theorem 3.12 suggest setting $\eta = 2$. The next remark additionally confirms that we may have $\mathrm{C_{qopr}} \to +\infty$ as $\eta \to +\infty$, thus pointing out the importance of explicitly knowing a safe value of the penalty parameter.

REMARK 3.13 (Locking effect). The penalization in a_h^{ub} imposes that the solution u_h^{ub} of (3.18) approaches the subspace $Z \cap Z_h^{\mathrm{ub}}$ for $\eta \to +\infty$, as a consequence of Proposition 3.4. This entails that the constant $\mathrm{C_{qopr}}$ in Theorem 3.12 remains bounded in the limit $\eta \to +\infty$ only if the equivalence

$$\inf_{z_h \in Z \cap Z_h^{\mathrm{ub}}} \|\nabla(z - z_h)\|_{L^2(\Omega)} \overset{!}{\eqsim} \inf_{w_h \in (\mathring{S}_\ell^1)^2} \|\nabla(z - w_h)\|_{L^2(\Omega)} \tag{3.19}$$

holds for all $z \in Z$. Conversely, if (3.19) holds, we can assume that the function z_h in the proof of Theorem 3.12 varies only in $Z \cap Z_h^{\mathrm{ub}}$. This, in turn, provides a robust upper bound of $\mathrm{C_{qopr}}$ in the limit $\eta \to +\infty$. Whenever condition (3.19) fails, a locking effect may occur in the sense of Babuška & Suri (1992). We illustrate this in Section 5.3 by means of a numerical experiment.

Theorem 3.12 states that the discretization (3.15) enjoys a better velocity H^1-error estimate than the standard one (3.6); cf. Remark 2.8. The next result additionally ensures that the two discretizations are actually comparable if one considers the sum of the velocity H^1-error times viscosity plus the pressure L^2-error. Thus, in other words, the modifications introduced in (3.15) do not impair the quasi-optimality of (3.6).

THEOREM 3.14 (Quasi-optimality). For all $\eta > 1$, problem (3.15) is a quasi-optimal discretization of (2.1) with constant $C_{\mathrm{qo}} \lesssim \eta^3/(\eta - 1)$.

Proof. Denote by (u,p) and (u_h,p_h) the solutions of problems (2.1) and (3.15), respectively, with load $f \in H^{-1}(\Omega)^2$ and viscosity $\mu > 0$. In view of Theorem 3.12, it suffices to bound the pressure error $\|p-p_h\|_{L^2(\Omega)}$. To this end, let $q_h \in \widehat{S}^0_{\ell-2}$ be arbitrary and recall that the discrete divergence $\underline{\mathrm{div}}_h$ is given by (2.7). The inf-sup stability of the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair and Proposition 3.4 yield

$$\|p_h - q_h\|_{L^2(\Omega)} \le c \sup_{v_h \in (\mathring{S}_\ell^1)^2} \frac{\int_\Omega (p_h - q_h)\mathrm{div}\, E_h^{\mathrm{ub}} v_h}{\|\nabla v_h\|_{L^2(\Omega)}}.$$

For all $v_h \in (\mathring{S}_\ell^1)^2$, a comparison of (2.1) and (3.15) entails

$$\int_\Omega (p_h - q_h)\mathrm{div}\, E_h^{\mathrm{ub}} v_h = \mu\left(a_h^{\mathrm{ub}}(u_h, v_h) - \int_\Omega \nabla u : \nabla E_h^{\mathrm{ub}} v_h\right) + \int_\Omega (p - q_h)\underline{\mathrm{div}}_h\, v_h,$$

where we have again made use of Proposition 3.4. The last summand on the right-hand side vanishes if we let q_h be the L^2-orthogonal projection of p. Hence, invoking Lemma 3.11 and proceeding as in the proof of Theorem 3.12 we infer

$$\|p_h - q_h\|_{L^2(\Omega)} \le c\mu\eta\|\nabla(u - u_h)\|_{L^2(\Omega)}. \tag{3.20}$$

The triangle inequality and Theorem 3.12 conclude the proof. □

REMARK 3.15 (Discrete pressure error). Let (u, p) and (u_h, p_h) be as in Theorem 3.14 and denote by $\pi_h p$ the L^2-orthogonal projection of p onto $\widehat{S}^0_{\ell-2}$. Inequality (3.20) and Theorem 3.12 yield the following pressure-robust estimate of the 'discrete pressure error' $(\pi_h p - p_h)$:

$$\|\pi_h p - p_h\|_{L^2(\Omega)} \lesssim \frac{\mu\eta^3}{\eta - 1} \inf_{w_h \in (\mathring{S}^1_\ell)^2} \|\nabla(u - w_h)\|_{L^2(\Omega)}.$$

This readily entails that the choice of p_h is optimal whenever $u \in Z \cap Z_h^{\mathrm{ub}}$. Moreover, the claim of Theorem 3.14 could be slightly improved, in that we actually have

$$\mu\|\nabla(u - u_h)\|_{L^2(\Omega)} + \|p - p_h\|_{L^2(\Omega)} \le \frac{c\mu\eta^3}{\eta - 1} \inf_{w_h \in (\mathring{S}^1_\ell)^2} \|\nabla(u - w_h)\|_{L^2(\Omega)} + \inf_{q_h \in \widehat{S}^0_{\ell-2}} \|p - q_h\|_{L^2(\Omega)}.$$

The same observation applies to the nonconforming discretizations in Section 4.1, but not to the conforming ones in Section 4.2. Estimates in the spirit of (3.15) are standard in pressure-robust discretizations; see John *et al.* (2017, Lemma 4.4 and Theorem 5.2).

3.5 *Inhomogeneous continuity equation*

It is worth having a look at the case when the incompressibility constraint $\operatorname{div} u = 0$ of (1.1) is replaced by the inhomogeneous continuity condition $\operatorname{div} u = g$ with $g \in L^2_0(\Omega)$. The corresponding weak formulation reads as follows. Find $u \in H^1_0(\Omega)^2$ and $p \in L^2_0(\Omega)$ such that

$$\begin{aligned} &\forall\, v \in H^1_0(\Omega)^2, \qquad && \mu \int_\Omega \nabla u : \nabla v - \int_\Omega p \operatorname{div} v = \langle f, v\rangle\,, \\ &\forall\, q \in L^2_0(\Omega), && \int_\Omega q \operatorname{div} u = \int_\Omega qg. \end{aligned} \tag{3.21}$$

A possible extension of the discretization (3.15) with the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair consists in finding $u_h \in (\mathring{S}^1_\ell)^2$ and $p_h \in \widehat{S}^0_{\ell-2}$ such that

$$\begin{aligned} &\forall\, v_h \in \left(\mathring{S}^1_\ell\right)^2, \qquad && \mu\, a_h^{\mathrm{ub}}(u_h, v_h) - \int_\Omega p_h \operatorname{div} v_h = \left\langle f, E_h^{\mathrm{ub}} v_h\right\rangle, \\ &\forall\, q_h \in \widehat{S}^0_{\ell-2}, && \int_\Omega q_h \operatorname{div} u_h = \int_\Omega q_h g. \end{aligned} \tag{3.22}$$

The second equations of (3.21) and (3.22) impose $u \in Z(g)$ and $u_h \in Z_h^{\mathrm{ub}}(g)$, respectively, where

$$Z(g) := \left\{z \in H_0^1(\Omega)^2 \mid \operatorname{div} z = g\right\}, \qquad Z_h^{\mathrm{ub}}(g) := \left\{z_h \in \left(\mathring{S}_\ell^1\right)^2 \mid \underline{\operatorname{div}}_h z_h = \Pi_{\ell-2} g\right\}$$

and $\Pi_{\ell-2}$ is the L^2-orthogonal projection onto $\widehat{S}_{\ell-2}^0$.

Lemma 3.11 states that the consistency error in the left-hand side of (3.17) vanishes whenever $z_h \in Z \cap Z_h^{\mathrm{ub}}$. If instead we assume $z_h \in Z(g) \cap Z_h^{\mathrm{ub}}(g)$ for some $g \in L_0^2(\Omega)$ with $g \neq \Pi_{\ell-2} g$, the consistency error may not vanish. In fact, we possibly have $R_h^{\mathrm{ub}} z_h \neq 0$ as a consequence of Proposition 3.4. This suggests that a bound of the consistency error solely in terms of the best approximation H^1-error to z_h by elements of $Z(g)$ is probably not possible. Therefore, we do not expect that the discrete velocity u_h solving (3.22) is a near-best approximation of the analytical velocity in $(\mathring{S}_\ell^1)^2$, with respect to the H^1-norm.

Still, combining the equivalence (3.12) and the L^2-orthogonality of $\Pi_{\ell-2}$, we obtain the following generalization of Lemma 3.11:

$$\left|\int_\Omega \nabla z_h : \nabla E_h^{\mathrm{ub}} v_h - a_h^{\mathrm{ub}}(z_h, v_h)\right| \leq c\eta \left(\inf_{z \in Z(g)} \|\nabla(z - z_h)\|_{L^2(\Omega)} + \inf_{q_h \in \widehat{S}_{\ell-2}^0} \|g - q_h\|_{L^2(\Omega)} \right) \|\nabla v_h\|_{L^2(\Omega)}$$

for all $z_h \in Z_h^{\mathrm{ub}}(g)$ and $v_h \in (\mathring{S}_\ell^1)^2$, with $g \in L_0^2(\Omega)$. Apart from the additional term on the right-hand side of this estimate, the technique in the proof of Theorem 3.12 can still be applied with the help of Boffi *et al.* (2013, Proposition 5.1.3), and we finally derive

$$\|\nabla(u - u_h)\|_{L^2(\Omega)} \lesssim \inf_{v_h \in (\mathring{S}_\ell^1)^2} \|\nabla(u - v_h)\|_{L^2(\Omega)} + \inf_{q_h \in \widehat{S}_{\ell-2}^0} \|g - q_h\|_{L^2(\Omega)} \tag{3.23}$$

for any fixed $\eta > 1$. Similarly to (1.3), here the approximation power of the discrete pressure space in the L^2-norm may impair the velocity H^1-error because the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair is unbalanced. We confirm this suspicion by means of a numerical experiment in Section 5.4. Still, we remark that this estimate, unlike (1.3), is pressure-robust, i.e. independent of the analytical pressure. A corresponding bound of the pressure error can be derived by arguing as in the proof of Theorem 3.14.

The nonconforming discretization proposed in Section 4.1 has the notable property that the consistency error can always be bounded solely in terms of the best velocity H^1-error; cf. Remark 4.3. Therefore, in that case we achieve quasi-optimality and pressure-robustness even if an inhomogeneous continuity condition is imposed.

4. Generalizations of the paradigmatic discretization

The idea illustrated in the previous section can be generalized in various directions. An immediate observation is that the same construction applies to any other conforming and inf-sup stable pair V_h/Q_h such that

(i) $\widehat{S}_0^0$ is a subset of Q_h and

(ii) the discrete divergence $\underline{\operatorname{div}}_h$ can be computed elementwise.

The first condition is needed in Proposition 3.4 to ensure that the smoothing operator E_h^{ub} fulfils (3.1). The second one guarantees that the divergence correction R^{ub} can be computed elementwise. As a consequence, the proposed discretization is computationally feasible; cf. Remark 3.6. Conditions (i) and (ii) are verified, for instance, by the following generalization of the $\mathbb{P}_\ell/\mathbb{P}_{\ell-2}$ pair:

$$V_h = \big(\mathring{S}^1_\ell\big)^d \qquad \text{and} \qquad Q_h = \widehat{S}^0_{\ell-k}, \qquad b_h(v_h, q_h) = -\int_\Omega q_h \operatorname{div} v_h,$$

where $d \leq k \leq \ell$ and $d \in \{2, 3\}$. Another possibility is to consider the conforming Crouzeix–Raviart pairs described in Boffi *et al.* (2013, Sections 8.6.2 and 8.7.2). Stable pairs with continuous pressure, i.e. $Q_h \subseteq C^0(\Omega)$, do not fulfil (i), while (ii) is violated, for instance, by the modified Hood–Taylor pairs of Boffi *et al.* (2012).

We now aim to address more substantial generalizations. We mainly focus on the necessary modifications and, in particular, we omit all proofs that are similar to those in the previous section.

4.1 *Nonconforming pairs*

Assume that V_h/Q_h is a nonconforming pair, i.e. $V_h \nsubseteq H^1_0(\Omega)^d$. In this case, it does not seem appropriate to define the smoothing operator E_h as in (3.11), because of the condition $E_h(V_h) \subseteq H^1_0(\Omega)^d$. A possible fix for this problem is to replace v_h with $M_h v_h$, where $M_h : V_h \to H^1_0(\Omega)^d$ is a linear operator. To make sure that a counterpart of Proposition 3.4 holds, we require that $\operatorname{div} M_h v_h$ has elementwise the same mean as $\underline{\operatorname{div}}_h\, v_h$ for all $v_h \in V_h$. Therefore, we resort to a elementwise 'mean-mass-preserving' operator; cf. Proposition 4.1.

As before we illustrate this idea by means of a model example, namely the two-dimensional nonconforming Crouzeix–Raviart pair of degree $\ell \geq 2$. We do not consider the lowest-order case $\ell = 1$, as it is rather specific and it is already covered by Verfürth & Zanotti (2019); cf. Remark 4.2. A similar technique can be applied, for instance, with the modified Crouzeix–Raviart pairs of Matthies & Tobiska (2005) or with the three-dimensional generalizations of the Kouhia–Stenberg pair from Hu & Schedensack (2019). The original two-dimensional pair of Kouhia & Stenberg (1995) can be treated as indicated in Remark 4.2.

Let the mesh $\mathcal{M}$ be as in Section 3 and denote by $\mathcal{F}$ the faces of $\mathcal{M}$. A subscript to $\mathcal{F}$ indicates that we consider only those faces that are contained in the set specified by the subscript. We orient each interior face $F \in \mathcal{F}_\Omega$ with a normal unit vector n_F. We denote by $[\![\cdot]\!]_{|F}$ the jump on F in the direction of n_F. For boundary faces $F \in \mathcal{F}_{\partial\Omega}$, we orient n_F so that it points outside Ω and let $[\![\cdot]\!]_{|F}$ coincide with the trace on F; cf. Di Pietro & Ern (2012, Section 1.2.3). We use the subscript $\mathcal{M}$ to indicate the broken version of a differential operator on $\mathcal{M}$. For instance, the broken gradient of an elementwise H^1-function v is given by $(\nabla_{\mathcal{M}} v)_{|K} := \nabla(v_{|K})$ for all $K \in \mathcal{M}$.

The nonconforming Crouzeix–Raviart space of degree $\ell \in \mathbb{N}$ on $\mathcal{M}$, with homogeneous boundary conditions, can be defined as

$$\mathring{\mathrm{CR}}_\ell := \Big\{ v \in S^0_\ell \mid \forall F \in \mathcal{F} \text{ and } r \in \mathbb{P}_{\ell-1}(F), \quad \textstyle\int_F [\![v]\!] r = 0 \Big\}.$$

Notice that the integral $\int_F v$ is well defined for all $v \in \mathring{\mathrm{CR}}_\ell$ and $F \in \mathcal{F}$ and vanishes if $F \in \mathcal{F}_{\partial\Omega}$. Yet the jumps on mesh faces are not vanishing in general.

We assume hereafter $\ell \geq 2$. The two-dimensional nonconforming Crouziex–Raviart pair of degree ℓ is

$$V_h = (\mathring{\mathrm{CR}}_\ell)^2 \qquad \text{and} \qquad Q_h = \widehat{S}^0_{\ell-1}, \qquad b_h(v_h, q_h) = -\int_\Omega q_h \mathrm{div}_\mathcal{M} v_h.$$

Results concerning the inf-sup stability can be found in Fortin & Soulie (1983), Crouzeix & Falk (1989) and Baran & Stoyan (2007). Since the broken divergence $\mathrm{div}_\mathcal{M}$ maps V_h into Q_h, it coincides with the discrete divergence from (2.7), i.e. $\underline{\mathrm{div}}_h = \mathrm{div}_\mathcal{M}$. We measure the velocity error in the broken H^1-norm, augmented with scaled jumps. Thus, in the notation of Section 2 we set

$$\|v\|_h^2 = \|v\|_{\mathrm{cr}}^2 := \|\nabla_\mathcal{M} v\|^2_{L^2(\Omega)} + \sum_{F \in \mathcal{F}} h_F^{-1} \|[\![v]\!]\|^2_{L^2(F)},$$

where h_F is the diameter of F. An equivalent alternative would be to consider only the broken H^1-norm. Both options extend the H^1-norm to $H^1_0(\Omega)^2 + (\mathring{\mathrm{CR}}_\ell)^2$.

Let $\mathcal{V}_{\ell,\Omega}$ be the set of interior Lagrange nodes of degree ℓ in $\mathcal{M}$. For all $\nu \in \mathcal{V}_{\ell,\Omega}$, we denote by Φ^ν_ℓ the Lagrange basis function of $\mathring{S}^1_\ell$ associated with the evaluation at ν, i.e. $\Phi^\nu_\ell(\nu') = \delta_{\nu\nu'}$ for all $\nu' \in \mathcal{V}_{\ell,\Omega}$. Also fix an element $K_\nu \in \mathcal{M}$ with $\nu \in K_\nu$. We define a "simplified nodal averaging" operator $A_h^{\mathrm{cr}} : (\mathring{\mathrm{CR}}_\ell)^2 \to (\mathring{S}^1_\ell)^2$ by

$$A_h^{\mathrm{cr}} v_h := \sum_{\nu \in \mathcal{V}_{\ell,\Omega}} v_{h|K_\nu}(\nu)\, \Phi^\nu_\ell.$$

Next let m_F be the midpoint of any interior face $F \in \mathcal{F}_\Omega$. Consider the bubble function $\Phi^F_2 := 3(2\,|F|)^{-1} \Phi_2^{m_F}$, where $\Phi_2^{m_F}$ is the Lagrange basis function of $\mathring{S}^1_2$ associated with the evaluation at m_F. The normalization implies $\int_{F'} \Phi^F_2 = \delta_{FF'}$ for all $F' \in \mathcal{F}$, according to the Simpson quadrature formula. We introduce a 'bubble' operator $B_h^{\mathrm{cr}} : (\mathring{\mathrm{CR}}_\ell)^2 \to (\mathring{S}^1_\ell)^2$ by

$$B_h^{\mathrm{cr}} v_h := \sum_{F \in \mathcal{F}_\Omega} \left(\int_F v_h \right) \Phi^F_2.$$

We combine A_h^{cr} and B_h^{cr} to obtain the announced elementwise mean-mass-preserving operator M_h^{cr}. Roughly speaking, we use B_h^{cr} to enforce the first part of (4.2) below, while A_h^{cr} is responsible for the second part.

PROPOSITION 4.1 (Elementwise mean-mass-preserving operator). The linear operator $M_h^{\mathrm{cr}} : (\mathring{\mathrm{CR}}_\ell)^2 \to (\mathring{S}^1_\ell)^2$ given by

$$M_h^{\mathrm{cr}} v_h := A_h^{\mathrm{cr}} v_h + B_h^{\mathrm{cr}}(v_h - A_h^{\mathrm{cr}} v_h) \tag{4.1}$$

is such that

$$\int_K \mathrm{div}\, M_h^{\mathrm{cr}} v_h = \int_K \mathrm{div}\, v_h \quad \text{and} \quad \|v_h - M_h^{\mathrm{cr}} v_h\|_{\mathrm{cr}} \leq c \inf_{v \in H^1_0(\Omega)^2} \|v - v_h\|_{\mathrm{cr}} \tag{4.2}$$

for all $v_h \in (\mathring{\mathrm{CR}}_\ell)^2$ and $K \in \mathcal{M}$.

Proof. Let $v_h \in (\mathring{\mathrm{CR}}_\ell)^2$ and $F' \in \mathcal{F}_\Omega$ be given. The normalization of the functions $\{\Phi_2^F\}_{F\in\mathcal{F}_\Omega}$ reveals

$$\int_{F'} B_h^{\mathrm{cr}}(v_h - A_h^{\mathrm{cr}} v_h) = \sum_{F\in\mathcal{F}_\Omega} \int_F (v_h - A_h^{\mathrm{cr}} v_h)\delta_{FF'} = \int_{F'} (v_h - A_h^{\mathrm{cr}} v_h).$$

The same identities hold also for boundary faces $F' \in \mathcal{F}_{\partial\Omega}$ in view of the boundary conditions in $\mathring{\mathrm{CR}}_\ell$ and $\mathring{S}_\ell^1$. Rearranging terms we obtain $\int_{F'} M_h^{\mathrm{cr}} v_h = \int_{F'} v_h$ for all $F' \in \mathcal{F}$. Then, for all $K \in \mathcal{M}$, the Gauss theorem yields the first part of (4.2):

$$\int_K \operatorname{div} M_h^{\mathrm{cr}} v_h = \sum_{F'\in\mathcal{F}_{\partial K}} \int_{F'} M_h^{\mathrm{cr}} v_h \cdot n_K = \sum_{F'\in\mathcal{F}_{\partial K}} \int_{F'} v_h \cdot n_K = \int_K \operatorname{div} v_h.$$

A detailed proof of the second part of (4.2) can be found in Veeser & Zanotti (2018b, Section 3), where a similar, actually more involved, operator is considered. For this reason, we only sketch the argument. Let $K \in \mathcal{M}$ be given. Owing to the triangle inequality, we initially bound $\|\nabla(v_h - A_h^{\mathrm{cr}} v_h)\|_{L^2(K)}$ and $\|\nabla B_h^{\mathrm{cr}}(v_h - A_h^{\mathrm{cr}} v_h)\|_{L^2(K)}$. The scaling of the functions $\{\Phi_2^F\}_{F\in\mathcal{F}_{\partial K}}$ and the trace inequality imply

$$\begin{aligned} &\|\nabla(v_h - A_h^{\mathrm{cr}} v_h)\|_{L^2(K)} + \|\nabla B_h^{cr}(v_h - A_h^{\mathrm{cr}} v_h)\|_{L^2(K)} \\ &\qquad \lesssim h_K^{-1}\|v_h - A_h^{\mathrm{cr}} v_h\|_{L^2(K)} + \|\nabla(v_h - A_h^{\mathrm{cr}} v_h)\|_{L^2(K)}, \end{aligned} \tag{4.3}$$

where h_K is the diameter of K. Next, for all $\nu \in \mathcal{V}_{\ell,K}$ we have $v_{h|K}(\nu) = A_h^{\mathrm{cr}} v_h(\nu)$ if $\nu \in \operatorname{int}(K)$, otherwise $\left|v_{h|K}(\nu) - A_h^{\mathrm{cr}} v_h(\nu)\right| \lesssim \sum_{F\ni\nu} h_F^{-1/2}\|[\![v_h]\!]\|_{L^2(F)}$, where F varies in $\mathcal{F}$. This estimate and the scaling of the Lagrange basis functions entail that the right-hand side of (4.3) is bounded by $\sum_{F\cap K\neq\emptyset} h_F^{-1/2}\|[\![v_h]\!]\|_{L^2(F)}$. Squaring and summing over all $K \in \mathcal{M}$, we finally obtain

$$\|\nabla_{\mathcal{M}}(v_h - M_h^{\mathrm{cr}} v_h)\|_{L^2(\Omega)}^2 \lesssim \sum_{F\in\mathcal{F}} h_F^{-1}\|[\![v_h]\!]\|_{L^2(F)}^2.$$

We conclude by recalling the definition of the norm $\|\cdot\|_{\mathrm{cr}}$. □

According to the first part of (4.2), we can now construct a smoothing operator similarly to E_h^{ub} in Proposition 3.4. Recalling the local operators R_ℓ^K introduced in Section 3.2, we define $E_h^{\mathrm{cr}} : (\mathring{\mathrm{CR}}_\ell)^2 \to H_0^1(\Omega)^2$ by

$$E_h^{\mathrm{cr}} v_h := M_h^{\mathrm{cr}} v_h + \sum_{K\in\mathcal{M}} R_\ell^K(\operatorname{div}_{\mathcal{M}} v_h - \operatorname{div} M_h^{\mathrm{cr}} v_h). \tag{4.4}$$

Owing to the identity $\underline{\operatorname{div}}_h = \operatorname{div}_{\mathcal{M}}$, we see that E_h^{cr} fulfils condition (3.1) as a consequence of Propositions 3.2 and 4.1. Moreover, the stability of the operators R_ℓ^K and the second part of (4.2) provide a strengthened counterpart of (3.12) in that, for all $v_h \in (\mathring{\mathrm{CR}}_\ell)^2$, we have

$$\|v_h - E_h^{\mathrm{cr}} v_h\|_{\mathrm{cr}} \eqsim \inf_{v\in H_0^1(\Omega)^2} \|v - v_h\|_{\mathrm{cr}}. \tag{4.5}$$

Next, inspired by the definition of a_h^{ub} in (3.14) as well as by identity (3.16), we introduce the following bilinear form a_h^{cr} on $(\mathring{\mathrm{CR}}_\ell)^2$:

$$a_h^{\mathrm{cr}}(w_h, v_h) := \int_\Omega \nabla E_h^{\mathrm{cr}} w_h : \nabla E_h^{\mathrm{cr}} v_h + (\eta - 1) \int_\Omega \nabla_{\mathcal{M}} R_h^{\mathrm{cr}} w_h : \nabla_{\mathcal{M}} R_h^{\mathrm{cr}} v_h,$$

where $R_h^{\mathrm{cr}} := (E_h^{\mathrm{cr}} - \mathrm{Id})$ and $\eta > 1$ is a penalty parameter. The above-mentioned properties of E_h^{cr} imply that the necessary conditions in Lemmas 2.6 and 2.7 are fulfilled if we set $a_h = a_h^{\mathrm{cr}}$ and $E_h = E_h^{\mathrm{cr}}$. In this setting, the abstract discretization (2.2) reads as follows. Find $u_h \in (\mathring{\mathrm{CR}}_\ell)^2$ and $p_h \in \widehat{S}^0_{\ell-1}$ such that

$$\begin{aligned} &\forall\, v_h \in (\mathring{\mathrm{CR}}_\ell)^2, \qquad && \mu\, a_h^{\mathrm{cr}}(u_h, v_h) - \int_\Omega p_h \operatorname{div}_{\mathcal{M}} v_h = \left\langle f, E_h^{\mathrm{cr}} v_h \right\rangle, \\ &\forall\, q_h \in \widehat{S}^0_{\ell-1}, && \int_\Omega q_h \operatorname{div}_{\mathcal{M}} u_h = 0. \end{aligned} \tag{4.6}$$

Similarly to a_h^{ub} in Lemma 3.5, the form a_h^{cr} is coercive on $(\mathring{\mathrm{CR}}_\ell)^2$ for $\eta > 1$, with constant $\geq (1 - \eta^{-1})$. Moreover, in view of (4.5), we can estimate the consistency error of (4.6) by the following counterpart of Lemma 3.11:

$$\left| \int_\Omega \nabla_{\mathcal{M}} w_h : \nabla E_h^{\mathrm{cr}} v_h - a_h^{\mathrm{cr}}(w_h, v_h) \right| \leq c\eta \inf_{w \in H_0^1(\Omega)^2} \|w - w_h\|_{\mathrm{cr}} \|v_h\|_{\mathrm{cr}} \tag{4.7}$$

for all $w_h, v_h \in (\mathring{\mathrm{CR}}_\ell)^2$. Hence, we conclude that (4.6) is a quasi-optimal and pressure-robust discretization of (2.1) in the norm $\|\cdot\|_{\mathrm{cr}}$ and the constant $\mathrm{C}_{\mathrm{qopr}}$ from Definition 2.5 solely depends on η and the shape parameter of $\mathcal{M}$.

Arguing as in Theorem 3.14, we can establish also an estimate of the pressure L^2-error, only in terms of the best velocity H^1-error and of the best pressure L^2-error, provided that the pair $(CR)2/S0 - 1$ is inf-sup stable. Thus, problem (4.6) is also a quasi-optimal discretization of (2.1).

Locally supported basis functions of $\mathring{\mathrm{CR}}_\ell$ are described in Baran & Stoyan (2006, Section 3). With this basis and the standard nodal basis of $S^0_{\ell-1}$, we see that (4.6) is computationally feasible in the sense of Remark 2.4; cf. Remark 3.6.

Remark 4.2 (The pair $(\mathring{\mathrm{CR}}_1)^2/\widehat{S}^0_0$). In principle, the approach described for $\ell \geq 2$ applies also with $\ell = 1$, up to observing that R_2^K (and not R_1^K) should be used in (4.4). The point is that in this case an elementwise integration by parts and the identity $\int_{F'} E_h^{\mathrm{cr}} v_h = \int_{F'} M_h^{\mathrm{cr}} v_h = \int_{F'} v_h$, with $F' \in \mathcal{F}$, reveal $\int_\Omega \nabla_{\mathcal{M}} w_h : \nabla R_h^{\mathrm{cr}} v_h = 0$ for all $w_h, v_h \in (\mathring{\mathrm{CR}}_1)^2$. Hence, the form a_h^{cr} is given by $a_h^{\mathrm{cr}}(w_h, v_h) = \int_\Omega \nabla_{\mathcal{M}} w_h : \nabla_{\mathcal{M}} v_h + \eta \int_\Omega \nabla_{\mathcal{M}} R_h^{\mathrm{cr}} w_h : \nabla_{\mathcal{M}} R_h^{\mathrm{cr}} v_h$, showing that the penalization is actually not needed. Setting $\eta = 0$ annihilates the consistency error and corresponds to the discretization proposed in Verfürth & Zanotti (2019).

Remark 4.3 (Inhomogeneous continuity equation). The infimum on the right-hand side of (4.7) is taken over $H_0^1(\Omega)^2$ and not only over Z, unlike Lemma 3.11. This prevents the issue pointed out in Section 3.5. Therefore, the nonconforming Crouzeix–Raviart pair can be used to design a quasi-optimal and pressure-robust discretization of problem (3.21) with the inhomogeneous continuity condition $g \neq 0$.

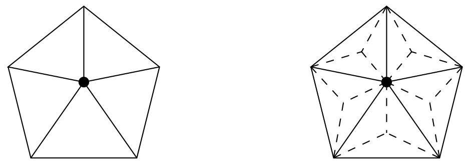

FIG. 2. Generic patch ω_ν (left) and Alfeld refinement $\mathcal{M}_\nu$ (right).

4.2 *Conforming pairs with continuous pressure*

Another class of pairs still not covered by our discussion are conforming pairs with continuous pressure. In fact, the following observations obstruct the construction of a smoothing operator as indicated in Proposition 3.4:

(i) Since $\widehat{S}_0^0$ is not a subspace of Q_h, the identity $\int_K \underline{\operatorname{div}}_h v_h = \int_K \operatorname{div} v_h$ may fail to hold for some $v_h \in V_h$ and $K \in \mathcal{M}$.

(ii) The computation of $\underline{\operatorname{div}}_h$ is probably unfeasible in the sense of Remark 2.4.

Item (i) entails that we cannot correct the divergence elementwise by means of the operators R_ℓ^K from Section 3.2. The shape functions of the lowest-order continuous space S_1^1 suggest working on patches of elements sharing a vertex instead. Item (ii) further indicates that we should never require a direct computation of $\underline{\operatorname{div}}_h$. The construction of a quasi-optimal and pressure-robust discretization of the Stokes equations is still possible under these constraints, but it is more involved than those in the previous sections. We mainly adapt ideas by Lederer *et al.* (2017).

As an example, we let the mesh $\mathcal{M}$ be as in Section 3 and consider the two-dimensional Hood–Taylor pair

$$V_h = (\mathring{S}_\ell^1)^2 \qquad \text{and} \qquad Q_h = \widehat{S}_{\ell-1}^1, \qquad b_h(v_h, q_h) = -\int_\Omega q_h \operatorname{div} v_h$$

with $\ell \geq 2$. The inf-sup condition (2.3) holds with $\beta^{-1} \leq c$ under mild assumptions on $\mathcal{M}$; see Boffi (1994). The discrete divergence coincides with the L^2-orthogonal projection of the analytical divergence onto $\widehat{S}_{\ell-1}^1$. We denote by Z_h^{ht} the discrete kernel.

Let $\mathcal{V} := \mathcal{V}_1$ denote the set of all vertices of $\mathcal{M}$. For each $\nu \in \mathcal{V}$, let Φ_1^ν be the Lagrange basis function of S_1^1 associated with the evaluation at ν, i.e. $\Phi_1^\nu(\nu') = \delta_{\nu\nu'}$ for all $\nu' \in \mathcal{V}$. Recall that Φ_1^ν is supported on the patch $\omega_\nu := \{K \in \mathcal{M} \mid \nu \in K\}$. Consider the Alfeld refinement $\mathcal{M}_\nu$ of ω_ν, i.e. the mesh obtained connecting the vertices and the barycentre of any triangle in ω_ν; cf. Fig. 2. The space $S_\ell^0(\mathcal{M}_\nu)$ and the subspaces

$$\mathring{S}_\ell^1(\mathcal{M}_\nu) \qquad \text{and} \qquad \widehat{S}_{\ell-1}^0(\mathcal{M}_\nu)$$

are defined on $\mathcal{M}_\nu$ analogously to S_ℓ^0 in (3.2) and $\mathring{S}_\ell^1$ and $\widehat{S}_{\ell-1}^0$ in (3.3), respectively. The elementwise local Lagrange interpolant $I_\ell^\nu : S_\ell^0(\mathcal{M}_\nu) \to S_{\ell-1}^0(\mathcal{M}_\nu)$ is given by

$$I_\ell^\nu v := \sum_{K \in \mathcal{M}_\nu} \sum_{\nu' \in \mathcal{V}_{\ell-1,K}} v_{|K}(\nu') \Phi_{\ell-1}^{\nu',K},$$

where $\mathcal{V}_{\ell-1,K}$ is the set of Lagrange nodes of degree $\ell-1$ in K and $\Phi_{\ell-1}^{\nu',K}$ is the Lagrange basis function of $\mathbb{P}_\ell(K)$ associated with the evaluation at ν' and extended to zero outside K. Consider also the simplified local averaging $A_\ell^\nu : S_\ell^0(\mathcal{M}_\nu) \to S_{\ell-1}^1$,

$$A_\ell^\nu v := \sum_{\nu \in \mathcal{V}_{\ell-1}} v_{|K_\nu}(\nu)\Phi_{\ell-1}^\nu,$$

where $K_\nu \in \mathcal{M}$ is a fixed element such that $\nu \in K_\nu$ and v is extended to zero outside ω_ν. As before, $\Phi_{\ell-1}^\nu$ denotes the Lagrange basis function of $\mathring{S}_{\ell-1}^1$ associated with the evaluation at ν.

We are now ready to define the operators $R_\ell^\nu : L^2(\Omega) \to H_0^1(\Omega)^2$ that will be used to correct the divergence in each patch ω_ν, $\nu \in \mathcal{V}$. Here R_ℓ^ν plays the same role as R_ℓ^K in Section 3.2. Given $q \in L^2(\Omega)$, let $u_\nu = u_\nu(q) \in \mathring{S}_\ell^1(\mathcal{M}_\nu)^2$ and $p_\nu = p_\nu(q) \in \widehat{S}_{\ell-1}^0(\mathcal{M}_\nu)$ be such that

$$\begin{aligned} &\forall\, v_\nu \in \mathring{S}_\ell^1(\mathcal{M}_\nu)^2, && \int_{\omega_\nu} \nabla u_\nu : \nabla v_\nu - \int_{\omega_\nu} p_\nu \mathrm{div}\, v_\nu = 0, \\ &\forall\, q_\nu \in \widehat{S}_{\ell-1}^0(\mathcal{M}_\nu), && \int_{\omega_\nu} q_\nu \mathrm{div}\, u_\nu = \int_{\omega_\nu} \left(A_\ell^\nu(q_\nu\Phi_1^\nu) - I_\ell^\nu(q_\nu\Phi_1^\nu)\right) q. \end{aligned} \tag{4.8}$$

This problem is uniquely solvable, according to Guzmán & Neilan (2018, Corollary 6.2). Then we set

$$R_\ell^\nu q := u_\nu \quad \text{in } \omega_\nu \qquad \text{and} \qquad R_\ell^\nu q := 0 \quad \text{in } \Omega \setminus \omega_\nu.$$

REMARK 4.4 (Local problems). The use of the Alfeld refinement $\mathcal{M}_\nu$ is the main difference compared to Lederer *et al.* (2017). This ensures that the pair $\mathring{S}_\ell^1(\mathcal{M}_\nu)^2/\widehat{S}_{\ell-1}^0(\mathcal{M}_\nu)$ is inf-sup stable. In fact, it is known that the stability of the Scott–Vogelius pair on ω_ν (without the Alfeld refinement) may be impaired if ν is a singular or nearly singular vertex; see Scott & Vogelius (1985). The partition of unity $\{\Phi_1^\nu\}_{\nu\in\mathcal{V}}$ and the interpolants $\{I_\ell^\nu\}_{\nu\in\mathcal{V}}$ account for the overlapping of the patches, while the averaging operators $\{A_\ell^\nu\}_{\nu\in\mathcal{V}}$ are used to avoid a direct computation of the discrete divergence in (4.9).

We define a global divergence correction $R_h^{\mathrm{ht}} : (\mathring{S}_\ell^1)^2 \to H_0^1(\Omega)^2$,

$$R_h^{\mathrm{ht}} v_h := \sum_{\nu\in\mathcal{V}} R_\ell^\nu \mathrm{div}\, v_h. \tag{4.9}$$

In contrast to E_h^{ub} and E_h^{cr} from (3.11) and (4.4), respectively, we now make use of a smoothing operator E_h^{ht} that is not guaranteed to be divergence preserving, i.e. (3.1) may fail to hold. We shall see, however, that it still satisfies the necessary conditions in Lemmas 2.6 and 2.7. In the following proposition we prove only a basic stability estimate, for the sake of simplicity.

PROPOSITION 4.5 (Smoothing operator for the Hood–Taylor pair). The linear operator $E_h^{\mathrm{ht}} : (\mathring{S}_\ell^1)^2 \to H_0^1(\Omega)^2$ given by

$$E_h^{\mathrm{ht}} v_h := v_h + R_h^{\mathrm{ht}} v_h$$

satisfies (2.12a) and (2.13a) and is such that, for all $v_h \in (\mathring{S}^1_\ell)^2$,

$$\|\nabla(v_h - E_h^{\mathrm{ht}} v_h)\|_{L^2(\Omega)} \le c\|\mathrm{div}\, v_h\|_{L^2(\Omega)}. \tag{4.10}$$

Proof. For all $v_h \in (\mathring{S}^1_\ell)^2$ and $q_h \in \widehat{S}^1_{\ell-1}$, we have

$$\int_\Omega q_h \mathrm{div}\, R_h^{\mathrm{ht}} v_h = \sum_{\nu\in\mathcal{V}} \int_{\omega_\nu} (A_\ell^\nu(q_h\Phi_1^\nu) - I_\ell^\nu(q_h\Phi_1^\nu))\mathrm{div}\, v_h = 0.$$

The first identity follows from the second equation of (4.8), which actually holds for all q_ν in $S^0_{\ell-1}(\mathcal{M}_\nu)$ (and not only in $\widehat{S}^0_{\ell-1}(\mathcal{M}_\nu)$), as both sides vanish if q_ν is constant. To check the second identity, observe that $A_\ell^\nu(q_h\Phi_1^\nu) = I_\ell^\nu(q_h\Phi_1^\nu)$ for all $\nu \in \mathcal{V}$, due to the continuity of $q_h\Phi_1^\nu$. Thus we derive the identity

$$\int_\Omega q_h \,\mathrm{div}\, E_h^{\mathrm{ht}} v_h = \int_\Omega q_h \,\mathrm{div}\, v_h$$

showing that condition (2.12a) holds. Next let $z_h \in Z^{\mathrm{ht}}$ be given and consider $q_h = \mathrm{div}\, E_h^{\mathrm{ht}} z_h$. Recall that $\{\Phi_1^\nu\}_{\nu\in\mathcal{V}}$ is a partition of unity and extend $I_\ell^\nu(q_h\Phi_1^\nu)$ to zero outside ω_ν. We infer $\sum_{\nu\in\mathcal{V}} I_\ell^\nu(q_h\Phi_1^\nu) = q_h$. Then, since z_h is discretely divergence-free, we have

$$\|q_h\|^2_{L^2(\Omega)} = \int_\Omega q_h \,\mathrm{div}\, z_h - \sum_{\nu\in\mathcal{V}} \int_\Omega I_\ell^\nu(q_h\Phi_1^\nu)\mathrm{div}\, z_h = 0.$$

This reveals $\mathrm{div}\, E_h^{\mathrm{ht}} z_h = 0$ and confirms that condition (2.13a) holds. Finally, owing to the stability of I_ℓ^ν and A_ℓ^ν in the $L^2(\omega_\nu)$-norm, we infer

$$\sup_{q_\nu \in \widehat{S}^0_{\ell-1}(\mathcal{M}_\nu)} \frac{\int_{\omega_\nu} \left(A_\ell^\nu(q_\nu\Phi_1^\nu) - I_\ell^\nu(q_\nu\Phi_1^\nu)\right) \mathrm{div}\, v_h}{\|q_\nu\|_{L^2(\omega_\nu)}} \le c\|\mathrm{div}\, v_h\|_{L^2(\omega_\nu)}$$

for all $\nu \in \mathcal{V}$ and $v_h \in (\mathring{S}^1_\ell)^2$. This entails $\|\nabla\, R_\ell^\nu \mathrm{div}\, v_h\|_{L^2(\omega_\nu)} \lesssim \|\mathrm{div}\, v_h\|_{L^2(\omega_\nu)}$, owing to Boffi *et al.* (2013, Corollary 4.2.1) and the inf-sup stability of the pair $\mathring{S}^1_\ell(\mathcal{M}_\nu)^2/\widehat{S}^0_{\ell-1}(\mathcal{M}_\nu)$ stated in Guzmán & Neilan (2018, Corollary 6.2). The definition of R_h^{ht} in (4.9) then implies

$$\|\nabla R_h^{\mathrm{ht}} v_h\|_{L^2(K)} \lesssim \sum_{K'\cap K\neq\emptyset} \|\mathrm{div}\, v_h\|_{L^2(K')}$$

for all $K \in \mathcal{M}$, where K' varies in $\mathcal{M}$. We conclude by summing over all elements of $\mathcal{M}$ and recalling the definition of E_h^{ht}. □

Next, for $\eta > 1$ we introduce the following bilinear form on $(\mathring{S}^1_\ell)^2$:

$$a_h^{\mathrm{ht}}(w_h, v_h) := \int_\Omega \nabla E_h^{\mathrm{ht}} w_h : \nabla E_h^{\mathrm{ht}} v_h + (\eta - 1) \int_\Omega \nabla R_h^{\mathrm{ht}} w_h : \nabla R_h^{\mathrm{ht}} v_h.$$

The abstract discretization (2.2) with $a_h = a_h^{\mathrm{ht}}$ and $E_h = E_h^{\mathrm{ht}}$ looks for $u_h \in (\mathring{S}^1_\ell)^2$ and $p_h \in \widehat{S}^1_{\ell-1}$ such that

$$\begin{aligned} &\forall\, v_h \in (\mathring{S}^1_\ell)^2, \qquad \mu\, a_h^{\mathrm{ht}}(u_h, v_h) - \int_\Omega p_h \operatorname{div} v_h = \left\langle f, E_h^{\mathrm{ht}} v_h \right\rangle, \\ &\forall\, q_h \in \widehat{S}^1_{\ell-1}, \qquad \int_\Omega q_h \operatorname{div} u_h = 0. \end{aligned} \tag{4.11}$$

This discretization is computationally feasible in the sense of Remark 2.4; cf. Remark 3.6. Yet the implementation is more costly than that of (3.15) and (4.6) because, in general, we cannot resort to one reference configuration for the solution of the local problems (4.8). The error analysis of (4.11) proceeds almost verbatim as in Section 3.4, with the help of Proposition 4.5. The only notable difference is that estimate (3.20) in the proof of Theorem 3.14 should be replaced by the weaker one $\|p_h - q_h\|_{L^2(\Omega)} \lesssim \mu\eta\|\nabla(u - u_h)\|_{L^2(\Omega)} + \|p - q_h\|_{L^2(\Omega)}$ because identity (3.1) may fail to hold.

5. Numerical experiments with the unbalanced $\mathbb{P}_2/\mathbb{P}_0$ pair

In this section we restrict our attention to the two-dimensional Stokes equations, with unit viscosity, posed in the unit square. In the notation of Section 2 this corresponds to

$$d = 2, \qquad \mu = 1, \qquad \Omega = (0,1)^2.$$

We investigate numerically the new discretization (3.15), based on the unbalanced $\mathbb{P}_2/\mathbb{P}_0$ pair, i.e.

$$V_h = (\mathring{S}^1_2)^2 \qquad \text{and} \qquad Q_h = \widehat{S}^0_0, \qquad b_h(v_h, q_h) = -\int_\Omega q_h \operatorname{div} v_h.$$

If not specified differently, the penalty parameter is set to

$$\eta = 2.$$

In particular, we shall emphasize the aspects that distinguish this discretization from the nonconforming one in Verfürth & Zanotti (2019). In that reference, one can find further experiments, illustrating the importance of quasi-optimality and pressure-robustness.

We shall consider the families $(\mathcal{M}_N^D)_{N\in\mathbb{N}_0}$ and $(\mathcal{M}_N^C)_{N\in\mathbb{N}_0}$ of triangular meshes of Ω. For $N \in \mathbb{N}_0$, we divide Ω into $2^N \times 2^N$ identical squares, with edges parallel to the x_1- and x_2-axes and with area 2^{-2N}. We obtain the 'diagonal mesh' $\mathcal{M}_N^D$ by dividing each square by the diagonal with positive slope. Similarly, we obtain the 'crisscross mesh' $\mathcal{M}_N^C$ by drawing both diagonals of each square; cf. Fig. 3. All experiments have been implemented in ALBERTA 3.0 (Heine *et al.*, 2019; Schmidt & Siebert, 2005).

5.1 *Smooth solution*

To illustrate the quasi-optimality and pressure-robustness of the new $\mathbb{P}_2/\mathbb{P}_0$ discretization, we first consider a test case with smooth analytical solution given by

$$u(x_1, x_2) = \operatorname{curl}(x_1^2(1-x_1)^2 x_2^2(1-x_2)^2), \qquad p(x_1, x_2) = \sin(2\pi x_1)\sin(2\pi x_2),$$

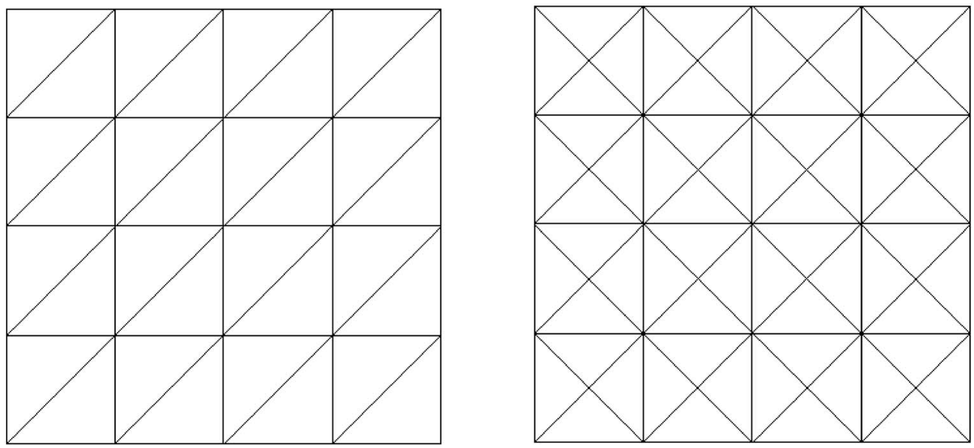

FIG. 3. Diagonal mesh $\mathcal{M}_N^D$ (left) and crisscross mesh $\mathcal{M}_N^C$ (right) with $N = 2$.

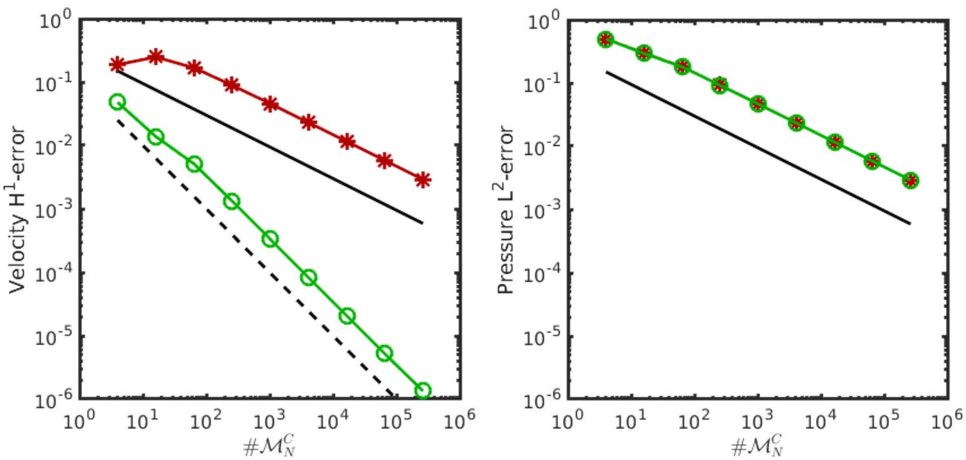

FIG. 4. Test case from Section 5.1. Velocity H^1-error (left) and pressure L^2-error (right) of standard ($*$) and new ($\circ$) $\mathbb{P}_2/\mathbb{P}_0$ discretizations. Plain and dashed lines indicate decay rates $(\#\mathcal{M}_N^C)^{-0.5}$ and $(\#\mathcal{M}_N^C)^{-1}$, respectively.

where $\mathrm{curl}(w) := (\partial_2 w, -\partial_1 w)$. We compare the performances of the standard $\mathbb{P}_2/\mathbb{P}_0$ discretization (3.6) and the new one (3.15) on the crisscross meshes $\mathcal{M}_N^C$ with $N = 0, \ldots, 8$. Figure 4 displays the respective balances of velocity H^1-error and pressure L^2-error versus $\#\mathcal{M}_N^C$, which is the number of triangles in the mesh.

We first observe that the pressure L^2-errors of both discretizations behave quite similarly and converge to zero with the maximum decay rate $(\#\mathcal{M}_N^C)^{-0.5}$. The velocity H^1-error of the standard discretization converges to zero with the same decay rate, as suggested by estimate (1.3), according to the approximation power of the discrete pressure space in the L^2-norm. Note, however, that such a rate is suboptimal with respect to the approximation power of the discrete velocity space in the H^1-norm. In contrast, the velocity H^1-error of the new discretization exhibits the maximum decay rate $(\#\mathcal{M}_N^C)^{-1}$,

TABLE 1 *Test case from Section 5.2. Velocity H^1-errors of the new $\mathbb{P}_2/\mathbb{P}_0$ discretization and corresponding EOCs with composite (left) or standard (right) quadrature rules for $\alpha \in \{1, 10^3\}$*

	$\alpha = 1$		$\alpha = 10^3$			$\alpha = 1$		$\alpha = 10^3$	
N	H^1-error	EOC	H^1-error	EOC	N	H^1-error	EOC	H^1-error	EOC
4	3.32e−04		3.32e−04		4	3.57e−04		1.29e−01	
5	8.31e−05	1.00	8.31e−05	1.00	5	1.07e−04	0.87	6.72e−02	0.47
6	2.08e−05	1.00	2.08e−05	1.00	6	4.01e−05	0.71	3.41e−02	0.49
7	5.19e−06	1.00	5.19e−06	1.00	7	1.80e−05	0.58	1.71e−02	0.50
8	1.30e−06	1.00	1.30e−06	1.00	8	8.72e−06	0.52	8.57e−03	0.50

as predicted by Theorem 3.12. The next experiments are intended to highlight some of the ingredients that contribute to making this optimal-order convergence possible.

5.2 *Composite numerical quadrature*

The evaluation of the duality $\langle f, E_h v_h \rangle$, $v_h \in (\mathring{S}^1_2)^2$ in the new $\mathbb{P}_2/\mathbb{P}_0$ discretization requires, in particular, the evaluation of $\langle f, \widetilde{v}_h \rangle$ for test functions $\widetilde{v}_h$ that are elementwise quadratic on the Alfeld refinement of the mesh at hand. This suggests that for each triangle K in the mesh, a composite quadrature rule based on the Alfeld refinement of K should be used. If one instead uses a standard quadrature rule in K, the resulting quadrature error could be not negligible, due to the low regularity of $\widetilde{v}_h$. Moreover, since the quadrature error is potentially not pressure-robust, as pointed out in Linke *et al.* (2018, Section 6.2), this may even affect the decay rate of the velocity H^1-error.

To illustrate such an effect, we consider a test case with analytical solution

$$u(x_1, x_2) = \operatorname{curl}(x_1^2(1-x_1)^2 x_2^2(1-x_2)^2), \qquad p(x_1, x_2) = \alpha \sin(2\pi x_1)\sin(2\pi x_2).$$

For $\alpha \in \{1, 10^3\}$ we apply the new $\mathbb{P}_2/\mathbb{P}_0$ discretization on the crisscross meshes $\mathcal{M}^C_N$ with $N = 0, \dots, 8$. We assemble the right-hand side both with a composite and with a standard quadrature rule of degree 6. For $N = 4, \dots, 8$, the corresponding velocity H^1-errors are reported in Table 1. In each case, we compute also the so-called experimental order of convergence (EOC), defined as

$$\mathrm{EOC}_N := \frac{\log(e_N/e_{N-1})}{\log(\#\mathcal{M}^C_{N-1}/\#\mathcal{M}^C_N)} = \frac{\log(e_{N-1}/e_N)}{\log 4},$$

where e_N denotes the H^1-error on $\mathcal{M}^C_N$.

When the composite quadrature rule is applied, the results seem insensitive to the parameter α and we observe the maximum decay rate $(\#\mathcal{M}^C_N)^{-1}$. In contrast, the use of the standard quadrature rule impairs the pressure-robustness stated in Theorem 3.12. In fact, for sufficiently large N, the velocity H^1-error is essentially proportional to α and exhibits the suboptimal decay rate $(\#\mathcal{M}^C_N)^{-0.5}$.

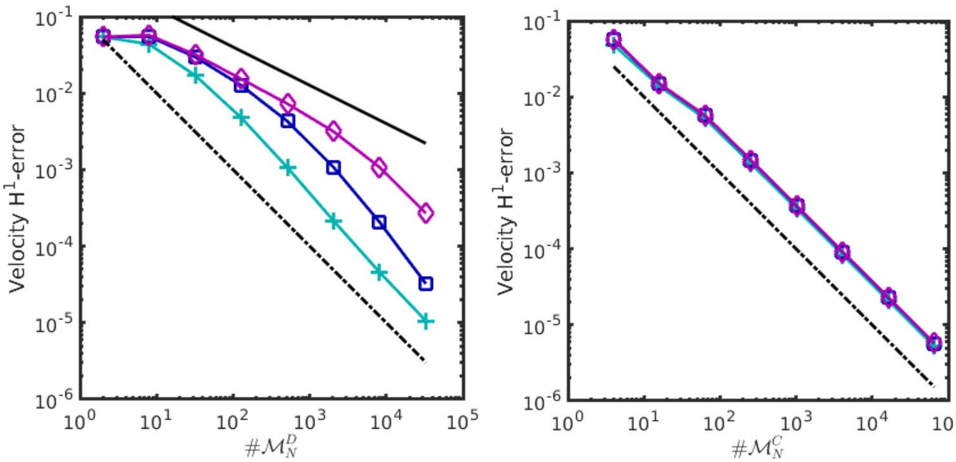

FIG. 5. Test case from Section 5.3. Velocity H^1-error of the new $\mathbb{P}_2/\mathbb{P}_0$ discretization on diagonal (left) and crisscross (right) meshes, for $\eta = 2$ (+), $\eta = 32$ (□) and $\eta = 512$ (◊). Plain and dashed lines indicate decay rates $(\#\mathcal{M}_N^*)^{-0.5}$ and $(\#\mathcal{M}_N^*)^{-1}$, with $* \in \{D, C\}$.

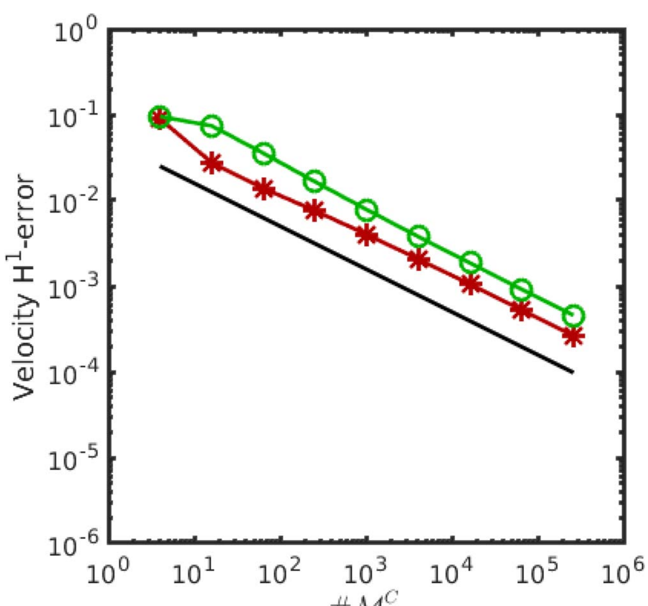

FIG. 6. Test case from Section 5.4. Velocity H^1-error of standard ($*$) and new ($\circ$) $\mathbb{P}_2/\mathbb{P}_0$ discretizations. The plain line indicates decay rate $(\#\mathcal{M}_N^C)^{-0.5}$.

5.3 *Locking*

As mentioned in Remark 3.9, the bilinear form a_h^{ub} in the new $\mathbb{P}_2/\mathbb{P}_0$ discretization has the same structure as the DG-SIP form of Arnold (1982). Still, one main difference is that Lemma 3.5 ensures the coercivity of the former for any penalty $\eta > 1$ (and not only for sufficiently large η). Moreover, the coercivity constant is ≥ 0.5 for $\eta = 2$. Having an explicit and safe choice of the penalty parameter is particularly useful in this context because we may have locking for large η, in view of Remark 3.13.

To illustrate this, we consider a test case with analytical solution

$$u(x_1, x_2) = \operatorname{curl}\big(x_1^2(1-x_1)^2 x_2^2(1-x_2)^2\big), \qquad p(x_1, x_2) = (x_1 - 0.5)(x_2 - 0.5).$$

We apply the new $\mathbb{P}_2/\mathbb{P}_0$ discretization for $\eta \in \{2, 32, 512\}$ both on diagonal meshes $\mathcal{M}_N^D$ and on crisscross meshes $\mathcal{M}_N^D$, with $N = 0, \dots, 7$. The velocity H^1-errors displayed in the right-hand panel of Fig. 5 indicate that the new discretization is robust with respect to η on crisscross meshes. This follows from the fact that condition (3.19) in Remark 3.13 holds for such meshes, as a consequence of Qin (1994, Theorem 4.3.1). In contrast, adopting the terminology of Babuška & Suri (1992), we observe in the left-hand panel of Fig. 5 locking of order $(\mathcal{M}_N^D)^{1/2}$ when diagonal meshes are used.

5.4 *Inhomogeneous continuity equation*

We finally point out that the quasi-optimality and pressure-robustness of the new $\mathbb{P}_2/\mathbb{P}_0$ discretization, as stated in Theorem 3.12, hinges on the homogeneity of the continuity equation in the Stokes problem (2.1); cf. Section 3.5.

To see this, we consider the more general problem (3.21) and approximate the analytical solution

$$u(x_1, x_2) = \begin{pmatrix} x_1(1-x_1)x_2(1-x_2) \\ x_1(1-x_1)x_2(1-x_2) \end{pmatrix}, \qquad p(x_1, x_2) = (x_1 - 0.5)(x_2 - 0.5)$$

on the crisscross meshes $\mathcal{M}_N^C$ with $N = 0, \dots, 8$. Note, in particular, that $\operatorname{div} u$ is not elementwise constant on $\mathcal{M}_N^C$.

Comparing the velocity H^1-errors of the standard $\mathbb{P}_2/\mathbb{P}_0$ discretization (3.6) and the new one (3.15), we see that the former is slightly smaller than the latter and that both errors converge to zero with decay rate $(\mathcal{M}_N^C)^{-0.5}$; cf. Fig. 6. This confirms that inequality (3.23) captures the correct behaviour of the new discretization. Thus, for this problem, we expect that the new discretization performs significantly better than the standard one only in the case of large pressure L^2-errors.

Acknowledgements

We wish to thank Rüdiger Verfürth for reading some preliminary versions of this manuscript and for suggesting several improvements in the presentation.

Funding

DFG (Deutsche Forschungsgemeinschaft) research (grant KR 3984/5-1) ‘Convergence Analysis for Adaptive Discontinuous Galerkin Methods’.

References

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

Arnold, D. N. (1993) On nonconforming linear-constant elements for some variants of the Stokes equations. *Istit. Lombardo Accad. Sci. Lett. Rend. A*, **127**, 83–93.

Babuška, I. & Suri, M. (1992) Locking effects in the finite element approximation of elasticity problems. *Numer. Math.*, **62**, 439–463.

Badia, S., Codina, R., Gudi, T. & Guzmán, J. (2014) Error analysis of discontinuous Galerkin methods for the Stokes problem under minimal regularity. *IMA J. Numer. Anal.*, **34**, 800–819.

Baran, Á. & Stoyan, G. (2006) Crouzeix-Velte decompositions for higher-order finite elements. *Comput. Math. Appl.*, **51**, 967–986.

Baran, Á. & Stoyan, G. (2007) Gauss-Legendre elements: a stable, higher order non-conforming finite element family. *Computing*, **79**, 1–21.

Boffi, D. (1994) Stability of higher order triangular Hood–Taylor methods for the stationary Stokes equations. *Math. Models Methods Appl. Sci.*, **4**, 223–235.

Boffi, D., Brezzi, F. & Fortin, M. (2013) *Mixed Finite Element Methods and Applications.* Springer Series in Computational Mathematics, vol. 44. Heidelberg: Springer.

Boffi, D., Cavallini, N., Gardini, F. & Gastaldi, L. (2012) Local mass conservation of Stokes finite elements. *J. Sci. Comput.*, **52**, 383–400.

Boffi, D. & Lovadina, C. (1997) Analysis of new augmented Lagrangian formulations for mixed finite element schemes. *Numer. Math.*, **75**, 405–419.

Brenner, S. C. (2004) Korn's inequalities for piecewise H^1 vector fields. *Math. Comp.*, **73**, 1067–1087.

Brezzi, F. (1974) On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, **8**, 129–151.

Buffa, A. & Christiansen, S. H. (2007) A dual finite element complex on the barycentric refinement. *Math. Comp.*, **76**, 1743–1769.

Crouzeix, M. & Falk, R. S. (1989) Nonconforming finite elements for the Stokes problem. *Math. Comp.*, **52**, 437–456.

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

Di Pietro, D. A. & Ern, A. (2012) *Mathematical Aspects of Discontinuous Galerkin Methods*. Mathématiques & Applications (Berlin) [Mathematics & Applications], vol. 69. Heidelberg: Springer, p. xviii+384.

Fortin, M. & Soulie, M. (1983) A nonconforming piecewise quadratic finite element on triangles. *Int. J. Numer. Methods Engrg.*, **19**, 505–520.

Fu, G., Guzmán, J. & Neilan, M. (2018) Exact smooth piecewise polynomial sequences on Alfeld splits. arXiv:1807.05883v1.

Gauger, N. R., Linke, A. & Schroeder, P. W. (2018) SMAI Journal of Computational Mathematics. **5**, 89–129.

Georgoulis, E. H. & Pryer, T. (2018) Recovered finite element methods. *Comput. Methods Appl. Mech. Engrg.*, **332**, 303–324.

Guzmán, J. & Neilan, M. (2014) Conforming and divergence-free Stokes elements on general triangular meshes. *Math. Comp.*, **83**, 15–36.

Guzmán, J. & Neilan, M. (2018) Inf-sup stable finite elements on barycentric refinements producing divergence–free approximations in arbitrary dimensions. *SIAM J. Numer. Anal.*, **56**, 2826–2844.

Heine, K.-J., Köster, D., Kriessl, O., Schmidt, A. & Siebert, K. (2019) ALBERTA—an adaptive hierachical finite element toolbox. Available at http:// www. alberta- fem. de.

Hu, J. & Schedensack, M. (2019) Two low-order nonconforming finite element methods for the Stokes flow in three dimensions. *IMA J. Numer. Anal.*, **39**, 1447–1470.

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

KOUHIA, R. & STENBERG, R. (1995) A linear nonconforming finite element method for nearly incompressible elasticity and Stokes flow. *Comput. Methods Appl. Mech. Engrg.*, **124**, 195–212.

LEDERER, P. L., LINKE, A., MERDON, C. & SCHÖBERL, J. (2017) Divergence-free reconstruction operators for pressure-robust Stokes discretizations with continuous pressure finite elements. *SIAM J. Numer. Anal.*, **55**, 1291–1314.

LINKE, A. (2014) On the role of the Helmholtz decomposition in mixed methods for incompressible flows and a new variational crime. *Comput. Methods Appl. Mech. Engrg.*, **268**, 782–800.

LINKE, A., MATTHIES, G. & TOBISKA, L. (2016) Robust arbitrary order mixed finite element methods for the incompressible Stokes equations with pressure independent velocity errors. *ESAIM Math. Model. Numer. Anal.*, **50**, 289–309.

LINKE, A., MERDON, C., NEILAN, M. & NEUMANN, F. (2018) Quasi-optimality of a pressure-robust nonconforming finite element method for the Stokes-problem. *Math. Comp.*, **87**, 1543–1566.

MATTHIES, G. & TOBISKA, L. (2005) Inf-sup stable non-conforming finite elements of arbitrary order on triangles. *Numer. Math.*, **102**, 293–309.

NOCHETTO, R. H. & PYO, J.-H. (2004) Optimal relaxation parameter for the Uzawa method. *Numer. Math.*, **98**, 695–702.

OLSHANSKII, M. A. & REUSKEN, A. (2004) Grad-div stabilization for Stokes equations. *Math. Comp.*, **73**, 1699–1718.

QIN, J. (1994) On the convergence of some simple finite elements for incompressible flows. *Ph.D. Thesis*, Penn State University.

SCHMIDT, A. & SIEBERT, K. G. (2005) *Design of Adaptive Finite Element Software.* Lecture Notes in Computational Science and Engineering, vol. 42. Berlin: Springer, p. xii+315. The Finite Element Toolbox ALBERTA.

SCOTT, L. R. & VOGELIUS, M. (1985) Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials. *RAIRO Modél. Math. Anal. Numér.*, **19**, 111–143.

VEESER, A. & ZANOTTI, P. (2018a) Quasi-optimal nonconforming methods for symmetric elliptic problems. I—abstract theory. *SIAM J. Numer. Anal.*, **56**, 1621–1642.

VEESER, A. & ZANOTTI, P. (2018b) Quasi-optimal nonconforming methods for symmetric elliptic problems. III—discontinuous Galerkin and other interior penalty methods. *SIAM J. Numer. Anal.*, **56**, 2871–2894.

VERFÜRTH, R. & ZANOTTI, P. (2019) A quasi-optimal Crouzeix–Raviart discretization of the Stokes equations. *SIAM J. Numer. Anal.*, **57**, 1082–1099.

ZHANG, S. (2005) A new family of stable mixed finite elements for the 3D Stokes equations. *Math. Comp.*, **74**, 543–554.

ZHANG, S. (2007) On the family of divergence-free finite elements on tetrahedral grids for the Stokes equations. Preprint, University of Delaware. Preprint at http:// www. math. udel. edu/~ szhang/ research/ p/ review. pdf.