**No. 675** **July 2024**

# INF-SUP STABLE DISCRETIZATION OF THE QUASI-STATIC BIOT'S EQUATIONS IN POROELASTICITY

**C. Kreuzer, P. Zanotti**

**ISSN: 2190-1767**

# INF-SUP STABLE DISCRETIZATION OF THE QUASI-STATIC BIOT'S EQUATIONS IN POROELASTICITY

CHRISTIAN KREUZER AND PIETRO ZANOTTI

Abstract. We propose a new full discretization of the Biot's equations in poroelasticity. The construction is driven by the inf-sup theory, which we recently developed. It builds upon the four-field formulation of the equations obtained by introducing the total pressure and the total fluid content. We discretize in space with Lagrange finite elements and in time with backward Euler. We establish inf-sup stability and quasi-optimality of the proposed discretization, with robust constants with respect to all material parameters. We further construct an interpolant showing how the error decays for smooth solutions.

## 1. Introduction

This paper is the second one in a series initiated by [24], regarding the analysis and the discretization of the quasi-static Biot's equations in poroelasticity. (See (2.1) below for the statement of the problem). The series centers around the use of the inf-sup theory for the stability and the error analysis, with the aim of highlighting the possible advantages stemming from the proposed approach, which appears to be new in this context.

The inf-sup theory is a framework for the analysis of general linear variational problems. The main result therein is the so-called Banach-Nečas theorem (see, e.g., [14, Section 25.3]), that characterizes the well-posedness of such problems. Successful applications of the Banach-Nečas theorem additionally provide a two-sided stability estimate, entailing that the space of all possible solutions is isomorphic to that of all possible data, cf. Theorem 2.4 below. Such an estimate is of special interest, as it is the starting point for the derivation of sharp a posteriori error estimates, since it establishes the equivalence of error and residual, cf. [43, Section 4.1.4]. Moreover, if the same approach applies also to the discretization at hand, then the resulting stability estimate is the starting point for the derivation of sharp a priori error estimates, see [2] for conforming discretizations and [4] for nonconforming ones. The inf-sup theory is useful also in the design of robust preconditioners [20, 30] and in the convergence analysis of some adaptive procedures [16].

The inf-sup theory is well-established for the analysis and the discretization of stationary linear equations. We refer to [14] for an overview with several examples. The situation is substantially different for evolution equations, that are usually analyzed by other techniques. While the application of the inf-sup theory to the heat equation has been recently considered by various authors (see, e.g., [15, Chapter 71]

2010 *Mathematics Subject Classification.* Primary 65M15, 65M60; Secondary 74F10, 76S05.

*Key words and phrases.* Inf-sup stability, quasi-static Biot's equations, poroelasticity, quasi-optimality, robustness, a priori analysis.

and the references therein), we are aware of only a few results for other prototypical problems, like the wave and the Schrödinger equation [39].

The quasi-static Biot's equations in poroelasticity are not an exception in this respect. In fact, on the one hand, some authors have used the inf-sup theory in the analysis and the discretization of the stationary problem resulting from the semi-discretization in time, see e.g. [21, 23, 27]. On the other hand, in the quasi-static case, we are not aware of any result obtained by this approach. For the a posteriori analysis, Li and Zikatanov [25] used, in a sense, an equivalent argument. For the a priori analysis, all papers we are aware of build upon a different argument, namely an energy estimate that seems to go back to a seminal contribution of Ženíšek [46]. A far by exhaustive list includes [3, 8, 9, 22, 26, 33, 34, 44].

To our best knowledge, our series is the first contribution making a systematic use of the inf-sup theory in the design and in the error analysis of a discretization of the quasi-static Biot's equations. Our first paper [24] is devoted to the analysis of the equations. It establishes the well-posedness as well as a two-sided stability estimate, which is robust with respect to all material parameters. The latter contributions and the functional setting distinguish our results from previous ones by Ženíšek [46] and Showalter [38]. In particular, we consider an equivalent four-field formulation of the equations, that is obtained from the original one by introducing the so-called total pressure and total fluid content. In [24] we prove also that, in certain circumstances, additional regularity in space of the data imply corresponding additional regularity in space of the solution. Establishing a similar result in time is more challenging, due to possible singularities at the initial time; compare e.g. with [31, 38] and the discussion in Remark (4.10) below. Further previous contributions to the regularity theory are in [6, 45].

In this paper we propose a backward Euler discretization in time and an abstract discretization in space of the Biot's equations. We establish the well-posedness and a two-sided stability estimate via the inf-sup theory, by mimicking the argument in [24]. Then we consider a simple realization of the abstract discretization in space, making use of Lagrange finite elements for all variables. We prove a quasi-optimal a priori error estimate, meaning that the error is equivalent to (i.e. bounded from above and below by) the best error. To our best knowledge, it is the first time that such a result is established for a discretization of the Biot's equations.

The error notion we consider is motivated by the inf-sup theory and it is relatively involved, as all variables are coupled in a nontrivial way. Therefore, we further elaborate on our error estimate, by showing that the best error can be bounded by a sum of decoupled best errors in standard norms, that are much easier to investigate. All constants in our results are robust with respect to all material parameters and do not require any additional regularity beyond the one guaranteed by the well-posedness of the equations. Finally, we establish first-order convergence, with respect to the space and time meshsize, for sufficiently smooth solutions.

**Organization.** In section 2 we state the equations and recall the main results from [24]. Section 3 establishes the stability of an abstract discretization. Section 4 is devoted to the a priori error analysis for an exemplary concrete discretization, which is also tested numerically in section 5.

**Notation.** Throughout the paper, we denote by $\|\cdot\|_{\mathbb{X}}$ the norm of a Hilbert space $\mathbb{X}$. The dual space $\mathbb{X}^*$ acts on $\mathbb{X}$ through the pairing $\langle\cdot,\cdot\rangle_{\mathbb{X}}$. We denote by $L^2(\mathbb{X})$,

$H^1(\mathbb{X})$ and $C^0(\mathbb{X})$, respectively, the Bochner spaces of all $L^2$, $H^1$ and $C^0$ functions mapping the time interval $[0,T]$ into $\mathbb{X}$. For a measurable set $\Omega \subseteq \mathbb{R}^d$, we adopt the simplified notation $(\cdot,\cdot)_\Omega$ and $\|\cdot\|_\Omega$ for the scalar product and the norm in $L^2(\Omega)$. We write $a \lesssim b$ and $a \eqsim b$ when there are constants $0 < \underline{c} \leq \overline{c}$, possibly different at different occurrences, such that $a \leq \overline{c}\, b$ and $\underline{c}\, b \leq a \leq \overline{c}\, b$, respectively. The hidden constants are independent of the material parameters involved in the equations. The dependence on other relevant quantities is addressed case by case, see e.g. Remark 4.1.

## 2. Inf-sup theory for the Biot's equations

This section introduces the Biot's equations and summarizes some results from [24], that are useful for the construction and the analysis of the discretization in the next sections.

2.1. **Initial-boundary value problem.** Let $\Omega \subseteq \mathbb{R}^d$, $1 \leq d \leq 3$, be a bounded domain with polyhedral and Lipschitz continuous boundary. The flow of a Newtonian fluid inside a linear elastic porous medium located in $\Omega$, in the time interval $(0,T)$ with $T > 0$, is modeled by the quasi-static Biot's equations as follows

$$\begin{aligned} -\mathrm{div}(2\mu\nabla_S u + (\lambda \mathrm{div} u - \alpha p)\mathrm{I}) &= f_u \qquad \text{in } \Omega \times (0,T) \\ \partial_t(\alpha \mathrm{div} u + \sigma p) - \mathrm{div}(\kappa \nabla p) &= f_p \qquad \text{in } \Omega \times (0,T). \end{aligned} \tag{2.1a}$$

The first equation states the momentum balance for the elastic porous medium, whereas the second one is the mass balance for the fluid.

The unknown functions in the equations are the displacement $u : \Omega \to \mathbb{R}^d$ of the elastic porous medium and the pressure $p : \Omega \to \mathbb{R}$ of the fluid. The differential operator $\nabla_S$ is the symmetric part of the gradient and I is $d \times d$ identity tensor. The material parameters, denoted by Greek letters, are the Lamé constants $\mu, \lambda > 0$, the Biot-Willis constant $\alpha > 0$, the constrained specific storage coefficient $\sigma \geq 0$ and the hydraulic conductivity $\kappa > 0$. Consistently with [24], we assume that all parameters are constant in $\overline{\Omega} \times [0,T]$ for simplicity. We refer to [24, Remark 2.1] for a discussion on possible extensions.

We are interested in the initial-boundary value problem obtained by prescribing also the initial condition

$$(\alpha \mathrm{div} u + \sigma p)_{|t=0} = \ell_0 \quad \text{in} \quad \Omega \tag{2.1b}$$

as well as the boundary conditions

$$\begin{aligned} u &= 0 \quad \text{on} \quad \Gamma_{u,E} \times (0,T) \\ (2\mu\nabla_S u + (\lambda \mathrm{div} u - \alpha p)I)\mathsf{n} &= g_u \quad \text{on} \quad \Gamma_{u,N} \times (0,T) \\ p &= 0 \quad \text{on} \quad \Gamma_{p,E} \times (0,T) \\ \kappa \nabla p \cdot \mathsf{n} &= g_p \quad \text{on} \quad \Gamma_{p,N} \times (0,T) \end{aligned} \tag{2.1c}$$

where $\Gamma_{u,E} \cup \Gamma_{u,N} = \partial\Omega = \Gamma_{p,E} \cup \Gamma_{p,N}$ and $\Gamma_{u,E} \cap \Gamma_{u,N} = \emptyset = \Gamma_{p,E} \cap \Gamma_{p,N}$. The letter $\mathsf{n}$ denotes the outward unit normal vector on $\partial\Omega$. The subscripts '$E$' and '$N$' are intended to assist the reader in distinguishing the portions of the boundary with essential and natural conditions.

We point out that different statements of the initial and of the boundary conditions are sometimes met in the literature. We refer to [24, Remark 2.2-2.3] for a more extensive discussion on this point.

2.2. **Weak Formulation and well-posedness.** For convenience, we introduce a compact notation for the differential operators in the Biot's equations (2.1), namely

$$\mathcal{E} := -\mathrm{div}(2\mu\nabla_S) \qquad \mathcal{D} := \mathrm{div} \qquad \mathcal{L} := -\mathrm{div}(\kappa\nabla). \tag{2.2}$$

Notice that $\mathcal{E}$ and $\mathcal{L}$ act on $u$ and $p$, respectively, in (2.1a). Comparing also with the boundary conditions, it looks reasonable that the regularity of $u$ in space in a weak formulation can be described via

$$\mathbb{U} := \begin{cases} H^1(\Omega)^d/\mathrm{RM} & \text{if } \Gamma_{u,N} = \partial\Omega \\ H^1_{\Gamma_{u,E}}(\Omega)^d & \text{otherwise} \end{cases} \tag{2.3}$$

with RM denoting the rigid body motions. Analogously, for the regularity of $p$ in space, we consider

$$\mathbb{P} := \begin{cases} H^1(\Omega) \cap L^2_0(\Omega) & \text{if } \Gamma_{p,N} = \partial\Omega \\ H^1_{\Gamma_{p,E}}(\Omega) \cap L^2_0(\Omega) & \text{if } \Gamma_{p,N} \neq \partial\Omega, \Gamma_{u,E} = \partial\Omega, \sigma = 0 \\ H^1_{\Gamma_{p,E}}(\Omega) & \text{otherwise} \end{cases} \tag{2.4}$$

where $L^2_0(\Omega) = \{q \in L^2(\Omega) \mid \int_\Omega q = 0\}$. We refer to [24, Remark 2.5] for a motivation of the nonstandard definition of $\mathbb{P}$ in the second case.

*Remark* 2.1 (Notation). We are aware of the fact, that the abstract notation introduced here (and the subsequent one for all related spaces and operators) makes the reading possibly harder. Still, it has the advantage that all combinations of the boundary conditions (as well as the critical case $\sigma = 0$) can be treated at the same time. In our view, this is quite important, because our approach to the analysis and the discretization of the Biot's equations is mainly the same in all cases, but each case may require subtle minor modifications.

The action of the divergence on $u$ in (2.1a) indicates that also the space

$$\mathbb{D} := \mathcal{D}(\mathbb{U}) = \begin{cases} L^2_0(\Omega) & \text{if } \Gamma_{u,E} = \partial\Omega \\ L^2(\Omega) & \text{otherwise} \end{cases} \tag{2.5}$$

plays a relevant role in the Biot's equations. Similarly, since $p$ is involved in (2.1a) also without the action of any differential operator in space, we repeatedly make use of

$$\overline{\mathbb{P}} = \begin{cases} L^2_0(\Omega) & \text{if } \Gamma_{p,N} = \partial\Omega \text{ or } \Gamma_{u,E} = \partial\Omega, \sigma = 0 \\ L^2(\Omega) & \text{otherwise,} \end{cases} \tag{2.6}$$

where the closure is taken with respect to the $L^2(\Omega)$-norm. An important point for our analysis is that both $\mathbb{D}$ and $\overline{\mathbb{P}}$ are subspaces of $L^2(\Omega)$, but their mutual relation depends on the boundary conditions and on the parameter $\sigma$. Therefore, it is useful introducing

$$\mathcal{P}_{\mathbb{D}} : L^2(\Omega) \to \mathbb{D} \qquad \text{and} \qquad \mathcal{P}_{\overline{\mathbb{P}}} : L^2(\Omega) \to \overline{\mathbb{P}},$$

the $L^2(\Omega)$-orthogonal projections onto $\mathbb{D}$ and $\overline{\mathbb{P}}$, respectively.

*Remark* 2.2 (Functional setting). The diagram in Figure 2.1 summarizes the relation among the spaces and the operators introduced up to this point. In addition, we denote by $i : \mathbb{P} \to \overline{\mathbb{P}}$ the inclusion operator and $\mathcal{D}^*$ and $i^*$ are the adjoint operators of $\mathcal{D}$ and $i$, respectively. The spaces $\mathbb{D}$ and $\overline{\mathbb{P}}$ are identified with their duals

via the $L^2(\Omega)$-scalar product. Thus, the square on the right side of the diagram involves the Hilbert triplet

$$\mathbb{P} \hookrightarrow \overline{\mathbb{P}} \equiv \overline{\mathbb{P}}^* \hookrightarrow \mathbb{P}^*. \tag{2.7}$$

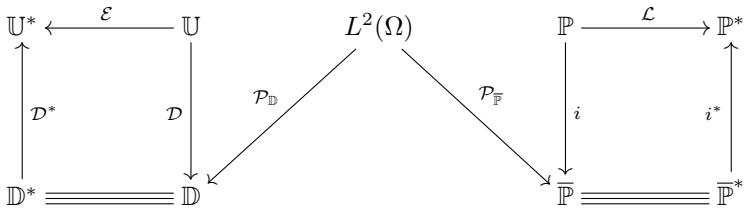


FIGURE 2.1. Spaces and operators describing the regularity in space for the weak formulation (2.8) of the Biot's equations. The triple lines on the bottom indicate identification via the $L^2(\Omega)$-scalar product.

With this preparation, we are in position to recall the weak formulation of the initial-boundary value problem (2.1) introduced in [24, section 2.3], namely

$$\begin{aligned}
\mathcal{E}u + \mathcal{D}^* p_{\mathrm{tot}} &= \ell_u && \text{in } L^2(\mathbb{U}^*)\\
\lambda \mathcal{D}u - p_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}} p &= 0 && \text{in } L^2(\mathbb{D})\\
\alpha \mathcal{P}_{\overline{\mathbb{P}}} \mathcal{D}u + \sigma p - m &= 0 && \text{in } L^2(\overline{\mathbb{P}})\\
\partial_t m + \mathcal{L} p &= \ell_p && \text{in } L^2(\mathbb{P}^*)\\
m(0) &= \ell_0 && \text{in } \mathbb{P}^*.
\end{aligned} \tag{2.8}$$

The loads $\ell_u \in L^2(\mathbb{U}^*)$ and $\ell_p \in L^2(\mathbb{P}^*)$ result from the data $f_u$ and $f_p$ in the equations (2.1a) as well as $g_u$ and $g_p$ in the boundary conditions (2.1c).

*Remark* 2.3 (Auxiliary variables). Compared to (2.1a), the weak formulation (2.8) involves two additional unknown variables, namely the total pressure $p_{\mathrm{tot}}$ and the total fluid content $m$ defined by the second and third equation, respectively. Introducing these variables is not strictly necessary for our analysis, but it substantially simplifies the definition of the space $\mathbb{Y}_1$ in (2.9) below. The use of the total pressure was observed in [27] to help also the design of robust linear solvers.

The main result in [24] states that (2.8) is uniquely solvable within the closure $\overline{\mathbb{Y}}_1$ of the space

$$\mathbb{Y}_1 := L^2(\mathbb{U}) \times L^2(\mathbb{D}) \times L^2(\mathbb{P}) \times \left(L^2(\overline{\mathbb{P}}) \cap H^1(\mathbb{P}^*)\right) \tag{2.9}$$

with respect to the norm

$$\begin{aligned}
&\|(\widetilde{u}, \widetilde{p}_{\mathrm{tot}}, \widetilde{p}, \widetilde{m})\|_1^2 :=\\
&\qquad \int_0^T \left( \|\widetilde{u}\|_{\mathbb{U}}^2 + \frac{1}{\mu} \|\widetilde{p}_{\mathrm{tot}}\|_\Omega^2 + \|\partial_t \widetilde{m} + \mathcal{L}\widetilde{p}\|_{\mathbb{P}^*}^2 \right) + \|\widetilde{m}(0)\|_{\mathbb{P}^*}^2\\
&\qquad + \int_0^T \left( \frac{1}{\mu + \lambda} \|\lambda \mathcal{D}\widetilde{u} - \widetilde{p}_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}} \widetilde{p}\|_\Omega^2 + \gamma \|\alpha \mathcal{P}_{\overline{\mathbb{P}}} \mathcal{D}\widetilde{u} + \sigma \widetilde{p} - \widetilde{m}\|_\Omega^2 \right).
\end{aligned} \tag{2.10}$$

Here $\mathbb{U}$ and $\mathbb{P}$ are equipped with the energy norm

$$\|\cdot\|_{\mathbb{U}}^2 = \langle \mathcal{E}\cdot,\cdot\rangle_{\mathbb{U}} \qquad \text{and} \qquad \|\cdot\|_{\mathbb{P}}^2 = \langle \mathcal{L}\cdot,\cdot\rangle_{\mathbb{P}} \tag{2.11}$$

and the parameter $\gamma$ is given by

$$\gamma := \begin{cases} \min\left\{\dfrac{\mu+\lambda}{\alpha^2}, \dfrac{1}{\sigma}\right\} & \text{if } \sigma > 0 \text{ and } \overline{\mathbb{P}} \subseteq \mathbb{D} \\ \dfrac{\mu+\lambda}{\alpha^2} + \dfrac{1}{\sigma} & \text{if } \sigma > 0 \text{ and } \overline{\mathbb{P}} \not\subseteq \mathbb{D} \\ \dfrac{\mu+\lambda}{\alpha^2} & \text{if } \sigma = 0. \end{cases} \tag{2.12}$$

Taking the closure is indeed necessary, because $\mathbb{Y}_1$ is not closed with respect to $\|\cdot\|_1$, cf. [24, Proposition 4.4].

**Theorem 2.4** (Well-posedness of the weak formulation)**.** *For all possible data $(\ell_u, \ell_p, \ell_0) \in L^2(\mathbb{U}^*)\times L^2(\mathbb{P}^*)\times \mathbb{P}^*$, the weak formulation* (2.8) *has a unique solution $y_1 = (u, p_{tot}, p, m) \in \overline{\mathbb{Y}}_1$, which satisfies the two-sided stability bound*

$$\|y_1\|_1^2 \eqsim \int_0^T \left(\|\ell_u\|_{\mathbb{U}^*}^2 + \|\ell_p\|_{\mathbb{P}^*}^2\right) + \|\ell_0\|_{\mathbb{P}^*}^2.$$

*Moreover, we have $m \in C^0(\mathbb{P}^*)$ as well as the norm equivalence*

$$\|y_1\|_1^2 \eqsim \|y_1\|_1^2 + \|m\|_{L^\infty(\mathbb{P}^*)}^2 + \int_0^T \left(\lambda\|\mathcal{D}u\|_\Omega^2 + \gamma^{-1}\|p\|_\Omega^2\right).$$

*All hidden constants depend only on $\Omega$ and $T$.*

*Proof.* Combine [24, Theorem 3.5] with [24, Proposition 4.1]. □

Although we omit the proof of Theorem 2.4, it is worth roughly summarizing how it is obtained. Indeed, we shall make use of a similar argument in order to verify the well-posedness of the discretization introduced in the next section, cf. Theorem 3.9 below.

The weak formulation (2.8) can be viewed as a special instance of the following linear variational problem: given $\ell \in \mathbb{Y}_2^*$, find $y_1 \in \overline{\mathbb{Y}}_1$ such that

$$b(y_1, y_2) = \langle \ell, y_2\rangle_{\mathbb{Y}_2} \qquad \forall y_2 \in \mathbb{Y}_2. \tag{2.13}$$

The test space is obtained by collecting all possible test functions for (2.8), namely

$$\mathbb{Y}_2 := L^2(\mathbb{U}) \times L^2(\mathbb{D}) \times L^2(\overline{\mathbb{P}}) \times L^2(\mathbb{P}) \times \mathbb{P}$$

and it is equipped with the norm

$$\begin{aligned} \|(v, q_{\mathrm{tot}}, q, n, n_0)\|_2^2 := &\int_0^T \left(\|v\|_{\mathbb{U}}^2 + \|n\|_{\mathbb{P}}^2\right) + \|n_0\|_{\mathbb{P}}^2 \\ &+ \int_0^T \left((\mu+\lambda)\|q_{\mathrm{tot}}\|_\Omega^2 + \gamma^{-1}\|q\|_\Omega^2\right). \end{aligned} \tag{2.14}$$

The bilinear form $b : \overline{\mathbb{Y}}_1 \times \mathbb{Y}_2 \to \mathbb{R}$ is defined according to the left-hand side of (2.8) by

$$\begin{aligned} b(\widetilde{y}_1, y_2) := &\int_0^T \Big( \langle \mathcal{E}\widetilde{u} + \mathcal{D}^*\widetilde{p}_{\mathrm{tot}}, v\rangle_{\mathbb{U}} + \langle \partial_t \widetilde{m} + \mathcal{L}\widetilde{p}, n\rangle_{\mathbb{P}} \Big) + \langle \widetilde{m}(0), n_0\rangle_{\mathbb{P}} \\ &+ \int_0^T \Big( (\lambda\mathcal{D}\widetilde{u} - \widetilde{p}_{\mathrm{tot}} - \alpha\mathcal{P}_{\mathbb{D}}\widetilde{p}, q_{\mathrm{tot}})_\Omega + \big(\alpha\mathcal{P}_{\overline{\mathbb{P}}}\mathcal{D}\widetilde{u} + \sigma\widetilde{p} - \widetilde{m}, q\big)_\Omega \Big) \end{aligned} \tag{2.15}$$

for $\widetilde{y}_1 = (\widetilde{u}, \widetilde{p}_{\mathrm{tot}}, \widetilde{p}, \widetilde{m}) \in \overline{\mathbb{Y}}_1$ and $(v, q_{\mathrm{tot}}, q, n, n_0) \in \mathbb{Y}_2$.

The so-called Banach-Nečas theorem characterizes the well-posedness of the linear variational problem in terms of boundedness, inf-sup stability and nondegeneracy of the form $b$, see e.g. [14, theorem 25.9]. These properties are verified in [24, section 3]. Their combination implies the well-posedness of (2.8) as well as the two-sided stability bound in Theorem 2.4.

*Remark* 2.5 (Trial functions). As in (2.10) and (2.15), we use hereafter the superscript '$\sim$' to distinguish a general trial function in $\overline{\mathbb{Y}}_1$ from the solution of the weak formulation (2.8).

*Remark* 2.6 (Functions and functionals). Owing to Remark 2.2, we often identify the functions in $\mathbb{D}$ and $\overline{\mathbb{P}}$ with their Riesz representatives in $\mathbb{D}^*$ and $\overline{\mathbb{P}}^*$ and vice versa. For instance, the term $\mathcal{D}^* p_{\mathrm{tot}}$ from the first equation in (2.8) and even the space $L^2(\overline{\mathbb{P}}) \cap H^1(\mathbb{P}^*)$ proposed for the total fluid content must be interpreted in this way. As usual, we omit to explicitly indicate the Riesz isometries, accepting some ambiguity, so as to alleviate the notation. We apply the same principles to the discretization in the next sections.

## 3. Abstract inf-sup stable discretization

In this section we design a discretization of the weak formulation (2.8) of the initial-boundary value problem for the Biot's equations (2.1). The space discretization is challenging, due to the nontrivial coupling of the variables and the various differential operators acting on them. Therefore, we initially work with a general discretization, so as to allow for the maximal flexibility. We discuss a concrete realization in section 4 below. The time discretization seems less problematic, because (2.8) involves only one time derivative. Hence, we directly make a concrete choice, namely the backward Euler scheme.

The space discretization in this section builds upon a number of assumptions that must be verified case by case. The set of our assumptions is identified by special tags with the dedicated enumeration (H1), (H2), etc. With a small abuse, we actually include among the assumptions also the definition of some relevant constants. In those cases, the size (and not just the existence) of the constants is the property to be verified in each concrete example.

The main result in this section is the well-posedness established in Theorem 3.9. We do not attempt at analyzing the error at this level of generality. Indeed, we believe that the estimates we would obtain either require too many assumptions or are too much abstract to be of interest. Thus, we prefer to discuss the error analysis for the concrete example in section 4.

*Remark* 3.1 (Notation for the discretization). In general, we mark all spaces and operators related to the discretization in space by the subscript '$\mathsf{s}$' and those related to the discretization in time by the subscript '$\mathsf{t}$'. The combination '$\mathsf{st}$' of the two subscripts identifies the full discretization in space and time. To alleviate the notation, we use capitol letters (in place of subscripts) to distinguish the functions involved in the discretization from those related to the original Biot's equations.

### 3.1. Abstract discretization in space.

The general concept of our discretization in space consists in replacing all spaces and operators in Figure 2.1 by finite-dimensional counterparts, while preserving the structure of the diagram therein, cf.

Figure 3.1. In order to discretize the displacement and the pressure, we consider two finite-dimensional linear spaces

$$\mathbb{U}_\mathsf{s} \qquad \text{and} \qquad \mathbb{P}_\mathsf{s} \tag{3.1}$$

i.e. discrete counterparts of the spaces $\mathbb{U}$ in (2.3) and $\mathbb{P}$ in (2.4). We replace the operators $\mathcal{E}$ and $\mathcal{L}$ in (2.2) by positive definite and self-adjoint linear operators

$$\mathcal{E}_\mathsf{s} : \mathbb{U}_\mathsf{s} \to \mathbb{U}_\mathsf{s}^* \qquad \text{and} \qquad \mathcal{L}_\mathsf{s} : \mathbb{P}_\mathsf{s} \to \mathbb{P}_\mathsf{s}^*. \tag{3.2}$$

In analogy with (2.11), we equip the two spaces with the energy norms

$$\|\cdot\|_{\mathbb{U}_\mathsf{s}}^2 := \langle \mathcal{E}_\mathsf{s} \cdot, \cdot \rangle_{\mathbb{U}_\mathsf{s}} \qquad \text{and} \qquad \|\cdot\|_{\mathbb{P}_\mathsf{s}}^2 := \langle \mathcal{L}_\mathsf{s} \cdot, \cdot \rangle_{\mathbb{P}_\mathsf{s}}. \tag{3.3}$$

Then, the dual norms on $\mathbb{U}_\mathsf{s}^*$ and $\mathbb{P}_\mathsf{s}^*$ are given by

$$\begin{aligned} \|\cdot\|_{\mathbb{U}_\mathsf{s}^*}^2 &:= \sup_{V \in \mathbb{U}_\mathsf{s}} \frac{\langle \cdot, V \rangle_{\mathbb{U}_\mathsf{s}}}{\|V\|_{\mathbb{U}_\mathsf{s}}} &&= \langle \cdot, \mathcal{E}_\mathsf{s}^{-1} \cdot \rangle_{\mathbb{U}_\mathsf{s}} \\ \|\cdot\|_{\mathbb{P}_\mathsf{s}^*}^2 &:= \sup_{N \in \mathbb{P}_\mathsf{s}} \frac{\langle \cdot, N \rangle_{\mathbb{P}_\mathsf{s}}}{\|N\|_{\mathbb{P}_\mathsf{s}}} &&= \langle \cdot, \mathcal{L}_\mathsf{s}^{-1} \cdot \rangle_{\mathbb{P}_\mathsf{s}}. \end{aligned} \tag{3.4}$$

Notice that $\mathbb{U}_\mathsf{s}$ and $\mathbb{P}_\mathsf{s}$ are not required to be conforming, i.e. subspaces of $\mathbb{U}$ and $\mathbb{P}$, respectively. Therefore, in order to define an error notion on the sums $\mathbb{U} + \mathbb{U}_\mathsf{s}$ and $\mathbb{P} + \mathbb{P}_\mathsf{s}$, we assume the following.

(H1) The norms $\|\cdot\|_\mathbb{U}$ and $\|\cdot\|_\mathbb{P}$ in (2.11) can be extended to $\mathbb{U} + \mathbb{U}_\mathsf{s}$ and $\mathbb{P} + \mathbb{P}_\mathsf{s}$ with $\|\cdot\|_\mathbb{U} \eqsim \|\cdot\|_{\mathbb{U}_\mathsf{s}}$ in $\mathbb{U}_\mathsf{s}$ and $\|\cdot\|_\mathbb{P} \eqsim \|\cdot\|_{\mathbb{P}_\mathsf{s}}$ in $\mathbb{P}_\mathsf{s}$.

In order to discretize the space $\mathbb{D}$ in (2.5), we consider a linear operator

$$\mathcal{D}_\mathsf{s} : \mathbb{U}_\mathsf{s} \to L^2(\Omega) \tag{3.5}$$

i.e. a discrete counterpart of the divergence $\mathcal{D}$ in (2.2). Then, we set

$$\mathbb{D}_\mathsf{s} := \mathcal{D}_\mathsf{s}(\mathbb{U}_\mathsf{s}). \tag{3.6}$$

The proof of Theorem 2.4 given in [24] exploits the norm equivalence $\mu\|\mathcal{D}^* \cdot\|_{\mathbb{U}^*}^2 \eqsim \|\cdot\|_\Omega^2$ in $\mathbb{D}$, that is nothing else than boundedness and surjectivity of $\mathcal{D}$, cf. [14, Lemma C40]. (Note that here $\mathbb{D}$ is identified with its dual via the $L^2(\Omega)$-scalar product, cf. Remark 2.2.) In order to reproduce this property at the discrete level, we assume the following.

(H2) There are constants $c = c(\mathbb{U}_\mathsf{s})$ and $C = C(\mathbb{U}_\mathsf{s})$ with $0 < c \leq C$ and such that $c\|\cdot\|_\Omega^2 \leq \mu\|\mathcal{D}_\mathsf{s}^* \cdot\|_{\mathbb{U}_\mathsf{s}^*}^2 \leq C\|\cdot\|_\Omega^2$ in $\mathbb{D}_\mathsf{s}$.

Actually, this prescribes that $\mathbb{U}_\mathsf{s}/\mathbb{D}_\mathsf{s}$ is a stable pair for the discretization of the Stokes equations. Note that also $\mathbb{D}_\mathsf{s}$ is identified here with its dual space.

The proof of Theorem 2.4 given in [24] exploits also the density of $\mathbb{P}$ in $\overline{\mathbb{P}}$, that gives rise to the Hilbert triplet in (2.7). Since $\mathbb{P}_\mathsf{s}$ is finite-dimensional, we are led to discretize $\overline{\mathbb{P}}$ by $\mathbb{P}_\mathsf{s}$ itself, giving rise to the triplet

$$\mathbb{P}_\mathsf{s} = \overline{\mathbb{P}}_\mathsf{s} \equiv \overline{\mathbb{P}}_\mathsf{s}^* = \mathbb{P}_\mathsf{s}^*. \tag{3.7}$$

Also in this case, the identification of $\overline{\mathbb{P}}_\mathsf{s}$ with its dual space is made via the $L^2(\Omega)$-scalar product. Hence, the pairing $\langle \cdot, \cdot \rangle_{\mathbb{P}_\mathsf{s}}$ coincides with $(\cdot, \cdot)_\Omega$ upon identifying the functionals in $\mathbb{P}_\mathsf{s}^*$ with their Riesz representative.

*Remark* 3.2 (Discretization of the Hilbert triplet). As $\mathbb{P}_\mathsf{s}$ coincides with its closure, the above Hilbert triplet is trivial from the algebraic viewpoint. Still, the spaces in it play different roles and are equipped with different norms, depending on their position, in analogy with (2.7). To alleviate the notation, we omit hereafter the symbol of the closure.

We denote by $\mathcal{P}_{\mathbb{D}_\mathsf{s}}$ and $\mathcal{P}_{\mathbb{P}_\mathsf{s}}$ the $L^2(\Omega)$-orthogonal projections onto $\mathbb{D}_\mathsf{s}$ and $\mathbb{P}_\mathsf{s}$, respectively. In particular, the adjoint $\mathcal{P}^*_{\mathbb{P}_\mathsf{s}}$ of the second projection maps functionals on $\mathbb{P}_\mathsf{s}$ into functionals on $\mathbb{P}$. This observation is important for the definition of the error notion in section 3.4 below, because the trial norm $\|\cdot\|_1$ in (2.10) involves the dual norm $\|\cdot\|_{\mathbb{P}^*}$. Thus we assume the following, in order to keep the norm of $\mathcal{P}^*_{\mathbb{P}_\mathsf{s}}$ under control.

(H3) There are constants $c = c(\mathbb{P}_\mathsf{s})$ and $C = C(\mathbb{P}_\mathsf{s})$ with $0 < c \le C$ and such that $c\|\cdot\|_{\mathbb{P}^*_\mathsf{s}} \le \|\mathcal{P}^*_{\mathbb{P}_\mathsf{s}} \cdot\|_{\mathbb{P}^*} \le C\|\cdot\|_{\mathbb{P}^*_\mathsf{s}}$ in $\mathbb{P}^*_\mathsf{s}$.

The upper bound here is equivalent to the $\mathbb{P}$-stability of the projection $\mathcal{P}_{\mathbb{P}_\mathsf{s}}$. The lower bound can be formulated as an inf-sup condition and it is equivalent to the existence of a bounded right inverse of $\mathcal{P}_{\mathbb{P}_\mathsf{s}}$, cf. [23, Proposition 3.5].

Finally, the proof of the inf-sup stability in Lemma 3.8 below for vanishing $\sigma$ makes use of the following assumption.

(H4) The inclusion $\mathbb{P}_\mathsf{s} \subseteq \mathbb{D}_\mathsf{s}$ holds true if $\sigma = 0$.

Notice that the spaces $\overline{\mathbb{P}}$ and $\mathbb{D}$ in (2.6) and (2.5), respectively, satisfy the same inclusion.

*Remark* 3.3 (Spurious pressure oscillations). The combination of the assumptions (H2) and (H4) prescribes that $\mathbb{U}_\mathsf{s}/\mathbb{P}_\mathsf{s}$ is a stable pair for the discretization of the Stokes equations if $\sigma = 0$. This property was observed both numerically [18] and theoretically [29] to be important to prevent from spurious pressure oscillations in certain regimes. Indeed, forgetting the time derivative for a moment (or, more precisely, discretizing it in time) the Biot's equations (2.1a) are close to the Stokes equations for vanishing $\sigma$ and small $\kappa$.

Figure 3.1 summarizes the relation among the spaces and the operators introduced in this section. As announced, the structure is exactly as in Figure 2.1. We refer to Remark 2.2 for the details of the notation.

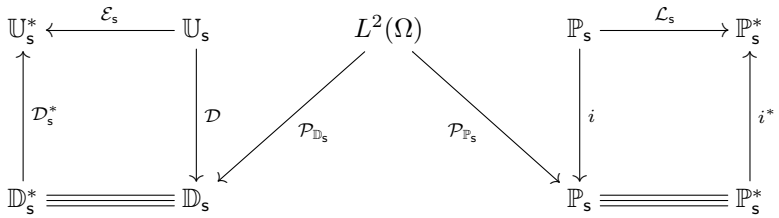


FIGURE 3.1. Spaces and operators describing the regularity in space for the discretization of the Biot's equations. The triple lines on the bottom indicate identification via the $L^2(\Omega)$-scalar product.

3.2. **Discretization in time.** As announced, we consider a simple first-order discretization in time, namely backward Euler. To this end, we introduce a partition

$$0 = t_0 < t_1 < \cdots < t_J = T \tag{3.8}$$

of the time interval $[0,T]$ with $J \geq 1$. We denote the local time intervals and their length by

$$I_j := [t_{j-1}, t_j] \qquad \text{and} \qquad |I_j| := t_j - t_{j-1} \tag{3.9}$$

respectively, for $j = 1, \dots, J$.

For $\mathbb{X} \in \{\mathbb{U}_\mathsf{s}, \mathbb{D}_\mathsf{s}, \mathbb{P}_\mathsf{s}\}$, we consider the space of piecewise time-constant functions on the above partition with values in $\mathbb{X}$

$$\mathbb{S}^0_\mathsf{t}(\mathbb{X}) := \big\{ X \in L^\infty(\mathbb{X}) \mid X_{|I_j} =: X_j \in \mathbb{X},\ j = 1, \dots, J \big\}.$$

Whenever useful, we identify a function $X \in \mathbb{S}^0_\mathsf{t}(\mathbb{X})$ with the sequence of its values $(X_j)_{j=1}^J \subseteq \mathbb{X}$.

The space $\mathbb{S}^0_\mathsf{t}(\mathbb{X})$ is suitable for a first-order discretization in time of $L^2(\mathbb{X})$, i.e. of the first three components in the trial (2.9) and of the first four components in the test (2.2). The last component in the trial space (the fluid content) is different, as it involves $H^1$-regularity in time. We discretize it in time by

$$\mathbb{S}^0_\mathsf{t}(\mathbb{P}_\mathsf{s}) \times \mathbb{P}_\mathsf{s}$$

where the second component plays the role of the initial value. Then we introduce a discrete counterpart $\mathrm{d}_\mathsf{t} : \mathbb{S}^0_\mathsf{t}(\mathbb{P}_\mathsf{s}) \times \mathbb{P}_\mathsf{s} \to \mathbb{S}^0_\mathsf{t}(\mathbb{P}^*_\mathsf{s})$ of the time derivative $\partial_t$, in the vein of the backward Euler scheme, i.e.

$$\mathrm{d}_\mathsf{t}(\widetilde{M}, \widetilde{M}_0)_{I_j} := \frac{\widetilde{M}_j - \widetilde{M}_{j-1}}{|I_j|} \tag{3.10}$$

for $(\widetilde{M}, \widetilde{M}_0) \in \mathbb{S}^0_\mathsf{t}(\mathbb{P}_\mathsf{s}) \times \mathbb{P}_\mathsf{s}$ and for all $j = 1, \dots, J$.

The proof of Theorem 2.4 given in [24] and, in particular, the control of the point values of the total fluid content hinge on the following integration by parts rule $2\int_0^T \big\langle \partial_t \widetilde{m}, \mathcal{L}^{-1}\widetilde{m} \big\rangle_{\mathbb{P}} = \|\widetilde{m}(T)\|^2_{\mathbb{P}^*} - \|\widetilde{m}(0)\|^2_{\mathbb{P}^*}$, which holds true for $\widetilde{m} \in H^1(\mathbb{P}^*)$, cf. [15, Lemma 64.40]. The operator $\mathrm{d}_\mathsf{t}$ defined above satisfies a similar relation.

**Lemma 3.4** (Time discrete integration by parts rule)**.** *We have*

$$2\int_0^{t_j} \langle \mathrm{d}_\mathsf{t}(\widetilde{M}, \widetilde{M}_0), \mathcal{L}^{-1}_\mathsf{s}\widetilde{M} \rangle_{\mathbb{P}_\mathsf{s}} \geq \|\widetilde{M}_j\|^2_{\mathbb{P}^*_\mathsf{s}} - \|\widetilde{M}_0\|^2_{\mathbb{P}^*_\mathsf{s}}$$

*for all* $(\widetilde{M}, \widetilde{M}_0) \in \mathbb{S}^0_\mathsf{t}(\mathbb{P}_\mathsf{s}) \times \mathbb{P}_\mathsf{s}$ *and* $j = 1, \dots, J$.

*Proof.* We exploit (3.10), rearrange terms and recall the second part of (3.4). It results

$$\begin{aligned} 2\int_0^{t_j} \langle \mathrm{d}_\mathsf{t}(\widetilde{M}, \widetilde{M}_0), \widetilde{\mathcal{L}}^{-1}_\mathsf{s} M \rangle_{\mathbb{P}_\mathsf{s}} &= 2\sum_{k=1}^{j} \langle \widetilde{M}_k - \widetilde{M}_{k-1}, \mathcal{L}^{-1}_\mathsf{s}\widetilde{M}_k \rangle_{\mathbb{P}_\mathsf{s}} \\ &= \|\widetilde{M}_j\|^2_{\mathbb{P}^*_\mathsf{s}} + \sum_{k=1}^{j} \|\widetilde{M}_k - \widetilde{M}_{k-1}\|^2_{\mathbb{P}^*_\mathsf{s}} - \|\widetilde{M}_0\|^2_{\mathbb{P}^*_\mathsf{s}} \end{aligned}$$

cf. [15, eq. (67.9)]. This readily implies the claimed inequality. □

3.3. **Full discretization.** Combining the discretizations in space and time from the previous sections, we are in position to propose an abstract full discretization of the initial-boundary value problem (2.1) for the Biot's equations.

We consider the trial space

$$\mathbb{Y}_{1,\mathsf{st}} := \mathbb{S}^0_{\mathsf{t}}(\mathbb{U}_{\mathsf{s}}) \times \mathbb{S}^0_{\mathsf{t}}(\mathbb{D}_{\mathsf{s}}) \times \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}}) \times \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}}) \times \mathbb{P}_{\mathsf{s}}. \tag{3.11}$$

Inspired by (2.10) and taking the assumption (H1) in section 3.1 into account, we equip $\mathbb{Y}_{1,\mathsf{st}}$ with the norm

$$\begin{aligned} &\|(\widetilde{U}, \widetilde{P}_{\mathrm{tot}}, \widetilde{P}, \widetilde{M}, \widetilde{M}_0)\|^2_{1,\mathsf{st}} := \\ &\qquad \int_0^T \Big( \|\widetilde{U}\|^2_{\mathbb{U}} + \frac{1}{\mu}\|\widetilde{P}_{\mathrm{tot}}\|^2_{\Omega} + \|\mathrm{d}_{\mathsf{t}}(\widetilde{M}, \widetilde{M}_0) + \mathcal{L}_{\mathsf{s}}\widetilde{P}\|^2_{\mathbb{P}^*_{\mathsf{s}}} \Big) + \|\widetilde{M}_0\|^2_{\mathbb{P}^*_{\mathsf{s}}} \\ &\quad + \int_0^T \Big( \frac{1}{\mu+\lambda}\|\lambda \mathcal{D}_{\mathsf{s}}\widetilde{U} - \widetilde{P}_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}_{\mathsf{s}}}\widetilde{P}\|^2_{\Omega} + \gamma \|\alpha \mathcal{P}_{\mathbb{P}_{\mathsf{s}}}\mathcal{D}_{\mathsf{s}}\widetilde{U} + \sigma\widetilde{P} - \widetilde{M}\|^2_{\Omega} \Big). \end{aligned} \tag{3.12}$$

*Remark* 3.5 (Equivalent trial norm). According to the assumption (H3) in section 3.1, we could equivalently replace the discrete dual norm $\|\cdot\|_{\mathbb{P}^*_{\mathsf{s}}}$ in the definition of $\|\cdot\|_{1,\mathsf{st}}$ by $\|\mathcal{P}^*_{\mathbb{P}_{\mathsf{s}}}\cdot\|_{\mathbb{P}^*}$. All the results stated in section 3.4 hold true also in this case, with the only difference that the hidden constants additionally depend on the constants in (H3). This observation is important in view of the definition of the error notion in section 4.2.

For the test space, we proceed similarly and set

$$\mathbb{Y}_{2,\mathsf{st}} := \mathbb{S}^0_{\mathsf{t}}(\mathbb{U}_{\mathsf{s}}) \times \mathbb{S}^0_{\mathsf{t}}(\mathbb{D}_{\mathsf{s}}) \times \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}}) \times \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}}) \times \mathbb{P}_{\mathsf{s}}. \tag{3.13}$$

Recalling again the assumption (H1) in section 3.1, we equip $\mathbb{Y}_{2,\mathsf{st}}$ with the norm $\|\cdot\|_2$ in (2.14).

Let $(L_u, L_p, L_0) \in \mathbb{S}^0_{\mathsf{t}}(\mathbb{U}^*_{\mathsf{s}}) \times \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}^*_{\mathsf{s}}) \times \mathbb{P}^*_{\mathsf{s}}$ be a discretization of the corresponding data $(\ell_u, \ell_p, \ell_0) \in L^2(\mathbb{U}^*) \times L^2(\mathbb{P}^*) \times \mathbb{P}^*$ in the weak formulation (2.8) of the Biot's equations. We consider the following full discretization of (2.8): find $(U, P_{\mathrm{tot}}, P, M, M_0) \in \mathbb{Y}_{1,\mathsf{st}}$ such that

$$\begin{aligned} \mathcal{E}_{\mathsf{s}}U + \mathcal{D}^*_{\mathsf{s}}P_{\mathrm{tot}} &= L_u && \text{in } \mathbb{S}^0_{\mathsf{t}}(\mathbb{U}^*_{\mathsf{s}}) \\ \lambda\mathcal{D}_{\mathsf{s}}U - P_{\mathrm{tot}} - \alpha\mathcal{P}_{\mathbb{D}_{\mathsf{s}}}P &= 0 && \text{in } \mathbb{S}^0_{\mathsf{t}}(\mathbb{D}_{\mathsf{s}}) \\ \alpha\mathcal{P}_{\mathbb{P}_{\mathsf{s}}}\mathcal{D}_{\mathsf{s}}U + \sigma P - M &= 0 && \text{in } \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}}) \\ \mathrm{d}_{\mathsf{t}}(M, M_0) + \mathcal{L}_{\mathsf{s}}P &= L_p && \text{in } \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}^*_{\mathsf{s}}) \\ M_0 &= L_0 && \text{in } \mathbb{P}^*_{\mathsf{s}}. \end{aligned} \tag{3.14}$$

Note that, also in this case, there are some ambiguities between functions and functionals, that can be clarified in the vein of Remark 2.6.

*Remark* 3.6 (Two- and four-field formulation). The weak formulation 2.8 involves four unknown variables but it can be equivalently rewritten as a two-field weak formulation of the initial-boundary value problem (2.1), by eliminating the total pressure $p_{\mathrm{tot}}$ and the total fluid content $m$, cf. [24, Remark 2.7]. Analogously, we could eliminate the discrete total pressure $P_{\mathrm{tot}}$ and the discrete fluid content $M$ from (3.14). In this way, we would equivalently obtain a discretization of the two-field weak formulation, with trial and test spaces given by $\mathbb{S}^0_{\mathsf{t}}(\mathbb{U}_{\mathsf{s}}) \times \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}}) \times \mathbb{P}_{\mathsf{s}}$.

Since we aim at establishing the well-posedness of (3.14) via the inf-sup theory, it is convenient viewing it as an instance of the following linear variational problem: given $L \in \mathbb{Y}_{2,\mathsf{st}}^*$, find $Y_1 \in \mathbb{Y}_{1,\mathsf{st}}$ such that

$$b_{\mathsf{st}}(Y_1, Y_2) = \langle L, Y_2 \rangle_{\mathbb{Y}_{2,\mathsf{st}}} \qquad \forall Y_2 \in \mathbb{Y}_{2,\mathsf{st}}. \tag{3.15}$$

Of course, this can be seen as a discretization of (2.13). Here, the bilinear form $b_{\mathsf{st}} : \mathbb{Y}_{1,\mathsf{st}} \times \mathbb{Y}_{2,\mathsf{st}} \to \mathbb{R}$ is defined by

$$\begin{aligned} b_{\mathsf{st}}(\widetilde{Y}_1, Y_2) := \\ \int_0^T \Big( \langle \mathcal{E}_{\mathsf{s}} \widetilde{U} + \mathcal{D}_{\mathsf{s}}^* \widetilde{P}_{\mathrm{tot}}, V \rangle_{\mathbb{U}_{\mathsf{s}}} + \langle \mathrm{d}_{\mathsf{t}}(\widetilde{M}, \widetilde{M}_0) + \mathcal{L}_{\mathsf{s}} \widetilde{P}, N \rangle_{\mathbb{P}_{\mathsf{s}}} \Big) + \langle \widetilde{M}_0, N_0 \rangle_{\mathbb{P}_{\mathsf{s}}} \\ + \int_0^T \Big( (\lambda \mathcal{D}_{\mathsf{s}} \widetilde{U} - \widetilde{P}_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}_{\mathsf{s}}} \widetilde{P}, Q_{\mathrm{tot}})_\Omega + (\alpha \mathcal{P}_{\mathbb{P}_{\mathsf{s}}} \mathcal{D}_{\mathsf{s}} \widetilde{U} + \sigma \widetilde{P} - \widetilde{M}, Q)_\Omega \Big) \end{aligned} \tag{3.16}$$

for $\widetilde{Y}_1 = (\widetilde{U}, \widetilde{P}_{\mathrm{tot}}, \widetilde{P}, \widetilde{M}, \widetilde{M}_0) \in \mathbb{Y}_{1,\mathsf{st}}$ and $Y_2 = (V, Q_{\mathrm{tot}}, Q, N, N_0) \in \mathbb{Y}_{2,\mathsf{st}}$.

A remarkable difference between (2.13) and (3.15) is that, in the latter one, we do not have to explicitly take the closure of the trial space. Indeed, $\mathbb{Y}_{1,\mathsf{st}}$ is certainly closed, being finite-dimensional.

3.4. **Well-posedness.** The goal of this section is to establish the well-posedness of the discretization (3.14) by means of the inf-sup theory. Since the trial space $\mathbb{Y}_{1,\mathsf{st}}$ and the test space $\mathbb{Y}_{2,\mathsf{st}}$ are finite-dimensional with equal dimension, we can use a simplified version of the Banach-Nečas theorem, which does not require to verify the nondegeneracy of the form $b_{\mathsf{st}}$, see e.g. [14, Theorem 26.6]. Therefore, the well-posedness is equivalent to the properties verified in the next two lemmas.

**Lemma 3.7** (Boundedness)**.** *The bilinear form $b_{\mathsf{st}}$ in* (3.16) *satisfies*

$$\sup_{Y_2 \in \mathbb{Y}_{2,\mathsf{st}}} \frac{b_{\mathsf{st}}(\widetilde{Y}_1, Y_2)}{\|Y_2\|_2} \lesssim \|\widetilde{Y}_1\|_{1,\mathsf{st}}$$

*for all $\widetilde{Y}_1 \in \mathbb{Y}_{1,\mathsf{st}}$. The hidden constant depends only on the constants in the assumptions* (H1) *and* (H2) *in section 3.1.*

*Proof.* The claimed bound follows from the Cauchy-Schwarz inequality applied to each term in the definition of $b_{\mathsf{st}}$, in combination with (3.3) and with the norm equivalences in the assumptions (H1) and (H2). □

**Lemma 3.8** (Inf-sup stability)**.** *The bilinear form $b_{\mathsf{st}}$ in* (3.16) *satisfies*

$$\begin{aligned} \sup_{Y_2 \in \mathbb{Y}_{2,\mathsf{st}}} \frac{b_{\mathsf{st}}(\widetilde{Y}_1, Y_2)}{\|Y_2\|_2} \gtrsim \\ (1+T)^{-\frac{1}{2}} \left( \|\widetilde{Y}_1\|_{1,\mathsf{st}}^2 + \|\widetilde{M}\|_{L^\infty(\mathbb{P}_{\mathsf{s}}^*)}^2 + \int_0^T \left( \lambda \|\mathcal{D}_{\mathsf{s}} \widetilde{U}\|_\Omega^2 + \gamma^{-1} \|\widetilde{P}\|_\Omega^2 \right) \right)^{\frac{1}{2}}. \end{aligned}$$

*for all $\widetilde{Y}_1 = (\widetilde{U}, \widetilde{P}_{tot}, \widetilde{P}, \widetilde{M}, \widetilde{M}_0) \in \mathbb{Y}_{1,\mathsf{st}}$. The hidden constant depends only on the constants in the assumptions* (H1) *and* (H2) *in section 3.1.*

*Proof.* See section 3.5. □

The combination of the two above lemmas implies the main result of this section.

**Theorem 3.9** (Well-posedness of the full discretization)**.** *For all possible loads* $(L_u, L_p, L_0) \in \mathbb{S}^0_{\mathsf{t}}(\mathbb{U}^*_{\mathsf{s}}) \times \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}^*_{\mathsf{s}}) \times \mathbb{P}^*_{\mathsf{s}}$*, the equations* (3.14) *have a unique solution* $Y_1 = (U, P_{tot}, P, M, M_0) \in \mathbb{Y}_{1,\mathsf{st}}$*, which satisfies the two-sided stability bound*

$$\|Y_1\|^2_{1,\mathsf{st}} \eqsim \int_0^T \left( \|L_u\|^2_{\mathbb{U}^*_{\mathsf{s}}} + \|L_p\|^2_{\mathbb{P}^*_{\mathsf{s}}} \right) + \|L_0\|^2_{\mathbb{P}^*_{\mathsf{s}}}. \tag{3.17}$$

*Moreover, we have the norm equivalence*

$$\|Y_1\|^2_{1,\mathsf{st}} \eqsim \|Y_1\|^2_{1,\mathsf{st}} + \|M\|^2_{L^\infty(\mathbb{P}^*_{\mathsf{s}})} + \int_0^T \left( \lambda \|\mathcal{D}_{\mathsf{s}} U\|^2_\Omega + \gamma^{-1} \|P\|^2_\Omega \right). \tag{3.18}$$

*The hidden constants depend only on the final time* $T$ *and on the constants in the assumptions* (H1) *and* (H2) *in section 3.1.*

*Proof.* The equations (3.14) are equivalent to the linear variational problem (3.15) with load $L = (L_u, 0, 0, L_p, L_0) \in \mathbb{Y}^*_{2,\mathsf{st}}$. Then, the simplified version of the Banach-Nečas theorem for finite-dimensional linear variational problems [14, Theorem 26.6] implies existence and uniqueness of the solution, in combination with Lemmas 3.7 and 3.8. The identity $b_{\mathsf{st}}(Y_1, \cdot) = L$ in $\mathbb{Y}^*_{2,\mathsf{st}}$ further implies

$$\begin{aligned} \|Y_1\|^2_{1,\mathsf{st}} + \|M\|^2_{L^\infty(\mathbb{P}^*_{\mathsf{s}})} + \int_0^T \left( \lambda \|\mathcal{D}_{\mathsf{s}} U\|^2_\Omega + \gamma^{-1} \|P\|^2_\Omega \right) \\ \lesssim \|(L_u, 0, 0, L_p, L_0)\|^2_{2,*} \lesssim \|Y_1\|^2_{1,\mathsf{st}} \end{aligned}$$

according to estimates in Lemmas 3.7 and 3.8. The hidden constants depend only on the ones therein. Here $\|\cdot\|_{2,*}$ denotes the dual of the test norm. This readily implies the second claimed equivalence. The first one follows by recalling the definition of the norm $\|\cdot\|_2$ in (2.14). □

*Remark* 3.10 (Jumps of the discrete fluid content). Recall that the component $M$ of the solution $Y_1$ in Theorem 3.9 represents the discretization of the total fluid content $m$ in Theorem 2.4. While the latter one is guaranteed to be continuous in time, the former one is not, (indeed it is piecewise constant). Still, a careful inspection at the proofs of Lemmas 3.4 and 3.8 reveals that the left-hand side in the stability bound (3.17) controls the jumps of $M$ in time, namely

$$\sum_{j=1}^{J} \|M_j - M_{j-1}\|^2_{\mathbb{P}^*_{\mathsf{s}}} \lesssim \|Y_1\|^2_{1,\mathsf{st}}.$$

This can be viewed as a discrete counterpart of the continuity of $m$.

3.5. **Inf-sup stability.** This section is devoted to the proof of Lemma 3.8. The argument is similar to the one in [24, section 3.1], which establishes the inf-sup stability of the form $b$ in (2.15). Nevertheless, some subtleties exist with regard to the discretization and we therefore include a detailed proof so as to keep the discussion as much complete and self-contained as possible.

Let $\widetilde{Y}_1 = (\widetilde{U}, \widetilde{P}_{\mathrm{tot}}, \widetilde{P}, \widetilde{M}, \widetilde{M}_0) \in \mathbb{Y}_{1,\mathsf{st}}$ be given and set

$$\widetilde{H}_{\mathbb{D}_{\mathsf{s}}} := \lambda \mathcal{D}_{\mathsf{s}} \widetilde{U} - \widetilde{P}_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}_{\mathsf{s}}} \widetilde{P} \qquad \text{and} \qquad \widetilde{H}_{\mathbb{P}_{\mathsf{s}}} := \alpha \mathcal{P}_{\mathbb{P}_{\mathsf{s}}} \mathcal{D}_{\mathsf{s}} \widetilde{U} + \sigma \widetilde{P} - \widetilde{M}$$

for shortness. Since the projections $\mathcal{P}_{\mathbb{D}_\mathsf{s}}$ and $\mathcal{P}_{\mathbb{P}_\mathsf{s}}$ map onto $\mathbb{D}_\mathsf{s}$ and $\mathbb{P}_\mathsf{s}$, respectively, we have $\widetilde{H}_{\mathbb{D}_\mathsf{s}} \in \mathbb{S}^0_\mathsf{t}(\mathbb{D}_\mathsf{s})$ as well as $\widetilde{H}_{\mathbb{P}_\mathsf{s}} \in \mathbb{S}^0_\mathsf{t}(\mathbb{P}_\mathsf{s})$. We consider the test function

$$Y_{2,j} := \Big( (\widetilde{U} + \mathcal{E}_\mathsf{s}^{-1}\mathcal{D}_\mathsf{s}^*\widetilde{P}_{\mathrm{tot}})\chi_j,\ \frac{4\max\{1,C\}}{\mu+\lambda}\widetilde{H}_{\mathbb{D}_\mathsf{s}}\chi_j, \frac{4\gamma}{\min\{1,c\}}\widetilde{H}_{\mathbb{P}_\mathsf{s}}\chi_j, \\ \mathcal{L}_\mathsf{s}^{-1}(2\widetilde{M} + \mathrm{d}_\mathsf{t}(\widetilde{M},\widetilde{M}_0) + \mathcal{L}_\mathsf{s}\widetilde{P})\chi_j,\ \mathcal{L}_\mathsf{s}^{-1}\widetilde{M}_0 \Big)$$

with $j = 1, \dots, J$, where $\chi_j : [0,T] \to \mathbb{R}$ denotes the indicator function of the interval $[0, t_j]$. Here the constants $c$ and $C$ are as in the assumption (H2) in section 3.1. We refer to [24, Remark 3.9] for a motivation of the test function.

First of all, notice that we indeed have that $Y_{2,j} \in \mathbb{Y}_{2,\mathsf{st}}$, i.e. it is an admissible test function. Indeed, recall the definition of the spaces $\mathbb{Y}_{1,\mathsf{st}}$ and $\mathbb{Y}_{2,\mathsf{st}}$ in (3.11) and (3.13). The operator $\mathcal{E}_\mathsf{s}^{-1}\mathcal{D}_\mathsf{s}^*$ maps $\mathbb{D}_\mathsf{s}$ into $\mathbb{U}_\mathsf{s}$ and $\mathcal{L}_\mathsf{s}$ is an isometry between $\mathbb{P}_\mathsf{s}$ and $\mathbb{P}_\mathsf{s}^*$. Moreover, the indicator function $\chi_j$ is piecewise constant on the partition (3.8) of $[0,T]$. Hence, the multiplication by $\chi_j$ preserves the inclusion in $\mathbb{Y}_{2,\mathsf{st}}$.

The proof consists of two main steps. First, we establish the lower bound

$$b_\mathsf{st}(\widetilde{Y}_1, Y_{2,j} + Y_{2,J}) \gtrsim \|\widetilde{Y}_1\|^2_{1,\mathsf{st}} + \|\widetilde{M}\|^2_{L^\infty(\mathbb{P}_\mathsf{s}^*)} + \int_0^T \Big(\lambda\|\mathcal{D}_\mathsf{s}\widetilde{U}\|^2_\Omega + \gamma^{-1}\|\widetilde{P}\|^2_\Omega\Big) \tag{3.19}$$

for a suitable index $1 \le j \le J$. Then, we estimate the norm of the test function

$$\|Y_{2,j}\|_2^2 \lesssim \|\widetilde{Y}_1\|^2_{1,\mathsf{st}} + (1+T)\|\widetilde{M}\|^2_{L^\infty(\mathbb{P}_\mathsf{s}^*)} \tag{3.20}$$

for some hidden constant independent of $j$. The combination of these inequalities readily implies the inf-sup stability claimed in Lemma 3.8.

We start with the proof of (3.19). Using the test function $Y_{2,j}$ in the definition (3.16) of the form $b_\mathsf{st}$ yields

$$\begin{aligned} b_\mathsf{st}(\widetilde{Y}_1, Y_{2,j}) &= \int_0^{t_j} \langle \mathcal{E}_\mathsf{s}\widetilde{U} + \mathcal{D}_\mathsf{s}^*\widetilde{P}_{\mathrm{tot}}, \widetilde{U} + \mathcal{E}_\mathsf{s}^{-1}\mathcal{D}_\mathsf{s}^*\widetilde{P}_{\mathrm{tot}}\rangle_{\mathbb{U}_\mathsf{s}} \qquad (=: \mathfrak{I}_1) \\ &\quad + 2\int_0^{t_j} \langle \mathrm{d}_\mathsf{t}(\widetilde{M},\widetilde{M}_0) + \mathcal{L}_\mathsf{s}\widetilde{P}, \mathcal{L}_\mathsf{s}^{-1}\widetilde{M}\rangle_{\mathbb{P}_\mathsf{s}} \qquad (=: \mathfrak{I}_2) \\ &\quad + \|\widetilde{M}_0\|^2_{\mathbb{P}_\mathsf{s}^*} + \int_0^{t_j} \|\mathrm{d}_\mathsf{t}(\widetilde{M},\widetilde{M}_0) + \mathcal{L}_\mathsf{s}\widetilde{P}\|^2_{\mathbb{P}_\mathsf{s}^*} \\ &\quad + \int_0^{t_j} \frac{4\max\{1,C\}}{\mu+\lambda}\|\widetilde{H}_{\mathbb{D}_\mathsf{s}}\|^2_\Omega \\ &\quad + \int_0^{t_j} \frac{4\gamma}{\min\{1,c\}}\|\widetilde{H}_{\mathbb{P}_\mathsf{s}}\|^2_\Omega. \end{aligned} \tag{3.21}$$

Notice that the dual norms on the third line are obtained thanks to the second part of (3.4). We are led to analyze the terms $\mathfrak{I}_1$ and $\mathfrak{I}_2$ on the right-hand side.

Regarding the first term, we have

$$\mathfrak{I}_1 = \int_0^{t_j} \Big( \langle \mathcal{E}_\mathsf{s}\widetilde{U}, \widetilde{U}\rangle_{\mathbb{U}_\mathsf{s}} + 2\langle \mathcal{D}_\mathsf{s}^*\widetilde{P}_{\mathrm{tot}}, \widetilde{U}\rangle_{\mathbb{U}_\mathsf{s}} + \langle \mathcal{D}_\mathsf{s}^*\widetilde{P}, \mathcal{E}_\mathsf{s}^{-1}\mathcal{D}_\mathsf{s}^*\widetilde{P}_{\mathrm{tot}}\rangle_{\mathbb{U}_\mathsf{s}} \Big).$$

We rewrite the first summand with the help of the first identity in (3.3). Analogously, we estimate the third summand from below by means of the first identity in (3.4) and the assumption (H2) in section 3.1

$$\langle \mathcal{E}_\mathsf{s}\widetilde{U}, \widetilde{U}\rangle_{\mathbb{U}_\mathsf{s}} = \|\widetilde{U}\|^2_{\mathbb{U}_\mathsf{s}} \quad \text{and} \quad \langle \mathcal{D}_\mathsf{s}^*\widetilde{P}_{\mathrm{tot}}, \mathcal{E}_\mathsf{s}^{-1}\mathcal{D}_\mathsf{s}^*\widetilde{P}_{\mathrm{tot}}\rangle_{\mathbb{U}_\mathsf{s}} \ge \frac{c}{\mu}\|\widetilde{P}_{\mathrm{tot}}\|^2_\Omega.$$

We estimate the remaining term by observing that the assumption (H2) implies also $\mu\|\mathcal{D}_{\mathsf{s}}\cdot\|_\Omega^2 \le C\|\cdot\|_{\mathbb{U}_{\mathsf{s}}}^2$ in $\mathbb{U}_{\mathsf{s}}$. This bound and a Young's inequality imply

$$\begin{aligned}\langle \mathcal{D}_{\mathsf{s}}^*\widetilde{P}_{\mathrm{tot}},\widetilde{U}\rangle_{\mathbb{U}_{\mathsf{s}}} &= -(\mathcal{D}_{\mathsf{s}}\widetilde{U},\widetilde{H}_{\mathbb{D}_{\mathsf{s}}})_\Omega + \lambda\|\mathcal{D}_{\mathsf{s}}\widetilde{U}\|_\Omega^2 - \alpha(\mathcal{D}_{\mathsf{s}}\widetilde{U},\mathcal{P}_{\mathbb{D}_{\mathsf{s}}}\widetilde{P})_\Omega\\ &\ge \frac{3\lambda}{4}\|\mathcal{D}_{\mathsf{s}}\widetilde{U}\|_\Omega^2 - \frac{1}{4}\|\widetilde{U}\|_{\mathbb{U}_{\mathsf{s}}}^2 - \frac{\max\{1,C\}}{\mu+\lambda}\|\widetilde{H}_{\mathbb{D}_{\mathsf{s}}}\|_\Omega^2 - \alpha(\mathcal{D}_{\mathsf{s}}\widetilde{U},\widetilde{P})_\Omega.\end{aligned}$$

Notice that we were able to omit the projection $\mathcal{P}_{\mathbb{D}_{\mathsf{s}}}$ in the last term thanks to the inclusion $\mathcal{D}_{\mathsf{s}}\widetilde{U}\in\mathbb{S}_{\mathsf{t}}^0(\mathbb{D}_{\mathsf{s}})$. Combining this bound with the identities above reveals

$$\begin{aligned}\mathfrak{J}_1 \ge \int_0^{t_j}\Big(&\frac{1}{2}\|\widetilde{U}\|_{\mathbb{U}_{\mathsf{s}}}^2 + \frac{c}{\mu}\|\widetilde{P}_{\mathrm{tot}}\|_\Omega^2 + \frac{3\lambda}{2}\|\mathcal{D}_{\mathsf{s}}\widetilde{U}\|_\Omega^2\\ &- \frac{2\max\{1,C\}}{\mu+\lambda}\|\widetilde{H}_{\mathbb{D}_{\mathsf{s}}}\|_\Omega^2 - 2\alpha(\mathcal{D}_{\mathsf{s}}\widetilde{U},\widetilde{P})_\Omega\Big).\end{aligned}$$

Regarding the other critical term in (3.21), we have

$$\mathfrak{J}_2 = 2\int_0^{t_j}\Big(\langle \mathrm{d}_{\mathsf{t}}(\widetilde{M},\widetilde{M}_0),\mathcal{L}_{\mathsf{s}}^{-1}\widetilde{M}\rangle_{\mathbb{P}_{\mathsf{s}}} + (\widetilde{M},\widetilde{P})_\Omega\Big).$$

The use of the $L^2(\Omega)$-scalar product in the second summand is justified by the discussion after (3.7). We estimate the first summand from below by invoking Lemma 3.4

$$2\int_0^{t_j}\langle \mathrm{d}_{\mathsf{t}}(\widetilde{M},\widetilde{M}_0),\mathcal{L}_{\mathsf{s}}^{-1}\widetilde{M}\rangle_{\mathbb{P}_{\mathsf{s}}} \ge \|\widetilde{M}_j\|_{\mathbb{P}_{\mathsf{s}}^*}^2 - \|\widetilde{M}_0\|_{\mathbb{P}_{\mathsf{s}}^*}^2.$$

To investigate the second summand, assume first $\sigma>0$. A Young's inequality reveals

$$\begin{aligned}(\widetilde{M},\widetilde{P})_\Omega &= (\widetilde{H}_{\mathbb{P}_{\mathsf{s}}},\widetilde{P})_\Omega + \alpha(\mathcal{P}_{\mathbb{P}_{\mathsf{s}}}\mathcal{D}_{\mathsf{s}}\widetilde{U},\widetilde{P})_\Omega + \sigma\|\widetilde{P}\|_\Omega^2\\ &\ge \alpha(\mathcal{D}_{\mathsf{s}}\widetilde{U},\widetilde{P})_\Omega + \frac{\sigma}{2}\|\widetilde{P}\|_\Omega^2 - \frac{1}{2\sigma}\|\widetilde{H}_{\mathbb{P}_{\mathsf{s}}}\|_\Omega^2.\end{aligned}$$

As before, we omit the projection $\mathcal{P}_{\mathbb{P}_{\mathsf{s}}}$ thanks to the inclusion $\widetilde{P}\in\mathbb{S}_{\mathsf{t}}^0(\mathbb{P}_{\mathsf{s}})$. Alternatively, for general $\sigma\ge 0$, it holds that

$$\begin{aligned}(\widetilde{M},\widetilde{P})_\Omega &= (\widetilde{H}_{\mathbb{P}_{\mathsf{s}}},\widetilde{P})_\Omega + \alpha(\mathcal{P}_{\mathbb{P}_{\mathsf{s}}}\mathcal{D}_{\mathsf{s}}\widetilde{U},\widetilde{P})_\Omega + \sigma\|\widetilde{P}\|_\Omega^2\\ &= -\frac{1}{\alpha}(\widetilde{H}_{\mathbb{P}_{\mathsf{s}}},\widetilde{H}_{\mathbb{D}_{\mathsf{s}}})_\Omega + \frac{\lambda}{\alpha}(\widetilde{H}_{\mathbb{P}_{\mathsf{s}}},\mathcal{D}_{\mathsf{s}}\widetilde{U})_\Omega - \frac{1}{\alpha}(\widetilde{H}_{\mathbb{P}_{\mathsf{s}}},\widetilde{P}_{\mathrm{tot}})_\Omega\\ &\quad + (\widetilde{H}_{\mathbb{P}_{\mathsf{s}}},\widetilde{P}-\mathcal{P}_{\mathbb{D}_{\mathsf{s}}}\widetilde{P})_\Omega + \alpha(\mathcal{D}_{\mathsf{s}}\widetilde{U},\widetilde{P})_\Omega + \sigma\|\widetilde{P}\|_\Omega^2.\end{aligned}$$

Bounding the term $(\widetilde{H}_{\mathbb{P}_{\mathsf{s}}},\widetilde{P}-\mathcal{P}_{\mathbb{D}_{\mathsf{s}}}\widetilde{P})_\Omega$ conveniently is subtle. Whenever $\mathbb{P}_{\mathsf{s}}\subseteq\mathbb{D}_{\mathsf{s}}$, we have $\mathcal{P}_{\mathbb{D}_{\mathsf{s}}}\widetilde{P}=\widetilde{P}$, hence $(\widetilde{H}_{\mathbb{P}_{\mathsf{s}}},\widetilde{P}-\mathcal{P}_{\mathbb{D}_{\mathsf{s}}}\widetilde{P})_\Omega=0$. When the above inclusion fails, we have $\sigma>0$ due to the assumption (H4) in section 3.1. Then, we apply Young's inequality to obtain $(\widetilde{H}_{\mathbb{P}_{\mathsf{s}}},\widetilde{P}-\mathcal{P}_{\mathbb{D}_{\mathsf{s}}}\widetilde{P})_\Omega \le \|\widetilde{H}_{\mathbb{P}_{\mathsf{s}}}\|_\Omega^2/(2\sigma)+\sigma\|\widetilde{P}\|_\Omega^2/2$. Combining this observation with other applications of Young's inequality, the previous lower bound and the definition (2.12) of $\gamma$, we arrive at

$$\begin{aligned}(\widetilde{M},\widetilde{P})_\Omega \ge &-\frac{\max\{1,C\}}{2(\mu+\lambda)}\|\widetilde{H}_{\mathbb{D}_{\mathsf{s}}}\|_\Omega^2 - \frac{c}{4\mu}\|\widetilde{P}_{\mathrm{tot}}\|_\Omega^2 - \frac{\lambda}{2}\|\mathcal{D}_{\mathsf{s}}\widetilde{U}\|_\Omega^2\\ &+\frac{\sigma}{2}\|\widetilde{P}\|_\Omega^2 - \frac{3\gamma}{2\min\{1,c\}}\|\widetilde{H}_{\mathbb{P}_{\mathsf{s}}}\|_\Omega^2 + \alpha(\mathcal{D}_{\mathsf{s}}\widetilde{U},\widetilde{P})_\Omega.\end{aligned}$$

By combining the last two bounds with the above identity for $\mathfrak{I}_2$, we obtain

$$\begin{aligned}\mathfrak{I}_2 \geq{}& \|\widetilde{M}_j\|_{\mathbb{P}_\mathsf{s}^*}^2 - \|\widetilde{M}_0\|_{\mathbb{P}_\mathsf{s}^*}^2 \\ &+ \int_0^{t_j} \Big( -\frac{\max\{1,C\}}{\mu+\lambda}\|\widetilde{H}_{\mathbb{D}_\mathsf{s}}\|_\Omega^2 - \frac{c}{2\mu}\|\widetilde{P}_{\mathrm{tot}}\|_\Omega^2 - \frac{\lambda}{2}\|\mathcal{D}_\mathsf{s}\widetilde{U}\|_\Omega^2 + \sigma\|\widetilde{P}\|_\Omega^2 \\ &\qquad - \frac{3\gamma}{\min\{1,c\}}\|\widetilde{H}_{\mathbb{P}_\mathsf{s}}\|_\Omega^2 + 2\alpha(\mathcal{D}_\mathsf{s}\widetilde{U},\widetilde{P})_\Omega \Big).\end{aligned}$$

After this preparation, we are in position to choose a specific test function in order to establish (3.19). We let $\overline{j} \in \{1,\dots,J\}$ be such that $\|\widetilde{M}_{\overline{j}}\|_{\mathbb{P}_\mathsf{s}^*} = \|\widetilde{M}\|_{L^\infty(\mathbb{P}_\mathsf{s}^*)}$. Then, we insert the previous lower bounds of $\mathfrak{I}_1$ and $\mathfrak{I}_2$ into (3.21), resulting in

$$\begin{aligned}&b_{\mathsf{st}}(\widetilde{Y}_1, Y_{2,\overline{j}} + Y_{2,J}) \geq \\ &\quad \int_0^T \Big( \frac{1}{2}\|\widetilde{U}\|_{\mathbb{U}_\mathsf{s}}^2 + \frac{c}{2\mu}\|\widetilde{P}_{\mathrm{tot}}\|_\Omega^2 + \|\mathrm{d}_\mathsf{t}(\widetilde{M},\widetilde{M}_0) + \mathcal{L}_\mathsf{s}\widetilde{P}\|_{\mathbb{P}_\mathsf{s}^*}^2 \Big) \\ &+ \int_0^T \Big( \frac{1}{\mu+\lambda}\|\widetilde{H}_{\mathbb{D}_\mathsf{s}}\|_\Omega^2 + \gamma\|\widetilde{H}_{\mathbb{P}_\mathsf{s}}\|_\Omega^2 + \frac{\lambda}{2}\|\mathcal{D}_\mathsf{s}\widetilde{U}\|_\Omega^2 + \sigma\|\widetilde{P}\|_\Omega^2 \Big) + \|\widetilde{M}\|_{L^\infty(\mathbb{P}_\mathsf{s}^*)}^2.\end{aligned}$$

In combination with the definition (3.12) of the norm $\|\cdot\|_{1,\mathsf{st}}$ and the assumption (H1) in section 3.1, this almost establishes (3.19). Indeed, we have

$$b_{\mathsf{st}}(\widetilde{Y}_1, Y_{2,\overline{j}} + Y_{2,J}) \gtrsim \|\widetilde{Y}_1\|_{1,\mathsf{st}}^2 + \|\widetilde{M}\|_{L^\infty(\mathbb{P}_\mathsf{s}^*)}^2 + \int_0^T \Big( \lambda\|\mathcal{D}_\mathsf{s}\widetilde{U}\|_\Omega^2 + \sigma\|\widetilde{P}\|_\Omega^2 \Big).$$

The proof that we can indeed replace $\sigma$ by $\gamma^{-1}$ in the last summand follows verbatim the argument in the proof of [24, Proposition 4.1].

The last step of the proof consists in establishing (3.20). According to the definition (2.14) of the test norm $\|\cdot\|_2$, we have for all $j = 1,\dots,J$, that

$$\begin{aligned}\|Y_{2,j}\|_2^2 = {}& \int_0^{t_j} \Big( \|\widetilde{U} + \mathcal{E}_\mathsf{s}^{-1}\mathcal{D}_\mathsf{s}^*\widetilde{P}_{\mathrm{tot}}\|_{\mathbb{U}}^2 + \|\mathcal{L}_\mathsf{s}^{-1}(2\widetilde{M} + \mathrm{d}_\mathsf{t}(\widetilde{M},\widetilde{M}_0) + \mathcal{L}_\mathsf{s}\widetilde{P})\|_{\mathbb{P}}^2 \Big) \\ &+ \int_0^{t_j} \Big( \frac{16\max\{1,C^2\}}{\mu+\lambda}\|\widetilde{H}_{\mathbb{D}_\mathsf{s}}\|_\Omega^2 + \frac{16\gamma}{\min\{1,c^2\}}\|\widetilde{H}_{\mathbb{P}_\mathsf{s}}\|_\Omega^2 \Big) + \|\mathcal{L}_\mathsf{s}^{-1}\widetilde{M}_0\|_{\mathbb{P}}^2.\end{aligned}$$

We exploit the identities (3.3) and (3.4) and the norm equivalences in the assumptions (H1) and (H2) from section 3.1. We extend also the integrals from the interval $[0,t_j]$ to $[0,T]$. Hence, we obtain

$$\begin{aligned}\|Y_{2,j}\|_2^2 \lesssim{}& \int_0^T \Big( \|\widetilde{U}\|_{\mathbb{U}}^2 + \frac{1}{\mu}\|\widetilde{P}_{\mathrm{tot}}\|_\Omega^2 + \|\widetilde{M}\|_{\mathbb{P}_\mathsf{s}^*}^2 + \|\mathrm{d}_\mathsf{t}(\widetilde{M},\widetilde{M}_0) + \mathcal{L}_\mathsf{s}\widetilde{P}\|_{\mathbb{P}_\mathsf{s}^*}^2 \Big) \\ &+ \int_0^T \Big( \frac{1}{\mu+\lambda}\|\widetilde{H}_{\mathbb{D}_\mathsf{s}}\|_\Omega^2 + \gamma\|\widetilde{H}_{\mathbb{P}_\mathsf{s}}\|_\Omega^2 \Big) + \|\widetilde{M}_0\|_{\mathbb{P}_\mathsf{s}^*}^2.\end{aligned}$$

Finally, we recall the definition (3.12) of the norm $\|\cdot\|_{1,\mathsf{st}}$ and exploit the upper bound $\int_0^T \|\cdot\|_{\mathbb{P}_\mathsf{s}^*}^2 \leq T\|\cdot\|_{L^\infty(\mathbb{P}_\mathsf{s}^*)}^2$. This establishes (3.20) for some hidden constant independent of $j$ and concludes the proof.

*Remark* 3.11 (Time-dependent space discretization). In our framework the space discretization does not change in time and this is important for the proof of the inf-sup stability in this section. Indeed, a change in the space discretization would imply the change of the operator $\mathcal{L}_\mathsf{s}$ in time and this would have an effect on our use of the integration by parts formula from Lemma 3.4. For the same reason,

we argued in [24, Remark 2.1] that the parameter $\kappa$ in the Biot's equations (2.1a) could be allowed to vary in space but not in time.

## 4. Discretization with Lagrange elements in space

In this section we propose an exemplary concrete space discretization, based on $H^1$-conforming Lagrange finite elements for all variables. Hence, we first verify the assumptions (H1)-(H4) from section 3.1, in order to infer the well-posedness. Then, we discuss the a priori error analysis.

Our notation and assumptions for the finite elements are as follows. We denote by $\mathfrak{T}$ a face-to-face simplicial mesh of $\Omega$. The shape constant of $\mathfrak{T}$ is given by

$$\max_{\mathsf{T}\in\mathfrak{T}} \frac{\mathsf{h}_\mathsf{T}}{\mathsf{r}_\mathsf{T}} \tag{4.1}$$

where $\mathsf{h}_\mathsf{T}$ is the diameter of a $d$-simplex $\mathsf{T}\in\mathfrak{T}$ and $\mathsf{r}_\mathsf{T}$ is the diameter of the largest ball inscribed in $\mathsf{T}$.

We denote by $\mathfrak{F}$ the set of all faces of $\mathfrak{T}$ and by $\mathfrak{F}^i$ the interior faces. We assume that the mesh is compatible with the boundary conditions (2.1c). This means that each face $\mathsf{F}\in\mathfrak{F}\setminus\mathfrak{F}^i$ (i.e. each boundary face) satisfies

$$\text{either } \mathsf{F}\subseteq\Gamma_{u,N} \text{ or } \mathsf{F}\subseteq\Gamma_{u,E} \qquad \text{and} \qquad \text{either } \mathsf{F}\subseteq\Gamma_{p,N} \text{ or } \mathsf{F}\subseteq\Gamma_{p,E}.$$

Hence, we have two partitions of the boundary faces $\mathfrak{F}_{u,E}\cup\mathfrak{F}_{u,N}$ and $\mathfrak{F}_{p,E}\cup\mathfrak{F}_{p,N}$, where each set $\mathfrak{F}_*$ is defined as

$$\mathfrak{F}_* := \{\mathsf{F}\in\mathfrak{F} \mid \mathsf{F}\subseteq\Gamma_*\}. \tag{4.2}$$

We let the meshsize $\mathsf{h}$ and the normal $\mathsf{n}$ be the piecewise constant functions on $\mathfrak{T}$ and on the skeleton of $\mathfrak{T}$ (i.e. the union of all faces) defined as

$$\mathsf{h}_{|\mathsf{T}} := \mathsf{h}_\mathsf{T} \qquad \text{and} \qquad \mathsf{n}_{|\mathsf{F}} := \mathsf{n}_\mathsf{F}$$

for all $\mathsf{T}\in\mathfrak{T}$ and $\mathsf{F}\in\mathfrak{F}$, respectively. Here $\mathsf{n}_\mathsf{F}$ is a fixed unit normal vector of $\mathsf{F}$, pointing outside $\Omega$ if $\mathsf{F}$ is a boundary face.

The space $P_\mathsf{k}(S)$, $\mathsf{k}\geq 0$, consist of all polynomials of total degree $\leq\mathsf{k}$ on an $n$-simplex $S\subseteq\mathbb{R}^d$ with $1\leq n\leq d$. The corresponding space of (possibly discontinuous) piecewise polynomials on the mesh $\mathfrak{T}$ is denoted by $P_\mathsf{k}(\mathfrak{T})$.

### 4.1. Concrete space discretization.

A concrete realization of the abstract space discretization from section 3.1 requires a specific choice of the spaces $\mathbb{U}_\mathsf{s}$ and $\mathbb{P}_\mathsf{s}$ and of the operators $\mathcal{E}_\mathsf{s}$ and $\mathcal{L}_\mathsf{s}$ in (3.1) and (3.2), respectively. We have to prescribe also the operator $\mathcal{D}_\mathsf{s}$ in (3.5) and to characterize its range $\mathbb{D}_\mathsf{s}$ in (3.6). Finally, we need to verify the assumptions (H1)-(H4).

For the discretization of the displacement, we consider $H^1$-conforming Lagrange finite elements of degree $\mathsf{k}+1$ with $\mathsf{k}\geq 1$, i.e., we set

$$\mathbb{U}_\mathsf{s} := P_{\mathsf{k}+1}(\mathfrak{T})^d\cap\mathbb{U}. \tag{4.3}$$

The intersection with the space $\mathbb{U}$ from (2.3) enforces the global continuity (hence the $H^1$-conformity) as well as the boundary conditions. As usual in conforming finite elements, we discretize the operator $\mathcal{E}$ in (2.2) by $\mathcal{E}_\mathsf{s}:\mathbb{U}_\mathsf{s}\to\mathbb{U}_\mathsf{s}^*$ defined as

$$\langle\mathcal{E}_\mathsf{s}\widetilde{U},V\rangle_{\mathbb{U}_\mathsf{s}} := \langle\mathcal{E}\widetilde{U},V\rangle_{\mathbb{U}} \tag{4.4}$$

for all $\widetilde{U},V\in\mathbb{U}_\mathsf{s}$.

Recall that the operator $\mathcal{D}_\mathsf{s}$ should be a discretization of the divergence and that the pair $\mathbb{U}_\mathsf{s}/\mathbb{D}_\mathsf{s}$ with $\mathbb{D}_\mathsf{s} = \mathcal{D}_\mathsf{s}(\mathbb{U}_\mathsf{s})$ should enjoy a discrete Stokes inf-sup condition in order to satisfy assumption (H2). Given the above definition of $\mathbb{U}_\mathsf{s}$, one option for $\mathbb{D}_\mathsf{s}$ is suggested by the so-called Hood-Taylor pair, cf. [14, section 54.4]. Hence, we consider $H^1$-conforming Lagrange finite elements of degree $\mathsf{k}$

$$\mathbb{D}_\mathsf{s} := P_\mathsf{k}(\mathfrak{T}) \cap H^1(\Omega) \cap \mathbb{D} \tag{4.5}$$

and we define $\mathcal{D}_\mathsf{s} : \mathbb{U}_\mathsf{s} \to \mathbb{D}_\mathsf{s}$ by

$$(\mathcal{D}_\mathsf{s}\widetilde{U}, Q_{\mathrm{tot}})_\Omega = (\mathcal{D}\widetilde{U}, Q_{\mathrm{tot}})_\Omega \tag{4.6}$$

for all $\widetilde{U} \in \mathbb{U}_\mathsf{s}$ and $Q_{\mathrm{tot}} \in \mathbb{D}_\mathsf{s}$. According to (2.5), the intersection with $\mathbb{D}$ in (4.5) simply enforces the vanishing mean value when $\Gamma_{u,E} = \partial\Omega$. Note also that $\mathcal{D}_\mathsf{s}\widetilde{U}$ is nothing else than the $\mathbb{H}$-orthogonal projection of $\mathcal{D}\widetilde{U}$ onto $\mathbb{D}_\mathsf{s}$.

Finally, the assumption (H4) suggests that also the space $\mathbb{P}_\mathsf{s}$, for the discretization of the pressure and of the total fluid content, should consist of $H^1$-conforming Lagrange finite elements of degree $\mathsf{k}$. Therefore, we set

$$\mathbb{P}_\mathsf{s} := P_\mathsf{k}(\mathfrak{T}) \cap \mathbb{P}. \tag{4.7}$$

The intersection with the space $\mathbb{P}$ from (2.4) enforces global continuity, the boundary conditions and, possibly, the vanishing mean value. In this case, we define $\mathcal{L}_\mathsf{s} : \mathbb{P}_\mathsf{s} \to \mathbb{P}_\mathsf{s}^*$ in terms of $\mathcal{L}$ in (2.2) by

$$\langle \mathcal{L}_\mathsf{s}\widetilde{P}, Q\rangle_{\mathbb{P}_\mathsf{s}} := \langle \mathcal{L}\widetilde{P}, Q\rangle_{\mathbb{P}} \tag{4.8}$$

for all $\widetilde{P}, Q \in \mathbb{P}_\mathsf{s}$.

*Remark* 4.1 (Hidden constants). In order to simplify the statement of the next results, it is implicitly understood hereafter that all hidden constants in our estimates potentially depend on the final time $T$, the shape constant (4.1) of $\mathfrak{T}$ and the polynomial degree $\mathsf{k}$. In particular, the latter is arbitrary but fixed in our setting. The possible dependence on other relevant quantities is addressed case by case.

Having introduced all the spaces and the operators required for the space discretization in section 3.1, we can verify the validity of the assumptions (H1)-(H4).

**Proposition 4.2** (Verification of the assumptions)**.** *Let the spaces $\mathbb{U}_\mathsf{s}$, $\mathbb{D}_\mathsf{s}$ and $\mathbb{P}_\mathsf{s}$ and the operators $\mathcal{E}_\mathsf{s}$, $\mathcal{D}_\mathsf{s}$ and $\mathcal{L}_\mathsf{s}$ be defined by* (4.3)-(4.8)*.*

1. *The assumption* (H1) *holds true and the equivalences therein are actually identities.*
2. *The assumption* (H2) *holds true and the constants therein depend only on the quantities mentioned in Remark 4.1, provided that each $d$-simplex in $\mathfrak{T}$ has at least one vertex in the interior of $\Omega$.*
3. *The assumption* (H3) *holds true and the constants therein depend only on the quantities mentioned in Remark 4.1, provided that the grading of $\mathfrak{T}$, defined as in* [12]*, is strictly less than $(\sqrt{2\mathsf{k}+d} + \sqrt{\mathsf{k}})/(\sqrt{2\mathsf{k}+d} - \sqrt{\mathsf{k}})$.*
4. *The assumption* (H4) *holds true.*

*Proof.* (1) Owing to the inclusions $\mathbb{U}_\mathsf{s} \subseteq \mathbb{U}$ and $\mathbb{P}_\mathsf{s} \subseteq \mathbb{P}$, there is no need to extend the norms $\|\cdot\|_\mathbb{U}$ and $\|\cdot\|_\mathbb{P}$. Moreover, the combination of (3.3) with (4.4) and (4.8) readily implies $\|\cdot\|_{\mathbb{U}_\mathsf{s}} = \|\cdot\|_\mathbb{U}$ in $\mathbb{U}_\mathsf{s}$ as well as $\|\cdot\|_{\mathbb{P}_\mathsf{s}} = \|\cdot\|_\mathbb{P}$ in $\mathbb{P}_\mathsf{s}$.

(2) The upper bound in (H2) follows from the boundedness of $\mathcal{D}_\mathsf{s}$, which satisfies $\|\mathcal{D}_\mathsf{s}\cdot\|_\Omega^2 \leq \|\mathcal{D}\cdot\|_\Omega^2 \lesssim \mu^{-1}\|\cdot\|_\mathbb{U}^2$ in $\mathbb{U}_\mathsf{s}$, according to (2.2), (2.11) and (4.6). The lower bound can be rephrased as

$$c\|\cdot\|_\Omega^2 \leq \mu\left(\sup_{V\in\mathbb{U}_\mathsf{s}}\frac{(\mathcal{D}V,\cdot)_\Omega}{\|V\|_\mathbb{U}}\right)^2$$

in $\mathbb{D}_\mathsf{s}$. The definitions (2.2) and (2.11) reveal that this is equivalent to the discrete Stokes inf-sup stability of the pair $\mathbb{U}_\mathsf{s}/\mathbb{D}_\mathsf{s}$, i.e. of the Hood-Taylor pair. The latter condition is known to hold true under the above assumption on the mesh; see [14, sections 54.3 and 54.4].

(3) We have

$$\|\mathcal{P}_{\mathbb{P}_\mathsf{s}}^*\cdot\|_{\mathbb{P}^*} = \sup_{w\in\mathbb{P}}\frac{\langle\cdot,\mathcal{P}_{\mathbb{P}_\mathsf{s}}w\rangle_{\mathbb{P}_\mathsf{s}}}{\|w\|_\mathbb{P}} = \sup_{w\in\mathbb{P}}\frac{(\cdot,\mathcal{P}_{\mathbb{P}_\mathsf{s}}w)_\Omega}{\|w\|_\mathbb{P}}$$

in $\mathbb{P}_\mathsf{s}^*$. The inclusion $\mathbb{P}_\mathsf{s}\subseteq\mathbb{P}$ implies the lower bound in (H3) with $c=1$, because $\mathcal{P}_{\mathbb{P}_\mathsf{s}}$ is a projection onto $\mathbb{P}_\mathsf{s}$. The upper bound is equivalent to the $\mathbb{P}$-stability of $\mathcal{P}_{\mathbb{P}_\mathsf{s}}$; cf. [41]. Owing to (2.11) and (4.7), this is the $H^1(\Omega)$-stability of the $L^2(\Omega)$-orthogonal projection onto $H^1$-conforming Lagrange finite element spaces. Such a condition is known to hold under the asserted grading assumption thanks to [11, Theorem 4.14(ii)] together with [11, Theorem 4.4] and [11, Remark 4.4].

(4) Let $\sigma=0$. The combination of (4.5) and (4.7) with (2.4) and (2.5) implies, for $\Gamma_{u,E}=\partial\Omega$ that

$$\mathbb{P}_\mathsf{s} = P_\mathsf{k}(\mathfrak{T})\cap H^1_{\Gamma_{p,E}}(\Omega)\cap L^2_0(\Omega) \quad\text{and}\quad \mathbb{D}_\mathsf{s} = P_\mathsf{k}(\mathfrak{T})\cap H^1(\Omega)\cap L^2_0(\Omega).$$

For $\Gamma_{u,E}\neq\partial\Omega$ we have instead,

$$\mathbb{P}_\mathsf{s} \subseteq P_\mathsf{k}(\mathfrak{T})\cap H^1_{\Gamma_{p,E}}(\Omega) \quad\text{and}\quad \mathbb{D}_\mathsf{s} = P_\mathsf{k}(\mathfrak{T})\cap H^1(\Omega).$$

In both cases we have $\mathbb{P}_\mathsf{s}\subseteq\mathbb{D}_\mathsf{s}$, i.e., assumption (H4) is verified. □

*Remark* 4.3 (Assumptions for (H3)). We refer to [12, Definitions 1.1] for the exact definition of mesh grading. Clearly the grading of quasi-uniform meshes equals 1 and thus satisfies the grading condition in Proposition 4.2(3). It follows from [12, Theorem 1.3], that the grading condition is also satisfied for all polynomial degrees $\mathsf{k}\geq 1$ and dimensions $d\leq 6$ provided $\mathfrak{T}$ is obtained by adaptive bisection of a colored initial mesh; see [12, Assumption 3.1] for the notion of colored mesh. Note that assumption (H3) can be verified under different assumptions. Indeed, the $H^1(\Omega)$-stability of the $L^2(\Omega)$-orthogonal projection onto finite element spaces has been extensively analyzed by various authors; we refer [11] for an overview of the existing results.

Having verified all assumptions in section 3.1, we deduce the well-posedness of the discretization (3.14) with the spaces and the operators proposed in this section. In particular, the inclusions $\mathbb{U}_\mathsf{s}\subseteq\mathbb{U}$ and $\mathbb{P}_\mathsf{s}\subseteq\mathbb{P}$ suggest to define the data in the discretization by restriction of the data in the weak formulation (2.8)

$$L_u = \ell_{u|\mathbb{S}^0_\mathsf{t}(\mathbb{U}_\mathsf{s})}, \qquad L_p = \ell_{p|\mathbb{S}^0_\mathsf{t}(\mathbb{P}_\mathsf{s})} \qquad\text{and}\qquad L_0 = \ell_{0|\mathbb{P}_\mathsf{s}}. \tag{4.9}$$

**Theorem 4.4** (Well-posedness with Lagrange elements)**.** *Let the spaces $\mathbb{U}_\mathsf{s}$, $\mathbb{D}_\mathsf{s}$ and $\mathbb{P}_\mathsf{s}$ and the operators $\mathcal{E}_\mathsf{s}$, $\mathcal{D}_\mathsf{s}$ and $\mathcal{L}_\mathsf{s}$ be defined by* (4.3)-(4.8)*. Assume that $\mathfrak{T}$ is obtained from a colored initial mesh by newest vertex bisection and that each $d$-simplex*

*in $\mathfrak{T}$ has at least one vertex in the interior of $\Omega$. Then, the discretization* (3.14) *with the data* (4.9) *has a unique solution $Y_1 = (U, P_{tot}, P, M, M_0) \in \mathbb{Y}_{1,\mathsf{st}}$ with*

$$
\begin{aligned}
\|Y_1\|_{1,\mathsf{st}}^2 &\eqsim \|Y_1\|_{1,\mathsf{st}}^2 + \|M\|_{L^\infty(\mathbb{P}_\mathsf{s}^*)}^2 + \int_0^T \big(\lambda \|\mathcal{D}_\mathsf{s} U\|_\Omega^2 + \gamma^{-1}\|P\|_\Omega^2\big) \\
&\eqsim \int_0^T \Big(\|L_u\|_{\mathbb{U}_\mathsf{s}^*}^2 + \|L_p\|_{\mathbb{P}_\mathsf{s}^*}^2\Big) + \|L_0\|_{\mathbb{P}_\mathsf{s}^*}^2 \\
&\lesssim \int_0^T \big(\|\ell_u\|_{\mathbb{U}^*}^2 + \|\ell_p\|_{\mathbb{P}^*}^2\big) + \|\ell_0\|_{\mathbb{P}^*}^2 .
\end{aligned}
\tag{4.10}
$$

*All hidden constants depend only on the quantities mentioned in Remark 4.1.*

*Proof.* The combination of Theorem 3.9 with Proposition 4.2 implies the existence and the uniqueness of the solution and guarantees that the two equivalences in (4.10) hold true. Then, the upper bound follows from the discretization (4.9) of the data and the inclusions $\mathbb{U}_\mathsf{s} \subseteq \mathbb{U}$ and $\mathbb{P}_\mathsf{s} \subseteq \mathbb{P}$. □

4.2. **Quasi-optimality.** In order to quantify the accuracy of the proposed discretization, we need an error notion that is related to both the trial norm $\|\cdot\|_1$ in (2.10) and to the discrete trial norm $\|\cdot\|_{1,\mathsf{st}}$ in (3.12), cf. (4.12) below. A major issue is that the former one involves the dual norm in $\mathbb{P}$, whereas the latter one involves the dual norm in $\mathbb{P}_\mathsf{s}$.

In order to compare functionals defined on the two spaces, we can use the adjoint $\mathcal{P}^* : \mathbb{P}_\mathsf{s}^* \to \mathbb{P}^*$ of a bounded projection $\mathcal{P} : \mathbb{P} \to \mathbb{P}_\mathsf{s}$. The specific choice of $\mathcal{P}$ is critical because of the term $\partial_t \widetilde{m} + \mathcal{L}\widetilde{p}$ in the norm $\|\cdot\|_1$. Since this term acts in space via the pairing $\langle\cdot,\cdot\rangle_\mathbb{P}$, one may want to employ the $L^2(\Omega)$-orthogonal projection, because the pivot space in the Hilbert triplet (2.7) is equipped with the $L^2(\Omega)$-norm. Still, according to (2.11), the action of $\mathcal{L}$ induces the $\mathbb{P}$-norm, thus suggesting to make use of the $\mathbb{P}$-orthogonal projection. Of course, similar comments apply to the counterpart of $\partial_t \widetilde{m} + \mathcal{L}\widetilde{p}$ in $\|\cdot\|_{1,\mathsf{st}}$.

Both options have advantages and disadvantages. The latter one, being related to the $\mathbb{P}$-norm, suggests the use of piecewise polynomials of degree $\mathsf{k}+1$ for the discretization of $\mathbb{P}$, but this would be critical for the assumption (H4), as $\mathbb{D}_\mathsf{s}$ consists of piecewise polynomials of degree $\mathsf{k}$, cf. (4.5)-(4.7). The former option does not have this issue, therefore we go for it. Still, the derivation of higher-order decay rates in space appears critical in this case, cf. Remark 4.12.

In accordance with the above discussion and with the inclusions $\mathbb{U}_\mathsf{s} \subseteq \mathbb{U}$ and $\mathbb{P}_\mathsf{s} \subseteq \mathbb{P}$, we consider the error notion $\mathtt{Err} : \overline{\mathbb{Y}}_1 \times \mathbb{Y}_{1,\mathsf{st}} \to [0,+\infty)$ defined as

$$
\begin{aligned}
\mathtt{ERR}(\widetilde{y}_1, \widetilde{Y}_1)^2 := &\int_0^T \left(\|u - \widetilde{U}\|_\mathbb{U}^2 + \frac{1}{\mu}\|p_{\mathrm{tot}} - \widetilde{P}_{\mathrm{tot}}\|_\Omega^2\right) \\
&+ \int_0^T \|\partial_t \widetilde{m} + \mathcal{L}\widetilde{p} - \mathcal{P}_{\mathbb{P}_\mathsf{s}}^*(\mathrm{d}_\mathsf{t}(\widetilde{M}, \widetilde{M}_0) + \mathcal{L}_\mathsf{s}\widetilde{P})\|_{\mathbb{P}^*}^2 \\
&+ \|\widetilde{m}(0) - \mathcal{P}_{\mathbb{P}_\mathsf{s}}^* \widetilde{M}_0\|_{\mathbb{P}^*}^2 \\
&+ \int_0^T \frac{1}{\lambda+\mu}\|\lambda\mathcal{D}\widetilde{u} - \widetilde{p}_{\mathrm{tot}} - \alpha\mathcal{P}_\mathbb{D}\widetilde{p} - (\lambda\mathcal{D}_\mathsf{s}\widetilde{U} - \widetilde{P}_{\mathrm{tot}} - \alpha\mathcal{P}_{\mathbb{D}_\mathsf{s}}\widetilde{P})\|_\Omega^2 \\
&+ \int_0^T \gamma\|\alpha\mathcal{P}_{\overline{\mathbb{P}}}\mathcal{D}\widetilde{u} + \sigma\widetilde{p} - \widetilde{m} - (\alpha\mathcal{P}_{\mathbb{P}_\mathsf{s}}\mathcal{D}_\mathsf{s}\widetilde{U} + \sigma\widetilde{P} - \widetilde{M})\|_\Omega^2
\end{aligned}
\tag{4.11}
$$

for $\widetilde{y}_1 = (\widetilde{u}, \widetilde{p}_{\mathrm{tot}}, \widetilde{p}, \widetilde{m}) \in \overline{\mathbb{Y}}_1$ and $\widetilde{Y}_1 = (\widetilde{U}, \widetilde{P}_{\mathrm{tot}}, \widetilde{P}, \widetilde{M}, \widetilde{M}_0) \in \mathbb{Y}_{1,\mathsf{st}}$. Notice that this is not a norm on the sum $\overline{\mathbb{Y}}_1 + \mathbb{Y}_{1,\mathsf{st}}$. For instance, in general, we have $\mathcal{D}\widetilde{U} \neq \mathcal{D}_{\mathsf{s}}\widetilde{U}$ for $\widetilde{U} \in \mathbb{S}^0_{\mathsf{t}}(\mathbb{U}_{\mathsf{s}}) \subseteq L^2(\mathbb{U})$. Still, the above error notion measures the accuracy in the approximation of all functions and functionals involved in the trial norm $\|\cdot\|_1$ and it holds that

$$\mathtt{ERR}(\widetilde{y}_1, 0) = \|\widetilde{y}_1\|_1 \qquad \text{and} \qquad \mathtt{ERR}(0, \widetilde{Y}_1) \eqsim \|\widetilde{Y}_1\|_{1,\mathsf{st}}. \tag{4.12}$$

Indeed, the second equivalence follows from (1) and (3) in Proposition 4.2.

The equations (3.14) with the spaces and the operators defined by (4.3)-(4.8) are not a conforming Petrov-Galerkin discretization of the weak formulation (2.8). Indeed, the trial space $\mathbb{Y}_{1,\mathsf{st}}$ in the former problem is not a subspace of its counterpart $\overline{\mathbb{Y}}_1$ in the latter one. Nevertheless, we have for the test spaces the inclusion

$$\mathbb{Y}_{2,\mathsf{st}} \subseteq \mathbb{Y}_2.$$

This property and the definition (4.9) of the load in (3.14) by restriction of the one in (2.8) ensure that we can still guarantee the fundamental *quasi-optimality* property of inf-sup stable conforming Petrov-Galerkin methods by a standard argument; cf. [2].

**Theorem 4.5** (Quasi-optimality)**.** *Let all assumptions in Theorem 4.4 be verified. Denote by $y_1 \in \overline{\mathbb{Y}}_1$ and $Y_1 \in \mathbb{Y}_{1,\mathsf{st}}$, respectively, the solutions of* (2.8) *and* (3.14) *with* (4.9)*. Then, we have*

$$\mathtt{ERR}(y_1, Y_1) \lesssim \inf_{\widetilde{Y}_1 \in \mathbb{Y}_{1,\mathsf{st}}} \mathtt{ERR}(y_1, \widetilde{Y}_1). \tag{4.13}$$

*The hidden constant depends only on the quantities mentioned in Remark 4.1.*

*Proof.* For $\widetilde{Y}_1 \in \mathbb{Y}_{1,\mathsf{st}}$, the triangle inequality and (4.12) imply

$$\mathtt{ERR}(y_1, Y_1) \leq \mathtt{ERR}(y_1, \widetilde{Y}_1) + \mathtt{ERR}(0, Y_1 - \widetilde{Y}_1) \lesssim \mathtt{ERR}(y_1, \widetilde{Y}_1) + \|Y_1 - \widetilde{Y}_1\|_{1,\mathsf{st}}.$$

According to the inf-sup stability in Lemma 3.8, we have

$$\|Y_1 - \widetilde{Y}_1\|_{1,\mathsf{st}} \lesssim \sup_{Y_2 \in \mathbb{Y}_{2,h}} \frac{b_{\mathsf{st}}(Y_1 - \widetilde{Y}_1, Y_2)}{\|Y_2\|_2} = \sup_{Y_2 \in \mathbb{Y}_2^h} \frac{b(y_1, Y_2) - b_{\mathsf{st}}(\widetilde{Y}_1, Y_2)}{\|Y_2\|_2},$$

where the identity follows by comparing problems (2.8) and (3.14). We exploit the definitions (2.15), (3.16), and (4.4)-(4.8), as well as the invariance of the projection $\mathcal{P}_{\mathbb{P}_{\mathsf{s}}}$ onto $\mathbb{P}_{\mathsf{s}}$. For $y_1 = (u, p_{\mathrm{tot}}, p, m)$ and $\widetilde{Y}_1 = (\widetilde{U}, \widetilde{P}_{\mathrm{tot}}, \widetilde{P}, \widetilde{M}, \widetilde{M}_0)$, this results in

$$\begin{aligned}
b(y_1, Y_2) - b_{\mathsf{st}}(\widetilde{Y}_1, Y_2) :=& \\
&\int_0^T \Big( \big\langle \mathcal{E}(u - \widetilde{U}), V \big\rangle_{\mathbb{U}} + \big(p_{\mathrm{tot}} - \widetilde{P}_{\mathrm{tot}}, \mathcal{D}V\big)_\Omega \Big) \\
&+ \int_0^T \big\langle \partial_t m + \mathcal{L}\widetilde{P} - \mathcal{P}^*_{\mathbb{P}_{\mathsf{s}}}(\mathrm{d}_{\mathsf{t}}(\widetilde{M}, \widetilde{M}_0) + \mathcal{L}_{\mathsf{s}}\widetilde{P}), N \big\rangle_{\mathbb{P}} \\
&+ \big\langle m - \mathcal{P}^*_{\mathbb{P}_{\mathsf{s}}}\widetilde{M}_0, N_0 \big\rangle_{\mathbb{P}} \\
&+ \int_0^T \big(\lambda \mathcal{D}u - p_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}} p - (\lambda \mathcal{D}_{\mathsf{s}}\widetilde{U} - \widetilde{P}_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}_{\mathsf{s}}}\widetilde{P}), Q_{\mathrm{tot}}\big)_\Omega \\
&+ \int_0^T \big(\alpha \mathcal{P}_{\overline{\mathbb{P}}} \mathcal{D}u + \sigma p - m - (\alpha \mathcal{P}_{\mathbb{P}_{\mathsf{s}}} \mathcal{D}_{\mathsf{s}}\widetilde{U} + \sigma \widetilde{P} - \widetilde{M}), Q\big)_\Omega.
\end{aligned} \tag{4.14}$$

Then Cauchy-Schwarz inequalities, the bound $\|\mathcal{D}V\|_\Omega^2 \lesssim \mu^{-1}\|V\|_{\mathbb{U}}^2$ and the definition (2.14) of the test norm $\|\cdot\|_2$ yield

$$\|Y_1 - \widetilde{Y}_1\|_{1,\mathsf{st}} \lesssim \mathtt{ERR}(y_1, \widetilde{Y}_1).$$

Inserting this estimate into the first inequality above concludes the proof. □

*Remark* 4.6 (Augmented error notion). Our definition of the error notion is in a sense minimal, because we have included only the terms that arise by applying the Cauchy-Schwarz inequality in (4.14). Still, the first equivalence in (4.10) clarifies that that the statement and the proof of Theorem 4.5 will remain unchanged if we augment $\mathtt{ERR}(\cdot,\cdot)$ by adding any of the following terms

$$\|\widetilde{m} - \mathcal{P}_{\mathbb{P}_\mathsf{s}}^*\widetilde{M}\|_{L^\infty(\mathbb{P}^*)}^2, \qquad \int_0^T \lambda\|\mathcal{D}\widetilde{u} - \mathcal{D}_\mathsf{s}\widetilde{U}\|_\Omega^2, \qquad \text{and} \qquad \int_0^T \gamma^{-1}\|\widetilde{p} - \widetilde{P}\|_\Omega^2.$$

4.3. **A priori error estimates.** The quasi-optimality stated above has two remarkable properties and (at least) one clear disadvantage. One the one hand, the estimate (4.13) holds true for any solution of the weak formulation (2.8) (i.e. no additional regularity is required) and the hidden constant is robust with respect to all material parameters. On the other hand, it is not immediate how (and even if) the error decays to zero, because the definition (4.11) of $\mathtt{ERR}(\cdot,\cdot)$ involves a nontrivial coupling of the various components of the solution.

In this section, we investigate the latter aspect. Roughly speaking, we aim at showing that the best error in the right-hand side of (4.13) is equivalent to a sum of best errors. In order to preserve the two above-mentioned properties, we make sure also that the constants in the equivalence are independent of the material parameters and that no additional regularity of the solution of (2.8) is required beyond the minimal one, namely $y_1 \in \overline{\mathbb{Y}}_1$. When such a result is available, the decay of the error to zero can be easily discussed in terms of classical results from approximation theory.

The precise statement of our main result is as follows. Notice that the last term on the right-hand side of (4.15) does not appear in our definition of the error notion, but we could equivalently include it according to Remark 4.6.

**Theorem 4.7** (Best error decoupling)**.** *Let all assumptions in Theorem 4.4 be verified and denote by $y_1 = (u, p_{tot}, p, m) \in \overline{\mathbb{Y}}_1$ the solution of* (2.8)*. Then we have that*

$$\begin{aligned}
&\inf_{\widetilde{Y}_1 \in \mathbb{Y}_{1,\mathsf{st}}} \mathtt{ERR}(y_1, \widetilde{Y}_1)^2 \lesssim \\
&\quad \inf_{\widehat{U}\in\mathbb{S}_\mathsf{t}^0(\mathbb{U}_\mathsf{s})} \int_0^T \|u - \widehat{U}\|_{\mathbb{U}}^2 + \inf_{\widehat{P}_{tot}\in\mathbb{S}_\mathsf{t}^0(\mathbb{D})} \int_0^T \frac{1}{\mu}\|p_{tot} - \widehat{P}_{tot}\|_\Omega^2 \\
&\quad + \inf_{\widehat{W}\in\mathbb{S}_\mathsf{t}^0(\mathbb{P}_\mathsf{s})} \int_0^T \|\partial_t m + \mathcal{L}p - \mathcal{P}_{\mathbb{P}_\mathsf{s}}^*\widehat{W}\|_{\mathbb{P}^*}^2 + \inf_{\widehat{M}_0\in\mathbb{P}_\mathsf{s}} \|m(0) - \mathcal{P}_{\mathbb{P}_\mathsf{s}}^*\widehat{M}_0\|_{\mathbb{P}^*}^2 \\
&\quad + \varepsilon_\mathsf{s}^{-1} \inf_{\widehat{P}\in\mathbb{S}_\mathsf{t}^0(\mathbb{P}_\mathsf{s})} \int_0^T \frac{1}{\gamma}\|p - \widehat{P}\|_\Omega^2.
\end{aligned} \tag{4.15}$$

*The hidden constant depends on the quantities mentioned in Remark 4.1 and on the constant in* (4.20) *and $\varepsilon_\mathsf{s}$ is defined in* (4.24) *below.*

*Proof.* See sections 4.4-4.6. □

A first consequence of Theorem 4.7 is plain convergence: the error converges to zero as the mesh-size converges to zero, both in space and time. This holds true irrespective of the regularity of the solution. Moreover, first-order convergence can be established under additional regularity assumptions.

**Corollary 4.8** (First-order convergence)**.** *Let all assumptions in Theorem 4.4 be verified. Denote by* $y_1 = (u, p_{tot}, p, m) \in \overline{\mathbb{Y}}_1$ *and* $Y_1 \in \mathbb{Y}_{1,\mathsf{st}}$*, respectively, the solutions of* (2.8) *and* (3.14) *with* (4.9)*. Assume additionally*

$$\begin{aligned} &u \in H^1(\mathbb{U}) \cap L^2(H^2(\Omega)^d) && p_{tot} \in H^1(\mathbb{D}) \cap L^2(H^1(\Omega)) \\ &p \in H^1(\overline{\mathbb{P}}) \cap L^2(H^1(\Omega)) && (\partial_t m + \mathcal{L}p) \in H^1(\mathbb{P}^*) \cap L^2(L^2(\Omega)) \\ &m(0) \in L^2(\Omega). \end{aligned} \tag{4.16}$$

*Then, the error can be bounded from above as follows*

$$\begin{aligned} \mathit{ERR}(y_1, Y_1)^2 \lesssim \Big(\max_{\Omega} \mathsf{h}\Big)^2 \Bigg\{ &\frac{1}{\kappa}\|m(0)\|_\Omega^2 \\ &+ \int_0^T \left(\mu\|\nabla^2 u\|_\Omega^2 + \frac{1}{\mu}\|\nabla p_{tot}\|_\Omega^2 + \frac{1}{\varepsilon_{\mathsf{s}}\gamma}\|\nabla p\|_\Omega^2 + \frac{1}{\kappa}\|\partial_t m + \mathcal{L}p\|_\Omega^2\right) \Bigg\} \\ &+ \Big(\max_{j=1,\dots,J} |I_j|\Big)^2 \int_0^T \left(\|\partial_t u\|_{\mathbb{U}}^2 + \frac{1}{\mu}\|\partial_t p_{tot}\|_\Omega^2 + \frac{1}{\varepsilon_{\mathsf{s}}\gamma}\|\partial_t p\|_\Omega^2 + \|\partial_t(\partial_t m + \mathcal{L}p)\|_{\mathbb{P}^*}^2\right). \end{aligned}$$

*The hidden constant depends on the quantities mentioned in Remark 4.1 and on the constant in* (4.20) *and* $\varepsilon_{\mathsf{s}}$ *is defined in* (4.24) *below.*

*Proof.* Combine Theorem 4.5 with Theorem 4.7. Then, use standard Bramble-Hilbert-like estimates (see, e.g., [40, Chapter 7]) in order to bound each term on the right-hand side of (4.15). □

*Remark* 4.9 (Regularity in space). According to [24, Theorem 5.4], the solution $y_1$ of (2.8) satisfies the space regularity in (4.16) at least in some circumstances. To be more precise, we have

$$\begin{aligned} &u \in L^2(H^2(\Omega)^2) && p_{\mathrm{tot}} \in L^2(H^1(\Omega)) \\ &p \in L^2(H^1(\Omega)) && (\partial_t m + \mathcal{L}p) \in L^2(L^2(\Omega)) \end{aligned}$$

when the data are such that

$$\ell_u \in L^2(L^2(\Omega)^2), \qquad \ell_p \in L^2(L^2(\Omega)), \qquad \ell_0 \in L^2(\Omega),$$

upon additionally assuming that $\Omega \subseteq \mathbb{R}^2$ is convex, the boundary conditions (2.1c) are posed on $\Gamma_{u,E} = \partial\Omega = \Gamma_{p,N}$, and $\lambda \gtrsim \mu$.

*Remark* 4.10 (Regularity in time). According to [31], the time regularity in (4.16) may fail to hold even for smooth data. For a rough explanation, assume the data are smooth in time, the solution enjoys the time regularity in (4.16) and, in addition, we have $m \in C^1(\mathbb{P}^*)$, i.e. we have one more time derivative than in Theorem 2.4. Then, on the one-hand, the fourth equation in (2.8) implies $p \in C^0(\mathbb{P})$. On the other hand, the evaluation of the other equations at $t = 0$ implies that the pair $(u(0), p(0))$ solves the Stokes-like problem

$$\begin{aligned} (\mathcal{E} + \lambda\mathcal{D}^*\mathcal{D})u(0) + \alpha\mathcal{D}^*\mathcal{P}_{\mathbb{D}}p(0) &= f_u(0) \quad \text{in } \mathbb{U}^* \\ \alpha\mathcal{P}_{\overline{\mathbb{P}}}\mathrm{div}u(0) + \sigma p(0) &= \ell_0 \qquad \text{in } \overline{\mathbb{P}}. \end{aligned}$$

In general, the solution of this problem is in $\mathbb{U}\times\overline{\mathbb{P}}$. Hence the condition $p(0)\in\mathbb{P}$ can hold true only for compatible data, otherwise some singularity must be expected at the initial time. We refer to the Terzaghi test case discussed in section 5 below for a numerical illustration.

*Remark* 4.11 (Grading of the time partition). The constant in Theorem 4.7 depends on the one in (4.20), which measures the grading of the partition employed for the discretization in time. The latter constants enters into play because Theorem 4.7 does not assume additional regularity in time of the solution. This calls for a Scott-Zhang-like interpolation in time, see section 4.4 below for the details. When all components of the solution are continuous in time, a Lagrange-like interpolation is possible and the constant in (4.20) does not enter into the error estimation, cf. [40, Theorem 4.5]. Still, according to Remark 4.10, the continuity at $t=0$ is not obvious.

*Remark* 4.12 (Decay rate). The space discretization considered in this section is actually of higher-order, because we make use of $H^1$-conforming Lagrange finite elements of degree $\mathsf{k}+1$ for the displacement and of degree $\mathsf{k}$ for the other components of the solution with $\mathsf{k}\geq 1$, cf. (4.3)-(4.7). Therefore, the question arises if a higher-order decay rate with respect to $\mathsf{h}$ can be obtained under higher regularity assumptions on the solution. On the one hand, the $\mathbb{P}^*$-norm errors on the right-hand side of (4.15) appear to be critical in this respect, because the space $\mathbb{P}_{\mathsf{s}}$, possibly incorporating boundary conditions, is used to approximate functional from $\mathbb{P}^*$, which have no prescribed boundary conditions. On the other hand, it is unclear to us if the above-mentioned regularity of the solution can be expected in general. Indeed, the regularity result mentioned in Remark 4.9, assumes $\Gamma_{p,E}=\emptyset$. Due to the subtlety of this matter, we postpone any further investigation on this point to future work.

The remaining part of this section is devoted to the proof of Theorem 4.7. For this purpose, we aim at constructing a bounded interpolant $\mathcal{I}:\overline{\mathbb{Y}}_1\to\mathbb{Y}_{1,\mathsf{st}}$. The operator $\mathcal{I}$ cannot be obtained by just approximating each component of the solution irrespective of the others, because of the nontrivial coupling of the components in the error notion (4.11). Thus, to make sure that $\mathcal{I}$ is bounded and robust with respect to the material parameters, we must guarantee that it is compatible with the coupling. Roughly speaking this means that, when the error notion involves some combination of the components of the solution, it must be possible to accurately approximate it by the corresponding combination of the components of the interpolant. We achieve this goal with the help of a number of commutative diagrams, cf. Figures 4.1 and 4.2.

We divide our construction into three parts. In Section 4.4 we first discuss the interpolation in time. In Section 4.5 we discuss the interpolation in space. Finally, in Section 4.6, we combine the interpolation in time and space in order to prove Theorem 4.7.

4.4. **Time interpolation.** The error notion $\mathtt{ERR}(\cdot,\cdot)$ involves several terms with $L^2$ regularity in time, so we initially consider the problem of approximating function from $L^2(\mathbb{X})$ into $\mathbb{S}^0_{\mathsf{t}}(\mathbb{X})$ for some Hilber space $\mathbb{X}$. Since we cannot control the point values of a function in $L^2(\mathbb{X})$, we cannot use Lagrange interpolation; cf. remark 4.11. Thus, we resort to Scott-Zhang-like [37] interpolation in the vein of [40, Section 4].

Recall (3.8)-(3.9). For $j \in \{1, \dots, J\}$, the polynomial $\psi_j \in P_1(I_j)$ defined as

$$\psi_j(t) := 6\frac{t - t_{j-1}}{|I_j|^2} - \frac{2}{|I_j|} \tag{4.17}$$

is such that,

$$\int_{I_j} Q\psi_j = Q(t_j) \qquad \forall Q \in P_1(I_j). \tag{4.18}$$

We define $\mathcal{J} : L^2(\mathbb{X}) \to \mathbb{S}^0_{\mathfrak{t}}(\mathbb{X})$ by

$$(\mathcal{J}x)_{|I_j} := \int_{I_j} x\psi_j$$

for $x \in L^2(\mathbb{X})$ and $j = 1, \dots J$.

Since the error notion in (4.11) involves the discrete time derivative $\mathrm{d}_{\mathfrak{t}}$, it is useful to introduce another interpolant $\widetilde{\mathcal{J}} : L^2(\mathbb{X}) \to \mathbb{S}^0_{\mathfrak{t}}(\mathbb{X})$, defined as

$$(\widetilde{\mathcal{J}}x)_{|I_j} := \frac{\displaystyle\int_{I_j} \left( \int_{t_{j-1}}^{t} x(s)\,\mathrm{d}s \right) \psi_j(t)\,\mathrm{d}t + \int_{I_{j-1}} \left( \int_{t}^{t_{j-1}} x(s)\,\mathrm{d}s \right) \psi_{j-1}(t)\,\mathrm{d}t}{|I_j|}$$

for $j = 2, \dots J$ and

$$(\widetilde{\mathcal{J}}x)_{|I_1} := \frac{\displaystyle\int_{I_1} \left( \int_{0}^{t} x(s)\mathrm{d}s \right) \psi_1(t)}{|I_1|}.$$

The first part of Lemma 4.13 below ensures that both $\mathcal{J}x$ and $\widetilde{\mathcal{J}}x$ are near best approximations of $x$ in the $L^2(\mathbb{X})$-norm. The second part additionally clarifies that the two interpolants are related, through the weak and the discrete time derivative, as summarized by the commuting diagram in Figure 4.1.

**Lemma 4.13** (Time interpolation)**.** *The operators $\mathcal{J}$ and $\widetilde{\mathcal{J}}$ defined above are such that*

$$\int_0^T \|x - \mathcal{J}x\|_{\mathbb{X}}^2 \le 4 \inf_{X \in \mathbb{S}^0_{\mathfrak{t}}(\mathbb{X})} \int_0^T \|x - X\|_{\mathbb{X}}^2 \tag{4.19a}$$

$$\int_0^T \|x - \widetilde{\mathcal{J}}x\|_{\mathbb{X}}^2 \lesssim \inf_{X \in \mathbb{S}^0_{\mathfrak{t}}(\mathbb{X})} \int_0^T \|x - X\|_{\mathbb{X}}^2 \tag{4.19b}$$

*for all $x \in L^2(\mathbb{X})$. Moreover, for $x \in H^1(\mathbb{X})$, we have*

$$\mathrm{d}_{\mathfrak{t}}(\mathcal{J}x, x(0)) = \widetilde{\mathcal{J}}\partial_t x. \tag{4.19c}$$

*The hidden constant in* (4.19b) *is an increasing function of*

$$\max_{j=2,\dots,J} \frac{|I_{j-1}|}{|I_j|}. \tag{4.20}$$

*Proof.* According to [40, Remark 4.1], the operator $\mathcal{J}$ is invariant on $\mathbb{S}^0_{\mathfrak{t}}(\mathbb{X})$ and it is bounded with

$$\int_0^T \|\mathcal{J}x\|_{\mathbb{X}}^2 \le 4 \int_0^T \|x\|_{\mathbb{X}}^2$$

for all $x \in L^2(\mathbb{X})$. The combination of the two properties implies (4.19a).

Invariance of $\widetilde{\mathcal{J}}$ on $\mathbb{S}^0_{\mathsf{t}}(\mathbb{X})$ follows by elementary calculations employing (4.18). In order to prove stability for $\widetilde{\mathcal{J}}$, we observe for $x \in L^2(\mathbb{X})$ and $j \geq 2$ that

$$\begin{aligned}\|(\widetilde{\mathcal{J}}x)_{|I_j}\|_{\mathbb{X}} &\leq \frac{1}{|I_j|}\int_{I_j}\|x\|_{\mathbb{X}}\int_{I_j}|\psi_j| + \frac{1}{|I_j|}\int_{I_{j-1}}\|x\|_{\mathbb{X}}\int_{I_{j-1}}|\psi_{j-1}| \\ &\leq \left(\int_{I_j}\|x\|^2_{\mathbb{X}}\right)^{\frac{1}{2}}\left(\int_{I_j}|\psi_j|^2\right)^{\frac{1}{2}} + \frac{|I_{j-1}|}{|I_j|}\left(\int_{I_{j-1}}\|x\|^2_{\mathbb{X}}\right)^{\frac{1}{2}}\left(\int_{I_{j-1}}|\psi_{j-1}|^2\right)^{\frac{1}{2}}.\end{aligned}$$

By recalling (4.17)-(4.18), we arrive at

$$\int_{I_j}\|\widetilde{\mathcal{J}}x\|^2_{\mathbb{X}} \leq 4\left(1+\frac{|I_{j-1}|}{|I_j|}\right)\int_{I_{j-1}\cup I_j}\|x\|^2_{\mathbb{X}}.$$

The same argument applies for $j = 1$. Summing over $j = 1, \dots, J$ proves (4.19b).

Finally (4.19c) follows from the fundamental theorem of calculus and (4.18). □

$$\begin{array}{ccc} H^1(\mathbb{X}) & \xrightarrow{\quad\partial_t\quad} & L^2(\mathbb{X}) \\ \Big\downarrow{\scriptstyle (\,\mathcal{J},\,(\cdot)(0)\,)} & & \Big\downarrow{\scriptstyle \widetilde{\mathcal{J}}} \\ \mathbb{S}^0_{\mathsf{t}}(\mathbb{X})\times\mathbb{X} & \xrightarrow{\quad\mathrm{d}_{\mathrm{t}}\quad} & \mathbb{S}^0_{\mathsf{t}}(\mathbb{X}) \end{array}$$

FIGURE 4.1. Commutative diagram representing the relation between the time interpolants. The second component in the operator $(\,\mathcal{J},\,(\cdot)(0)\,)$ on the left is the evaluation at $t = 0$.

4.5. **Space interpolation.** Regarding the approximation in space, the error notion $\texttt{ERR}(\cdot,\cdot)$ leads to the problem of interpolating functions from $\mathbb{U}$ into $\mathbb{U}_{\mathsf{s}}$, from $\mathbb{D}$ into $\mathbb{D}_{\mathsf{s}}$ as well as from $\overline{\mathbb{P}}$ and $\mathbb{P}^*$ into $\mathbb{P}_{\mathsf{s}}$. Moreover, the spaces are related via the operators $\mathcal{D}$ and $\mathcal{L}$ and their discrete counterparts and the error notion involves various $L^2(\Omega)$-orthogonal projections. Therefore, commutative diagrams like the one in Figure 4.1 are of interest also in this context.

In order to deal with all these tasks, we invoke the existence of an operator which maps $H^1$-conforming piecewise polynomials into $H^2$-conforming ones and enjoys several other properties, whose importance is made clear along the proof of the next results in this section. The operator is defined on the Lagrange space of degree $\mathsf{k}$ without boundary conditions, namely

$$S^1_{\mathsf{k}} := P_{\mathsf{k}}(\mathfrak{T}) \cap H^1(\Omega).$$

Moreover, we denote the jumps and the averages across faces by the usual symbols $\{\!\{\cdot\}\!\}$ and $[\![\cdot]\!]$, cf. [14, Section 38.2.1].

**Lemma 4.14** (Smoothing operator)**.** *There is a linear operator $\mathcal{S} : S^1_{\mathsf{k}} \to H^2(\Omega)$ which satisfies the following properties*

$$\mathcal{S}(\mathbb{P}_{\mathsf{s}}) \subseteq \mathbb{P} \tag{4.21a}$$

$$\nabla \mathcal{S}Q \cdot \mathsf{n} = 0 \qquad \text{on } \Gamma_{p,N} \tag{4.21b}$$

$$\int_{\mathsf{T}} (\mathcal{S}Q)N = \int_{\mathsf{T}} QN \qquad \forall N \in P_{\mathsf{k}}(\mathsf{T}),\ \mathsf{T} \in \mathfrak{T} \tag{4.21c}$$

$$\int_{\mathsf{F}} (\mathcal{S}Q)N = \int_{\mathsf{F}} QN \qquad \forall N \in P_{k-1}(\mathsf{F}),\ F \in \mathfrak{F}^i \cup \mathfrak{F}_{p,N} \tag{4.21d}$$

*as well as the stability estimate*

$$\|D^j(Q - \mathcal{S}Q)\|^2_{\mathsf{T}} \lesssim \sum_{\mathsf{F} \in \mathfrak{F}^i \cup \mathfrak{F}_{p,N},\ \mathsf{F} \cap \mathsf{T} \neq \emptyset} \int_{\mathsf{F}} \{\!\{\mathsf{h}\}\!\}^{3-2j} |[\![\nabla Q]\!] \cdot \mathsf{n}|^2 \tag{4.21e}$$

*for all $Q \in S^1_{\mathsf{k}}$ and $j \in \{0,1,2\}$. The hidden constant depends only on the quantities mentioned in Remark 4.1.*

*Proof.* See Appendix A. □

Having the operator $\mathcal{S}$ at hand, we define $\mathcal{I}_\Omega : L^2(\Omega) \to S^1_{\mathsf{k}}$ via the problem

$$(\mathcal{I}_\Omega \widetilde{p}_{\mathrm{tot}}, Q)_\Omega = (\widetilde{p}_{\mathrm{tot}}, \mathcal{S}Q)_\Omega \qquad \forall Q \in S^1_{\mathsf{k}} \tag{4.22a}$$

for $\widetilde{p}_{\mathrm{tot}} \in L^2(\Omega)$. Analogously, we define $\mathcal{I}_{\mathbb{P}^*} : \mathbb{P}^* \to \mathbb{P}_{\mathsf{s}}$ via the problem

$$(\mathcal{I}_{\mathbb{P}^*} \widetilde{m}, N)_\Omega = \langle \widetilde{m}, \mathcal{S}N \rangle_{\mathbb{P}} \qquad \forall N \in \mathbb{P}_{\mathsf{s}} \tag{4.22b}$$

for $\widetilde{m} \in \mathbb{P}^*$. Note that the main difference between $\mathcal{I}_\Omega$ and $\mathcal{I}_{\mathbb{P}^*}$ is in the range. We introduce also $\mathcal{I}_{\mathbb{U}} : \mathbb{U} \to \mathbb{U}_{\mathsf{s}}$ by

$$\mathcal{I}_{\mathbb{U}} \widetilde{u} := \widehat{U} + \mathcal{R}_{\mathsf{s}}(\mathcal{I}_\Omega \mathcal{D} \widetilde{u} - \mathcal{D}_{\mathsf{s}} \widehat{U}) \tag{4.22c}$$

for $\widetilde{u} \in \mathbb{U}$, where $\widehat{U}$ is the $\mathbb{U}$-orthogonal projection of $u$ onto $\mathbb{U}_{\mathsf{s}}$ and $\mathcal{R}_{\mathsf{s}} : \mathbb{D}_{\mathsf{s}} \to \mathbb{U}_{\mathsf{s}}$ is a right inverse of $\mathcal{D}_{\mathsf{s}}$. The existence and the boundedness of the latter operator are equivalent to Proposition 4.2(2), cf. [14, Lemma C.42]. Notice also that we have $\mathcal{I}_\Omega \mathcal{D} \widetilde{u} \in \mathbb{D}_{\mathsf{s}}$ in view of Lemma 4.15(2) below.

Let us clarify the logic behind the definitions in (4.22). The interpolant $\mathcal{I}_\Omega$ is intended for the approximation of the total pressure, $\mathcal{I}_{\mathbb{P}^*}$ for the total fluid content and the pressure, and $\mathcal{I}_{\mathbb{U}}$ for the displacement. It is important for our purpose that $\mathcal{I}_{\mathbb{P}^*}$ is $\mathbb{P}^*$-stable and this requires the use of an operator $\mathcal{S}$ in the right-hand side of (4.22b), mapping the test functions (at least) into $\mathbb{P}$. (We mention, in passing, that $\mathcal{S}$ is actually required to map into even more regular functions, in order to enforce the $L^2(\Omega)$-stability of the interpolant $\mathcal{I}_{\mathbb{P}}$ introduced below, cf. Lemma 4.17.) In principle, the use of $\mathcal{S}$ is not needed in (4.22a), when only stability in the $L^2(\Omega)$-norm is required. Still, it is important at some point relating $\mathcal{I}_\Omega$ and $\mathcal{I}_{\mathbb{P}^*}$ via a commutative diagram, cf. Figure 4.2 below. Therefore, we use $\mathcal{S}$ also in (4.22a). Finally, the interpolant $\mathcal{I}_{\mathbb{U}}$ is nothing else than the $\mathbb{U}$-orthogonal projection plus a correction, that is necessary for another commutative diagram.

**Lemma 4.15** (Space interpolation – Part 1)**.** *The operators $\mathcal{I}_\Omega$, $\mathcal{I}_{\mathbb{P}^*}$ and $\mathcal{I}_{\mathbb{U}}$ defined in* (4.22) *enjoy the following properties.*

*(1) $\mathcal{I}_\Omega$ maps $\mathbb{D}$ into $\mathbb{D}_{\mathsf{s}}$ and, for all $\widetilde{p}_{tot} \in L^2(\Omega)$, we have*

$$\|\widetilde{p}_{tot} - \mathcal{I}_\Omega \widetilde{p}_{tot}\|_\Omega \lesssim \inf_{\widehat{P}_{tot} \in S^1_{\mathsf{k}}} \|\widetilde{p}_{tot} - \widehat{P}_{tot}\|_\Omega.$$

*(2) For all* $\widetilde{m} \in \mathbb{P}^*$ *and* $\widetilde{p} \in \overline{\mathbb{P}}$*, we have, respectively,*

$$\|\widetilde{m} - \mathcal{P}^*_{\mathbb{P}_\mathsf{s}} \mathcal{I}_{\mathbb{P}^*} \widetilde{m}\|_{\mathbb{P}^*} \lesssim \inf_{\widehat{M} \in \mathbb{P}_\mathsf{s}} \|\widetilde{m} - \mathcal{P}^*_{\mathbb{P}_\mathsf{s}} \widehat{M}\|_{\mathbb{P}^*}$$

*and*

$$\|\widetilde{p} - \mathcal{I}_{\mathbb{P}^*} \widetilde{p}\|_\Omega \lesssim \inf_{\widehat{P} \in \mathbb{P}_\mathsf{s}} \|\widetilde{p} - \widehat{P}\|_\Omega.$$

*(3) For all* $\widetilde{u} \in \mathbb{U}$*, we have* $\mathcal{D}_\mathsf{s} \mathcal{I}_\mathbb{U} \widetilde{u} = \mathcal{I}_\Omega \mathcal{D} \widetilde{u}$ *as well as*

$$\|\widetilde{u} - \mathcal{I}_\mathbb{U} \widetilde{u}\|_\mathbb{U} \lesssim \inf_{\widehat{U} \in \mathbb{U}_\mathsf{s}} \|\widetilde{u} - \widehat{U}\|_\mathbb{U}.$$

*(4) The following identities hold true in* $L^2(\Omega)$

$$\mathcal{P}_{\mathbb{D}_\mathsf{s}} \mathcal{I}_\Omega = \mathcal{I}_\Omega \mathcal{P}_\mathbb{D} \qquad \text{and} \qquad \mathcal{P}_{\mathbb{P}_\mathsf{s}} \mathcal{I}_\Omega = \mathcal{I}_{\mathbb{P}^*} \mathcal{P}_{\overline{\mathbb{P}}}.$$

*The hidden constants depend only on the quantities mentioned in Remark 4.1.*

*Proof.* (1) We first check that $\mathcal{I}_\Omega$ indeed maps $\mathbb{D}$ into $\mathbb{D}_\mathsf{s}$. Recall (2.5) and (4.5). For $\mathbb{D} = L^2(\Omega)$ there is nothing to prove. If $\mathbb{D} = L^2_0(\Omega)$, then (4.22a) implies, for $\widetilde{p}_\mathrm{tot} \in \mathbb{D}$, that

$$(\mathcal{I}_\Omega \widetilde{p}_\mathrm{tot}, 1)_\Omega = (\widetilde{p}_\mathrm{tot}, \mathcal{S}1)_\Omega = (\widetilde{p}_\mathrm{tot}, 1)_\Omega = 0.$$

Note that the identity $\mathcal{S}1 = 1$ follows from (4.21e) with $j = 1$. This confirms the inclusion $\mathcal{I}_\Omega \widetilde{p}_\mathrm{tot} \in \mathbb{D}_\mathsf{s}$. A similar argument reveals that $\mathcal{I}_\Omega$ is invariant on $S^1_\mathsf{k}$. In fact, owing to (4.21c), we have, for $\widetilde{p}_\mathrm{tot} \in S^1_\mathsf{k}$, that

$$(\mathcal{I}_\Omega \widetilde{p}_\mathrm{tot}, Q)_\Omega = (\widetilde{p}_\mathrm{tot}, \mathcal{S}Q)_\Omega = (\widetilde{p}_\mathrm{tot}, Q)_\Omega \qquad \forall Q \in S^1_\mathsf{k}.$$

Finally, for general $\widetilde{p}_\mathrm{tot} \in \mathbb{H}$, the boundedness (4.21e) of $\mathcal{S}$ with $j = 0$, combined with a discrete trace inequality [13, Lemma 12.8] and an inverse estimate [13, Lemma 12.1], yields

$$\|\mathcal{I}_\Omega \widetilde{p}_\mathrm{tot}\|_\Omega = \sup_{Q \in S^1_\mathsf{k}} \frac{(\widetilde{p}_\mathrm{tot}, \mathcal{S}Q)_\Omega}{\|Q\|_\Omega} \lesssim \|\widetilde{p}_\mathrm{tot}\|_\Omega.$$

The claimed estimate follows by the invariance and the boundedness of $\mathcal{I}_\Omega$.

(2) For $\widetilde{m} \in \mathbb{P}^*$ and $\widehat{M} \in \mathbb{P}_\mathsf{s}$, we have

$$\widetilde{m} - \mathcal{P}^*_{\mathbb{P}_\mathsf{s}} \mathcal{I}_{\mathbb{P}^*} \widetilde{m} = \widetilde{m} - \mathcal{P}^*_{\mathbb{P}_\mathsf{s}} \widehat{M} + \mathcal{P}^*_{\mathbb{P}_\mathsf{s}} (\widehat{M} - \mathcal{I}_{\mathbb{P}^*} \widetilde{m}).$$

We bound the norm of the second summand on the right-hand side with the help of Proposition 4.2(3), the inclusion $\mathbb{P}_\mathsf{s} \subseteq L^2(\Omega)$ and (4.21e) with $j = 1$ and a discrete trace inequality

$$\begin{aligned} \|\mathcal{P}^*_{\mathbb{P}_\mathsf{s}} (\widehat{M} - \mathcal{I}_{\mathbb{P}^*} \widetilde{m})\|_{\mathbb{P}^*} \eqsim \|\widehat{M} - \mathcal{I}_{\mathbb{P}^*} \widetilde{m}\|_{\mathbb{P}^*_\mathsf{s}} &= \sup_{N \in \mathbb{P}_\mathsf{s}} \frac{\langle \widehat{M} - \mathcal{I}_{\mathbb{P}^*} \widetilde{m}, N \rangle_{\mathbb{P}_\mathsf{s}}}{\|N\|_\mathbb{P}} \\ &= \sup_{N \in \mathbb{P}_\mathsf{s}} \frac{\langle \mathcal{P}^*_{\mathbb{P}_\mathsf{s}} \widehat{M} - \widetilde{m}, \mathcal{S}N \rangle_\mathbb{P}}{\|N\|_\mathbb{P}} \lesssim \|\mathcal{P}^*_{\mathbb{P}_\mathsf{s}} \widehat{M} - \widetilde{m}\|_{\mathbb{P}^*}. \end{aligned}$$

The first claimed estimate follows by combining this bound with the above identity. The other estimate follows by establishing invariance on $\mathbb{P}_\mathsf{s}$ as well as boundedness in the $L^2(\Omega)$-norm exactly as in (1).

(3) Recall that the operator $\mathcal{R}_\mathsf{s}$ in (4.22c) is a right inverse of $\mathcal{D}_\mathsf{s}$, i.e. $\mathcal{D}_\mathsf{s} \mathcal{R}_\mathsf{s}$ is the identity on $\mathbb{D}_\mathsf{s}$. Then, the first part of the statement directly follows from the definition of $\mathcal{I}_\mathbb{U}$. We verify the second part as in (1), i.e. we prove that $\mathcal{I}_\mathbb{U}$ is

invariant on $\mathbb{U}_{\mathsf{s}}$ and bounded in the $\mathbb{U}$-norm. For $\widetilde{u} \in \mathbb{U}_{\mathsf{s}}$, we have $\widehat{U} = \widetilde{u}$ in (4.22c). Then, according to (4.21c) and (4.6), we infer

$$(\mathcal{I}_\Omega \mathcal{D}\widetilde{u}, Q_{\mathrm{tot}})_\Omega = (\mathcal{D}\widetilde{u}, \mathcal{S}Q_{\mathrm{tot}})_\Omega = (\mathcal{D}\widetilde{u}, Q_{\mathrm{tot}})_\Omega = (\mathcal{D}_{\mathsf{s}}\widetilde{u}, Q_{\mathrm{tot}})_\Omega \qquad \forall Q_{\mathrm{tot}} \in \mathbb{D}_{\mathsf{s}}.$$

This implies $\mathcal{I}_\Omega \mathcal{D}\widetilde{u} = \mathcal{D}_{\mathsf{s}}\widetilde{u}$ and, in turn, $\mathcal{I}_{\mathbb{U}}\widetilde{u} = \widetilde{u}$. Next, for general $\widetilde{u} \in \mathbb{U}$, we have

$$\|\mathcal{I}_{\mathbb{U}}\widetilde{u}\|_{\mathbb{U}} \lesssim \|\widehat{U}\|_{\mathbb{U}} + \|\mathcal{I}_\Omega \mathcal{D}\widetilde{u} - \mathcal{D}_{\mathsf{s}}\widehat{U}\|_\Omega.$$

Indeed, the boundedness of $\mathcal{R}_{\mathsf{s}}$ is equivalent to 4.2(3); see [14, Lemma C.42]. We conclude the proof of (3) with the help of the boundedness of $\mathcal{I}_\Omega$, $\mathcal{D}$ and $\mathcal{D}_{\mathsf{s}}$ in the respective norms and by the definition of $\widehat{U}$ as the $\mathbb{U}$-orthogonal projection of $\widetilde{u}$.

(4) Recall that $\mathcal{P}_{\mathbb{D}}$ and $\mathcal{P}_{\mathbb{D}_{\mathsf{s}}}$ are the $L^2(\Omega)$-orthogonal projections onto the spaces $\mathbb{D}$ and $\mathbb{D}_{\mathsf{s}}$, defined in (2.5) and (4.5), respectively. For $\widetilde{p}_{\mathrm{tot}} \in L^2(\Omega)$, we have $\mathcal{I}_\Omega \mathcal{P}_{\mathbb{D}} \widetilde{p}_{\mathrm{tot}} \in \mathbb{D}_{\mathsf{s}}$, in view of (1) above. Then, for all $Q_{\mathrm{tot}} \in \mathbb{D}_{\mathsf{s}}$, it holds that

$$(\mathcal{I}_\Omega \mathcal{P}_{\mathbb{D}} \widetilde{p}_{\mathrm{tot}}, Q_{\mathrm{tot}})_\Omega = (\mathcal{P}_{\mathbb{D}} \widetilde{p}_{\mathrm{tot}}, \mathcal{S}Q_{\mathrm{tot}})_\Omega = (\widetilde{p}_{\mathrm{tot}}, \mathcal{S}Q_{\mathrm{tot}})_\Omega = (\mathcal{I}_\Omega \widetilde{p}_{\mathrm{tot}}, Q_{\mathrm{tot}})_\Omega.$$

In particular, we can remove $\mathcal{P}_{\mathbb{D}}$ after the second identity, because $\mathcal{S}$ maps $\mathbb{D}_{\mathsf{s}}$ into $\mathbb{D}$, thanks to (4.21c). Hence, the first claimed identity is verified. The proof of the other one is similar. Recall that $\mathcal{P}_{\overline{\mathbb{P}}}$ and $\mathcal{P}_{\mathbb{P}_{\mathsf{s}}}$ are the $L^2(\Omega)$-orthogonal projections onto the spaces $\overline{\mathbb{P}}$ and $\mathbb{P}_{\mathsf{s}}$, defined in (2.4) and (4.7), respectively. For $\widetilde{p} \in L^2(\Omega)$ and $N \in \mathbb{P}_{\mathsf{s}}$, we have

$$(\mathcal{I}_{\mathbb{P}^*} \mathcal{P}_{\overline{\mathbb{P}}} \widetilde{p}, N)_\Omega = (\mathcal{P}_{\overline{\mathbb{P}}} \widetilde{p}, \mathcal{S}N)_\Omega = (\widetilde{p}, \mathcal{S}N)_\Omega = (\mathcal{I}_\Omega \widetilde{p}, N)_\Omega.$$

This time we could remove $\mathcal{P}_{\overline{\mathbb{P}}}$ after the second identity thanks to (4.21a). Thus, also the second claimed identity is verified. □

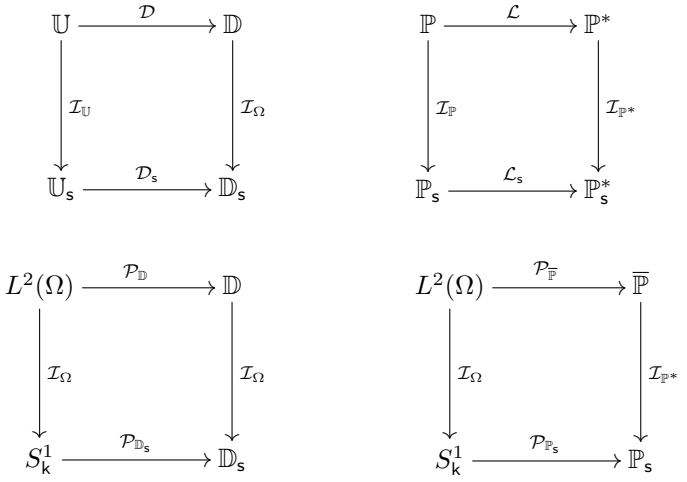


FIGURE 4.2. Commutative diagrams representing the relations among the space interpolants.

The interpolants defined up to this point are compatible with the divergence, i.e. with the operators $\mathcal{D}$ and $\mathcal{D}_{\mathsf{s}}$, and with the various $L^2(\Omega)$-orthogonal projections, as summarized in Figure 4.2. Still, the compatibility with the Laplacian, i.e. with

the operators $\mathcal{L}$ and $\mathcal{L}_{\mathsf{s}}$, is not guaranteed. Since $\mathcal{L}$ and $\mathcal{L}_{\mathsf{s}}$ are one-to-one, the diagram on the upper right corner of Figure 4.2 uniquely defines $\mathcal{I}_{\mathbb{P}} : \mathbb{P} \to \mathbb{P}_{\mathsf{s}}$ as

$$\mathcal{I}_{\mathbb{P}} := \mathcal{L}_{\mathsf{s}}^{-1} \mathcal{I}_{\mathbb{P}^*} \mathcal{L}. \tag{4.23}$$

The next lemma reveals that also $\mathcal{I}_{\mathbb{P}}$ has favorable approximation properties. To this end, we introduce the following broken $H^2$-norm

$$\|\cdot\|_{H^2(\mathfrak{T})}^2 := \kappa \left( \sum_{\mathsf{T} \in \mathfrak{T}} \|D^2 \cdot\|_{L^2(\mathsf{T})}^2 + \sum_{\mathsf{F} \in \mathfrak{F}^i \cup \mathfrak{F}_{p,N}} \|\{\!\{\mathsf{h}\}\!\}^{-1/2} [\![\nabla \cdot]\!]\|_{L^2(\mathsf{F})}^2 \right)$$

as well as the constant

$$\varepsilon_{\mathsf{s}} := \inf_{N \in \mathbb{P}_{\mathsf{s}}} \frac{\|\mathcal{L}_{\mathsf{s}} N\|_{\Omega}}{\|N\|_{H^2(\mathfrak{T})}}. \tag{4.24}$$

*Remark* 4.16 (Discrete elliptic regularity). The constant $\varepsilon_{\mathsf{s}}$ in (4.24) is certainly finite, because $\mathbb{P}_{\mathsf{s}}$ is finite-dimensional. The size of the constant is related to the control of a $H^2$-like norm by the $L^2$-norm of $\mathcal{L}_{\mathsf{s}}$, i.e. our discretization of the Laplacian. Therefore, the constant is a discrete measure of the elliptic regularity. Note also that $\varepsilon_{\mathsf{s}}$ can be equivalently interpreted as an inf-sup constant

$$\varepsilon_{\mathsf{s}} = \inf_{N \in \mathbb{P}_{\mathsf{s}}} \sup_{\widetilde{P} \in \mathbb{P}_{\mathsf{s}}} \frac{\langle \mathcal{L} N, \widetilde{P} \rangle_{\mathbb{P}}}{\|N\|_{H^2(\mathfrak{T})} \|\widetilde{P}\|_{\Omega}},$$

where the stability of the standard weak formulation of the Laplacian is analyzed in a nonstandard (i.e. nonsymmetric) setting, cf. (4.8). We are aware of only a few results regarding the size of $\varepsilon_{\mathsf{s}}$. For $H^1$-conforming Lagrange finite elements and convex $\Omega$, Makridakis established a lower bound in terms of the shape parameter (4.1) of $\mathfrak{T}$, provided that the mesh is not too much graded, see [28]. It is unclear to us whether the latter condition is necessary or not. When $\Omega$ is not convex, the connection with the elliptic regularity suggests that a lower bound of $\varepsilon_{\mathsf{s}}$ only in terms of shape regularity might be not possible. As for the question raised in Remark 4.12, we postpone further investigation on this point to future work.

**Lemma 4.17** (Space interpolation – Part 2)**.** *The operator $\mathcal{I}_{\mathbb{P}}$ defined above has a unique bounded extension (denoted by the same symbol) to $\overline{\mathbb{P}}$ which satisfies for all $\widetilde{p} \in \overline{\mathbb{P}}$ the estimate*

$$\|\widetilde{p} - \mathcal{I}_{\mathbb{P}} \widetilde{p}\|_{\Omega} \lesssim \varepsilon_{\mathsf{s}}^{-1} \inf_{\widehat{P} \in \mathbb{P}_{\mathsf{s}}} \|\widetilde{p} - \widehat{P}\|_{\Omega}.$$

*The hidden constant depends only on the quantities mentioned in Remark 4.1.*

*Proof.* As for Lemma 4.15(2)-(3), we verify the claimed estimate by showing that $\mathcal{I}_{\mathbb{P}}$ is invariant on $\mathbb{P}_{\mathsf{s}}$ and that it is bounded in the $L^2(\Omega)$-norm. Note that the latter property implies also the existence of a unique bounded extension of $\mathcal{I}_{\mathbb{P}}$ to $\overline{\mathbb{P}}$, because $\mathbb{P}$ is a dense subspace thereof. Assume first $\widetilde{p} \in \mathbb{P}_{\mathsf{s}}$. Owing to (4.23) and (4.21a), we have

$$\langle \mathcal{I}_{\mathbb{P}} \widetilde{p}, \mathcal{L}_{\mathsf{s}} N \rangle_{\mathbb{P}_{\mathsf{s}}} = \langle \mathcal{L} \widetilde{p}, \mathcal{S} N \rangle_{\mathbb{P}} = \langle \widetilde{p}, \mathcal{L} \mathcal{S} N \rangle_{\mathbb{P}} \qquad N \in \mathbb{P}_{\mathsf{s}}.$$

We recall (2.2) and integrate by parts element-wise [1, eq. (3.6)]

$$\langle \mathcal{I}_{\mathbb{P}} \widetilde{p}, \mathcal{L}_{\mathsf{s}} N \rangle_{\mathbb{P}_{\mathsf{s}}} = \kappa \left( - \sum_{\mathsf{T} \in \mathfrak{T}} \int_{\mathsf{T}} (\Delta \widetilde{p}) \mathcal{S} N + \sum_{\mathsf{F} \in \mathfrak{F}^i \cup \mathfrak{F}_{p,N}} \int_{\mathsf{F}} ([\![\nabla \widetilde{p}]\!] \cdot \mathsf{n}) \mathcal{S} N \right).$$

Then, we make use of (4.21c)-(4.21d) and we integrate by parts back. It results

$$\langle \mathcal{I}_{\mathbb{P}}\widetilde{p}, \mathcal{L}_{\mathsf{s}}N\rangle_{\mathbb{P}_{\mathsf{s}}} = \langle \widetilde{p}, \mathcal{L}N\rangle_{\mathbb{P}} = \langle \widetilde{p}, \mathcal{L}_{\mathsf{s}}N\rangle_{\mathbb{P}_{\mathsf{s}}}$$

where the last identity is obtained with the help of (4.8). Since $\mathcal{L}_{\mathsf{s}}$ is one-to-one, we conclude $\mathcal{I}_{\mathbb{P}}\widetilde{p} = \widetilde{p}$.

Next, for general $\widetilde{p} \in \mathbb{P}$, a similar argument as before yields

$$\|\mathcal{I}_{\mathbb{P}}\widetilde{p}\|_{\Omega} = \sup_{N\in\mathbb{P}_{\mathsf{s}}} \frac{\langle \mathcal{I}_{\mathbb{P}}\widetilde{p}, \mathcal{L}_{\mathsf{s}}N\rangle_{\mathbb{P}_{\mathsf{s}}}}{\|\mathcal{L}_{\mathsf{s}}N\|_{\Omega}} = \sup_{N\in\mathbb{P}_{\mathsf{s}}} \frac{\langle \widetilde{p}, \mathcal{L}\mathcal{S}N\rangle_{\mathbb{P}}}{\|\mathcal{L}_{\mathsf{s}}N\|_{\Omega}}.$$

Recall the inclusion $\mathcal{S}N \in H^2(\Omega)$ and (4.21a)-(4.21b). We integrate by parts, then we invoke (4.21e) with $j = 2$ and (4.24)

$$\|\mathcal{I}_{\mathbb{P}}\widetilde{p}\|_{\mathbb{H}} = \kappa(\widetilde{p}, -\Delta\mathcal{S}N)_{\Omega} \lesssim \|\widetilde{p}\|_{\Omega}\|N\|_{H^2(\mathfrak{T})} \leq \varepsilon_{\mathsf{s}}^{-1}\|\widetilde{p}\|_{\Omega}\|\mathcal{L}_{\mathsf{s}}N\|_{\Omega}.$$

The combination of this bound with the previous identity implies the announced boundedness of $\mathcal{I}_{\mathbb{P}}$. □

4.6. **Space-time interpolation.** We are in position to combine the interpolants from the two previous sections in order to conclude the proof of Theorem 4.7. To this end, let us preliminary recall that the operators acting in time commute with those acting in space. Our main device is the space-time interpolant $\mathcal{I} : \mathbb{Y}_1 \to \mathbb{Y}_{1,\mathsf{st}}$ defined as

$$\mathcal{I}\widetilde{y}_1 := \big(\mathcal{I}_{\mathbb{U}}\mathcal{J}\widetilde{u},\ \mathcal{I}_{\Omega}\mathcal{J}\widetilde{p}_{\mathrm{tot}},\ \mathcal{I}_{\mathbb{P}}\widetilde{\mathcal{J}}\widetilde{p},\ \mathcal{I}_{\mathbb{P}^*}\mathcal{J}\widetilde{m},\ \mathcal{I}_{\mathbb{P}^*}\widetilde{m}(0)\big)$$

for $\widetilde{y}_1 = (\widetilde{u}, \widetilde{p}_{\mathrm{tot}}, \widetilde{p}, \widetilde{m}) \in \mathbb{Y}_1$.

Our strategy to verify Theorem 4.7 is as follows. We first establish (4.15) for $y_1 \in \mathbb{Y}_1$. Then we extend the result by density, because the right-hand side of (4.15) is bounded with respect to the norm $\|\cdot\|_1$, cf. Theorem 2.4.

First of all, we recall the definitions of $\mathbb{Y}_1$ and $\mathbb{Y}_{1,\mathsf{st}}$ in (2.9) and (3.11), respectively, and notice that $\mathcal{I}$ is well-defined. In fact, the discussion in the previous sections shows that the interpolation of each component acts as follows

$$\begin{aligned}
L^2(\mathbb{U}) &\xrightarrow{\mathcal{J}} \mathbb{S}^0_{\mathsf{t}}(\mathbb{U}) \xrightarrow{\mathcal{I}_{\mathbb{U}}} \mathbb{S}^0_{\mathsf{t}}(\mathbb{U}_{\mathsf{s}}) && \text{for } \widetilde{u}\\
L^2(\mathbb{D}) &\xrightarrow{\mathcal{J}} \mathbb{S}^0_{\mathsf{t}}(\mathbb{D}) \xrightarrow{\mathcal{I}_{\Omega}} \mathbb{S}^0_{\mathsf{t}}(\mathbb{D}_{\mathsf{s}}) && \text{for } \widetilde{p}_{\mathrm{tot}}\\
L^2(\mathbb{P}) &\xrightarrow{\widetilde{\mathcal{J}}} \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}) \xrightarrow{\mathcal{I}_{\mathbb{P}}} \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}}) && \text{for } \widetilde{p}\\
H^1(\mathbb{P}^*) &\xrightarrow{\mathcal{J}} \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}^*) \xrightarrow{\mathcal{I}_{\mathbb{P}^*}} \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}}) && \text{for } \widetilde{m}\\
\mathbb{P}^* &\xrightarrow{\mathcal{I}_{\mathbb{P}^*}} \mathbb{P}_{\mathsf{s}} && \text{for } \widetilde{m}(0).
\end{aligned}$$

In particular, the second line makes use of Lemma 4.15(1) and the last one exploits the inclusion $\widetilde{m} \in H^1(\mathbb{P}^*) \subseteq C^0(\mathbb{P}^*)$.

In order to verify Theorem 4.7, we bound the left-hand side of (4.15) by

$$\inf_{\widetilde{Y}_1\in\mathbb{Y}_{1,\mathsf{st}}} \mathtt{ERR}(\widetilde{y}_1, \widetilde{Y}_1)^2 \leq \mathtt{ERR}(\widetilde{y}_1, \mathcal{I}\widetilde{y}_1)^2. \tag{4.25}$$

Then, we recall the definition (4.11) of the error notion and we bound the six terms therein (denoted by *(i)*, *(ii)*, ... for brevity) one by one. Lemma 4.13 and Lemma 4.15(3) state that $\mathcal{J}$ and $\mathcal{I}_{\mathbb{U}}$ are bounded linear projections. Therefore, their combination $\mathcal{J}\mathcal{I}_{\mathbb{U}}$ is a $L^2(\mathbb{U})$-bounded linear projection onto $\mathbb{S}^0_{\mathsf{t}}(\mathbb{U}_{\mathsf{s}})$. This implies

$$(i) \lesssim \inf_{\widehat{U}\in\mathbb{S}^0_{\mathsf{t}}(\mathbb{U}_{\mathsf{s}})} \int_0^T \|\widetilde{u} - \widehat{U}\|^2_{\mathbb{U}}.$$

The same reasoning with Lemma 4.15(1) implies

$$(ii) \lesssim \inf_{\widehat{P}_{\mathrm{tot}} \in \mathbb{S}^0_{\mathsf{t}}(\mathbb{D}_{\mathsf{s}})} \int_0^T \frac{1}{\mu} \|\widetilde{p}_{\mathrm{tot}} - \widehat{P}_{\mathrm{tot}}\|^2_\Omega.$$

Regarding the third term, we first recall (4.19c)

$$\mathrm{d}_{\mathsf{t}}(\mathcal{I}_{\mathbb{P}^*} \mathcal{J} \widetilde{m},\ \mathcal{I}_{\mathbb{P}^*} \widetilde{m}(0)) = \mathcal{I}_{\mathbb{P}^*} \mathrm{d}_{\mathsf{t}}(\mathcal{J} \widetilde{m},\ \widetilde{m}(0)) = \mathcal{I}_{\mathbb{P}^*} \widetilde{\mathcal{J}} \partial_t \widetilde{m},$$

then, we make use of (4.23)

$$\mathcal{L}_{\mathsf{s}} \mathcal{I}_{\mathbb{P}} \widetilde{\mathcal{J}} \widetilde{p} = \mathcal{I}_{\mathbb{P}^*} \widetilde{\mathcal{J}} \mathcal{L} \widetilde{p}$$

and, combining the two identities, we arrive at

$$\mathrm{d}_{\mathsf{t}}(\mathcal{I}_{\mathbb{P}^*} \mathcal{J} \widetilde{m},\ \mathcal{I}_{\mathbb{P}^*} \widetilde{m}(0)) + \mathcal{L}_{\mathsf{s}} \mathcal{I}_{\mathbb{P}} \widetilde{\mathcal{J}} \widetilde{p} = \mathcal{I}_{\mathbb{P}^*} \widetilde{\mathcal{J}} (\partial_t \widetilde{m} + \mathcal{L} \widetilde{p}).$$

The same argument used in the above bounds for *(i)* and *(ii)*, with Lemma 4.15(2), implies

$$(iii) \lesssim \inf_{\widehat{W} \in \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}})} \int_0^T \|\partial_t \widetilde{m} + \mathcal{L} \widetilde{p} - \mathcal{P}^*_{\mathbb{P}_{\mathsf{s}}} \widehat{W}\|^2_{\mathbb{P}^*}.$$

For the fourth term, we directly use Lemma 4.15(2) to obtain

$$(iv) \lesssim \inf_{\widehat{M}_0 \in \mathbb{P}_{\mathsf{s}}} \|\widetilde{m}(0) - \mathcal{P}^*_{\mathbb{P}_{\mathsf{s}}} \widehat{M}_0\|^2_{\mathbb{P}^*}.$$

Concerning the fifth term, we invoke Lemma 4.15(3)

$$\begin{aligned} &\lambda \mathcal{D}_{\mathsf{s}} \mathcal{I}_{\mathbb{U}} \mathcal{J} \widetilde{u} - \mathcal{I}_\Omega \mathcal{J} \widetilde{p}_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}_{\mathsf{s}}} \mathcal{I}_{\mathbb{P}} \widetilde{\mathcal{J}} \widetilde{p} = \\ &\qquad \mathcal{I}_\Omega \mathcal{J} (\mathcal{D} \widetilde{u} - \widetilde{p}_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}} \widetilde{p}) + \alpha \mathcal{I}_\Omega \mathcal{J} \mathcal{P}_{\mathbb{D}} \widetilde{p} - \alpha \mathcal{P}_{\mathbb{D}_{\mathsf{s}}} \mathcal{I}_{\mathbb{P}} \widetilde{\mathcal{J}} \widetilde{p} \end{aligned}$$

The reasoning from the bound of *(ii)* shows that the first summand in the right-hand side is a near best approximation of $(\mathcal{D} \widetilde{u} - \widetilde{p}_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}} \widetilde{p})$ in $\mathcal{S}^0_{\mathsf{t}}(\mathbb{D}_{\mathsf{s}})$. The other two summands can be rewritten as $\alpha \mathcal{P}_{\mathbb{D}_{\mathsf{s}}} (\mathcal{I}_\Omega \mathcal{J} - \mathcal{I}_{\mathbb{P}} \widetilde{\mathcal{J}}) \widetilde{p}$, thanks to Lemma 4.15(4). Arguing as before, we see that both $\mathcal{I}_\Omega \mathcal{J} \widetilde{p}$ and $\mathcal{I}_{\mathbb{P}} \widetilde{\mathcal{J}} \widetilde{p}$ are near best approximations of $\widetilde{p}$ in $\mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}})$, the latter one in view of Lemma 4.17. These observations, the triangle inequality and the definition (2.12) of $\gamma$ reveal

$$(v) \lesssim \inf_{\widehat{Q}_{\mathrm{tot}} \in \mathbb{S}^0_{\mathsf{t}}(\mathbb{D}_{\mathsf{s}})} \int_0^T \frac{1}{\lambda + \mu} \|(\mathcal{D} \widetilde{u} - \widetilde{p}_{\mathrm{tot}} - \alpha \mathcal{P}_{\mathbb{D}} \widetilde{p}) - \widehat{Q}_{\mathrm{tot}}\|^2_\Omega + \varepsilon_{\mathsf{s}}^{-1} \inf_{\widehat{P} \in \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}})} \int_0^T \frac{1}{\gamma} \|\widetilde{p} - \widehat{P}\|^2_\Omega.$$

The sixth and last term can be treated similarly. We invoke again Lemma 4.15(3) and Lemma 4.15(4), so as to obtain

$$\begin{aligned} &\alpha \mathcal{P}_{\mathbb{P}_{\mathsf{s}}} \mathcal{D}_{\mathsf{s}} \mathcal{I}_{\mathbb{U}} \mathcal{J} \widetilde{u} + \sigma \mathcal{I}_{\mathbb{P}} \widetilde{\mathcal{J}} \widetilde{p} - \mathcal{I}_{\mathbb{P}^*} \mathcal{J} \widetilde{m} \\ &\qquad = \mathcal{I}_{\mathbb{P}^*} \mathcal{J} (\alpha \mathcal{P}_{\mathbb{P}} \mathcal{D} \widetilde{u} + \sigma \widetilde{p} - \widetilde{m}) + \sigma (\mathcal{I}_{\mathbb{P}} \widetilde{\mathcal{J}} - \mathcal{I}_{\mathbb{P}^*} \mathcal{J}) \widetilde{p}. \end{aligned}$$

Again, all operators yield near best approximation in the respective spaces. In particular, for $\mathcal{I}_{\mathbb{P}^*} \mathcal{J}$, we make use of the second part of Lemma 4.15(2). Thus we arrive at the following estimate

$$(vi) \lesssim \inf_{\widehat{Q} \in \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}})} \int_0^T \gamma \|(\mathcal{P}_{\mathbb{P}} \mathcal{D} \widetilde{u} + \sigma \widetilde{p} - \widetilde{m}) - \widehat{Q}\|^2_\Omega + \varepsilon_{\mathsf{s}}^{-1} \inf_{\widehat{P} \in \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}})} \int_0^T \frac{1}{\gamma} \|\widetilde{p} - \widehat{P}\|^2_\Omega$$

with the help of the definition (2.12) of $\gamma$. We insert the above bounds of *(i)*-*(vi)* into (4.25). This establishes (4.15) for $\widetilde{y}_1 \in \mathbb{Y}_1$ and the right-hand side in the resulting estimate is bounded in terms of the norm $\|\cdot\|_1$, cf. Theorem 2.4. Thus,

we can extend the estimate by density from $\mathbb{Y}_1$ to $\overline{\mathbb{Y}}_1$. We conclude by noticing that the bounds of *(v)* and *(vi)* simplify to

$$(v) + (vi) \lesssim \varepsilon_{\mathsf{s}}^{-1} \inf_{\widehat{P} \in \mathbb{S}^0_{\mathsf{t}}(\mathbb{P}_{\mathsf{s}})} \int_0^T \frac{1}{\gamma} \|\widetilde{p} - \widehat{P}\|_\Omega^2$$

if $\widetilde{y}_1 = y_1$ is the solution of (2.8).

## 5. Numerical results

In this section we test the performance of the concrete space discretization proposed in Section 4 on two well-established benchmarks for the Biot's equations. Owing to Corollary 4.8 and Remark 4.12, we restrict ourselves to the lowest order case $\mathsf{k} = 1$, corresponding to quadratic (resp. linear) $H^1$-conforming Lagrange finite elements for the displacement (resp. for the other variables). All experiments have been implemented with the help of the library ALBERTA 3.0, see [19, 35].

5.1. **Terzaghi's problem.** The consolidation of a vertical soil column of total depth $H > 0$ is a classical one-dimensional test case in poroelasticity. The column is subject to compression and it is drained on top, whereas it is impermeable and no displacement occurs at the bottom. Moreover, there is no other force acting in the interior of the column, the initial fluid content equals zero and there are no sources nor sinks. Therefore, in this case, the initial-boundary value problem (2.1) for the Biot's equations reads

$$\begin{aligned} -(2\mu+\lambda)u_{zz} + \alpha p_z &= 0, & \partial_t(\alpha u_z + \sigma p) - \kappa p_{zz} &= 0, && \text{in } (0,H)\times(0,T) \\ -(2\mu+\lambda)u_z + \alpha p &= F, & p &= 0, && \text{on } \{0\}\times(0,T) \\ u &= 0, & p_z &= 0, && \text{on } \{H\}\times(0,T) \\ & & \alpha u_z + \sigma p &= 0, && \text{in } (0,H)\times\{0\} \end{aligned} \tag{5.1}$$

with the subscripts $(\cdot)_z$ and $(\cdot)_{zz}$ denoting the partial derivatives with respect to the depth $z \in (0, H)$ and $F$ the modulus of the force acting on top of the column.

Remarkably, the exact solution of (5.1) is explicitly known, cf. [32, Section 4.1.1]. In particular, the pressure equals

$$p(z,t) = p_0 \sum_{m=0}^{+\infty} \frac{4}{(2m+1)\pi} \sin\Big(\frac{(2m+1)\pi z}{2H}\Big) \exp\Big(-\frac{(2m+1)^2\pi^2\widetilde{\gamma}kt}{4H^2}\Big) \tag{5.2}$$

with the auxiliary parameters

$$\widetilde{\gamma} := \Big(\frac{\alpha^2}{2\mu+\lambda} + \sigma\Big)^{-1} \qquad \text{and} \qquad p_0 := \frac{\alpha\widetilde{\gamma}F}{2\mu+\lambda}.$$

The displacement can be derived via the relation

$$u_z = \frac{\alpha p - F}{2\mu+\lambda}$$

and the other variables are obtained through their definition, cf. Remark 2.3.

In analogy with [32, Section 4.1.1], we set

$$H = 1, \qquad T = 1, \qquad F = 10^3$$

as well as [1]

$$\mu = 41667, \quad \lambda = 27778, \quad \alpha = 1, \quad \sigma = 0.1, \quad \kappa = 10^{-6}.$$

The explicit knowledge of the exact solution makes this test case particularly suited to observe the error decay rate. Owing to (2.12), (4.11), Remark 4.6 and Theorem 4.7, we consider the following error notion

$$\int_0^T \left( \|u - U\|_{\mathbb{U}}^2 + \frac{1}{\mu} \|p_{\mathrm{tot}} - P_{\mathrm{tot}}\|_\Omega^2 + \sigma \|p - P\|_\Omega^2 \right) \tag{5.3}$$

where the pressure $L^2(L^2(\Omega))$-error is included for a better comparison with the literature. For the evaluation of the exact solution, we take into account the first 5000 summands in the series (5.2).

We first consider a fixed time discretization with $J = 5000$ intervals of equal size and observe the error decay rate with respect to the space discretization. For this purpose, we begin with a mesh in space consisting only of the interval $(0, H)$ and refine it 10 times by dividing all intervals each time. We report on the left side of Table 1 below the evaluation of the error (5.3) and the experimental order of convergence (EOC) with respect to the number of degrees of freedom in the space discretization, namely

$$\#\mathrm{DOFs} = \dim(\mathbb{U}_{\mathsf{s}} \times \mathbb{D}_{\mathsf{s}} \times \mathbb{P}_{\mathsf{s}} \times \mathbb{P}_{\mathsf{s}}) \tag{5.4}$$

cf. (3.11). By increasing the number of refinements, the EOC appears to converge to 1.5, with only a slight decrease in the last refinement, which is likely due to an insufficient number of intervals in the time discretization.

| #DOFS | error | EOC |
|---|---|---|
| 9 | $1.42 \times 10^{-2}$ | |
| 14 | $1.08 \times 10^{-2}$ | 0.63 |
| 24 | $7.65 \times 10^{-3}$ | 0.63 |
| 44 | $5.41 \times 10^{-3}$ | 0.57 |
| 84 | $3.82 \times 10^{-3}$ | 0.54 |
| 164 | $2.54 \times 10^{-3}$ | 0.61 |
| 324 | $1.33 \times 10^{-3}$ | 0.95 |
| 644 | $5.61 \times 10^{-4}$ | 1.26 |
| 1284 | $2.14 \times 10^{-4}$ | 1.40 |
| 2564 | $7.86 \times 10^{-5}$ | 1.45 |
| 5124 | $2.94 \times 10^{-5}$ | 1.42 |

| $J$ | error | EOC |
|---|---|---|
| 5 | $1.17 \times 10^{-4}$ | |
| 10 | $7.23 \times 10^{-5}$ | 0.70 |
| 20 | $4.41 \times 10^{-5}$ | 0.71 |
| 40 | $2.67 \times 10^{-5}$ | 0.72 |
| 80 | $1.61 \times 10^{-5}$ | 0.73 |
| 160 | $9.70 \times 10^{-6}$ | 0.73 |
| 320 | $5.85 \times 10^{-6}$ | 0.73 |

TABLE 1. Error decay in Terzaghi's problem with respect to the space discretization (left) and the time discretization (right).

The decay rate 1.5 is consistent with the one observed, e.g., in [32, Section 4.1.1]. There the EOC is actually 0.5 and the difference is explained by the different space discretization and error notion. Indeed, our discretization is of second-order (cf. Remark 4.12) whereas the one in [32] is of first-order only. For a theoretical justification note that the exact pressure (5.2) is singular, because the initial value

[1] More precisely [32] sets the Young's modulus to $E = 10^5$ and the Poisson's ratio to $\nu = 0.2$.

$p(\cdot,0)=p_0$ does not satisfy the Dirichlet boundary condition on top of the column. More precisely, for $s\in(0,1)$, we have

$$\|p\|^2_{L^2(H^{1+s}(\Omega))}\approx\sum_{m=0}^{+\infty}m^{2s}\int_0^T\exp(-m^2t)\approx\sum_{m=0}^{+\infty}m^{2s-2},$$

where the symbol $\approx$ indicates that we neglect all multiplicative constants apart of those depending on $m$. Therefore, we have $p\in L^2(H^{1+s}(\Omega))$ if and only if $s<0.5$.

Second, we consider a fixed space discretization with the mesh obtained by 13 global refinements of the interval $(0,H)$. In other words, the mesh is two levels finer than the last one on the left side of Table 1 and it is such that #DOFs = 40964. We use time discretizations with $J$ intervals of equal size and observe the error decay for increasing $J$. According to the data on the right side of Table 1, the EOC is somehow close to 0.75. Although we were not able to find a similar convergence test in the literature, the observed value is in line with our expectation. Indeed, by arguing as before (see also [36, Section 7.1]), we observe that

$$\|p\|^2_{H^s(L^2(\Omega))}\approx\sum_{m=0}^{+\infty}m^{4s-2}\int_0^T\exp(-m^2t)\approx\sum_{m=0}^{+\infty}m^{4s-4}$$

for $s\in(0,1)$. Therefore, we have $p\in H^s(L^2(\Omega))$ if and only if $s<0.75$.

5.2. **Cantilever bracket problem.** We consider the extension to poroelasticity of a well-established two-dimensional test case in linear elasticity. The elastic material initially occupies the region $\Omega=(0,1)^2$, it is clamped on the left side, a uniformly distributed load is applied on the top side and the entire boundary is assumed to be impermeable. Moreover, there is no other force acting in the interior of $\Omega$, the initial fluid content equals zero and there are no sources nor sinks. Therefore, in this case, the initial-boundary value problem (2.1) for the Biot's equations reads

$$\begin{aligned}
-\mathrm{div}(2\mu\nabla_S u+(\lambda\mathrm{div}u-\alpha p)\mathrm{I})=0,\quad \partial_t(\alpha\mathrm{div}u+\sigma p)-\kappa\Delta p=0,&\quad\text{in }\Omega\times(0,T)\\
u=0,\qquad\qquad\qquad\qquad\qquad \nabla p\cdot\mathsf{n}=0,&\quad\text{on }\Gamma_{\mathtt{L}}\times(0,T)\\
(2\mu\nabla_S u+(\lambda\mathrm{div}u-\alpha p)\mathrm{I})\mathsf{n}=-F,\qquad \nabla p\cdot\mathsf{n}=0,&\quad\text{on }\Gamma_{\mathtt{T}}\times(0,T)\\
(2\mu\nabla_S u+(\lambda\mathrm{div}u-\alpha p)\mathrm{I})\mathsf{n}=0,\qquad \nabla p\cdot\mathsf{n}=0,&\quad\text{on }\Gamma_{\mathtt{RB}}\times(0,T)\\
\alpha\mathrm{div}u+\sigma p=0,&\quad\text{in }\Omega\times\{0\}
\end{aligned}$$

with $\Gamma_{\mathtt{L}}$, $\Gamma_{\mathtt{T}}$ and $\Gamma_{\mathtt{RB}}$ as on the left side of Figure 5.1.

In analogy with [32, Section 10.1] and [44], we set

$$T=0.005,\qquad F=1$$

as well as [2]

$$\mu=3571.4,\quad \lambda=14286,\quad \alpha=0.93,\quad \sigma=0,\quad \kappa=10^{-7}.$$

For the time discretization, we consider $J=5$ intervals of equal size. For the space discretization, we use the mesh on the right side of Figure 5.1 involving 5861 degrees of freedom, whose number is computed as in (5.4). Since the exact solution is unknown in this case, we confine ourselves to a qualitative investigation. More

[2]More precisely [32, 44] set the Young's modulus to $E=10^4$ and the Poisson's ratio to $\nu=0.4$.

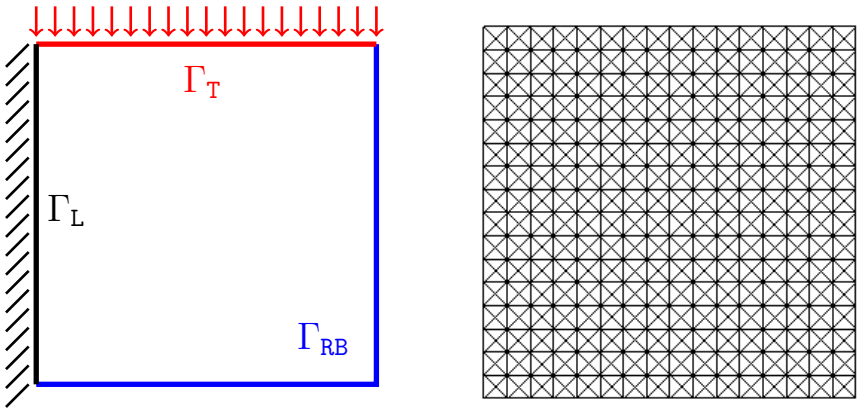


FIGURE 5.1. Boundary conditions (left) and computational mesh (right) for the cantilever bracket problem.

precisely, we plot the approximate pressure at the final time along four vertical lines at the following abscissas

$$x_1 = 0.26, \qquad x_2 = 0.33, \qquad x_3 = 0.4, \qquad x_4 = 0.45.$$

The plot in Figure 5.2 is qualitatively similar to the corresponding ones in [32, 44] and, in particular, no pressure oscillations are observed.

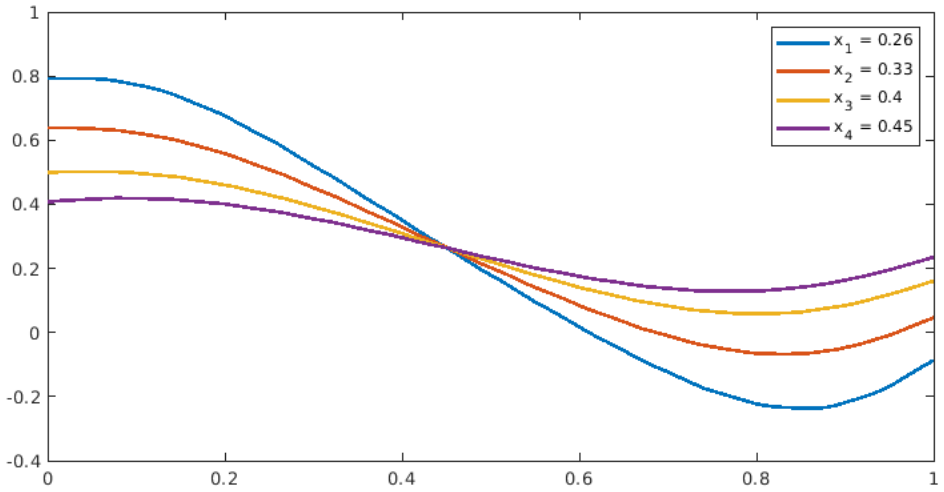


FIGURE 5.2. Pressure profile in the cantilever bracket problem along the vertical lines $x = x_j$, $j = 1, \ldots, 4$.

**Funding.** Christian Kreuzer gratefully acknowledges support by the DFG research grant KR 3984/5-2. Pietro Zanotti was supported by the PRIN 2022 PNRR project "Uncertainty Quantification of coupled models for water flow and contaminant

transport" (No. P2022LXLYY), financed by the European Union – Next Generation EU and by the GNCS-INdAM project CUP E53C23001670001.


## References

[1] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2001/02), pp. 1749–1779.

[2] I. Babuška, *Error-bounds for finite element method*, Numer. Math., 16 (1970/1971), pp. 322–333.

[3] L. Berger, R. Bordas, D. Kay, and S. Tavener, *Stabilized lowest-order finite element approximation for linear three-field poroelasticity*, SIAM J. Sci. Comput., 37 (2015), pp. A2222–A2245.

[4] A. Berger, R. Scott, and G. Strang, *Approximate boundary conditions in the finite element method*, Symposia Mathematica, Vol. X (Convegno di Analisi Numerica, INDAM, Rome), (1972), pp. 295–313.

[5] D. Boffi, M. Botti, and D. A. Di Pietro, *A nonconforming high-order method for the Biot problem on general meshes*, SIAM J. Sci. Comput., 38 (2016), pp. A1508–A1537.

[6] L. Botti, M. Botti, and D. A. Di Pietro, *An abstract analysis framework for monolithic discretisations of poroelasticity with application to hybrid high-order methods*, Comput. Math. Appl., 91 (2021), pp. 150–175.

[7] S. Brenner, and L.Y. Sung, *Virtual enriching operators*, Calcolo, 56 (2018), 25 pp.

[8] R. Bürger, S. Kumar, D. Mora, R. Ruiz-Baier, and N. Verma, *Virtual element methods for the three-field formulation of time-dependent linear poroelasticity*, Adv. Comput. Math., 47 (2021), Paper No. 2, 37 pp.

[9] Y. Chen, Y. Luo, and M. Feng, *Analysis of a discontinuous Galerkin method for the Biot's consolidation problem*, Appl. Math. Comput., 219 (2013), pp. 9043–9056.

[10] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. Studies in Mathematics and its Applications, Vol. 4.

[11] L. Diening, J. Storn, and T. Tscherpel, *On the Sobolev and $L^p$-stability of the $L^2$-projection*, SIAM J. Numer. Anal., 59 (2021), pp. 2571–2607.

[12] L. Diening, J. Storn, and T. Tscherpel, *Grading of Triangulations Generated by Bisection*, arXiv:2305.05742

[13] A. Ern and J.-L. Guermond, *Finite elements I—Approximation and interpolation*, vol. 72 of Texts in Applied Mathematics, Springer, Cham, 2021.

[14] A. Ern and J.-L. Guermond, *Finite elements II—Galerkin approximation, elliptic and mixed PDEs*, vol. 73 of Texts in Applied Mathematics, Springer, Cham, 2021.

[15] A. Ern and J.-L. Guermond, *Finite elements III—first-order and time-dependent PDEs*, vol. 74 of Texts in Applied Mathematics, Springer, Cham, 2021.

[16] M. Feischl, *Inf-sup stability implies quasi-orthogonality*, Math. Comp., 91 (2022), pp. 2059–2094.

[17] E. H. Georgoulis, P. Houston, and J. Virtanen, *An a posteriori error indicator for discontinuous Galerkin approximations of fourth-order elliptic problems*, IMA J. Numer. Anal., 31 (2011), pp. 281–298.

[18] J. B. Haga, H. Osnes, H. P. Lantangen, *On the causes of pressure oscillations in low-permeable and low-compressible porous media*, Int. J. Numer. Anal. Meth. Geomech. 36 (2012)

[19] K.-J. Heine, D. Köster, O. Kriessl, A. Schmidt, and K. Siebert, *Alberta - an adaptive hierachical finite element toolbox*. http://www.alberta-fem.de.

[20] R. Hiptmair, *Operator preconditioning*, Comput. Math. Appl., 52 (2006), pp. 699–706.
[21] Q. Hong and J. Kraus, *Parameter-robust stability of classical three-field formulation of Biot's consolidation model*, Electron. Trans. Numer. Anal., 48 (2018), pp. 202–226.
[22] G. Kanschat and B. Riviere, *A finite element method with strong mass conservation for Biot's linear consolidation model*, J. Sci. Comput., 77 (2018), pp. 1762–1779.
[23] A. Khan and P. Zanotti, *A nonsymmetric approach and a quasi-optimal and robust discretization for the Biot's model*, Math. Comp., 91 (2022), pp. 1143–1170.
[24] C. Kreuzer and P. Zanotti, *Inf-sup theory for the quasi-static Biot's equations in poroelasticity*, arXiv preprint arXiv:2407.02932, (2024).
[25] Y. Li and L. T. Zikatanov, *Residual-based a posteriori error estimates of mixed methods for a three-field Biot's consolidation model*, IMA J. Numer. Anal., 42 (2022), pp. 620–648.
[26] J. J. Lee, *Robust error analysis of coupled mixed methods for Biot's consolidation model*, J. Sci. Comput., 69 (2016), pp. 610–632.
[27] J. J. Lee, K.-A. Mardal, and R. Winther, *Parameter-robust discretization and preconditioning of Biot's consolidation model*, SIAM J. Sci. Comput., 39 (2017), pp. A1–A24.
[28] C. G. Makridakis, *On the Babuška–Osborn approach to finite element analysis: $L^2$ estimates for unstructured meshes*, Numer. Math., 139 (2018), pp. 831–844.
[29] K.-A. Mardal, M. E. Rognes, and T. B. Thompson, *Accurate discretization of poroelasticity without Darcy stability: Stokes-Biot stability revisited*, BIT, 61 (2021), pp. 941–976.
[30] K.-A. Mardal and R. Winther, *Preconditioning discretizations of systems of partial differential equations*, Numer. Linear Algebra Appl., 18 (2011), pp. 1–40.
[31] M. A. Murad, V. Thomée, and A. F. D. Loula, *Asymptotic behaviour of semidiscrete finite-element approximations of Biot's consolidation problem*, SIAM Numer. Anal., 3 (1996), pp. 1065–1083.
[32] P. J. Phillips, *Finite element methods in linear poroelasticity: theoretical and computational results*, Ph.D. thesis, University of Texas at Austin (2005).
[33] P. J. Phillips and M. F. Wheeler, *A coupling of mixed and discontinuous Galerkin finite-element methods for poroelasticity*, Comput. Geosci., 12 (2008), pp. 417–435.
[34] C. Rodrigo, F. J. Gaspar, X. Hu, and L. T. Zikatanov, *Stability and monotonicity for some discretizations of the Biot's consolidation model*, Comput. Methods Appl. Mech. Engrg., 298 (2016), pp. 183–204.
[35] A. Schmidt and K. G. Siebert, *Design of Adaptive Finite Element Software*, vol. 42 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin, 2005. The Finite Element Toolbox ALBERTA.
[36] D. Schötzau and C. Schwab, *Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method*, SIAM J. Numer. Anal., 38 (2000), pp. 837–875.
[37] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[38] R. E. Showalter, *Diffusion in poro-elastic media*, J. Math. Anal. Appl., 251 (2000), pp. 310–340.
[39] O. Steinbach and M. Zank, *A generalized inf-sup stable variational formulation for the wave equation*, J. Math. Anal. Appl., 505 (2022), Paper No. 125457, 24 pp.
[40] F. Tantardini, *Quasi-optimality in the backward Euler-Galerkin method for linear parabolic problems*, PhD thesis, Università degli Studi di Milano, Italy, 2014.
[41] F. Tantardini and A. Veeser, *The $L^2$-projection and quasi-optimality of Galerkin methods for parabolic equations*, SIAM J. Numer. Anal., 54 (2016), pp. 317–340.
[42] A. Veeser and P. Zanotti, *Quasi-optimal nonconforming methods for symmetric elliptic problems. II—Overconsistency and classical nonconforming elements*, SIAM J. Numer. Anal., 57 (2019), pp. 266–292.
[43] R. Verfürth, *A Posteriori Error Estimation Techniques for Finite Element Methods*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2013.
[44] S.-Y. Yi, *A coupling of nonconforming and mixed finite element methods for Biot's consolidation model*, Numer. Methods Partial Differential Equations, 29 (2013), pp. 1749–1777.
[45] S.-Y. Yi, *A study of two modes of locking in poroelasticity*, SIAM J. Numer. Anal., 55 (2017), pp. 1915–1936.
[46] A. Ženíšek, *The existence and uniqueness theorem in Biot's consolidation theory*, Aplikace matematiky, 29 (1984), pp. 194–211.

## Appendix A. Proof of Lemma 4.14

The construction of a linear operator $\mathcal{S} : S^1_{\mathsf{k}} \to H^2(\Omega)$ satisfying all properties listed in Lemma 4.14 is rather technical but it makes use only of classical tools from finite element analysis. To simplify the discussion as much as possible, we detail the construction only for the lowest degree and dimension, namely $\mathsf{k} = 2$ and $d = 2$. This case minimizes the technical aspects but, at the same time, it illustrates all relevant issues. We address the extension to the other cases at the end of the appendix, in Remark A.4.

Our construction builds upon two preliminary steps. In the first one, we exhibit an operator mapping $S^1_{\mathsf{k}}$ to $H^2(\Omega)$, which satisfies the first, second and last conditions in Lemma 4.14. This can be done by some sort of averaging into a $H^2(\Omega)$-conforming finite element space. Similar operators exist in the literature but typically differ in the boundary conditions, see e.g. [7, 17]. In the second step we show that the third and fourth conditions in Lemma 4.14 can be enforced by combinations of smooth bubble functions. Both techniques are common in the a posteriori analysis of (nonconforming) discretizations of fourth-order equations.

Let us begin with the first step. For the sake of presentation, we use the abbreviations

$$\Gamma_E = \overline{\Gamma}_{p,E}, \quad \mathfrak{F}_E = \mathfrak{F}_{p,E} \qquad \text{and} \qquad \Gamma_N = \overline{\Gamma}_{p,N}, \quad \mathfrak{F}_N = \mathfrak{F}_{p,N}$$

in what follows. The condition (4.21a) can be rephrased as

$$S^1_2 \ni Q = 0 \quad \text{on } \Gamma_E \qquad \Longrightarrow \qquad \mathcal{S}Q = 0 \quad \text{on } \Gamma_E \tag{A.1}$$

cf. (2.4) and (4.7). The condition (4.21a) actually requires also the inclusion in $L^2_0(\Omega)$ in some cases, but that can be obtained by enforcing (4.21c), so we do not care about it here. Hence (A.1) and (4.21b) are nothing else than Dirichlet and Neumann boundary conditions on $\Gamma_E$ and $\Gamma_N$, respectively.

We construct a first operator $\mathcal{S}_1 : S^1_2 \to H^2(\Omega)$ by mapping into the so-called HCT space, which is obtained by the finite element shown in Figure A.1, see also [10, Section 6.1]. The degrees of freedom in this space are the normal derivative at the midpoint $m_{\mathsf{F}}$ of each edge $\mathsf{F} \in \mathfrak{F}$ as well as the evaluation of the function and of its gradient at each vertex. For convenience, we denote by $\mathfrak{V}$ the set of all vertices.

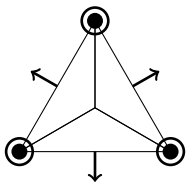

Figure A.1. Schematic representations of the HCT finite element

For the normal derivative at the edges, we set

$$\nabla \mathcal{S}_1 Q(m_{\mathsf{F}}) \cdot \mathsf{n} = \begin{cases} 0 & \text{if } \mathsf{F} \in \mathfrak{F}_N, \\ \nabla Q_{|\mathsf{T}_{\mathsf{F}}}(m_{\mathsf{F}}) \cdot \mathsf{n} & \text{otherwise,} \end{cases} \qquad \mathsf{F} \in \mathfrak{F} \tag{A.2a}$$

where $\mathsf{T}_{\mathsf{F}} \in \mathfrak{T}$ is a (arbitrary but fixed) triangle in the mesh containing $\mathsf{F}$. For the point values, we just exploit the continuity of $Q$ and take

$$\mathcal{S}_1 Q(\mathsf{v}) = Q(\mathsf{v}), \qquad \mathsf{v} \in \mathfrak{V}. \tag{A.2b}$$

We set the gradient at the interior vertices in analogy with the second case in (A.2a) to

$$\nabla \mathcal{S}_1 Q(\mathsf{v}) = \nabla Q_{|\mathsf{T}_\mathsf{v}}(\mathsf{v}), \qquad \mathsf{v} \in \mathfrak{V} \cap \Omega \tag{A.2c}$$

where $\mathsf{T}_\mathsf{v} \in \mathfrak{T}$ is a (arbitrary but fixed) triangle in the mesh containing $\mathsf{v}$.

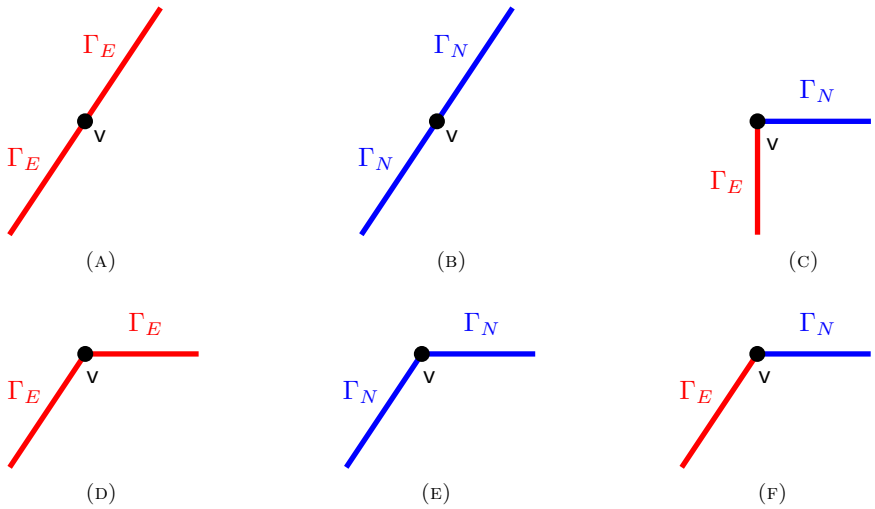


FIGURE A.2. Different cases for the definition of the gradient of $\mathcal{S}_1 Q$ at a boundary vertex $\mathsf{v}$. Note that the highlighted edges are aligned in (A)-(B) but not in (D)-(E) and that they meet at a right angle in (C) but not in (F)

The definition of the gradient at the other vertices is more involved. Indeed, the boundary conditions suggest to treat the normal and the tangential components differently. Of course, this is especially critical when $\mathsf{v}$ is a corner of $\Omega$ and/or if it lies at the intersection of $\Gamma_E$ and $\Gamma_N$. Figure A.2 summarizes all possible cases. We treat all cases simultaneously by Algorithm 1. The underlying concept is that we first set the gradient to zero in the normal direction(s), so as to comply with (4.21b). Then, we treat the tangent direction(s) to enforce (A.1). Moreover, special attention is needed to make sure that the gradient is neither over- nor under-determined.

Within Algorithm 1, we denote by $\mathsf{n}_\mathsf{F}$ and $\mathsf{t}_\mathsf{F}$, respectively, the outward normal unit vector and a tangent unit vector of a boundary edge $\mathsf{F} \in \mathfrak{F} \cap \partial\Omega$. We make use also of the unique triangle $\mathsf{T}_\mathsf{F} \in \mathfrak{T}$ which contains $\mathsf{F}$. The set $\mathfrak{L}$ is the linear space spanned by the directions used in the definition of the gradient, so it grows from $\{0\}$ to $\mathbb{R}^2$.

*Remark* A.1 (Simplified averaging). In (A.2a), (A.2c) and in Algorithm 1 we define $\mathcal{S}_1 Q$ in terms of the restriction of $Q$ to only one triangle in mesh. Therefore, we refer to $\mathcal{S}_1$ as a *simplified averaging*, so as to distinguish it from classical averaging operators, which take averages over all triangles containing the support of the degree of freedom under examination. Apart of the different definition, the simplified averagings reproduce all relevant properties of the classical ones, cf. [42, Section 3].

**Algorithm 1** Gradient of $\mathcal{S}_1 Q$ at boundary vertices

**Require:** $\mathsf{v} \in \mathfrak{V} \cap \partial\Omega$ boundary vertex
**Provide**: $\nabla \mathcal{S}_1 Q(\mathsf{v})$
$\mathfrak{L} \leftarrow \{0\}$
**for all** $\mathsf{F} \in \mathfrak{F}_N$ with $\mathsf{v} \in \mathsf{F}$ **do**
  $\nabla \mathcal{S}_1 Q(\mathsf{v}) \cdot \mathsf{n}_\mathsf{F} \leftarrow 0$
  $\mathfrak{L} \leftarrow \mathfrak{L} + \operatorname{span}\{\mathsf{n}_\mathsf{F}\}$
**end for**
**for all** $\mathsf{F} \in \mathfrak{F}_E$ with $\mathsf{v} \in \mathsf{F}$ **do**
  **if** $\mathsf{t}_\mathsf{F} \notin \mathfrak{L}$ **then**
    $\nabla \mathcal{S}_1 Q(\mathsf{v}) \cdot \mathsf{t}_\mathsf{F} \leftarrow \nabla Q_{|\mathsf{T}_\mathsf{F}}(\mathsf{v}) \cdot \mathsf{t}_\mathsf{F}$
    $\mathfrak{L} \leftarrow \mathfrak{L} + \operatorname{span}\{\mathsf{t}_\mathsf{F}\}$
  **end if**
**end for**
**if** $\dim(\mathfrak{L}) = 1$ **then**
  choose $\mathsf{w} \in \mathbb{R}^2$ with $\mathsf{w} \perp \mathfrak{L}$
  choose $\mathsf{T} \in \mathfrak{T}$ with $\mathsf{v} \in \mathsf{T}$
  $\nabla \mathcal{S}_1 Q(\mathsf{v}) \cdot \mathsf{w} \leftarrow \nabla Q_{|\mathsf{T}}(\mathsf{v}) \cdot \mathsf{w}$
**end if**

**Lemma A.2** (Simplified averaging)**.** *The linear operator $\mathcal{S}_1 : S_2^1 \to H^2(\Omega)$ defined by* (A.2) *and Algorithm 1 satisfies* (A.1)*,* (4.21b) *and* (4.21e)*. In particular, the hidden constant in* (4.21e) *depends only on the quantities mentioned in Remark 4.1.*

*Proof.* Let $\mathsf{F} \in \mathfrak{F} \cap \partial\Omega$ be a boundary face. For $Q \in S_2^1$, the restriction of $\mathcal{S}_1 Q$ to $\mathsf{F}$ is a third-order polynomial, owing to the definition of the HCT element. If $Q = 0$ on $\Gamma_E$ and $\mathsf{F} \in \mathfrak{F}_E$, then both $\mathcal{S}_1 Q$ and its tangential derivative along $\mathsf{F}$ vanish at the endpoints of $\mathsf{F}$, thanks to (A.2b) and Algorithm 1. This proves that $\mathcal{S}_1 Q$ vanishes on $\mathsf{F}$ and verifies (A.1). Similarly, note that the normal derivative of $\mathcal{S}_1 Q$ on $\mathsf{F}$ is a second-order polynomial. If $\mathsf{F} \in \mathfrak{F}_N$, then the normal derivative vanishes on $\mathsf{F}$, because it vanishes at the midpoint and at the endpoints of $\mathsf{F}$, in view of (A.2a) and Algorithm 1. This verifies (4.21b).

Regarding the claimed stability estimate, let $\mathsf{T} \in \mathfrak{T}$. The scaling properties of the HCT basis functions (see e.g. [42, Lemma 3.14]) imply

$$\|D^j(Q - \mathcal{S}_1 Q)\|_\mathsf{T}^2 \eqsim \mathsf{h}_\mathsf{T}^{2-2j} \sum_{\mathsf{v} \in \mathfrak{V} \cap \mathsf{T}} |(Q_{|\mathsf{T}} - \mathcal{S}_1 Q)(\mathsf{v})|^2 + \mathsf{h}_\mathsf{T}^{4-2j} \Big( \sum_{\mathsf{v} \in \mathfrak{V} \cap \mathsf{T}} |\nabla(Q_{|\mathsf{T}} - \mathcal{S}_1 Q)(\mathsf{v})|^2 + \sum_{\mathsf{F} \in \mathfrak{F} \cap \mathsf{T}} |\nabla(Q_{|\mathsf{T}} - \mathcal{S}_1 Q)(m_\mathsf{F}) \cdot \mathsf{n}|^2 \Big)$$

for $j \in \{0, 1, 2\}$. The first summand on the right-hand side vanishes because of (A.2b), while the other ones are bounded in terms of jumps. Indeed, the definition (A.2a) immediately yields

$$|\nabla(Q_{|\mathsf{T}} - \mathcal{S}_1 Q)(m_\mathsf{F}) \cdot \mathsf{n}_{|\mathsf{F}}| \le \begin{cases} 0 & \text{if } \mathsf{F} \in \mathfrak{F}_E, \\ |[\![\nabla Q]\!]_{|\mathsf{F}} \cdot \mathsf{n}| & \text{otherwise,} \end{cases}$$

for all edges $\mathsf{F} \in \mathfrak{F} \cap \mathsf{T}$. Similarly, according to Algorithm 1, we have

$$|\nabla(Q_{|\mathsf{T}} - \mathcal{S}_1 Q)(\mathsf{v})| \le \sum_{\mathsf{F} \in \mathfrak{F}^i \cup \mathfrak{F}_N, \mathsf{v} \in \mathsf{F}} |[\![\nabla Q]\!]_{|\mathsf{F}} \cdot \mathsf{n}|.$$

for all vertices $\mathsf{v} \in \mathfrak{V} \cap \mathsf{T}$. The proof of this estimate is tedious, as it must be verified for each case in Figure A.2, but it builds only upon a classical argument for (simplified) averaging operators, see e.g. [42, Lemma 3.1]. We insert the two bounds above into the previous equivalence. Then, we obtain (4.21e) by an inverse estimate on the edges, noticing that $\mathsf{h}_\mathsf{T} \eqsim \{\!\{h\}\!\}_{|\mathsf{F}}$ for all edges $\mathsf{F}$ touching $\mathsf{T}$. □

For the second preliminary step in our construction, we make use of bubble functions. In particular, we use 'element' bubbles to enforce (4.21c). For $\mathsf{T} \in \mathfrak{T}$, let $b_\mathsf{T} \in P_6(\mathsf{T})$ be the unique polynomial of degree 6 in $\mathsf{T}$ such that both $b_\mathsf{T}$ and $\nabla b_\mathsf{T}$ vanish on $\partial\mathsf{T}$ and $\int_\mathsf{T} b_\mathsf{T} = 1$, cf. [43, Section 3.2.5]. We extend $b_\mathsf{T}$ to an $H^2(\Omega)$ function by zero.

Similarly, we use 'edge' bubbles to enforce (4.21d) on interior edges. Thus, for $\mathsf{F} \in \mathfrak{F}^i$, we denote by $\mathsf{T}_1$ and $\mathsf{T}_2$ the two triangles in the mesh containing $\mathsf{F}$. Let $b_\mathsf{F} \in P_8(\mathsf{T}_1 \cup \mathsf{T}_2)$ be the unique polynomial of degree 8 on $\mathsf{T}_1 \cup \mathsf{T}_2$ such that both $b_\mathsf{F}$ and $\nabla b_\mathsf{F}$ vanish on $\partial(\mathsf{T}_1 \cup \mathsf{T}_2)$ and $\int_\mathsf{F} b_\mathsf{F} = 1$, cf. [43, Section 3.2.5]. As before, we extend $b_\mathsf{F}$ to an $H^2(\Omega)$ function by zero.

The condition (4.21d) must be enforced also on the edges in $\mathfrak{F}_N$. Here we must employ bubble functions with vanishing normal derivative on the respective edge, because of (4.21b). Let $\mathsf{T} \in \mathfrak{T}$ be the unique triangle in the mesh containing $\mathsf{F}$ and define $\mathsf{T}'$ by mirroring $\mathsf{T}$ through $\mathsf{F}$. We define $b_\mathsf{F}$ as before on $\mathsf{T} \cup \mathsf{T}'$. Hence, the symmetry of the support implies $\nabla b_\mathsf{F} \cdot \mathsf{n} = 0$ on $\mathsf{F}$. Then, we extend $b_{\mathsf{F}|\mathsf{T}}$ to $H^2(\Omega)$ by zero.

Having all these bubble functions at hand, we define $\mathcal{S}_2 : H^1(\Omega) \to H^2(\Omega)$ by

$$\mathcal{S}_2 Q := \mathcal{S}_\mathfrak{F} Q + \mathcal{S}_\mathfrak{T}(Q - \mathcal{S}_\mathfrak{F} Q) \tag{A.3}$$

for $Q \in H^1(\Omega)$, where

$$\mathcal{S}_\mathfrak{F} Q := \sum_{\mathsf{F} \in \mathfrak{F} \cup \mathfrak{F}_N} \left( \int_\mathsf{F} Q \right) b_\mathsf{F} \qquad \text{and} \qquad \mathcal{S}_\mathfrak{T} Q := \sum_{\mathsf{T} \in \mathfrak{T}} \left( \int_\mathsf{T} Q \right) b_\mathsf{T}.$$

The next lemma summarizes the properties of $\mathcal{S}_2$ that are relevant to our purpose.

**Lemma A.3** (Bubble operator)**.** *The operator $\mathcal{S}_2 : H^1(\Omega) \to H^2(\Omega)$ defined by* (A.3) *satisfies* (4.21b),(4.21c) *and* (4.21d)*. Moreover $\mathcal{S}_2 Q$ vanishes on $\Gamma_E$ and satisfies the estimate*

$$\|\mathcal{S}_2 Q\|_\mathsf{T} \lesssim \|Q\|_\mathsf{T} + \mathsf{h}_\mathsf{T} \|\nabla Q\|_\mathsf{T}$$

*for all $Q \in H^1(\Omega)$ and $\mathsf{T} \in \mathfrak{T}$. The hidden constant depends only on the quantities mentioned in Remark 4.1.*

*Proof.* The properties (4.21b),(4.21c) and (4.21d) and the identity $\mathcal{S}_2 Q = 0$ on $\Gamma_E$ for $Q \in H^1(\Omega)$ hold true by construction. The local stability estimate follows by the scaling of the bubble functions. We refer to [42, Lemma 3.8] where a similar result is proved. □

We conclude the construction of the operator $\mathcal{S} : S_2^1 \to H^2(\Omega)$ announced in Lemma 4.14 by combining the operators $\mathcal{S}_1$ and $\mathcal{S}_2$ introduced in the two steps above. More precisely, we set

$$\mathcal{S} Q := \mathcal{S}_1 Q + \mathcal{S}_2(Q - \mathcal{S}_1 Q)$$

for $Q \in S_2^1$. The properties of $\mathcal{S}_1$ and $\mathcal{S}_2$ listed in Lemmas A.2-A.3 imply that (4.21a)-(4.21d) hold true. Regarding the local estimate (4.21e), note that we have

$$\|D^j(Q - \mathcal{S} Q)\|_\mathsf{T} \lesssim \|D^j(Q - \mathcal{S}_1 Q)\|_\mathsf{T} + \mathsf{h}_\mathsf{T}^{-j} \|Q - \mathcal{S}_1 Q\|_\mathsf{T}.$$

for $j \in \{0,1,2\}$ and $\mathsf{T} \in \mathfrak{T}$, by Lemma A.3 and inverse estimates. We conclude by invoking the stability of $\mathcal{S}_1$ established in Lemma A.2.

*Remark* A.4 (Higher degree/dimension). Our construction extends to higher polynomial degree $\mathsf{k} \geq 2$ and/or to higher space dimension $d \geq 2$ up to some additional technicalities. The construction of the simplified averaging $\mathcal{S}_1$ requires $H^2(\Omega)$-conforming spaces, that can be obtained by, e.g., higher-order HCT finite elements for $d = 2$ (cf. [17, Section 3] ) or the virtual elements in [7] for $d = 2, 3$. The operator $\mathcal{S}_2$ employs the same bubble functions mentioned above, but one has to define it via the solution of local problems on the simplices and on the faces of the mesh, in the vein of [42, Lemma 3.8].

TU Dortmund, Fakultät für Mathematik, D-44221 Dortmund, Germany
*Email address*: christian.kreuzer@tu-dortmund.de

Università degli Studi di Pavia, Dipartimento di Matematica,, 27100 Pavia, Italy
*Email address*: pietro.zanotti@unipv.it