



# An energy-efficient GMRES–multigrid solver for space-time finite element computation of dynamic poroelasticity

Mathias Anselmann<sup>1</sup> · Markus Bause<sup>1</sup> · Nils Margenberg<sup>1</sup> · Pavel Shamko<sup>1</sup>

Received: 25 March 2023 / Accepted: 10 February 2024  
© The Author(s) 2024

## Abstract

We present and analyze computationally Geometric MultiGrid (GMG) preconditioning techniques for Generalized Minimal RESidual (GMRES) iterations to space-time finite element methods (STFEMs) for a coupled hyperbolic–parabolic system modeling, for instance, flow in deformable porous media. By using a discontinuous temporal test basis, a time marching scheme is obtained. Higher order approximations that offer the potential to inherit most of the rich structure of solutions to the continuous problem on computationally feasible grids increase the block partitioning dimension of the algebraic systems, comprised of generalized saddle point blocks. Our V-cycle GMG preconditioner uses a local Vanka-type smoother. Its action is defined in an exact mathematical way. Due to nonlocal coupling mechanisms of 348 unknowns, the smoother is applied on patches of elements. This ensures damping of higher order error frequencies. By numerical experiments of increasing complexity, the efficiency of the solver for STFEMs of different polynomial order is illustrated and confirmed. Its parallel scalability is analyzed. Beyond this study of classical performance engineering, the solver’s energy efficiency is investigated as an additional and emerging dimension in the design and tuning of algorithms on the hardware.

**Keywords** Poroelasticity · Dynamic Biot model · Space-time finite element approximation · Geometric multigrid method · Vanka-type smoother · Computational efficiency · Energy consumption

## 1 Introduction

### 1.1 Mathematical model

We present and investigate numerically a geometric multigrid preconditioning technique, based on a local Vanka-type smoother, for solving by GMRES iterations the linear systems that arise from space-time finite element discretizations of the coupled hyperbolic–parabolic system of dynamic poroelasticity

$$\rho \partial_t^2 \mathbf{u} - \nabla \cdot (\mathbf{C}\boldsymbol{\varepsilon}(\mathbf{u})) + \alpha \nabla p = \rho \mathbf{f}, \quad \text{in } \Omega \times (0, T], \quad (1.1a)$$

$$c_0 \partial_t p + \alpha \nabla \cdot \partial_t \mathbf{u} - \nabla \cdot (\mathbf{K} \nabla p) = g, \quad \text{in } \Omega \times (0, T], \quad (1.1b)$$

$$\mathbf{u}(0) = \mathbf{u}_0, \quad \partial_t \mathbf{u}(0) = \mathbf{u}_1,$$

$$p(0) = p_0, \quad \text{in } \Omega \times \{0\}, \quad (1.1c)$$

$$\mathbf{u} = \mathbf{u}_D, \quad \text{on } \Gamma_u^D \times (0, T], \quad (1.1d)$$

$$-(\mathbf{C}\boldsymbol{\varepsilon}(\mathbf{u}) - \alpha p \mathbf{E})\mathbf{n} = \mathbf{t}_N, \quad \text{on } \Gamma_u^N \times (0, T], \quad (1.1e)$$

$$p = p_D, \quad \text{on } \Gamma_p^D \times (0, T], \quad (1.1f)$$

$$-\mathbf{K} \nabla p \cdot \mathbf{n} = p_N, \quad \text{on } \Gamma_p^N \times (0, T]. \quad (1.1g)$$

In (1.1),  $\Omega \subset \mathbb{R}^d$ , with  $d \in \{2, 3\}$ , is an open bounded Lipschitz domain with outer unit normal vector  $\mathbf{n}$  to the boundary  $\partial\Omega$  and  $T > 0$  is the final time point. We let  $\partial\Omega = \overline{\Gamma_u^D} \cup \overline{\Gamma_u^N}$  and  $\partial\Omega = \overline{\Gamma_p^D} \cup \overline{\Gamma_p^N}$  with (open) portions  $\Gamma_u^D$  and  $\Gamma_p^D$  of non-zero measure. Important applications of the model (1.1), that is studied as a prototype system, arise in poroelasticity; cf. [66] and [16–18]. In poroelasticity, Eq. (1.1) are referred to as the dynamic Biot model. The system (1.1) is used to describe flow of a slightly compressible viscous fluid through a deformable porous matrix. The small deformations of the matrix are described by the Navier equations of linear elasticity, and the diffusive fluid flow is described by Duhamel’s equation. The unknowns are the effective solid phase displacement  $\mathbf{u}$  and the effective fluid pressure  $p$ . The quantity

✉ Mathias Anselmann  
anselmann@hsu-hh.de

<sup>1</sup> Faculty of Mechanical and Civil Engineering, Helmut Schmidt University, Holstenhofweg 85, 22043 Hamburg, Germany

$\boldsymbol{\varepsilon}(\mathbf{u}) := (\nabla \mathbf{u} + (\nabla \mathbf{u})^\top)/2$  denotes the symmetrized gradient or strain tensor and  $\mathbf{E} \in \mathbb{R}^{d,d}$  is the identity matrix. Further,  $\rho$  is the effective mass density,  $\mathbf{C}$  is Gassmann's fourth order effective elasticity tensor,  $\alpha$  is Biot's pressure-storage coupling tensor,  $c_0$  is the specific storage coefficient and  $\mathbf{K}$  is the permeability field. For brevity, the positive quantities  $\rho > 0$ ,  $\alpha > 0$  and  $c_0 > 0$  as well as the tensors  $\mathbf{C}$  and  $\mathbf{K}$  are assumed to be constant in space and time. The tensors  $\mathbf{C}$  and  $\mathbf{K}$  are assumed to be symmetric and positive definite,

$$\exists k_0 > 0 \forall \boldsymbol{\xi} = \boldsymbol{\xi}^\top \in \mathbb{R}^{d,d} : \quad \sum_{i,j,k,l=1}^d \xi_{ij} C_{ijkl} \xi_{kl} \geq k_0 \sum_{j,k=1}^d |\xi_{jk}|^2, \quad (1.2a)$$

$$\exists k_1 > 0 \forall \boldsymbol{\xi} \in \mathbb{R}^d : \quad \sum_{i,j=1}^d \xi_i K_{ij} \xi_j \geq k_1 \sum_{i=1}^d |\xi_i|^2. \quad (1.2b)$$

Well-posedness of (1.1) is ensured; cf., e.g. [43, 64, 67]. This can be shown by different mathematical techniques, by semigroup methods [43, Thm. 2.2], Rothe's method and compactness arguments [67, Thm. 4.18 and Cor. 4.33] and Picard's theorem [64, Thm. 6.2.1]. In these works, boundary conditions different to the ones in (1.1d) to (1.1g) are partly used. System (1.1) is studied as a prototype poromechanical model. In order to enhance physical realism, generalizations of the model (1.1) have been developed and investigated in, e.g. [19, 23, 58]. We note that the system (1.1) is also formally equivalent to the classical coupled thermoelasticity system which describes the flow of heat through an elastic structure; cf. [22, 43, 52].

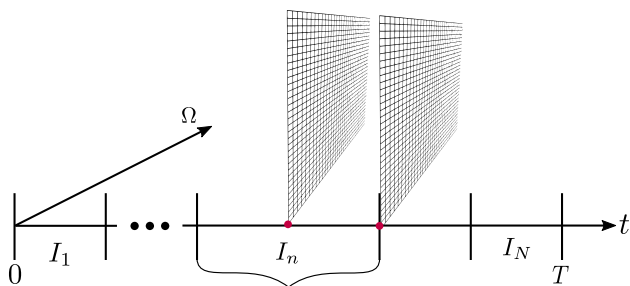
## 1.2 Space-time finite element and multigrid techniques

The coupled hyperbolic–parabolic structure of the system (1.1) of partial differential equations adds a facet of complexity onto its numerical simulation. A natural and promising approach for the numerical approximation of coupled systems is obtained by STFEMs that are based on an uniform treatment of the space and time discretization by variational techniques. STFEMs enable the discretization of even complex coupling terms that involve combinations of temporal and spatial derivatives as in (1.1b). Moreover, STFEMs offer the natural construction of higher order schemes that achieve accurate results on computationally feasible grids with a minimum of numerical costs. Time discretizations of higher regularity can be designed by combining variational and collocation techniques; cf., e.g. [6]. Finally, space-time adaptivity based on a-posteriori error control by duality concepts

and multi-rate in time approaches become feasible; cf., e.g. [9, 10, 21, 70].

STFEMs have been constructed in different ways. Holistic space-time methods on completely unstructured space-times meshes have been proposed and analyzed; cf., e.g. [51, 69] and the references therein. They aim at exploiting efficiently the enormous compute power of modern massively parallel high performance architectures. Time-parallel time-integrations methods like PARAREAL [32] are closely related to these methods. A further class of STFEMs is based on time-marching schemes that are constructed by the choice of a discontinuous temporal test basis, usage of a tensor product space-time mesh and discretization of the resulting problems in the spatial variables; cf., e.g. [1, 3, 40, 41] and the references therein. Such methods offer high flexibility for the finite element discretization of the temporal and spatial variables. The existing technology of iterative linear solver can be reused or adapted to the resulting linear systems that are built from blocks mimicing lower order time discretizations; cf. (4.6) and (4.7). Combinations of either approaches also exist. Therein, tensor product space-time meshes are used, but all time steps are assembled in a global system matrix and computed fully coupled without any sequential progressing; cf., e.g. [27, 33]. Within the classes of schemes, their members can differ by the application of continuous and discontinuous finite element techniques. For second-order hyperbolic problems further approaches are addressed in [12, 29, 74].

STFEMs lead to large linear systems of equations. Their solution demands for highly efficient and robust iterative solvers, in particular if three space variables are involved. Different Algebraic Multigrid (AMG) and GMG methods have been considered and investigated for the solution of STFEM algebraic systems, either in holistic or time marching form. Also, multigrid methods have been used as preconditioners for Krylov subspace iterations, like GMRES, to enhance their robustness. For the application of multigrid techniques in the STFEM context we refer, for instance, to [3, 27, 31, 33, 39, 40, 42, 51, 59, 63, 70–73]. In [33, 59], block Jacobi smoothing factors and two-grid convergence factors for arbitrary order discontinuous Galerkin time discretizations of a holistic approach are investigated for parabolic problems by using exponential local Fourier mode analysis. Instead of an adaptive coarsening as proposed in [33, 39], a space-time multigrid method, using an adaptive smoothing strategy in combination with standard coarsening in both temporal and spatial domains, was proposed and investigated by local Fourier analysis for the heat equation in [31]. Therein, the multigrid method is robust for both first-order Euler and second-order Crank-Nicolson temporal discretization schemes. In general, GMG techniques are widely used and employed in many variants. Flow and saddle point problems are prominent applications; cf. [28, 47, 77]. Massively



**Fig. 1** Space-time mesh for a piecewise linear ( $k = 1$ ) discontinuous Galerkin time discretization and a Lagrange basis w.r.t. the  $k + 1$  Gauss–Radau quadrature points of  $I_n$

parallel implementations of GMG methods on modern architectures show excellent scalability properties and their high efficiency has been recognized in [34, 35, 54]. Analyses of GMG methods (cf., e.g. [28, 38, 55]) have been done in particular for linear systems in saddle point form, arising from mixed discretizations of the Stokes problem.

In this work, we use discontinuous Galerkin time discretization [75] with arbitrary polynomial order  $k \in \mathbb{N}$  (for short  $dG(k)$ ), recasted as a time marching scheme. Time interpolation on each subinterval  $I_n = (t_{n-1}, t_n]$  of the time mesh  $\mathcal{M}_\tau := \{I_1, \dots, I_N\}$ , is done in terms of a Lagrangian basis with respect to the Gauss–Radau quadrature points of  $I_n$ ; cf. Fig. 1. For the space discretization, inf-sup stable pairs of finite element spaces are applied. Alternative approaches are presented, for instance, in [37, 48]. Dirichlet boundary conditions are implemented in weak form. Two discrete systems differing by the treatment of the term  $\nabla \cdot \partial_t \mathbf{u}$  in (1.1b) are proposed. Well-posedness of the discrete problems is proved for arbitrary polynomial order in space and time. On each subinterval  $I_n$ , this discretization leads to a linear system of equations with a  $(k + 1) \times (k + 1)$  block matrix (cf. (4.6)), where each of the blocks  $A_{ab}$ , for  $a, b = 1, \dots, k + 1$ , exhibits the structure

$$A_{a,b} = \begin{pmatrix} \mathbf{A} & \mathbf{B}^\top \\ -\mathbf{B} & \mathbf{C} \end{pmatrix} \tag{1.3}$$

with suitably defined submatrices  $\mathbf{A}$ ,  $\mathbf{B}$  and  $\mathbf{C}$  in (1.3) and  $\mathbf{A}$  itself being of the form in (1.3) again. We note that  $A_{a,b}$  has a generalized saddle point form and is positive stable under certain conditions; cf. [14]. The block structure (1.3) of  $dG(k)$  time discretizations imposes an facet of complexity on the iterative solution of the systems. For the solver we propose and analyze numerically GMRES iterations that are preconditioned by a V-cycle GMG method. To the best of our knowledge, theoretical analyses of GMG methods for SFFEM block partitioned systems are still missing in the literature.

GMG methods exploit different mesh levels of the underlying problem in order to reduce different frequencies of the

error by employing a relatively cheap smoother on each grid level. Different iterative methods have been proposed in the literature as smoothing procedure; cf. [28] and the references therein. They range from low-cost methods like Richardson, Jacobi, and SOR applied to the normal equation of the linear system to collective smoothers, that are based on the solution of small local problems. Here, we use a Vanka-type smoother [47, 55, 80] of the family of collective methods. Numerical computations have shown that an elementwise application of the Vanka smoother fails to reduce the high frequencies of the error on the multigrid levels. The reason for this comes through interelement couplings of spatial degrees of freedom of the scalar variable  $p$  in (1.1). As a remedy, we propose the application of the Vanka-type smoother on cell patches that are linked to the grid nodes and built from four neighbored cells in two dimensions and eight neighbored elements in three dimensions, with appropriate adaptations for grid nodes close to or on the domain’s boundary. Further, an averaging of the patchwise updates and relaxation strategy in the smoothing steps are employed. Then an efficient damping of frequencies in the error on the multigrid hierarchy is obtained. This Vanka-type smoother is presented in a mathematically exact way and its performance properties are investigated by numerical experiments of increasing complexity. Our numerical experiments confirm that GMRES iterations that are preconditioned by the proposed GMG method converge at a desired rate that is (nearly) independent of the mesh sizes in space and time; cf. also [3, 4].

### 1.3 Energy efficiency

In the past, performance engineering and hardware engineering for large scale simulations of physical phenomena have been eclipsed by the longing for ever more performance where faster seemed to be the only paradigm. ”Classical” performance engineering has been applied to enhance, firstly, the efficiency of the current method on the target hardware or to find numerical alternatives that might better fit to the hardware in use and/or, secondly, to develop other numerical methods can be found to improve the numerical efficiency. Tuning both simultaneously is called hardware-oriented numerics in the literature; cf. [78, 79]. Since recently, a growing awareness of energy consumption in computational science, particularly, in extreme scale computing with a view to exascale computing has raised; cf., e.g. [62]. It has been observed that as a consequence of decades of performance-centric hardware development there is a huge gap between pure performance and energy efficiency. An analysis of our algorithm’s parallel scaling and energy consumption properties by performance models exceeds the scope of this work and would overburden it. However, since energy consumption of application codes on the available hardware is of growing awareness and a key

for future improvements, we study the energy consumption and parallel scaling properties of our algorithm and its implementation by three-dimensional numerical experiments. The development of a proper model that quantifies performance and energy efficiency in some appropriate metric and can be used for a code optimization still deserves research and is left as a work for the future.

### 1.4 Outline of the work

This work is organized as follows. In Sect. 2 we introduce our notation. In Sect. 3 the space-time finite element approximation of arbitrary order of (1.1) is derived and well-posedness of the fully discrete problem is proved. Our GMRES–GMG solver is introduced in Sect. 4. In Sect. 5 our performed numerical computations for analyzing the performance properties of the overall approach are presented. In Sect. 6 we end with a summary and conclusions. In the appendix, supplementary results are summarized.

## 2 Basic notation

In this work, standard notation is used. We denote by  $H^1(\Omega)$  the Sobolev space of  $L^2(\Omega)$  functions with first-order weak derivatives in  $L^2(\Omega)$ . Further,  $H^{-1}(\Omega)$  is the dual space of  $H_0^1(\Omega)$ , with the standard modification if the Dirichlet condition is prescribed on a part  $\Gamma^D \subset \Gamma$  of the boundary  $\partial\Omega$  only; cf. (1.1). The latter is not explicitly borne out by the notation of  $H^{-1}(\Omega)$ . It is always clear from the context. For vector-valued functions we write those spaces bold. By  $\langle \cdot, \cdot \rangle_S$  we denote the  $L^2(S)$  inner product for a domain  $S$ . For  $S = \Omega$ , we simply write  $\langle \cdot, \cdot \rangle$ . For the norms of the Sobolev spaces the notation is

$$\|\cdot\| := \|\cdot\|_{L^2}, \quad \|\cdot\|_1 := \|\cdot\|_{H^1}.$$

For short, we put

$$Q := L^2(\Omega) \quad \text{and} \quad \mathbf{V} := \left(H^1(\Omega)\right)^d.$$

For a Banach space  $B$ , we let  $L^2(0, T; B)$  be the Bochner space of  $B$ -valued functions, equipped with its natural norm. For a subinterval  $J \subseteq [0, T]$ , we will use the notation  $L^2(J; B)$  for the corresponding Bochner space. In what follows, the constant  $c$  is generic and independent of the size of the space and time meshes.

For the time discretization, we decompose the time interval  $I := (0, T]$  into  $N$  subintervals  $I_n = (t_{n-1}, t_n]$ ,  $n = 1, \dots, N$ , where  $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T$  such that  $I = \bigcup_{n=1}^N I_n$ . We put  $\tau := \max_{n=1, \dots, N} \tau_n$  with  $\tau_n = t_n - t_{n-1}$ . Further, the set  $\mathcal{M}_\tau := \{I_1, \dots, I_N\}$  of time

intervals is called the time mesh. For a Banach space  $B$  and any  $k \in \mathbb{N}_0$ , we let

$$\mathbb{P}_k(I_n; B) := \left\{ w_\tau : I_n \rightarrow B, \right. \\ \left. w_\tau(t) = \sum_{j=0}^k W^j t^j \quad \forall t \in I_n, W^j \in B \quad \forall j \right\}. \tag{2.1}$$

For  $k \in \mathbb{N}_0$  we define the space of piecewise polynomial functions in time with values in  $B$  by

$$Y_\tau^k(B) := \{ w_\tau : \bar{I} \rightarrow B \mid w_\tau|_{I_n} \in \mathbb{P}_k(I_n; B) \\ \forall I_n \in \mathcal{M}_\tau, w_\tau(0) \in B \} \subset L^2(I; B). \tag{2.2}$$

For any function  $w : \bar{I} \rightarrow B$  that is piecewise sufficiently smooth with respect to the time mesh  $\mathcal{M}_\tau$ , for instance for  $w \in Y_\tau^k(B)$ , we define the right-hand sided and left-hand sided limit at a mesh point  $t_n$  by

$$w^+(t_n) := \lim_{t \rightarrow t_n+0} w(t), \quad \text{for } n < N, \text{ and} \\ w^-(t_n) := \lim_{t \rightarrow t_n-0} w(t), \quad \text{for } n > 0. \tag{2.3}$$

For the integration in time of a discontinuous Galerkin approach it is natural to use the right-sided  $(k + 1)$ -point Gauss–Radau quadrature formula. On the subinterval  $I_n$ , it reads as

$$Q_n(w) := \frac{\tau_n}{2} \sum_{\mu=1}^{k+1} \hat{\omega}_\mu^{\text{GR}} w(t_{n,\mu}^{\text{GR}}) \approx \int_{I_n} w(t) dt, \tag{2.4}$$

where  $t_{n,\mu}^{\text{GR}} = T_n(\hat{t}_\mu^{\text{GR}})$ , for  $\mu = 1, \dots, k + 1$ , are the Gauss–Radau quadrature points on  $I_n$  and  $\hat{\omega}_\mu^{\text{GR}}$  the corresponding weights. Here,  $T_n(\hat{t}) := (t_{n-1} + t_n)/2 + (\tau_n/2)\hat{t}$  is the affine transformation from  $\hat{I} = [-1, 1]$  to  $I_n$  and  $\hat{t}_\mu^{\text{GR}}$  are the Gauss–Radau quadrature points on  $\hat{I}$ . Formula (2.4) is exact for all polynomials  $w \in \mathbb{P}_{2k}(I_n; \mathbb{R})$ . In particular, there holds that  $t_{n,k+1}^{\text{GR}} = t_n$ .

For the space discretization, let  $\{\mathcal{T}_l\}_{l=0}^L$  be the decomposition on every multigrid level of  $\Omega$  into (open) quadrilaterals or hexahedrals, with  $\mathcal{T}_l = \{K_i \mid i = 1, \dots, N_l^{\text{el}}\}$ , for  $l = 0, \dots, L$ . These element types are chosen for our implementation (cf. Sect. 5) that is based on the deal.II library [7]. The finest partition is  $\mathcal{T}_h = \mathcal{T}_L$ . We assume that all the partitions  $\{\mathcal{T}_l\}_{l=0}^L$  are quasi-uniform with characteristic mesh size  $h_l$  and  $h_l = \gamma h_{l-1}$ ,  $\gamma \in (0, 1)$  and  $h_0 = \mathcal{O}(1)$ . On the actual mesh level, the finite element spaces used for approximating the unknowns  $\mathbf{u}$  and  $p$  of (1.1) are of the form ( $l \in \{0, \dots, L\}$ )

$$V_{h_l}^l := \{v_h \in V \cap C(\bar{\Omega})^d : v_h|_K \in V(K) \text{ for all } K \in \mathcal{T}_l\}, \tag{2.5a}$$

$$Q_{h_l}^{l,\text{cont}} := \{q_h \in Q \cap C(\bar{\Omega}) : q_h|_K \in Q(K) \text{ for all } K \in \mathcal{T}_l\}, \tag{2.5b}$$

$$Q_{h_l}^{l,\text{disc}} := \{q_h \in Q : q_h|_K \in Q(K) \text{ for all } K \in \mathcal{T}_l\}. \tag{2.5c}$$

By an abuse of notation, we skip the index  $l$  of the mesh level when it is clear from the context and put

$$V_h := V_h^l, \tag{2.6}$$

$$Q_h := Q_h^l \text{ with } Q_h^l \in \{Q_h^{l,\text{cont}}, Q_h^{l,\text{disc}}\}.$$

For the local spaces  $V(K)$  and  $Q(K)$  we employ mapped versions of the pairs  $Q_r^d/Q_{r-1}$  and  $Q_r^d/\mathbb{P}_{r-1}^{\text{disc}}$ , for  $r \geq 2$ . The pair  $Q_r^d/Q_{r-1}$  with a (globally) continuous approximation of the scalar variable  $p$  in  $Q_h^{l,\text{cont}}$  is the well-known Taylor–Hood family of finite element spaces. The pair  $Q_r^d/\mathbb{P}_{r-1}^{\text{disc}}$  comprises a discontinuous approximation of  $p$  in the broken polynomial space  $Q_h^{l,\text{disc}}$ . For the Navier–Stokes equations, the multigrid method has shown to work best for higher-order finite element spaces with discontinuous discrete pressure; cf. [46] and [3]. For a further discussion of mapped and unmapped versions of the pair  $Q_r^d/\mathbb{P}_{r-1}^{\text{disc}}$  we refer to [44, Subsec. 3.6.4]. For an analysis of stability properties of (spatial) discretizations for the quasi-static Biot system we refer to, e.g. [61]. Both choices of the local finite element spaces,  $Q_r^d/Q_{r-1}$  and  $Q_r^d/\mathbb{P}_{r-1}^{\text{disc}}$ , satisfy under some restrictions (cf. [82]) the inf–sup stability condition,

$$\inf_{q_h \in Q_h \setminus \{0\}} \sup_{v_h \in V_h \setminus \{0\}} \frac{b(v_h, q_h)}{\|v_h\|_1 \|q_h\|} \geq \beta > 0, \tag{2.7}$$

for some constant  $\beta$  independent of  $h$ ; cf. [44, 57]. In [8, 56] optimal interpolation error estimates for mapped finite elements on quadrilaterals and hexahedra are studied. It turned out that the optimality is given for special families of triangulations. In two and three dimensions, families of meshes, which are obtained by a regular uniform refinement of an initial coarse grid, are among these special families. Such a regular refinement that is natural for the construction of the multigrid hierarchy is used in our computations. Thus, for  $v \in H^{r+1}(\Omega)$  and  $q \in H^r(\Omega)$  there exist approximations  $i_h v \in V_h$  and  $j_h q \in Q_h$  such that, with some generic constant  $c > 0$  independent of  $h$ ,

$$\|v - i_h v\| + h \|\nabla(v - i_h v)\| \leq ch^{r+1}, \tag{2.8a}$$

$$\|q - j_h q\| \leq ch^r. \tag{2.8b}$$

### 3 Space-time finite element approximation

For the discretization we rewrite (1.1) as a first-order in time system by introducing the new variable  $v := \partial_t u$ . Then, we recover (1.1a) and (1.1b) as

$$\partial_t u - v = 0, \tag{3.1a}$$

$$\rho \partial_t v - \nabla \cdot (C \varepsilon(u)) + \alpha \nabla p = \rho f, \tag{3.1b}$$

$$c_0 \partial_t p + \alpha \nabla \cdot v - \nabla \cdot (K \nabla p) = g \tag{3.1c}$$

along with the initial and boundary conditions (1.1c) to (1.1g). For the approximation of (3.1) we use a monolithic approach to capture efficiently the dynamics of (3.1) and avoid additional consistency errors. An iterative coupling scheme for (3.1) is proposed, for instance, in [19]. We employ discontinuous Galerkin methods (cf. [75]) for the discretization of the time variable and inf-sup stable pairs of finite elements (cf. Sect. 2) for the approximation of the space variables in (3.1). The derivation of the discrete scheme, presented below in Problem B.1, is standard and not explicitly presented here. It follows the lines of [11], where continuous in time Galerkin methods are applied to (1.1), and [3, 4, 41, 42], where discontinuous in time Galerkin methods are used to discretize the Navier–Stokes system. In contrast to [11], Dirichlet boundary conditions are implemented here by Nitsche’s method [15, 30, 60]. This yields a strong link between two different families of inf-stable finite element pairs for the space discretization. The main reason for using Nitsche’s method here comes through our more general software framework. Nitsche’s method captures problems on evolving domains solved on fixed computational background grids (cf. [5]). We note that Nitsche’s method does not perturb the convergence behavior of the space-time discretization; cf. Sect. 5.

For the discrete scheme we need further notation. On the multigrid level  $l$  with decomposition  $\mathcal{T}_l$ , for  $w_h, \chi_h \in V_h$  and  $q_h, \psi_h \in Q_h$  we define

$$A_\gamma(w_h, \chi_h) := (C \varepsilon(w_h), \varepsilon(\chi_h)) - (C \varepsilon(w_h) n, \chi_h)_{\Gamma_u^D} + a_\gamma(w_h, \chi_h), \tag{3.2a}$$

$$C(\chi_h, q_h) := -\alpha (\nabla \cdot \chi_h, q_h) + \alpha (\chi_h \cdot n, q_h)_{\Gamma_u^D}, \tag{3.2b}$$

$$B_\gamma(q_h, \psi_h) := \begin{cases} (K \nabla q_h, \nabla \psi_h) - (K \nabla q_h \cdot n, \psi_h)_{\Gamma_p^D} + b_\gamma(q_h, \psi_h), \\ \text{for } Q_h = Q_h^{l,\text{cont}}, \\ \sum_{K \in \mathcal{T}_l} (K \nabla q_h, \nabla \psi_h)_K \\ - \sum_{F \in \mathcal{F}_h} (\{K \nabla q_h\} \cdot n, \{\psi_h\})_F \\ + (\{q_h\}, \{K \nabla \psi_h\} \cdot n)_F \\ + \sum_{F \in \mathcal{F}_h} \frac{\gamma}{h_F} (\{q_h\}, \{\psi_h\})_F, \\ \text{for } Q_h = Q_h^{l,\text{disc}}, \end{cases} \tag{3.2c}$$

where, for  $\mathbf{w} \in \mathbf{H}^{1/2}(\Gamma_u^D)$  and  $q \in H^{1/2}(\Gamma_p^D)$ ,

$$a_\gamma(\mathbf{w}, \boldsymbol{\chi}_h) := -\langle \mathbf{w}, \mathbf{C}\boldsymbol{\varepsilon}(\boldsymbol{\chi}_h)\mathbf{n} \rangle_{\Gamma_u^D} + \frac{\gamma_a}{h_F} \langle \mathbf{w}, \boldsymbol{\chi}_h \rangle_{\Gamma_u^D}, \quad (3.3a)$$

$$b_\gamma(q, \psi_h) := -\langle q, \mathbf{K}\nabla\psi_h \cdot \mathbf{n} \rangle_{\Gamma_p^D} + \frac{\gamma_b}{h_F} \langle q, \psi_h \rangle_{\Gamma_p^D}. \quad (3.3b)$$

The second of the options in (3.2c) amounts to a symmetric interior penalty discontinuous Galerkin discretization of the scalar variable  $p$ ; cf., e.g. [26, Sec. 4.2]. As usual, the average  $\{\cdot\}$  and jump  $\llbracket \cdot \rrbracket$  of a function  $w \in L^2(\Omega)$  on an interior face  $F$  between two elements  $K^+$  and  $K^-$ , such that  $F = \partial K^+ \cap \partial K^-$ , are

$$\{w\} := \frac{1}{2}(w^+ + w^-) \quad \text{and} \quad \llbracket w \rrbracket := w^+ - w^-.$$

For boundary faces  $F \subset \partial K \cap \partial\Omega$ , we set  $\{w\} := w|_K$  and  $\llbracket w \rrbracket := w|_K$ . The set of all faces (interior and boundary faces) on the multigrid level  $\mathcal{T}_l$  is denoted by  $\mathcal{F}_h$ . In the second of the options in (3.2c), the parameter  $\gamma$  of the last term has to be chosen sufficiently large, such that discrete coercivity on  $Q_h$  of  $B_\gamma$  is preserved. The local length  $h_F$  is chosen as  $h_F = \{\!|h_F|\!\} := \frac{1}{2}(|K^+|_d + |K^-|_d)$  with Hausdorff measure  $|\cdot|_d$ ; cf. [26, p. 125]. For boundary faces we set  $h_F := |K|_d$ . In (3.3), the quantities  $\gamma_a$  and  $\gamma_b$  are the algorithmic parameters of the stabilization terms in the Nitsche formulation. To ensure well-posedness of the discrete systems the parameters  $\gamma_a$  and  $\gamma_b$  have to be chosen sufficiently large; cf. Appendix A. Based on our numerical experiments we choose the algorithmic parameter  $\gamma$ ,  $\gamma_a$  and  $\gamma_b$  in (3.2c) and (3.3) as

$$\begin{aligned} \gamma_a &= 5 \cdot 10^4 \cdot r \cdot (r + 1) \quad \text{and} \\ \gamma &= \gamma_b = \frac{1}{2} \cdot r \cdot (r - 1), \end{aligned}$$

where  $r$  is the polynomial degree of the finite element space (2.5a) for the displacement variable.

Finally, for given  $\mathbf{f} \in \mathbf{H}^{-1}(\Omega)$ ,  $\mathbf{u}_D \in \mathbf{H}^{1/2}(\Gamma_u^D)$ ,  $\mathbf{t}_N \in \mathbf{H}^{-1/2}(\Gamma_u^N)$  and  $g \in H^{-1}(\Omega)$ ,  $p_D \in H^{1/2}(\Gamma_p^D)$ ,  $p_N \in H^{-1/2}(\Gamma_p^N)$  for  $Q_h = Q_h^{l,\text{cont}}$ , and properly fitted assumptions about the data for  $Q_h = Q_h^{l,\text{disc}}$ , we put

$$F_\gamma(\boldsymbol{\chi}_h) := \langle \mathbf{f}, \boldsymbol{\chi}_h \rangle - \langle \mathbf{t}_N, \boldsymbol{\chi}_h \rangle_{\Gamma_u^N} + a_\gamma(\mathbf{u}_D, \boldsymbol{\chi}_h), \quad (3.4a)$$

$$G_\gamma(\psi_h) := \begin{cases} \langle g, \psi_h \rangle - \alpha \langle \mathbf{v}_D \cdot \mathbf{n}, \psi_h \rangle_{\Gamma_u^D} \\ - \langle p_N, \psi_h \rangle_{\Gamma_p^N} + b_\gamma(p_D, \psi_h), \\ \text{for } Q_h = Q_h^{l,\text{cont}}, \\ \langle g, \psi_h \rangle - \sum_{F \in \mathcal{F}_h^{D,u}} \alpha \langle \mathbf{v}_D \cdot \mathbf{n}, \psi_h \rangle_F \\ - \sum_{F \in \mathcal{F}_h^{D,p}} \langle p_D, \llbracket \mathbf{K}\nabla\psi_h \rrbracket \cdot \mathbf{n} \rangle_F \\ + \sum_{F \in \mathcal{F}_h^{D,p}} \frac{\gamma}{h} \langle p_D, \{\!\|\psi_h\|\!\} \rangle_F \\ - \sum_{F \subset \mathcal{F}_h^{N,p}} \langle p_N, \{\!\|\psi_h\|\!\} \rangle_F, \\ \text{for } Q_h = Q_h^{l,\text{disc}}. \end{cases} \quad (3.4b)$$

In the second of the options in (3.4b), we denote by  $\mathcal{F}_h^{D,p} \subset \mathcal{F}_h$  and  $\mathcal{F}_h^{N,p} \subset \mathcal{F}_h$  the set of all element faces on the boundary parts  $\Gamma_p^D$  and  $\Gamma_p^N$ , respectively; cf. (1.1). The second of the terms on the right-hand side of (3.4b), with  $\mathbf{v}_D = \partial_t \mathbf{u}_D$ , is added to ensure consistency of the form (3.2b) in the fully discrete formulation (3.5c) of (1.1b), i.e., that the discrete equation (3.5c) is satisfied by the continuous solution to (1.1).

We use a temporal test basis that is supported on the subintervals  $I_n$ ; cf., e.g. [3, 41]. Then, a time marching process is obtained. In that, we assume that the trajectories  $\mathbf{u}_{\tau,h}$ ,  $\mathbf{v}_{\tau,h}$  and  $p_{\tau,h}$  have been computed before for all  $t \in [0, t_{n-1}]$ , starting with approximations  $\mathbf{u}_{\tau,h}(t_0) := \mathbf{u}_{0,h}$ ,  $\mathbf{v}_{\tau,h}(t_0) := \mathbf{u}_{1,h}$  and  $p_{\tau,h}(t_0) := p_{0,h}$  of the initial values  $\mathbf{u}_0$ ,  $\mathbf{v}_0$  and  $p_{0,h}$ . Then, we consider solving the following local problem on  $I_n$ .

**Problem 3.1** (Numerically integrated  $I_n$ -problem) For given  $\mathbf{u}_h^{n-1} := \mathbf{u}_{\tau,h}(t_{n-1}) \in \mathbf{V}_h$ ,  $\mathbf{v}_h^{n-1} := \mathbf{v}_{\tau,h}(t_{n-1}) \in \mathbf{V}_h$ , and  $p_h^{n-1} := p_{\tau,h}(t_{n-1}) \in Q_h$  with  $\mathbf{u}_{\tau,h}(t_0) := \mathbf{u}_{0,h}$ ,  $\mathbf{v}_{\tau,h}(t_0) := \mathbf{u}_{1,h}$  and  $p_{\tau,h}(t_0) := p_{0,h}$ , find  $(\mathbf{u}_{\tau,h}, \mathbf{v}_{\tau,h}, p_{\tau,h}) \in \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; Q_h)$  such that

$$\begin{aligned} Q_n(\langle \partial_t \mathbf{u}_{\tau,h}, \boldsymbol{\phi}_{\tau,h} \rangle - \langle \mathbf{v}_{\tau,h}, \boldsymbol{\phi}_{\tau,h} \rangle) \\ + \langle \mathbf{u}_{\tau,h}^+(t_{n-1}), \boldsymbol{\phi}_{\tau,h}^+(t_{n-1}) \rangle = \langle \mathbf{u}_h^{n-1}, \boldsymbol{\phi}_{\tau,h}^+(t_{n-1}) \rangle, \end{aligned} \quad (3.5a)$$

$$\begin{aligned} Q_n(\langle \rho \partial_t \mathbf{v}_{\tau,h}, \boldsymbol{\chi}_{\tau,h} \rangle + A_\gamma(\mathbf{u}_{\tau,h}, \boldsymbol{\chi}_{\tau,h}) + C(\boldsymbol{\chi}_{\tau,h}, p_{\tau,h})) \\ + \langle \rho \mathbf{v}_{\tau,h}^+(t_{n-1}), \boldsymbol{\chi}_{\tau,h}^+(t_{n-1}) \rangle \\ = Q_n(F_\gamma(\boldsymbol{\chi}_{\tau,h})) + \langle \rho \mathbf{v}_h^{n-1}, \boldsymbol{\chi}_{\tau,h}^+(t_{n-1}) \rangle, \end{aligned} \quad (3.5b)$$

$$\begin{aligned} Q_n(\langle c_0 \partial_t p_{\tau,h}, \psi_{\tau,h} \rangle - C(\mathbf{v}_{\tau,h}, \psi_{\tau,h}) + B_\gamma(p_{\tau,h}, \psi_{\tau,h})) \\ + \langle c_0 p_{\tau,h}^+(t_{n-1}), \psi_{\tau,h}^+(t_{n-1}) \rangle \\ = Q_n(G_\gamma(\psi_{\tau,h})) + \langle c_0 p_h^{n-1}, \psi_{\tau,h}^+(t_{n-1}) \rangle \end{aligned} \quad (3.5c)$$

for all  $(\boldsymbol{\phi}_{\tau,h}, \boldsymbol{\chi}_{\tau,h}, \psi_{\tau,h}) \in \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; Q_h)$ .

The trajectories defined by Problem 3.1, for  $n = 1, \dots, N$ , satisfy that  $\mathbf{u}_{\tau,h}, \mathbf{v}_{\tau,h} \in Y_\tau^k(\mathbf{V}_h)$  and  $p_{\tau,h} \in Y_\tau^k(Q_h)$ . The quadrature formulas on the left hand-side of (3.5) can be rewritten by time integrals since the Gauss–Radau formula (2.4) is exact for all polynomials  $w \in P_{2k}(I_n; \mathbb{R})$ . Well-posedness of Problem 3.1 is ensured.

**Lemma 3.2** (Existence and uniqueness of solutions to Problem 3.1) Problem 3.1 admits a unique solution.

**Proof** We prove Lem. 3.2 for  $Q_h = Q_h^{l,\text{cont}}$  only, thus assuming the first of the options in (3.2c) and (3.4b).

For  $Q_h = Q_h^{l, \text{disc}}$ , the proof can be done similarly by using, in addition, standard techniques of error analysis for discontinuous Galerkin methods; cf., e.g. [26, Sec. 4]. Since Problem (3.5) is finite-dimensional, it suffices to prove uniqueness of the solution. Existence of the solution then follows directly from its uniqueness. Let now  $(\mathbf{u}_{\tau,h}^{(1)}, \mathbf{v}_{\tau,h}^{(1)}, p_{\tau,h}^{(1)}) \in \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; Q_h)$  and  $(\mathbf{u}_{\tau,h}^{(2)}, \mathbf{v}_{\tau,h}^{(2)}, p_{\tau,h}^{(2)}) \in \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; Q_h)$  denote two triples of solutions to (3.5a). Their differences  $(\mathbf{u}_{\tau,h}, \mathbf{v}_{\tau,h}, p_{\tau,h}) = (\mathbf{u}_{\tau,h}^{(1)}, \mathbf{v}_{\tau,h}^{(1)}, p_{\tau,h}^{(1)}) - (\mathbf{u}_{\tau,h}^{(2)}, \mathbf{v}_{\tau,h}^{(2)}, p_{\tau,h}^{(2)})$  then satisfy the equations

$$Q_n(\langle \partial_t \mathbf{u}_{\tau,h}, \boldsymbol{\phi}_{\tau,h} \rangle - \langle \mathbf{v}_{\tau,h}, \boldsymbol{\phi}_{\tau,h} \rangle) + \langle \mathbf{u}_{\tau,h}^+(t_{n-1}), \boldsymbol{\phi}_{\tau,h}^+(t_{n-1}) \rangle = 0, \tag{3.6a}$$

$$Q_n(\langle \rho \partial_t \mathbf{v}_{\tau,h}, \boldsymbol{\chi}_{\tau,h} \rangle + A_\gamma(\mathbf{u}_{\tau,h}, \boldsymbol{\chi}_{\tau,h}) + C(\boldsymbol{\chi}_{\tau,h}, p_{\tau,h})) + \langle \rho \mathbf{v}_{\tau,h}^+(t_{n-1}), \boldsymbol{\chi}_{\tau,h}^+(t_{n-1}) \rangle = 0, \tag{3.6b}$$

$$Q_n(\langle c_0 \partial_t p_{\tau,h}, \psi_{\tau,h} \rangle - C(\mathbf{v}_{\tau,h}, \psi_{\tau,h}) + B_\gamma(p_{\tau,h}, \psi_{\tau,h})) + \langle c_0 p_{\tau,h}^+(t_{n-1}), \psi_{\tau,h}^+(t_{n-1}) \rangle = 0 \tag{3.6c}$$

for all  $(\boldsymbol{\phi}_{\tau,h}, \boldsymbol{\chi}_{\tau,h}, \psi_{\tau,h}) \in \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; Q_h)$ . We let  $A_\gamma : \mathbf{V}_h \mapsto \mathbf{V}_h$  be the discrete operator that is defined, for  $\mathbf{w}_h \in \mathbf{V}_h$  and all  $\boldsymbol{\phi}_h \in \mathbf{V}_h$ , by

$$\langle A_\gamma \mathbf{w}_h, \boldsymbol{\phi}_h \rangle = A_\gamma(\mathbf{w}_h, \boldsymbol{\phi}_h). \tag{3.7}$$

In (3.6) we choose  $\boldsymbol{\phi}_{\tau,h} = A_\gamma \mathbf{u}_{\tau,h}$ ,  $\boldsymbol{\chi}_{\tau,h} = \mathbf{v}_{\tau,h}$  and  $\psi_{\tau,h} = p_{\tau,h}$ . Adding the resulting equations yields that

$$Q_n(\langle \partial_t \mathbf{u}_{\tau,h}, A_\gamma \mathbf{u}_{\tau,h} \rangle + \langle \rho \partial_t \mathbf{v}_{\tau,h}, \mathbf{v}_{\tau,h} \rangle + \langle c_0 \partial_t p_{\tau,h}, p_{\tau,h} \rangle + B_\gamma(p_{\tau,h}, p_{\tau,h})) + \langle \mathbf{u}_{\tau,h}^+(t_{n-1}), A_\gamma \mathbf{u}_{\tau,h}^+(t_{n-1}) \rangle + \langle \rho \mathbf{v}_{\tau,h}^+(t_{n-1}), \mathbf{v}_{\tau,h}^+(t_{n-1}) \rangle + \langle c_0 p_{\tau,h}^+(t_{n-1}), p_{\tau,h}^+(t_{n-1}) \rangle = 0. \tag{3.8}$$

Recalling the exactness of the Gauss–Radau formula (2.4) for  $w \in \mathbb{P}_{2k}(I_n; \mathbb{R})$ , Eq. (3.8) yields that

$$\frac{1}{2} \int_{t_{n-1}}^{t_n} \frac{d}{dt} (\langle A_\gamma \mathbf{u}_{\tau,h}, \mathbf{u}_{\tau,h} \rangle + \langle \rho \mathbf{v}_{\tau,h}, \mathbf{v}_{\tau,h} \rangle + \langle c_0 p_{\tau,h}, p_{\tau,h} \rangle) dt + Q_n(B_\gamma(p_{\tau,h}, p_{\tau,h})) + \langle \mathbf{u}_{\tau,h}^+(t_{n-1}), A_\gamma \mathbf{u}_{\tau,h}^+(t_{n-1}) \rangle + \langle \rho \mathbf{v}_{\tau,h}^+(t_{n-1}), \mathbf{v}_{\tau,h}^+(t_{n-1}) \rangle + \langle c_0 p_{\tau,h}^+(t_{n-1}), p_{\tau,h}^+(t_{n-1}) \rangle = 0.$$

Using (3.7), this shows that

$$A_\gamma(\mathbf{u}_{\tau,h}(t_n), \mathbf{u}_{\tau,h}(t_n)) + \langle \rho \mathbf{v}_{\tau,h}(t_n), \mathbf{v}_{\tau,h}(t_n) \rangle + \langle c_0 p_{\tau,h}(t_n), p_{\tau,h}(t_n) \rangle + 2Q_n(B_\gamma(p_{\tau,h}, p_{\tau,h})) + A_\gamma(\mathbf{u}_{\tau,h}^+(t_{n-1}), \mathbf{u}_{\tau,h}^+(t_{n-1}))$$

$$+ \langle \rho \mathbf{v}_{\tau,h}^+(t_{n-1}), \mathbf{v}_{\tau,h}^+(t_{n-1}) \rangle + \langle c_0 p_{\tau,h}^+(t_{n-1}), p_{\tau,h}^+(t_{n-1}) \rangle = 0. \tag{3.9}$$

From (3.9) along with the discrete coercivity properties (A.3) of  $A_\gamma$  and (A.5) of  $B_\gamma$  we directly deduce that

$$\mathbf{u}_{\tau,h}(t_n) = \mathbf{u}_{\tau,h}^+(t_{n-1}) = \mathbf{0}, \mathbf{v}_{\tau,h}(t_n) = \mathbf{v}_{\tau,h}^+(t_{n-1}) = \mathbf{0}, p_{\tau,h}(t_n) = p_{\tau,h}^+(t_{n-1}) = 0, \tag{3.10}$$

as well as

$$p_{\tau,h}(t_{n,\mu}^{\text{GR}}) = 0, \text{ for } \mu = 1, \dots, k + 1. \tag{3.11}$$

In (3.11) we recall that  $t_{n,k+1}^{\text{GR}} = t_n$ . Relation (3.11) implies that  $p_{\tau,h} \equiv 0$  on  $I_n$ . For  $k = 0$ , the uniqueness of  $\mathbf{u}_{\tau,h}$  and  $\mathbf{v}_{\tau,h}$  is already proved by (3.10).

From now, let  $k \geq 1$ . To prove that  $\mathbf{u}_{\tau,h} \equiv \mathbf{0}$  and  $\mathbf{v}_{\tau,h} \equiv \mathbf{0}$ , by (3.10) it is sufficient to show that  $\mathbf{u}_{\tau,h}(t_{n,\mu}^{\text{G}}) = \mathbf{0}$  and  $\mathbf{v}_{\tau,h}(t_{n,\mu}^{\text{G}}) = \mathbf{0}$ , for  $\mu = 1, \dots, k$ , where  $t_{n,\mu}^{\text{G}}$ , for  $\mu = 1, \dots, k$ , are the nodes of the  $k$ -point Gauss quadrature formula on  $I_n$  that is exact for all polynomials in  $\mathbb{P}_{2k-1}(I_n; \mathbb{R})$ . This is done now. Recalling (3.10), we conclude from (3.6a) by a suitable choice of test functions that

$$\partial_t \mathbf{u}_{\tau,h}(t_{n,\mu}^{\text{GR}}) = \mathbf{v}_{\tau,h}(t_{n,\mu}^{\text{GR}}), \text{ for } \mu = 1, \dots, k. \tag{3.12}$$

Next, choosing  $\boldsymbol{\chi}_{\tau,h} = \mathbf{v}_{\tau,h}$  in (3.6b) and recalling (3.12) imply that

$$Q_n(\langle \rho \partial_t \mathbf{v}_{\tau,h}, \mathbf{v}_{\tau,h} \rangle + A_\gamma(\mathbf{u}_{\tau,h}, \partial_t \mathbf{u}_{\tau,h})) = 0 \tag{3.13}$$

By the exactness of the Gauss–Radau formula (2.4) for all  $w \in \mathbb{P}_{2k}(I_n; \mathbb{R})$  we have from (3.13) that

$$\int_{t_{n-1}}^{t_n} \langle \rho \partial_t \mathbf{v}_{\tau,h}, \mathbf{v}_{\tau,h} \rangle dt + \frac{1}{2} \int_{t_{n-1}}^{t_n} \frac{d}{dt} A_\gamma(\mathbf{u}_{\tau,h}, \mathbf{u}_{\tau,h}) dt = 0. \tag{3.14}$$

The second of the terms in (3.14) vanishes by (3.10). The stability result of [49, Lem. 2.1] then implies that

$$\mathbf{v}_{\tau,h}(t_{n,\mu}^{\text{G}}) = \mathbf{0}, \text{ for } \mu = 1, \dots, k. \tag{3.15}$$

From (3.15) along with (3.10) we then deduce that  $\mathbf{v}_{\tau,h} \equiv \mathbf{0}$  on  $I_n$ . Choosing the test function  $\boldsymbol{\phi}_{\tau,h} = \mathbf{u}_{\tau,h}$  in (3.6), using  $\mathbf{v}_{\tau,h} \equiv \mathbf{0}$  and applying the stability result [49, Lem. 2.1], it follows that

$$\mathbf{u}_{\tau,h}(t_{n,\mu}^{\text{G}}) = \mathbf{0}, \text{ for } \mu = 1, \dots, k. \tag{3.16}$$

From (3.16) along with (3.10) we then have that  $\mathbf{u}_{\tau,h} \equiv \mathbf{0}$  on  $I_n$ . Thus uniqueness of solutions to (3.5) and, thereby, well-posedness of Problem 3.1 is thus ensured.

In Appendix B an alternative formulation for the system (3.6) is still presented. It is based on using the time derivative  $\partial_t \mathbf{u}_{\tau,h}$  of the primal variable  $\mathbf{u}_{\tau,h}$  instead of the auxiliary variable  $\mathbf{v}_{\tau,h}$  in (3.5c). In this case, an additional surface integral has to be included; cf. Eq. (B.1c).

### 4 Algebraic solver by geometric multigrid preconditioned GMRES iterations

On the algebraic level, the variational problem (3.5) leads to linear systems of equations with complex block structure, in particular if higher order (piecewise) polynomial degrees  $k$  for the approximation of the temporal variable are involved. This demands for a robust and efficient linear solver, in particular in the three-dimensional case  $\Omega \subset \mathbb{R}^3$ . For solving (3.5) we consider using flexible GMRES iterations [65] that are preconditioned by a V-cycle geometric multigrid method (GMG) based on a local Vanka smoother. In [4], the GMG preconditioned GMRES solver is further embedded in a Newton iteration for solving space-time finite element discretizations of the Navier–Stokes system. Thus, nonlinear extensions of the prototype model (1.1) become feasible as well by our approach. For non-smooth nonlinearities, fixed point iterations, like the L-scheme [50], can be used instead of Newton’s method.

To derive the algebraic form of (3.5), the discrete functions  $\mathbf{u}_{\tau,h}$ ,  $\mathbf{v}_{\tau,h}$  and  $p_{\tau,h}$  are represented in a Lagrangian basis  $\{\chi_{n,m}\}_{m=1}^{k+1} \subset \mathbb{P}_k(I_n; \mathbb{R})$  with respect to the  $(k + 1)$  Gauss–Radau quadrature points of  $I_n$ , such that

$$\mathbf{u}_{\tau,h}|_{I_n}(\mathbf{x}, t) = \sum_{m=1}^{k+1} \mathbf{u}_{n,m}(\mathbf{x}) \chi_{n,m}(t), \tag{4.1a}$$

$$\mathbf{v}_{\tau,h}|_{I_n}(\mathbf{x}, t) = \sum_{m=1}^{k+1} \mathbf{v}_{n,m}(\mathbf{x}) \chi_{n,m}(t), \tag{4.1b}$$

$$p_{\tau,h}|_{I_n}(\mathbf{x}, t) = \sum_{m=1}^{k+1} p_{n,m}(\mathbf{x}) \chi_{n,m}(t). \tag{4.1c}$$

The resulting coefficient functions  $(\mathbf{u}_{n,m}, \mathbf{v}_{n,m}, p_{n,m}) \in V_h \times V_h \times Q_h$ , for  $m = 1, \dots, k + 1$ , are developed in terms of the finite element basis of  $V_h$  and  $Q_h$ , respectively. Letting  $V_h = \text{span}\{\psi_1, \dots, \psi_R\}$  and  $Q_h = \text{span}\{\xi_1, \dots, \xi_S\}$ , we get that

$$\mathbf{u}_{n,m}(\mathbf{x}) = \sum_{r=1}^R u_{n,m}^{(r)} \psi_r(\mathbf{x}), \tag{4.2a}$$

$$\mathbf{v}_{n,m}(\mathbf{x}) = \sum_{r=1}^R v_{n,m}^{(r)} \psi_r(\mathbf{x}), \tag{4.2b}$$

$$p_{n,m}(\mathbf{x}) = \sum_{s=1}^S p_{n,m}^{(s)} \xi_s(\mathbf{x}). \tag{4.2c}$$

For the coefficients of the expansions in (4.2a) we define the subvectors

$$\mathbf{U}_{n,m} = (u_{n,m}^{(1)}, \dots, u_{n,m}^{(R)})^\top, \quad \mathbf{V}_{n,m} = (v_{n,m}^{(1)}, \dots, v_{n,m}^{(R)})^\top, \tag{4.3a}$$

$$\mathbf{P}_{n,m} = (p_{n,m}^{(1)}, \dots, p_{n,m}^{(S)})^\top, \quad \text{for } m = 1, \dots, k + 1, \tag{4.3b}$$

of the degrees of freedom for all Gauss–Radau quadrature points and the global solution vector on  $I_n$  by

$$\mathbf{X}_n^\top = ((\mathbf{V}_{n,1})^\top, (\mathbf{U}_{n,1})^\top, (\mathbf{P}_{n,1})^\top, \dots, (\mathbf{V}_{n,k+1})^\top, (\mathbf{U}_{n,k+1})^\top, (\mathbf{P}_{n,k+1})^\top). \tag{4.4}$$

We note that  $\mathbf{X}_n$  comprises the (spatial) degrees of freedom for all  $(k + 1)$  Gauss–Radau nodes, representing the Lagrange interpolation points in time, of the subinterval  $I_n$ . The approximations at these time points will be computed simultaneously. Substituting (4.1) and (4.2a) into (3.5) and choosing in (3.5) the test basis  $\{\chi_{n,m} \psi_r, \chi_{n,m} \psi_r, \chi_{n,m} \xi_s\}$ , for  $m = 1, \dots, k + 1$ ,  $r = 1, \dots, R$  and  $s = 1, \dots, S$ , built from the trial basis in (4.1) and (4.2a), we obtain the following algebraic system.

**Problem 4.1** (Algebraic  $I_n$ -problem) *For the vector  $\mathbf{X}_n$ , defined in (4.4) along with (4.3), of the coefficients of the expansions (4.2a) solve*

$$\mathbf{A}_n \mathbf{X}_n = \mathbf{F}_n, \tag{4.5}$$

where the matrix  $\mathbf{A}_n$  exhibits the  $(k + 1) \times (k + 1)$  block structure

$$\mathbf{A}_n = (\mathbf{A}_{a,b})_{a,b=1}^{k+1} \tag{4.6}$$

with block submatrices  $\mathbf{A}_{a,b}$ , for  $a, b = 1, \dots, k + 1$ , defined by

$$\mathbf{A}_{a,b} = \begin{pmatrix} -\mathbf{M}_{a,b}^{0,V_h} & \mathbf{M}_{a,b}^{1,V_h} & \mathbf{0} \\ \mathbf{M}_{a,b}^{1,V_h} & \mathbf{S}_{a,b} + \mathbf{N}_{a,b}^A & \mathbf{C}_{a,b}^\top \\ \mathbf{0} & -\mathbf{C}_{a,b} & \mathbf{M}_{a,b}^{1,Q_h} + \mathbf{B}_{a,b} + \mathbf{N}_{a,b}^B \end{pmatrix}. \tag{4.7}$$



For the choice  $Q_h^l = Q_h^{l,\text{cont}}$  in (2.6), the explicit representation of the submatrices in (4.7) reads as

$$\begin{aligned} (M_{a,b}^{1,V_h})_{i,j} &:= Q_n(\langle \rho \partial_t \chi_{n,b} \psi_j, \chi_{n,a} \psi_i \rangle \\ &\quad + \langle \rho \chi_{n,b} (t_{n-1}^+) \psi_j, \chi_{n,a} (t_{n-1}^+) \psi_i \rangle), \\ (M_{a,b}^{0,V_h})_{i,j} &:= Q_n(\langle \rho \chi_{n,b} \psi_j, \chi_{n,a} \psi_i \rangle), \\ (S_{a,b})_{i,j} &:= Q_n(\langle C \boldsymbol{\varepsilon}(\chi_{n,b} \psi_j), \boldsymbol{\varepsilon}(\chi_{n,a} \psi_i) \rangle \\ &\quad + \langle C \boldsymbol{\varepsilon}(\chi_{n,b} \psi_j) \mathbf{n}, \chi_{n,a} \psi_i \rangle_{\Gamma_u^D}), \\ (N_{a,b}^A)_{i,j} &:= Q_n(a_\gamma(\chi_{n,b} \psi_j, \chi_{n,a} \psi_i)) \end{aligned}$$

as well as

$$\begin{aligned} (C_{a,b})_{r,j} &:= Q_n(-\alpha \langle \nabla \cdot (\chi_{n,b} \psi_j), \chi_{n,a} \xi_r \rangle \\ &\quad + \alpha \langle \chi_{n,b} \psi_j \cdot \mathbf{n}, \chi_{n,a} \xi_r \rangle_{\Gamma_u^D}), \\ (M_{a,b}^{1,Q_h})_{r,s} &:= Q_n(\langle c_0 \partial_t \chi_{n,b} \xi_s, \chi_{n,a} \xi_r \rangle \\ &\quad + \langle c_0 \chi_{n,b} (t_{n-1}^+) \xi_s, \chi_{n,a} (t_{n-1}^+) \xi_r \rangle) \end{aligned}$$

and

$$\begin{aligned} (B_{a,b})_{r,s} &:= Q_n(\langle \mathbf{K} \nabla(\chi_{n,b} \xi_s), \nabla(\chi_{n,a} \xi_r) \rangle \\ &\quad - \langle \mathbf{K} \nabla(\chi_{n,b} \xi_s) \cdot \mathbf{n}, \chi_{n,a} \xi_r \rangle_{\Gamma_p^D}), \\ (N_{a,b}^B)_{r,s} &:= Q_n(b_\gamma(\chi_{n,b} \xi_s, \chi_{n,a} \xi_r)), \end{aligned}$$

with  $a_\gamma(\cdot, \cdot)$  and  $b_\gamma(\cdot, \cdot)$  being defined in (3.3), for  $i, j = 1, \dots, R$  and  $r, s = 1, \dots, S$ . The vector  $\mathbf{F}_n$  in (4.5) is defined similarly, according to (3.5) along with (3.4). Its definition is skipped here for brevity. We note that (3.1a) is still multiplied by  $\rho > 0$  for the definition of the first row in (4.7). Multiplying the first block row in (4.7) by  $(-1)$  and recalling the symmetry of  $M_{a,b}^{1,V_h}$ , for  $a, b = 1, \dots, k$ , the upper left  $2 \times 2$  block subsystem in (4.7) itself admits the structure of the matrix (1.3). This might be exploited in future theoretical analyses of the solver or an improvement of the GMRES iterations (cf. [14, 36]), but is beyond the scope of the current work. For the family of finite element pairs with  $Q_h^l = Q_h^{l,\text{disc}}$  in (2.6), the definition of  $B_{a,b}$  has to be adjusted to the second of the options in (3.2c). Further, the contribution  $N_{a,b}^B$  has to be omitted.

Increasing values of the piecewise polynomial degree in time  $k$  enhance the complexity of the block structure of the system matrix  $\mathbf{A}_n$  in (4.6) along with (4.7). They impose an additional facet of challenge on the construction of efficient block preconditioners for (4.5). We solve the linear system (4.5) for the unknown  $\mathbf{X}_n$  on the subinterval  $I_n$  by flexible GMRES iterations [65] that are preconditioned by a V-cycle geometric multigrid (GMG) algorithm [76]. The ingredients of the GMRES–GMG approach are summarized in Alg. 4.1.

---

**Algorithm 4.1:** GMRES–GMG Solver

---

- Left preconditioning of each GMRES iteration by a single V-cycle GMG iteration.
  - Prolongation to grids by interpolation and application of the transpose operator for restriction.
  - $J_{\max}$  pre- and post-smoothing steps on each grid level  $\{T_l\}$  for  $l = 1, \dots, L$ .
  - Application of a (parallel) direct solver on the coarsest mesh partition  $T_0$ .
- 

In our computational studies of Sect. 5, the standard choice is  $J_{\max} = 4$  and the parallel direct solver is *SuperLU\_DIST* [53]. For the restriction and prolongation operators the deal.II classes *MultiGrid* and *MGTransferPrebuilt* are used. For the deal.II finite element library we refer to [7]. For details of the parallel implementation by the message passing interface (MPI) protocol of our GMG approach we also refer to [3].

The choice of the smoothing operator in the GMG method is extremely diverse; cf., e.g. [28, 47] for a further discussion. We use a collective smoother of Vanka type that is based on the solution of small local problems. Compared to the Navier–Stokes system [3], the dynamic Biot problem (1.1) requires adaptations in the construction of the local Vanka smoother. In inf-sup stable discretizations of the Navier–Stokes equations by the  $Q_r^d / P_{r-1}^{\text{disc}}$ ,  $r \geq 2$ , family of finite element spaces no coupling between the pressure degrees of freedom is involved due to the discontinuous approximation of the pressure variable. This feature leads to excellent performance properties of the Vanka smoother [3]. In contrast to this, the discretization (3.5c) of (1.1b) by  $P_{r-1}^{\text{disc}}$ ,  $r \geq 2$ , elements involves a coupling between degrees of freedom of the scalar variable  $p$  due to the presence of the face integrals over the average and jump in the second of the options in (3.2c) for the definition of the bilinear form  $B_\gamma$ . This coupling reduces the smoothing properties of elementwise Vanka operator. For the  $Q_{r-1}$ ,  $r \geq 2$ , family of elements for the scalar variable  $p$ , leading to the first of the options in (3.2c), the coupling of degrees of freedom of  $p$  by its continuous in space approximation and the loss of smoothing properties arise likewise. As a remedy, for both families of inf-sup stable approximation in (3.5) the local Vanka smoother is computed on overlapping patches of adjacent elements. In addition, an averaging of the patchwise updates is done after the Vanka smoother has been applied on all of them.

To construct the patchwise Vanka, we let on the mesh partition  $T_l$ , for  $l = 1, \dots, L$ , of the multigrid hierarchy the linear system, to be solved, be represented by

$$A_l d_l = b_l, \quad \text{for } l = 1, \dots, L. \tag{4.8}$$

To each grid node  $\xi_l^m$ , for  $m = 1, \dots, M_l$ , where  $M_l$  denotes the total number of grid nodes on the mesh partition  $\mathcal{T}_l$ , we built a patch of adjacent elements such that

$$P_l^m := \bigcup \{K \in \mathcal{T}_l \mid \xi_l^m \in \bar{K}\}, \quad \text{for } m = 1, \dots, M_l. \quad (4.9)$$

In two space dimensions  $P_l^m$  is built from four elements, if  $\xi_l^m \notin \partial\Omega$ . In three space dimensions,  $P_l^m$  has eight elements, if  $\xi_l^m \notin \partial\Omega$ . If  $\xi_l^m \in \partial\Omega$ , patches of less elements are obtained. On  $\mathcal{T}_l$ , let  $Z_l$  denote the index set of all global degrees of freedom with cardinality  $C_l$ ,

$$C_l := \text{card}(Z_l).$$

Let  $Z_l(P_l^m)$  denote the subset of  $Z_l$  all global degrees of freedom linked to the patch  $P_l^m$ , i.e. degrees of freedom of  $\mathbf{u}_{\tau,h}$ ,  $\mathbf{v}_{\tau,h}$  and  $p_{\tau,h}$  for the  $(k+1)$  Gauss–Radau time points of  $I_n$ . The cardinality of  $Z_l(P_l^m)$  is denoted by  $C_l^m$ ,

$$C_l^m := \text{card}(Z_l(P_l^m)), \quad \text{for } m = 1, \dots, M_l.$$

Further, we denote the index set of all local degrees of freedom on  $P_l^m$  by  $\hat{Z}_l(P_l^m) := \{0, \dots, C_l^m - 1\}$ . For the notation, we note that the set  $\hat{Z}_l(P_l^m)$  depends on the cardinality of the patch  $P_l^m$ . For a given patch  $P_l^m$  and a local degree of freedom with number  $\hat{\mu} \in \hat{Z}_l(P_l^m)$  let the mapping

$$\begin{aligned} \text{dof} : \mathcal{T}_l \times \hat{Z}_l(P_l^m) &\rightarrow Z_l, \\ \mu &= \text{dof}(P_l^m, \hat{\mu}) \in Z_l(P_l^m), \end{aligned} \quad (4.10)$$

yield the uniquely defined global number  $\mu \in Z_l$ . Finally, we put  $R = \dim \mathbf{V}_h$  and  $S = \dim Q_h$ .

We are now in a position to define the local Vanka operator for the patch  $P_l^m$  in an exact mathematical way; cf., e.g. [3, 40, 45, 80].

**Definition 4.2** (Patchwise Vanka smoother) For a patch  $P_l^m$ , for  $m \in \{1, \dots, M_l\}$ , let the  $P_l^m$ -local restriction operator  $\mathbf{R}_K : \mathbb{R}^{(k+1) \cdot (2R+S)} \rightarrow \mathbb{R}^{C_l^m}$  be defined by

$$\begin{aligned} (\mathbf{R}_{P_l^m} \mathbf{d})[\hat{\mu}] &= \mathbf{d}[\text{dof}(P_l^m, \hat{\mu})], \\ \text{for } \hat{\mu} &\in \hat{Z}_l(P_l^m), \end{aligned} \quad (4.11)$$

and, for the system matrix  $\mathbf{A}$  of (4.8), the patch system matrix  $\mathbf{A}_{P_l^m} \in \mathbb{R}^{C_l^m, C_l^m}$  by

$$\begin{aligned} \mathbf{A}_{P_l^m}[\hat{v}][\hat{\mu}] &:= \mathbf{A}_l[\text{dof}(P_l^m, \hat{v})][\text{dof}(P_l^m, \hat{\mu})], \\ \text{for } \hat{v}, \hat{\mu} &\in \hat{Z}_l(P_l^m). \end{aligned} \quad (4.12)$$

The local Vanka operator  $\mathbf{S}_{P_l^m} : \mathbb{R}^{(k+1) \cdot (2R+S)} \rightarrow \mathbb{R}^{C_l^m}$  is defined by

$$\mathbf{S}_{P_l^m}(\mathbf{d}) = \mathbf{R}_{P_l^m} \mathbf{d} + \omega \mathbf{A}_{P_l^m}^{-1} \mathbf{R}_{P_l^m} (\mathbf{b}_l - \mathbf{A}_l \mathbf{d}), \quad (4.13)$$

with some underrelaxation factor  $\omega > 0$ .

In the numerical experiments of Sect. 5 we choose  $\omega = 0.7$ . The  $P_l^m$ -local restriction operator (4.11) assigns to a global defect vector  $\mathbf{d} \in \mathbb{R}^{(k+1) \cdot (2R+S)}$  the local block vector  $\mathbf{R}_K \mathbf{d} \in \mathbb{R}^{C_l^m}$  that contains all components of  $\mathbf{d}$  that are associated with all degrees of freedom (for all  $(k+1)$  Gauss–Radau points of  $I_n$ ) belonging to the patch  $P_l^m$ . For the computation of the inverse  $(\mathbf{A}_{P_l^m})^{-1}$  in (4.13) we use LAPACK routines. The application of the smoother for (4.8) on the mesh partition  $\mathcal{T}_l$  is summarized in Algorithm 4.2.

**Algorithm 4.2:** Smoothing steps for (4.8) on mesh partition  $\mathcal{T}_l$ .

```

1 Initialize solution  $\mathbf{d}$  of (4.8) with  $\mathbf{0}$ ;
2 for  $j = 1, \dots, J_{\max}$  do
3   for  $(\mu = 0; \mu < C_l; \mu++)$  do
4      $p[\mu] = 0$ ;
5      $z[\mu] = 0.0$ ;
6   end
7   foreach patch  $P_l^m$  on  $\mathcal{T}_l$  do
8      $\mathbf{y} = \mathbf{S}_{P_l^m}(\mathbf{d})$ ;
9     for  $(\hat{\mu} = 0, \hat{\mu} < C_l^m; \hat{\mu}++)$  do
10       $z[\text{dof}(P_l^m, \hat{\mu})] += (\mathbf{E}_{P_l^m} \mathbf{y})[\text{dof}(P_l^m, \hat{\mu})]$ ;
11       $p[\text{dof}(P_l^m, \hat{\mu})]++$ ;
12    end
13  end
14  for  $(\mu = 0; \mu < C_l; \mu++)$  do
15     $d[\mu] = z[\mu]/p[\mu]$ ;
16  end
17 end

```

In line 1 of Alg. 4.2 the defect and solution vector is pre-initialized with  $\mathbf{0}$ . In line 4 the loop over all  $J_{\max}$  smoothing steps starts. In line 4 and 5 the counter vector  $\mathbf{p}$  for the number of updates of the degrees of freedom and the auxiliary vector  $\mathbf{z}$  are initialized with  $\mathbf{0}$ . In line 7 the loop over all patches  $P_l^m$ , for  $m = 1, \dots, M$ , starts. In line 8 the local Vanka smoother is applied on patch  $P_l^m$  to the current iterate  $\mathbf{d}$  of the defect vector and the image is stored in the local patch vector  $\mathbf{y}$ . In line 10 the local vector  $\mathbf{y}$  is assigned to an auxiliary global vector  $\mathbf{z}$  by the  $P_l^m$ -dependent extension operator  $\mathbf{E}_{P_l^m} : \mathbb{R}^{C_l^m} \rightarrow \mathbb{R}^{(k+1) \cdot (2R+S)}$ ,

$$(\mathbf{E}_{P_l^m} \mathbf{y})[\mu] = \begin{cases} y[\hat{\mu}], & \text{if } \exists \hat{\mu} \in \hat{Z}_l(P_l^m) : \mu = \text{dof}(K, \hat{\mu}), \\ 0, & \text{if } \mu \notin Z_l(P_l^m). \end{cases}$$

In line 11 the components of the counter  $\mathbf{p}$  are incremented for the (global) indices associated with the degrees of freedom of the patch  $P_l^m$  processed in the loop. Finally, in line 15 the arithmetic mean of the local (patchwise) updates  $\mathbf{z}$  is assigned to the update of the global defect vector  $\mathbf{d}$ .

Regarding the performance of Alg. 4.2 and the overall GMRES–GMG approach we note the following.

**Remark 4.3** • Averaging of the patchwise updates implemented in line 10 and 15 of Alg. 4.2, that is used instead of overwriting successively the (global) degrees of freedom within the patch loop starting in line 7, is essential and ensures the convergence and efficiency of the local Vanka smoother and, thereby, the desired performance of the overall GMRES–GMG linear solver. Without the averaging operation we encountered convergence problems of the GMRES–GMG solver for the experiments of Sect. 5.

- Likewise, the application of the Vanka smoother on the patches (4.9), instead of using an elementwise Vanka smoother on the element  $K$ , ensures its smoothing properties. The latter would lead to systems (4.13) of smaller dimension, however fails to smoothen errors. This is expected to be due coupling of the degrees of freedom of the scalar variable  $p$  in the spatial discretizations used here.

## 5 Numerical studies

In this section we study numerically the proposed space-time finite element and GMRES–GMG solver approach with respect to its computational and energy efficiency. Firstly, we demonstrate the accuracy of solutions in terms of convergence rates for a prescribed solution. Secondly, we analyze the convergence of the discretization for goal quantities of physical interest and the robustness of the GMRES–GMG solver in a two-dimensional test setting that is of interest in practice, for instance, in geomechanics for elucidating subsurface flow dynamics or in biomedical engineering for ultrasonic studies of bone or other calcified tissues to diagnose a variety of skeletal disorders. Finally, the investigations are extended to a challenging three-dimensional test case. Here a soft material with application in brain poromechanics [25] is used. The parallel scaling properties of our implementation are also investigated. Beyond these studies of classical performance engineering, the energy efficiency of the approach is considered further.

The implementation of the numerical scheme and the GMRES–GMG solver was done in an in-house high-performance frontend solver for the `deal.II` library [7]. For details of the parallel implementation of the geometric multigrid solver we refer to [3]. In all numerical experiments, the stopping criterion for the GMRES iterations is an absolute residual smaller than  $1e-8$ . The computations were performed on a Linux cluster with 571 nodes, each of them with 2 CPUs and 36 cores per CPU. The CPUs are *Intel Xeon Platinum 8360Y* with a base frequency of 2.4 GHz, a maximum turbo frequency of 3.5 GHz and a level 3 cache of 54 MB. Each node has 252 GB of main memory.

## 5.1 Accuracy of the discretization: experimental order of convergence

We study (1.1) for  $\Omega = (0, 1)^2$  and  $I = (1, 2]$  and the prescribed solution

$$\begin{aligned} \mathbf{u}(\mathbf{x}, t) &= \phi(\mathbf{x}, t)\mathbf{E}_2 \\ \text{and } p(\mathbf{x}, t) &= \phi(\mathbf{x}, t) \\ \text{with } \phi(\mathbf{x}, t) &= \sin(\omega_1 t^2) \sin(\omega_2 x_1) \sin(\omega_2 x_2) \end{aligned} \quad (5.1)$$

and  $\omega_1 = \omega_2 = \pi$ . We put  $\rho = 1.0$ ,  $\alpha = 0.9$ ,  $c_0 = 0.01$  and  $\mathbf{K} = \mathbf{E}_2$  with the identity  $\mathbf{E}_2 \in \mathbb{R}^{2,2}$ . For the fourth order elasticity tensor  $\mathbf{C}$ , isotropic material properties with Young’s modulus  $E = 100$  and Poisson’s ratio  $\nu = 0.35$ , corresponding to the Lamé parameters  $\lambda = 86.4$  and  $\mu = 37.0$ , are chosen. For an experiment with larger values of  $\lambda$  and  $\mu$  we refer to Table 10 in the appendix. In our experiments, the norm of  $L^\infty(I; L^2)$  is approximated by computing the function’s maximum value in the Gauss quadrature nodes  $t_{n,m}$  of  $I_n$ , i.e.,

$$\|w\|_{L^\infty(I; L^2)} \approx \max\{\|w|_{I_n}(t_{n,m})\| \mid m = 1, \dots, M, n = 1, \dots, N\}, \quad \text{with } M = 100.$$

We study the space-time convergence behavior of the scheme (3.5). For this, the domain  $\Omega$  is decomposed into a sequence of successively refined meshes of quadrilateral finite elements. The spatial and temporal mesh sizes are halved in each of the refinement steps. The step sizes of the coarsest space and time mesh are  $h_0 = 1/(2\sqrt{2})$  and  $\tau_0 = 0.1$ . We choose the polynomial degree  $k = 2$  and  $r = 3$ , such that discrete solutions  $\mathbf{u}_{\tau,h}, \mathbf{v}_\tau \in Y_\tau^2(\mathbf{V}_h)$ ,  $p_{\tau,h} \in Y_\tau^2(Q_h)$  with local spaces  $\mathbb{Q}_3^2/P_2^{\text{disc}}$  are obtained, as well as  $k = 3$  and  $r = 4$  with  $\mathbf{u}_{\tau,h}, \mathbf{v}_\tau \in Y_\tau^3(\mathbf{V}_h)$  and  $p_{\tau,h} \in Y_\tau^3(Q_h)$  and local spaces  $\mathbb{Q}_4^2/\mathbb{P}_3^{\text{disc}}$ ; cf. (2.2), (2.5) and (2.6). The calculated errors and corresponding experimental orders of convergence are summarized in Tables 1 and 2, respectively. The error is measured in the quantities associated with the energy of the system (1.1); cf. [43, p. 15] and [11]. Tables 1 and 2 nicely confirm the optimal rates of convergence with respect to the polynomial degrees in space and time of the overall approach. A notable difference in the convergence behavior between the pairs  $\mathbb{Q}_r^2/\mathbb{P}_{r-1}^{\text{disc}}$  and  $\mathbb{Q}_r^2/Q_{r-1}$  of local spaces for the discretization of the spatial variables is not observed. For completeness, we summarize in Appendix C the convergence results obtained for the pair  $\mathbb{Q}_r^2/Q_{r-1}$  of spaces of the Taylor–Hood family. A minor superiority of the pair  $\mathbb{Q}_r^2/\mathbb{P}_{r-1}^{\text{disc}}$  over the pair  $\mathbb{Q}_r^2/Q_{r-1}$  is only seen in the approximation of the scalar variable  $p$ . The coincidence of the convergence results also holds for the application of Problem B.1 instead of Problem 3.1. The two

**Table 1**  $L^2(L^2)$  and  $L^\infty(L^2)$  errors and experimental orders of convergence (EOC) for (5.1) with temporal polynomial degree  $k = 2$  and spatial degree  $r = 3$  for local spaces  $\mathbb{Q}_r^2/\mathbb{P}_{r-1}^{\text{disc}}$

$\tau$	$h$	$\ \nabla(\mathbf{u} - \mathbf{u}_{\tau,h})\ _{L^2(L^2)}$	EOC	$\ \mathbf{v} - \mathbf{v}_{\tau,h}\ _{L^2(L^2)}$	EOC	$\ p - p_{\tau,h}\ _{L^2(L^2)}$	EOC
$\tau_0/2^0$	$h_0/2^0$	1.2544218392e-02	–	3.4897282317e-02	–	2.4070118274e-03	–
$\tau_0/2^1$	$h_0/2^1$	1.5227995262e-03	3.04	3.9246006564e-03	3.15	2.8841669021e-04	3.06
$\tau_0/2^2$	$h_0/2^2$	1.8904870171e-04	3.01	4.8175203148e-04	3.03	3.5986044195e-05	3.00
$\tau_0/2^3$	$h_0/2^3$	2.3544020929e-05	3.01	5.9913957453e-05	3.01	4.4985485037e-06	3.00
$\tau_0/2^4$	$h_0/2^4$	2.9374127099e-06	3.00	7.4788750858e-06	3.00	5.6221925444e-07	3.00
$\tau$	$h$	$\ \nabla(\mathbf{u} - \mathbf{u}_{\tau,h})\ _{L^\infty(L^2)}$	EOC	$\ \mathbf{v} - \mathbf{v}_{\tau,h}\ _{L^\infty(L^2)}$	EOC	$\ p - p_{\tau,h}\ _{L^\infty(L^2)}$	EOC
$\tau_0/2^0$	$h_0/2^0$	8.1391652440e-02	–	1.6415428887e-01	–	1.0828068294e-02	–
$\tau_0/2^1$	$h_0/2^1$	1.1474006637e-02	2.83	2.8042308570e-02	2.55	1.7292127803e-03	2.65
$\tau_0/2^2$	$h_0/2^2$	1.5142126866e-03	2.92	3.6968566445e-03	2.92	2.3686467755e-04	2.87
$\tau_0/2^3$	$h_0/2^3$	1.9337108028e-04	2.97	4.6572185327e-04	2.99	3.0596110688e-05	2.95
$\tau_0/2^4$	$h_0/2^4$	2.4392244385e-05	2.99	5.8210228994e-05	3.00	3.8749300434e-06	2.98

**Table 2**  $L^2(L^2)$  and  $L^\infty(L^2)$  errors and experimental orders of convergence (EOC) with temporal polynomial degree  $k = 3$  and spatial degree  $r = 4$  for local spaces  $\mathbb{Q}_r^2/\mathbb{P}_{r-1}^{\text{disc}}$

$\tau$	$h$	$\ \nabla(\mathbf{u} - \mathbf{u}_{\tau,h})\ _{L^2(L^2)}$	EOC	$\ \mathbf{v} - \mathbf{v}_{\tau,h}\ _{L^2(L^2)}$	EOC	$\ p - p_{\tau,h}\ _{L^2(L^2)}$	EOC
$\tau_0/2^0$	$h_0/2^0$	2.3958455291e-03	–	1.7185653242e-02	–	7.0604463908e-04	–
$\tau_0/2^1$	$h_0/2^1$	1.0529085600e-04	4.51	5.4558622568e-04	4.98	3.4799927094e-05	4.34
$\tau_0/2^2$	$h_0/2^2$	5.9749777345e-06	4.14	1.9143193313e-05	4.83	1.2074793168e-06	4.85
$\tau_0/2^3$	$h_0/2^3$	3.6858164230e-07	4.02	9.6894400389e-07	4.30	6.1656313807e-08	4.29
$\tau_0/2^4$	$h_0/2^4$	2.2976498320e-08	4.00	5.6982903695e-08	4.09	3.6688337244e-09	4.07
$\tau$	$h$	$\ \nabla(\mathbf{u} - \mathbf{u}_{\tau,h})\ _{L^\infty(L^2)}$	EOC	$\ \mathbf{v} - \mathbf{v}_{\tau,h}\ _{L^\infty(L^2)}$	EOC	$\ p - p_{\tau,h}\ _{L^\infty(L^2)}$	EOC
$\tau_0/2^0$	$h_0/2^0$	1.3014842843e-02	–	1.1587851244e-01	–	4.9827180852e-03	–
$\tau_0/2^1$	$h_0/2^1$	8.7944275936e-04	3.89	4.1645282584e-03	4.80	2.6180886102e-04	4.25
$\tau_0/2^2$	$h_0/2^2$	5.6086043381e-05	3.97	2.0121085807e-04	4.37	9.4019466765e-06	4.80
$\tau_0/2^3$	$h_0/2^3$	3.5171465193e-06	4.00	1.1873567590e-05	4.08	5.0100645993e-07	4.23
$\tau_0/2^4$	$h_0/2^4$	2.1959929657e-07	4.00	7.3034599295e-07	4.02	3.0246305328e-08	4.05

discrete problems differ from each other by the discretization of the term  $\alpha \nabla \cdot \partial_t \mathbf{u}$  in (1.1b).

Next, we show numerically that the time discretization is even superconvergent of order  $2k + 1$  in the discrete time nodes  $t_n$ , for  $n = 1, \dots, N$ . For this, we introduce the time mesh dependent norm

$$\|w\|_{l^\infty(L^2)} := \max\{\|w(t_n)\| \mid n = 1, \dots, N\}. \tag{5.2}$$

We prescribe the solution

$$\begin{aligned} \mathbf{u}(\mathbf{x}, t) &= \begin{pmatrix} -2(x-1)^2x^2(y-1)y(2y-1)\sin(\omega_1t) \\ 2(x-1)x(2x-1)(y-1)^2y^2\sin(\omega_1t) \end{pmatrix}, \\ p(\mathbf{x}, t) &= -2(x-1)^2x^2(y-1)y(2y-1)\sin(\omega_2t) \end{aligned} \tag{5.3}$$

with  $\omega_1 = 40 \cdot \pi$  and  $\omega_2 = 10 \cdot \pi$ . We put  $\rho = 1.0$ ,  $\alpha = 0.9$ ,  $c_0 = 0.01$  and  $\mathbf{K} = \mathbf{E}_2$  with the identity  $\mathbf{E}_2 \in \mathbb{R}^{2,2}$ . For the elasticity tensor  $\mathbf{C}$ , isotropic material properties with Young’s modulus  $E = 100$  and Poisson’s ratio  $\nu = 0.35$  are used. For the local spaces we choose the pair  $\mathbb{Q}_5^2/\mathbb{Q}_4$  such that, for any  $t \in [0, T]$ , the solution (5.3) belongs to the discrete spaces  $\mathbf{V}_h$  and  $Q_h$ , respectively, and its spatial approximation is exact. This simplification is done here since we aim to study the convergence of the temporal discretization only. In the experiment, we choose  $k = 2$  such that discrete solutions  $\mathbf{u}_{\tau,h}, \mathbf{v}_\tau \in Y_\tau^2(\mathbf{V}_h)$  and  $p_{\tau,h} \in X_\tau^2(Q_h)$  are obtained. We use a spatial mesh that consists of 16 cells with  $h = 1/(2\sqrt{2})$  and set  $\tau_0 = 0.02$ . The calculated errors and corresponding experimental orders of convergence are summarized in Table 3. Superconvergence of order  $2k + 1$  in the discrete

**Table 3**  $L^2(L^2)$  and  $l^\infty(L^2)$  errors and experimental orders of convergence (EOC) for (5.3) with temporal polynomial degree  $k = 2$  and spatial degree  $r = 5$  for local spaces  $\mathbb{Q}_r^2/\mathbb{Q}_{r-1}$ , showing superconvergence of order  $2k + 1$  in the discrete time nodes  $t_n$ , i.e., w.r.t. the norm  $\|\cdot\|_{l^\infty(L^2)}$  defined in (5.2)

$\tau$	$\ \nabla(\mathbf{u} - \mathbf{u}_{\tau,h})\ _{L^2(L^2)}$	EOC	$\ \mathbf{v} - \mathbf{v}_{\tau,h}\ _{L^2(L^2)}$	EOC	$\ p - p_{\tau,h}\ _{L^2(L^2)}$	EOC
$\tau_0/2^0$	3.3600958819e-03	–	5.2037881317e-02	–	8.5492528946e-05	–
$\tau_0/2^1$	3.9552579243e-04	3.09	6.6291344116e-03	2.97	4.8852611508e-06	4.13
$\tau_0/2^2$	4.8913600397e-05	3.02	8.3404665964e-04	2.99	1.7945937694e-07	4.77
$\tau_0/2^3$	6.1123116212e-06	3.00	1.0446618892e-04	3.00	1.0554654134e-08	4.09
$\tau_0/2^4$	7.6411713541e-07	3.00	1.3065222159e-05	3.00	1.1603768561e-09	3.19
$\tau_0/2^5$	9.5518196224e-08	3.00	1.6333729266e-06	3.00	1.4370494853e-10	3.01
$\tau_0/2^6$	1.1939894351e-08	3.00	2.0417851944e-07	3.00	1.7952564965e-11	3.00
$\tau$	$\ \nabla(\mathbf{u} - \mathbf{u}_{\tau,h})\ _{l^\infty(L^2)}$	EOC	$\ \mathbf{v} - \mathbf{v}_{\tau,h}\ _{l^\infty(L^2)}$	EOC	$\ p - p_{\tau,h}\ _{l^\infty(L^2)}$	EOC
$\tau_0/2^0$	2.1302171198e-03	–	1.3451755096e-02	–	1.9152238265e-04	–
$\tau_0/2^1$	1.1003144753e-04	4.28	6.9871664304e-04	4.27	1.0401014806e-05	4.20
$\tau_0/2^2$	3.6416424630e-06	4.92	2.3145015656e-05	4.92	3.5819518738e-07	4.86
$\tau_0/2^3$	1.1784707720e-07	4.95	7.3594357548e-07	4.97	1.1336271728e-08	4.98
$\tau_0/2^4$	3.6966657143e-09	4.99	2.3096831287e-08	4.99	3.5849645590e-10	4.98
$\tau_0/2^5$	1.1560104141e-10	5.00	7.2294482894e-10	5.00	1.1225015472e-11	5.00
$\tau_0/2^6$	3.5626018095e-12	5.02	2.2323871214e-11	5.02	3.4673336542e-13	5.02

time nodes is clearly observed in the second of the arrays in Table 3.

### 5.2 Computational efficiency: accuracy of goal quantities and convergence of the GMRES–GMG solver in a 2d test case

For a two-dimensional test problem we study the potential of the proposed approach to compute reliably and efficiently goal quantities of physical interest. We also document the performance properties of the GMRES-GMG solver of Sect. 4 for the applied space-time finite element methods. Even though the test setting is still of academic nature, it is related to problems of practical interest in civil engineering (subsurface dynamics) or biomedical engineering (cf. [81]). In the numerical investigations a stiff material is assumed whereas in Sect. 5.3 a softer material will be studied. This is done for the sake of considering also a range of materials.

Beyond the boundary conditions (1.1d) and (1.1e), we also apply (homogeneous) directional (or componentwise) boundary conditions for  $\mathbf{u}$  on some part  $\Gamma_u^d \subset \partial\Omega$ . The directional boundary conditions read as

$$\mathbf{u} \cdot \mathbf{n} = 0 \quad \text{and} \quad (\boldsymbol{\sigma}(\mathbf{u})\mathbf{n}) \cdot \mathbf{t}_i = 0, \quad \text{for } i = 1, \dots, d - 1, \quad \text{on } \Gamma_u^d \times (0, T], \quad (5.4)$$

for the stress tensor  $\boldsymbol{\sigma}(\mathbf{u}) = \mathbf{C}\boldsymbol{\varepsilon}(\mathbf{u})$  and the unit basis vectors  $\mathbf{t}_i$ , for  $i = 1, \dots, d - 1$ , of the tangent space at  $\mathbf{x} \in \Gamma_u^d$ . In the definition of  $A_\gamma$  in (3.2a) and (3.3a), the conditions (5.4) still need to be implemented properly. By the second of the conditions in (5.4) we get for the second of the terms on the right-hand of (3.2a) that

$$\begin{aligned} & \langle \mathbf{C}\boldsymbol{\varepsilon}(\mathbf{u}_{\tau,h})\mathbf{n}, \boldsymbol{\chi}_{\tau,h} \rangle_{\Gamma_u^d} \\ &= \langle \mathbf{C}\boldsymbol{\varepsilon}(\mathbf{u}_{\tau,h})\mathbf{n} \cdot \mathbf{n}, \mathbf{n} \cdot \boldsymbol{\chi}_{\tau,h} \rangle_{\Gamma_u^d}. \end{aligned} \quad (5.5)$$

For the boundary part  $\Gamma_u^d$  we then put, similarly to (3.3a),

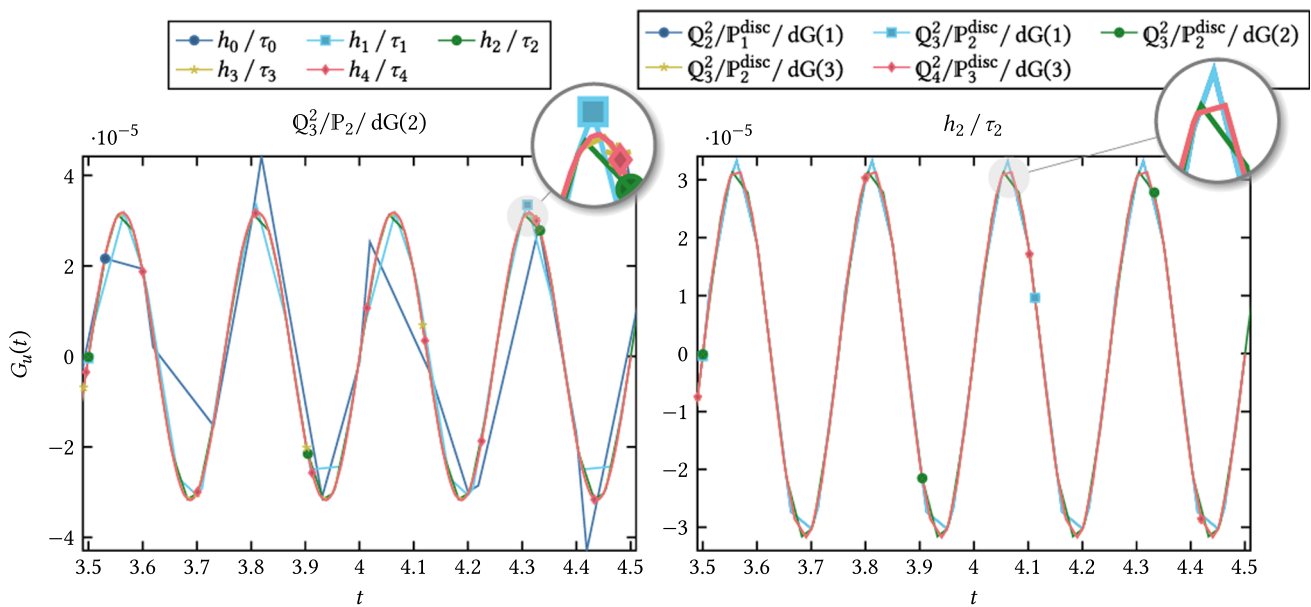
$$\begin{aligned} a_\gamma^d(\mathbf{w}, \boldsymbol{\chi}_h) &:= -\langle \mathbf{w} \cdot \mathbf{n}, \mathbf{C}\boldsymbol{\varepsilon}(\boldsymbol{\chi}_h)\mathbf{n} \cdot \mathbf{n} \rangle_{\Gamma_u^d} \\ &+ \gamma_a h^{-1} \langle \mathbf{w} \cdot \mathbf{n}, \boldsymbol{\chi}_h \cdot \mathbf{n} \rangle_{\Gamma_u^d}, \end{aligned}$$

while leaving  $a_\gamma$  unmodified for the part  $\Gamma_u^D$  where Dirichlet boundary conditions are prescribed for  $\mathbf{u}$ . In its entirety, we thus have that  $a_\gamma(\cdot, \cdot) := a_\gamma^D(\cdot, \cdot) + a_\gamma^d(\cdot, \cdot)$  with  $a_\gamma^D(\cdot, \cdot)$  being defined by the right-hand side of (3.3a). By the arguments of Appendix A, the coercivity of the resulting form  $A_\gamma$  is still ensured.

For our experiments, we consider the rectangular domain  $\Omega = (0, 0.5) \times (0, 1) \subset \mathbb{R}^2$  with boundary segments  $\Gamma_u^d = \{0\} \times (0, 1) \cup \{0.5\} \times (0, 1)$  and  $\Gamma_u^N = (0, 0.5) \times \{0\} \cup (0, 0.5) \times \{1\}$ . On the lower and upper part  $\Gamma_u^N$  of the boundary we impose in the boundary condition (1.1e) the traction force

$$\begin{aligned} \mathbf{t}_N &= \begin{pmatrix} 0 \\ s(t) \cdot 16x \cdot (x - 0.5) \cdot \sin(8\pi t) \end{pmatrix} \\ \text{with } s(t) &:= \begin{cases} 0.5 - 0.5 \cos(4\pi t^2), & \text{for } t < 0.5, \\ 1, & \text{else,} \end{cases} \end{aligned} \quad (5.6)$$

which amounts to applying a simultaneous compression or decompression force at the upper and lower boundary. For the scalar variable  $p$  we prescribe a homogeneous Dirichlet boundary condition (1.1f) on the lower and upper part



**Fig. 2** Goal quantity  $G_u$  of (5.7) for a sequence of successively refined meshes (left) and different polynomial orders of the STFEM on a fixed mesh (right)

$\Gamma_p^N = (0, 0.5) \times \{0\} \cup (0, 0.5) \times \{1\}$  of  $\partial\Omega$  and a homogeneous Neumann boundary condition (1.1g) else. We put  $\rho = 1.0$ ,  $\alpha = 0.9$ ,  $c_0 = 0.01$  and  $\mathbf{K} = \mathbf{E}_2$  with the identity  $\mathbf{E}_2 \in \mathbb{R}^{2,2}$ . For the elasticity tensor  $\mathbf{C}$ , isotropic material properties with Young’s modulus  $E = 20000$  and Poisson’s ratio  $\nu = 0.3$ ; cf. [81]. The final simulation time is  $T = 4.5$ . As goal quantities of this problem, we measure the magnitude of the displacement variable in normal direction as well as the pressure on a cross section plane  $\Gamma_m$ , given by

$$G_u = \int_{\Gamma_m} \mathbf{u} \cdot \mathbf{n} \, d\sigma,$$

$$G_p = \int_{\Gamma_m} p \, d\sigma, \quad \text{for } \Gamma_m := (0, 0.5) \times \{0.25\}. \quad (5.7)$$

We set the step sizes of the coarsest space-time mesh to  $h_0 = 0.125$  and  $\tau_0 = 0.2$ . Further mesh levels are obtained by a successive refinement by a factor of two such that  $h_i = h_0 \cdot 2^{-i}$  and  $\tau_i = \tau_0 \cdot 2^{-i}$  for  $i \in \mathbb{N}$ .

In Fig. 2 and Table 4 we illustrate the space-time convergence of the goal quantities and their maximum and minimum value in the final part  $t \in [3.5, 4.5]$  of the simulation time  $t \in [0, 4.5]$ . Various polynomial orders of the discretization in space and time are used. For brevity, only  $G_u$  is visualized in Fig. 2. Convergence of the goal quantity is clearly observed even though the differences are not strong in this two dimensional test case. In particular, in Table 4 we observe a dominating temporal discretization error and the gain in accuracy by higher order time discretization. Furthermore, in Table 5 we summarize the average number of GMRES iterations per time step needed to solve the resulting

linear system of equations. Since the GMG method with a single V-cycle is used as preconditioner and not as a system solver itself, the average number of GMRES iterations and the wall clock time are considered to be a reasonable measure for the performance of the GMRES–GMG solver. The robustness of the GMRES–GMG solver with respect to the refinement of the space-time mesh and the polynomial degrees in space and time is confirmed. In Table 6 we compare the contribution of assembling and solving to the wall clock time. Further, the impact of the number of pre- and post-smoothing steps  $J_{\max}$  in Alg. 4.2 on the performance of the GMRES–GMG solver is analyzed. The results show increasing wall clock time for higher numbers of smoothing steps. The relatively high computational costs of the GMRES–GMG suggests further performance tuning of the smoother. Nevertheless, the robustness of the GMG preconditioned GMRES iterations in Table 5 underline their potential as efficient black box solver for higher order STFEMs with complex block structures. In our numerical experiments, including the three-dimensional case (cf. Sect. 5.3) and further tests not documented here, we found that choosing  $J_{\max} = 4$  or  $J_{\max} = 5$  usually leads to a robust performance.

### 5.3 Computational efficiency: accuracy of goal quantities and convergence of the GMRES–GMG solver in a 3d test case

In this section we extend the numerical studies of the previous section to three space dimensions. For the geometry, we consider the pipe socket that is visualized in Fig. 3a. The pipe has a diameter of  $d = 2$  and consists in the  $x_1 - x_3$  plane of

**Table 4** Maximum and minimum of the goal quantities (5.7) in the subinterval  $t \in [3.5, 4.5]$  of the simulation time  $t \in [0, 4.5]$  for different space-time mesh refinements and approximation orders

		$\min\{G_u(t) \mid t \in [3.5, 4.5]\}$				
		$dG(1)$		$dG(2)$	$dG(3)$	
$h$	$\tau$	$Q_2^2/\mathbb{P}_1^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_4^2/\mathbb{P}_3^{\text{disc}}$
$h_0$	$\tau_0$	-3.3423e-5	-3.3422e-5	-4.2889e-5	-3.0480e-5	-3.0480e-5
$h_1$	$\tau_1$	-3.5801e-5	-3.5801e-5	-3.0650e-5	-3.1611e-5	-3.1611e-5
$h_2$	$\tau_2$	-3.0231e-5	-3.0231e-5	-3.1566e-5	-3.1715e-5	-3.1715e-5
$h_3$	$\tau_3$	-3.1797e-5	-3.1797e-5	-3.1653e-5	-3.1730e-5	-3.1731e-5
$h_4$	$\tau_4$	-3.1807e-5	-3.1807e-5	-3.1807e-5	-3.1821e-5	-3.1823e-5
$h_5$	$\tau_5$	-3.1803e-5	-3.1805e-5	-3.1811e-5	-	-
		$\max\{G_u(t) \mid t \in [3.5, 4.5]\}$				
		$dG(1)$		$dG(2)$	$dG(3)$	
$h$	$\tau$	$Q_2^2/\mathbb{P}_1^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_4^2/\mathbb{P}_3^{\text{disc}}$
$h_0$	$\tau_0$	-3.8341e-5	-3.8341e-5	4.4140e-5	-3.3208e-5	-3.3208e-5
$h_1$	$\tau_1$	-3.3593e-5	-3.3593e-5	-3.3484e-5	-3.1293e-5	-3.1293e-5
$h_2$	$\tau_2$	-3.3275e-5	-3.3275e-5	-3.1296e-5	-3.1293e-5	-3.1293e-5
$h_3$	$\tau_3$	-3.1797e-5	-3.1797e-5	-3.1653e-5	-3.1730e-5	-3.1731e-5
$h_4$	$\tau_4$	-3.1808e-5	-3.1807e-5	-3.1808e-5	-3.1808e-5	-3.1807e-5
$h_5$	$\tau_5$	-3.1803e-5	-3.1806e-5	-3.1838e-5	-	-
		$\min\{G_p(t) \mid t \in [3.5, 4.5]\}$				
		$dG(1)$		$dG(2)$	$dG(3)$	
$h$	$\tau$	$Q_2^2/\mathbb{P}_1^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_4^2/\mathbb{P}_3^{\text{disc}}$
$h_0$	$\tau_0$	-3.5401e-4	-3.5298e-4	4.9249e-4	-3.6870e-4	-3.6874e-4
$h_1$	$\tau_1$	-3.7277e-4	-3.7255e-4	-2.5756e-4	-2.6282e-4	-2.6285e-4
$h_2$	$\tau_2$	-2.5150e-4	-2.5153e-4	-2.6345e-4	-2.6518e-4	-2.6518e-4
$h_3$	$\tau_3$	-2.7118e-4	-2.7189e-4	-2.6918e-4	-2.6876e-4	-2.6887e-4
$h_4$	$\tau_4$	-2.6860e-4	-2.6911e-4	-2.6878e-4	-2.6916e-4	-2.6931e-4
$h_5$	$\tau_5$	-2.6816e-4	-2.6918e-4	-2.6995e-4	-	-
		$\max\{G_p(t) \mid t \in [3.5, 4.5]\}$				
		$dG(1)$		$dG(2)$	$dG(3)$	
$h$	$\tau$	$Q_2^2/\mathbb{P}_1^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_3^2/\mathbb{P}_2^{\text{disc}}$	$Q_4^2/\mathbb{P}_3^{\text{disc}}$
$h_0$	$\tau_0$	-3.7565e-4	-3.7451e-4	-4.9426e-4	-4.1165e-4	-4.1166e-4
$h_1$	$\tau_1$	-3.5279e-4	-3.5271e-4	-2.9835e-4	-2.9124e-4	-2.9128e-4
$h_2$	$\tau_2$	-2.9851e-4	-2.9828e-4	-2.7066e-4	-2.7177e-4	-2.7174e-4
$h_3$	$\tau_3$	-2.7164e-4	-2.7195e-4	-2.6866e-4	-2.6875e-4	-2.6887e-4
$h_4$	$\tau_4$	-2.6844e-4	-2.6911e-4	-2.6872e-4	-2.6931e-4	-2.6935e-4
$h_5$	$\tau_5$	-2.6817e-4	-2.6918e-4	-2.7036e-4	-	-

three parts: a quarter annulus with  $L_0 = \frac{\pi}{2}$ , an upper part, with  $L_1 = L_0$  and a lower part with  $L_2 = \frac{L_0}{2}$ . In contrast to Sect. 5.2, a soft material of brain poromechanics is now chosen; cf. [25]. We put  $\rho = 10^3$  [kg/m<sup>3</sup>],  $\alpha = 0.49$  [-],  $c_0 = 10^{-6}$  [m<sup>2</sup>/N] and  $\mathbf{K} = k_0 \mathbf{E}_3$ , with  $k_0 = 1.0$  [m<sup>2</sup>/Pa] and the identity matrix  $\mathbf{E}_3 \in \mathbb{R}^{3,3}$ . For the elasticity tensor  $\mathbf{C}$ , isotropic material properties with the Lamé parameter  $\lambda = 505$  [Pa] and  $\mu = 216$  [Pa], corresponding to Young's modulus  $E = 583.3$  [Pa] and Poisson's ratio  $\nu = 0.35$  [-],

are used. The geometry is supposed to mimic brain tissue or some section of an artery with neglecting the blood flow inside. On the curved surface area the directional boundary conditions (5.4) are prescribed. At the top and right outlet of the pipe socket a homogeneous Dirichlet condition for the pressure variable is used. For the displacement variable the traction force of (1.1e) is defined at the top ( $x_1 = L_1$ ) outlet by

**Table 5** Average number of performed GMG preconditioned GMRES iterations per time step

$h$	$\tau$	dG(1)		dG(2)	dG(3)	
		$\mathbb{Q}_2^2/\mathbb{P}_1^{\text{disc}}$	$\mathbb{Q}_3^2/\mathbb{P}_2^{\text{disc}}$	$\mathbb{Q}_3^2/\mathbb{P}_2^{\text{disc}}$	$\mathbb{Q}_3^2/\mathbb{P}_2^{\text{disc}}$	$\mathbb{Q}_4^2/\mathbb{P}_3^{\text{disc}}$
<i>(a) Average number of GMRES iterations per time step for different space-time polynomial degrees and mesh refinement levels</i>						
$h_0$	$\tau_0$	2.00	2.57	2.35	2.00	2.17
$h_1$	$\tau_1$	1.98	2.71	2.69	2.51	1.98
$h_2$	$\tau_2$	1.97	1.98	1.85	1.97	1.97
$h_3$	$\tau_3$	1.95	1.96	1.96	1.96	1.94
$h_4$	$\tau_4$	1.83	1.84	1.93	1.93	1.66
$h_5$	$\tau_5$	1.03	1.38	1.53	–	–
$h \setminus \tau$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$		
<i>(b) Average number of GMRES iterations for <math>\mathbb{Q}_3^2/\mathbb{P}_2^{\text{disc}}</math> and dG(2)</i>						
$h_2$	1.85	1.83	1.79	1.77		
$h_3$	1.80	1.96	1.71	1.69		
$h_4$	1.80	1.78	1.93	1.61		
$h_5$	1.80	1.75	1.60	1.53		

**Table 6** Wall clock Time (WT) accumulated over all time steps and Percentage (P) of total wall clock time for  $\mathbb{Q}_3/\mathbb{P}_2$  space and dG(2) time discretization on mesh ( $h_4, \tau_4$ ) for different numbers of patchwise Vanka smoothing steps  $J_{\max}$ ; cf. Alg. 4.2

Operation	$J_{\max} = 2$		$J_{\max} = 5$		$J_{\max} = 8$	
	WT [s]	P [%]	WT [s]	P [%]	WT [s]	P [%]
Assemble r.h.s.	58.1	2.0	58.3	1.3	58.4	0.99
Assemble matrix	277.0	9.4	276.0	6.1	274.0	4.6
Linear solver	2150.0	73.0	3720.0	83.0	5120.0	87.0
Setup Vanka smoother	352.0	12.0	363.0	8.1	365.0	6.2
Total wall time	2930		4500		5910	

$$t_N = \left( s(t) \sin(2\pi t) \left( \sqrt{x_2^2 + (x_3 - 2)^2} - 1 \right), 0, 0 \right)^\top$$

with  $s(t)$  being defined in (5.6), and at the right ( $x_3 = L_2$ ) outlet by

$$t_N = \left( 0, 0, s(t) \sin(2\pi t) \left( \sqrt{(x_1 - 2)^2 + x_2^2} - 1 \right) \right)^\top.$$

We measure the benchmark quantities defined in (5.7) on the cross section plane  $\Gamma_m : (\mathbf{x} - \mathbf{p}_m) \cdot \mathbf{n}_m = 0$ , with  $\mathbf{p}_m = \left( \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)^\top$  and  $\mathbf{n}_m = \left( \frac{\sqrt{3}}{2}, 0, -\frac{1}{\sqrt{2}} \right)^\top$ ; cf. Fig. 3a.

We put  $I = (0, T]$  with  $T = 7$  and set the time step size of the temporal discretization to  $\tau = \frac{0.02}{(ref_n - 3)}$ , where  $ref_n$  is the number of spatial refinement levels of the spatial grid, cf. Table 7. The spatial polynomial degree is fixed to  $r = 1$  for all simulations (cf. Sect. 2). The calculated profile for  $t = T$  of the modulus of the vectorial variable  $\mathbf{u}$  and the scalar variable  $p$  are illustrated in Fig. 3. In Table 7 we summarize characteristic quantities and results of our simulations for various

spatial and temporal resolutions and different temporal polynomial degrees of the STFEMs. In Fig. 4 we visualize the computed benchmark quantities (5.7) of some of the simulations over the temporal axis. The benchmark quantities on each refinement level are within the same range for temporal discretizations with polynomial degree  $k = 1$  and  $k = 2$ , but for  $k = 1$  oscillations on  $ref_4$  and  $ref_5$  are observed, which is not the case for  $k = 2$  or  $k = 1$  when using a finer time step size ( $ref_n = 6$ ). This indicates the superiority of higher order discretization schemes in the time domain.

Table 7 summarizes the results of our numerical convergence study for the goal quantities. The final row of Table 7 contains the results of the finest simulation that we could run on our hardware. Table 7 and Fig. 4 show that the solution (i.e., the goal quantities) is nearly fully converged. The final column of Table 7 summarizes the convergence statistics of the proposed GMRES–GMG solver. In terms of the average number of iterations per time step the solver is (almost) grid independent. This underlines its capability and robustness for solving efficiently the complex systems arising from space-



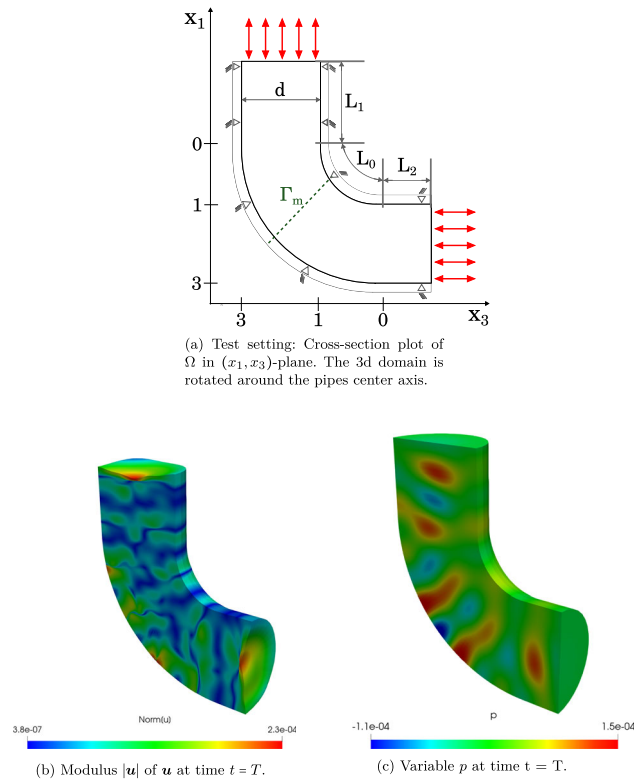


Fig. 3 Problem setting and profile of the solution at time  $t = T$

time finite element discretizations of the considered coupled hyperbolic–parabolic system.

### 5.4 Parallel scaling and energy efficiency

Here, we study briefly the parallel scaling and energy efficiency of our solver. By studying energy consumption, we’d like to draw attention to this emerging dimension in the tuning of algorithms. Energy efficiency broadens the classical hardware-oriented numerics that is applied to enhance the performance of the current method on the target hardware and/or to find other numerical methods to improve the numerical efficiency. For the longer term, energy and power consumption needs to be mapped into a rigorous performance model. Here, we restrict ourselves to illustrate numerically the parallel scaling and energy consumption properties of our implementation that uses Message Passing Interface(MPI) libraries and multi threading parallelism.

We perform a strong scaling benchmark for the test problem of Sect. 5.3 with  $k = 1, r = 3$  and  $ref_3 = 3$ , with 20996620 degrees of freedom in each subinterval  $I_n$  on the fine level  $\mathcal{T}_L$  and 45788 on  $\mathcal{T}_1$ . Throughout, we assign 36 MPI processes to each of the nodes used for the computations and vary the number  $n$  of nodes from  $n = 40$  to  $n = 200$ . For the evaluation of the parallel scaling properties, we compute the

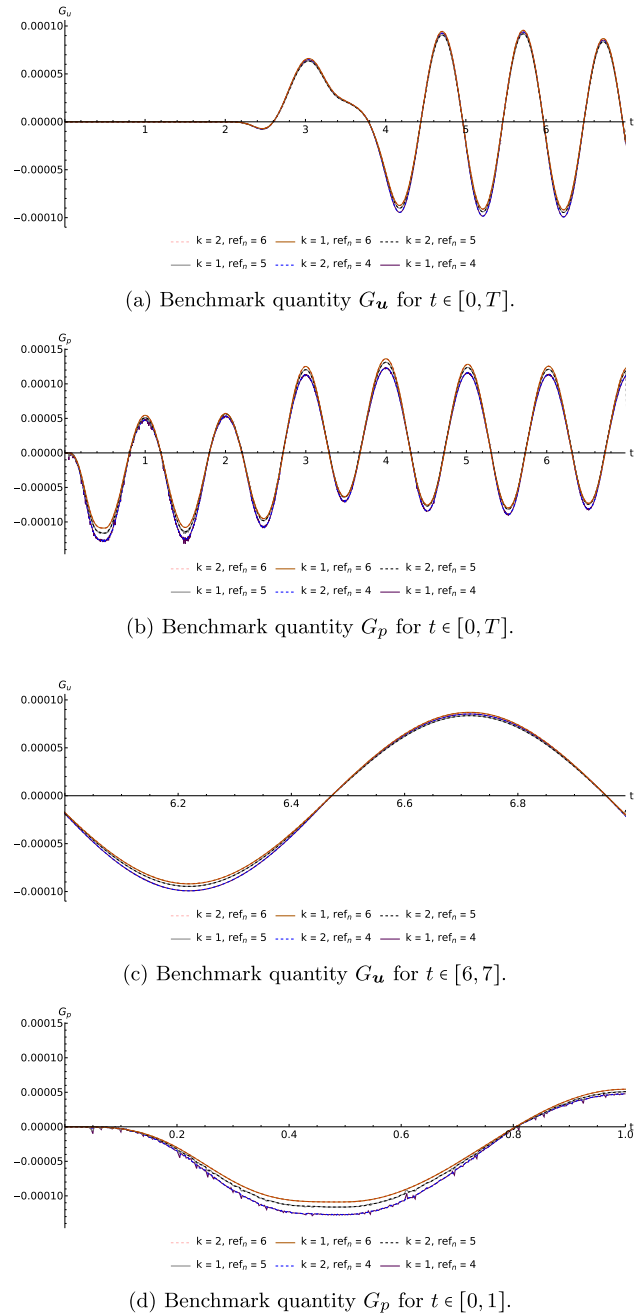


Fig. 4 Goal quantities  $G_u$  and  $G_p$  of (5.7) for different discretizations (spatial mesh resolution and polynomial degrees)

parallel speedup of the code (cf. [2]) that is approximated by

$$S = \frac{t_{\text{wall}}(n = n_{\text{min}})}{t_{\text{wall}}(n)}, \tag{5.8}$$

where  $t_{\text{wall}}(n)$  denotes the wall time of the simulation of fixed size on  $n$  compute nodes and  $t_{\text{wall}}(n = n_{\text{min}})$  is the wall time of the simulation on the minimum number of nodes involved in the scaling experiment. Secondly, we compute similarly

the energy ratio by means of

$$R = \frac{E(n)}{E(n = n_{\min})}, \tag{5.9}$$

where  $E(n)$  measures the total energy consumption of the simulation on  $n$  nodes. The energy consumption is determined by the Linux cluster workload manager slurm [68]. The energy consumption data is collected from hardware sensors using Intel’s Running Average Power Limit (RAPL) mechanism. It measures the energy consumption of the processor and memory. On our system, the sampling interval of energy consumption is determined by the value of 30 s. Figure 5 illustrates the results of the performance test. For  $n = 160$  and  $n = 200$ , the parallel scaling properties deteriorate. The reason for this is that due to the fixed problem size of the scaling test with 20996620 degrees of freedom the local problem size on each of the nodes is reduced for an increasing number of nodes such that the processor load decreases and communication increases. However, the (global) problem size is limited by the minimum number  $n = 40$  of nodes involved in the experiment and the memory (RAM) available on each of these nodes.

To quantify and evaluate the productivity or resource costs of the algorithm and its implementation, we use a simple model for the Productivity  $P = \frac{\text{Output } O}{\text{Input } I}$ ; cf. [24]. In an economic sense, all outputs should be the desired ones. Therefore, we use the reciprocal of the wall time  $t_{\text{wall}}$  as the output  $O = \frac{1}{t_{\text{wall}}}$ , such that a decrease in  $t_{\text{wall}}$  represents an increase of the (abstract) output. As the input  $I$  we use the total energy consumption  $E$  of the simulation. We scale the result by multiplying  $P$  with the constant factor  $E(n_{\min}) \cdot t_{\text{wall}}(n_{\min})$  such that the computation with  $n = n_{\min}$  has a productivity of  $P = 1.0$ :

$$P = \frac{1}{E(n)} \cdot E(n_{\min}) \cdot t_{\text{wall}}(n_{\min}) = \frac{S}{R}, \tag{5.10}$$

with  $S$  and  $R$  being defined in (5.8) and (5.9), respectively. The resulting productivity curve of our computations is presented in Fig. 6. In our simulations, the one on 120 compute nodes is the most productive one, that is, the ratio of output (low wall-time) to input (energy) is best. The quadratic interpolation predicts an even slightly increased productivity for 107 nodes ( $P = 1.09$ ). The characteristic quantities of our performance study are also summarized in Table 8.

## 6 Summary and outlook

In this work we presented and analyzed families of space-time finite element discretizations of the coupled hyperbolic–parabolic system (1.1) modeling, for instance, poroelasticity. The time discretization uses the discontinuous Galerkin method. The space discretization is based on inf-sup stable pairs of finite element spaces with continuous and discontinuous approximation of the scalar variable  $p$ . Well-posedness of the discrete problems is proved. For efficiently solving the arising algebraic systems with complex block structure in the case of increasing polynomial degrees of the time discretization a geometric multigrid preconditioner with a local Vanka smoother on patches of finite elements is proposed and studied. The overall approach is evaluated numerically. A convergence proof for our GMG methods to dynamic poroelasticity remains as a work for the future. Parallel scaling and energy consumption is also investigated. Multi-field formulations [12] of (1.1) with an explicit approximation of the stress tensor  $\sigma = C \epsilon$  and the flux vector  $q = -K \nabla p$  might be advantageous for applications of (1.1) in that their prediction are of interest; cf. [12, 29]. The design of tailored iterative solvers for suitable space-finite element approximations of such systems becomes even more challenging due to the increasing complexity of the system’s block structure. The feasibility of Vanka-type smoothers needs to be reconsidered.

**Table 7** Computed goal quantities (5.7) for different temporal and spatial approximations

$ref_n$	$k$	$DoF_{I_n}$	$G_{u_{\min}}$	$G_{u_{\max}}$	$G_{p_{\min}}$	$G_{p_{\max}}$	$\bar{n}_{\text{GMRES}}$
4	1	2741 516	− 9.93e−5	9.33e−5	−1.32e−4	1.23e−4	8
4	2	4,112,274	− 9.93e−5	9.37e−5	− 1.28e−4	1.23e−4	10
5	1	21,611,020	− 9.47e−5	9.18e−5	− 1.17e−4	1.31e−4	9
5	2	32,416,530	− 9.46e−5	9.18e−5	− 1.16e−4	1.31e−4	10
6	1	171,631,628	− 9.19e−5	9.55e−5	− 1.99e−4	1.36e−4	10
6	2	257,447,442	− 9.19e−5	9.54e−5	− 1.99e−4	1.36e−4	10

Here,  $ref_n$  is the number of spatial refinements of the initial grid and  $k$  is the polynomial degree of the time discretization (cf. Sect. 2).  $DoF_{I_n}$  is the total number of degrees of freedom on the fine grid level  $L$  of GMG per subinterval  $I_n$  (i.e. for all  $k + 1$  Gauss–Radau quadrature points on  $I_n$ ). The coarse grid level of GMG is  $ref_n = 1$  (with  $(k + 1) \cdot 3350$  degrees of freedom).  $G_{u_{\min}}$ ,  $G_{u_{\max}}$ ,  $G_{p_{\min}}$  and  $G_{p_{\max}}$  are the minimum and maximum of the benchmark quantities  $G_u$  and  $G_p$  on  $I$ ,  $\bar{n}_{\text{GMRES}}$  is the average number of GMRES iterations per time step

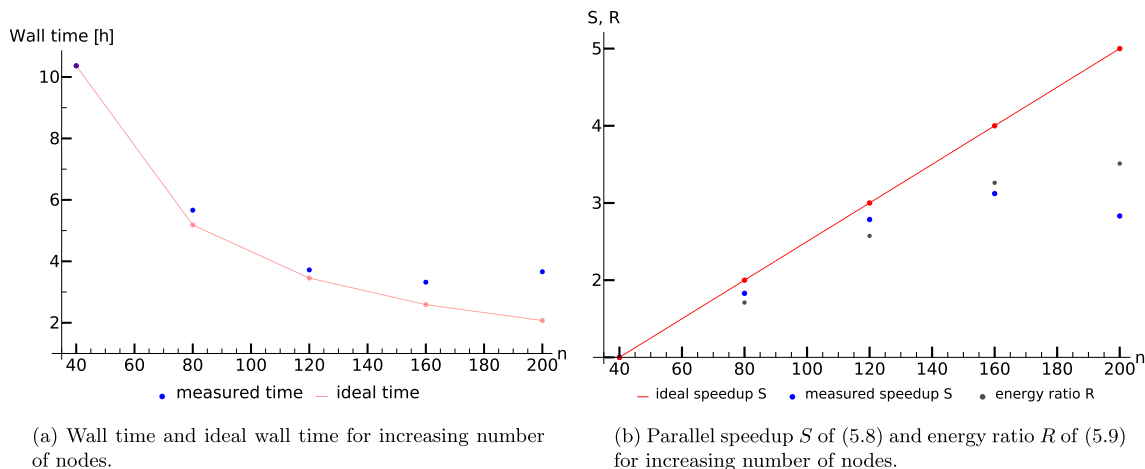


Fig. 5 Results of the strong scaling and energy consumption benchmark

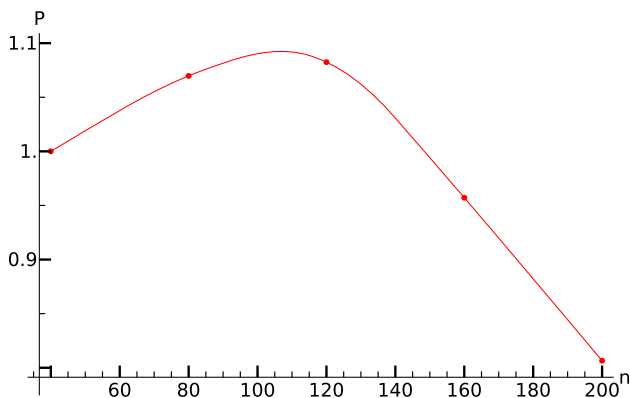


Fig. 6 Piecewise quadratic interpolation of the productivity function (5.10)

Table 8 Computed quantities of the scaling and performance benchmark

n	t <sub>wall</sub> (h)	E (kWh)	S	R	P
40	10.36	123.11	1.00	1.00	1.00
80	5.66	210.55	1.83	1.71	1.07
120	3.72	316.85	2.79	2.57	1.08
160	3.20	401.49	3.12	3.26	0.96
200	3.66	432.20	2.83	3.51	0.81

ered. Such type of approaches remain as a work for the future.

**Acknowledgements** Computational resources (HPC-cluster HSUper) were provided by the project hpc.bw, funded by dtcc.bw—Digitalization and Technology Research Center of the Bundeswehr. The authors thank M. Kronbichler (Ruhr University Bochum) for his hints for the implementation of the pipe socket geometry in deal.II.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix

### A Discrete coercivity

Here, we prove the discrete coercivity of the bilinear form  $A_\gamma$  defined in (3.2a) along with (3.3a). For this we introduce the mesh- and parameter-dependent norm

$$\|v_h\|_{h, \tilde{\gamma}_a} := \left( \|\varepsilon(v_h)\|^2 + \tilde{\gamma}_a h^{-1} \|v_h\|_{\Gamma_D}^2 \right)^{1/2}, \quad \text{for } v_h \in V_h, \tag{A.1}$$

with some parameter  $\tilde{\gamma}_a > 0$ . Its choice is addressed below. The norm property of (A.1) is ensured by a variant of Korn’s inequality; cf. [20, Eq. (1.19)]. We recall the well-known inverse inequality (cf. [75, p. 28])

$$\|\varepsilon(v_h)\mathbf{n}\|_{\Gamma_D} \leq ch^{-1/2} \|\varepsilon(v_h)\|, \quad \text{for } v_h \in V_h. \tag{A.2}$$

From (3.2a) along with (3.3a), it follows by the positive definiteness (1.2a) of  $C$ , the inequality of Cauchy–Young with

a sufficiently small constant  $\delta > 0$  and the trace inequality (A.2) that, for all  $\mathbf{v}_h \in \mathbf{V}_h$ ,

$$\begin{aligned} A_\gamma(\mathbf{v}_h, \mathbf{v}_h) &= \langle \mathbf{C}\boldsymbol{\varepsilon}(\mathbf{v}_h), \boldsymbol{\varepsilon}(\mathbf{v}_h) \rangle - \langle \mathbf{C}\boldsymbol{\varepsilon}(\mathbf{v}_h)\mathbf{n}, \mathbf{v}_h \rangle_{\Gamma_u^D} \\ &\quad - \langle \mathbf{v}_h, \mathbf{C}\boldsymbol{\varepsilon}(\mathbf{v}_h)\mathbf{n} \rangle_{\Gamma_u^D} + \gamma_a h^{-1} \langle \mathbf{v}_h, \mathbf{v}_h \rangle_{\Gamma_u^D} \\ &\geq c \|\boldsymbol{\varepsilon}(\mathbf{v}_h)\|^2 - c \|\boldsymbol{\varepsilon}(\mathbf{v}_h)\mathbf{n}\|_{\Gamma_u^D} \|\mathbf{v}_h\|_{\Gamma_u^D} \\ &\quad + \gamma_a h^{-1} \|\mathbf{v}_h\|_{\Gamma_u^D}^2 \\ &\geq c \|\boldsymbol{\varepsilon}(\mathbf{v}_h)\|^2 - c\delta h \|\boldsymbol{\varepsilon}(\mathbf{v}_h)\mathbf{n}\|_{\Gamma_u^D}^2 \\ &\quad - c\delta^{-1} h^{-1} \|\mathbf{v}_h\|_{\Gamma_u^D}^2 + \gamma_a h^{-1} \|\mathbf{v}_h\|_{\Gamma_u^D}^2 \\ &\geq c \|\boldsymbol{\varepsilon}(\mathbf{v}_h)\|^2 - c\delta \|\boldsymbol{\varepsilon}(\mathbf{v}_h)\|^2 \\ &\quad - c\delta^{-1} h^{-1} \|\mathbf{v}_h\|_{\Gamma_u^D}^2 + \gamma_a h^{-1} \|\mathbf{v}_h\|_{\Gamma_u^D}^2 \\ &= c \left( \|\boldsymbol{\varepsilon}(\mathbf{v}_h)\|^2 + \tilde{\gamma}_a h^{-1} \|\mathbf{v}_h\|_{\Gamma_u^D}^2 \right), \end{aligned}$$

with some constant  $\tilde{\gamma}_a > 0$  for a sufficiently large choice of the algorithmic parameter  $\gamma_a$  in (3.3a), such that  $\tilde{\gamma}_a := \gamma_a - c\delta^{-1} > 0$ . Thus, there holds for some constant  $c > 0$  that

$$A_\gamma(\mathbf{v}_h, \mathbf{v}_h) \geq c \|\mathbf{v}_h\|_{h, \tilde{\gamma}_a}, \quad \text{for all } \mathbf{v}_h \in \mathbf{V}_h. \quad (\text{A.3})$$

Secondly, we prove the discrete coercivity of the bilinear form  $B_\gamma$  defined in (3.2c) along with (3.3b). For brevity, we study the case  $Q_h = Q_h^{l, \text{cont}}$  only, such that  $Q_h \subset H^1(\Omega)$  is satisfied. The case  $Q_h = Q_h^{l, \text{disc}}$  can be captured similarly. We introduce the mesh- and parameter-dependent norm

$$\begin{aligned} \|q_h\|_{h, \tilde{\gamma}_b} &:= \left( \|\nabla q_h\|^2 + \tilde{\gamma}_b h^{-1} \|q_h\|_{\Gamma_p^D}^2 \right)^{1/2}, \\ \text{for } q_h \in Q_h &\subset H^1(\Omega), \end{aligned}$$

with some parameter  $\tilde{\gamma}_b > 0$ , where  $Q_h = Q_h^{l, \text{cont}}$ ; cf. (2.5). Further, we recall the well-known inverse inequality

$$\|\nabla q_h \cdot \mathbf{n}\|_{\Gamma_p^D} \leq ch^{-1/2} \|\nabla q_h\|, \quad \text{for } q_h \in Q_h. \quad (\text{A.4})$$

From (3.2c) along with (3.3b), it follows by the positive definiteness (1.2b) of  $\mathbf{K}$ , the inequality of Cauchy–Young with a sufficiently small constant  $\delta > 0$  and the trace inequality (A.4) that, for all  $q_h \in Q_h$ ,

$$\begin{aligned} B_\gamma(q_h, q_h) &= \langle \mathbf{K}\nabla q_h, \nabla q_h \rangle - \langle \mathbf{K}\nabla q_h \cdot \mathbf{n}, q_h \rangle_{\Gamma_p^D} \\ &\quad - \langle q_h, \mathbf{K}\nabla q_h \cdot \mathbf{n} \rangle_{\Gamma_p^D} + \gamma_b h^{-1} \langle q_h, q_h \rangle_{\Gamma_p^D} \\ &\geq c \|\nabla q_h\|^2 - c \|\nabla q_h \cdot \mathbf{n}\|_{\Gamma_p^D} \|q_h\|_{\Gamma_p^D} + \gamma_b h^{-1} \|q_h\|_{\Gamma_p^D}^2 \\ &\geq c \|\nabla q_h\|^2 - c\delta h \|\nabla q_h \cdot \mathbf{n}\|_{\Gamma_p^D}^2 - c\delta^{-1} h^{-1} \|q_h\|_{\Gamma_p^D}^2 \\ &\quad + \gamma_b h^{-1} \|q_h\|_{\Gamma_p^D}^2 \\ &\geq c \|\nabla q_h\|^2 - c\delta \|\nabla q_h\|^2 - c\delta^{-1} h^{-1} \|\mathbf{v}_h\|_{\Gamma_p^D}^2 \end{aligned}$$

$$\begin{aligned} &+ \gamma_b h^{-1} \|q_h\|_{\Gamma_p^D}^2 \\ &= c \left( \|\nabla q_h\|^2 + \tilde{\gamma}_b h^{-1} \|q_h\|_{\Gamma_p^D}^2 \right), \end{aligned}$$

with some constant  $\tilde{\gamma}_b > 0$  for a sufficiently large choice of the algorithmic parameter  $\gamma_b$  in (3.3b), such that  $\tilde{\gamma}_b := \gamma_b - c\delta^{-1} > 0$ . Thus, there holds for some constant  $c > 0$  that

$$B_\gamma(q_h, q_h) \geq c \|q_h\|_{h, \tilde{\gamma}_b}, \quad \text{for all } q_h \in Q_h. \quad (\text{A.5})$$

## B Alternative formulation of the fully discrete problem

Here, we present an alternative formulation for the fully discrete system to (1.1). The difference to Problem 3.1 comes through using  $\partial_t \mathbf{u}_{\tau, h}$  instead of  $\mathbf{v}_{\tau, h}$  in (3.5c) by means of (3.1a) and (3.5a), respectively. However, this modification requires the inclusion of an additional boundary integral in the resulting equation (B.1c). This non-obvious adaptation results from the proof of well-posedness of the discrete problem. The following problem is thus considered.

**Problem B.1** (Numerically integrated  $I_n$ -problem with  $\partial_t \mathbf{u}_{\tau, h}$ )  
For given  $\mathbf{u}_h^{n-1} := \mathbf{u}_{\tau, h}(t_{n-1}) \in \mathbf{V}_h$ ,  $\mathbf{v}_h^{n-1} := \mathbf{v}_{\tau, h}(t_{n-1}) \in \mathbf{V}_h$ , and  $p_h^{n-1} := p_{\tau, h}(t_{n-1}) \in Q_h$  with  $\mathbf{u}_{\tau, h}(t_0) := \mathbf{u}_{0, h}$ ,  $\mathbf{v}_{\tau, h}(t_0) := \mathbf{u}_{1, h}$  and  $p_{\tau, h}(t_0) := p_{0, h}$ , find  $(\mathbf{u}_{\tau, h}, \mathbf{v}_{\tau, h}, p_{\tau, h}) \in \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; Q_h)$  such that

$$\begin{aligned} Q_n \left( \langle \partial_t \mathbf{u}_{\tau, h}, \boldsymbol{\phi}_{\tau, h} \rangle - \langle \mathbf{v}_{\tau, h}, \boldsymbol{\phi}_{\tau, h} \rangle \right) \\ + \langle \mathbf{u}_{\tau, h}^+(t_{n-1}), \boldsymbol{\phi}_{\tau, h}^+(t_{n-1}) \rangle = \langle \mathbf{u}_h^{n-1}, \boldsymbol{\phi}_{\tau, h}^+(t_{n-1}) \rangle \end{aligned} \quad (\text{B.1a})$$

$$\begin{aligned} Q_n \left( \langle \rho \partial_t \mathbf{v}_{\tau, h}, \boldsymbol{\chi}_{\tau, h} \rangle + A_\gamma(\mathbf{u}_{\tau, h}, \boldsymbol{\chi}_{\tau, h}) \right. \\ \left. + C(\boldsymbol{\chi}_{\tau, h}, p_{\tau, h}) \right) + \langle \rho \mathbf{v}_{\tau, h}^+(t_{n-1}), \boldsymbol{\chi}_{\tau, h}^+(t_{n-1}) \rangle \\ = Q_n \left( F_\gamma(\boldsymbol{\chi}_{\tau, h}) \right) + \langle \rho \mathbf{v}_h^{n-1}, \boldsymbol{\chi}_{\tau, h}^+(t_{n-1}) \rangle, \end{aligned} \quad (\text{B.1b})$$

$$\begin{aligned} Q_n \left( \langle c_0 \partial_t p_{\tau, h}, \psi_{\tau, h} \rangle - C(\partial_t \mathbf{u}_{\tau, h}, \psi_{\tau, h}) + B_\gamma(p_{\tau, h}, \psi_{\tau, h}) \right) \\ + \langle c_0 p_{\tau, h}^+(t_{n-1}) + \alpha \nabla \cdot \mathbf{u}_{\tau, h}^+(t_{n-1}), \psi_{\tau, h}^+(t_{n-1}) \rangle \\ - \alpha \langle \mathbf{u}_{\tau, h}^+(t_{n-1}) \cdot \mathbf{n}, \psi_{\tau, h}^+(t_{n-1}) \rangle_{\Gamma_u^D} \\ = Q_n \left( G_\gamma(\psi_{\tau, h}) \right) + \langle c_0 p_h^{n-1} + \alpha \nabla \cdot \mathbf{u}_h^{n-1}, \psi_{\tau, h}^+(t_{n-1}) \rangle \\ - \alpha \langle \mathbf{u}_D(t_{n-1}) \cdot \mathbf{n}, \psi_{\tau, h}^+(t_{n-1}) \rangle_{\Gamma_u^D} \end{aligned} \quad (\text{B.1c})$$

for all  $(\boldsymbol{\phi}_{\tau, h}, \boldsymbol{\chi}_{\tau, h}, \psi_{\tau, h}) \in \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; \mathbf{V}_h) \times \mathbb{P}_k(I_n; Q_h)$ .

**Lemma B.2** (Existence and uniqueness of solutions to Problem B.1) Problem B.1 admits a unique solution.

**Proof** The proof follows basically the lines of the proof of Lem. 3.2 and is kept short. Only differences to Lem. 3.2 are depicted. For the differences  $(\mathbf{u}_{\tau,h}, \mathbf{v}_{\tau,h}, p_{\tau,h})$  of two triples satisfying (B.1) and the test functions  $\boldsymbol{\phi}_{\tau,h} = A_\gamma \mathbf{u}_{\tau,h}$ ,  $\boldsymbol{\chi}_{\tau,h} = \mathbf{v}_{\tau,h}$  and  $\psi_{\tau,h} = p_\tau$  there holds that

$$\begin{aligned} & \frac{1}{2} \int_{t_{n-1}}^{t_n} \frac{d}{dt} (A_\gamma(\mathbf{u}_{\tau,h}, \mathbf{u}_{\tau,h}) \\ & + \langle \rho \mathbf{v}_{\tau,h}, \mathbf{v}_{\tau,h} \rangle + \langle c_0 p_{\tau,h}, p_{\tau,h} \rangle) dt + Q_n(B_\gamma(p_{\tau,h}, p_{\tau,h})) \\ & + Q_n(C_\gamma(\mathbf{v}_{\tau,h} - \partial_t \mathbf{u}_{\tau,h}, p_{\tau,h})) \\ & + \alpha \langle \nabla \cdot \mathbf{u}_{\tau,h}^+(t_{n-1}), p_{\tau,h}^+(t_{n-1}) \rangle \\ & - \alpha \langle \mathbf{u}_{\tau,h}^+(t_{n-1}) \cdot \mathbf{n}, p_{\tau,h}^+(t_{n-1}) \rangle_{\Gamma_\mu^D} \\ & + A_\gamma(\mathbf{u}_{\tau,h}^+(t_{n-1}), \mathbf{u}_{\tau,h}^+(t_{n-1})) + \langle \rho \mathbf{v}_{\tau,h}^+(t_{n-1}), \boldsymbol{\chi}_{\tau,h}^+(t_{n-1}) \rangle \\ & + \langle c_0 p_{\tau,h}^+(t_{n-1}), p_{\tau,h}^+(t_{n-1}) \rangle = 0. \end{aligned} \tag{B.2}$$

Now, let  $l \in \{1, \dots, k+1\}$  be arbitrary but fixed and  $\boldsymbol{\phi}_{\tau,h} \in \mathbb{P}_k(I_n; \mathbf{V}_h)$  be chosen as

$$\begin{aligned} \boldsymbol{\phi}_{\tau,h}(t) & := \xi_{n,l}(t) \boldsymbol{\phi}_h \quad \text{with} \\ \xi_{n,l}(t) & := \left( \prod_{\substack{i=1 \\ i \neq l}}^{k+1} (t - t_{n,i}^{\text{GR}}) \right) \left( \prod_{\substack{i=1 \\ i \neq l}}^{k+1} (t_{n,l}^{\text{GR}} - t_{n,i}^{\text{GR}}) \right)^{-1} \\ & \in \mathbb{P}_k(I_n; \mathbb{R}), \quad \boldsymbol{\phi}_h \in \mathbf{V}_h, \end{aligned}$$

and the Gauss–Radau quadrature nodes  $t_{n,\mu}^{\text{GR}}$ , for  $\mu = 1, \dots, k+1$ ; cf. (2.4). By the exactness of the Gauss–Radau quadrature formula (2.4) for all polynomials in  $\mathbb{P}_{2k}(I_n; \mathbb{R})$  we deduce from (B.1a) that

$$\begin{aligned} 0 & = \frac{\tau_n}{2} \sum_{\mu=1}^{k+1} \hat{\omega}_\mu^{\text{GR}} (\langle \partial_t \mathbf{u}_{\tau,h}(t_{n,\mu}^{\text{GR}}), \boldsymbol{\phi}_{\tau,h}(t_{n,\mu}^{\text{GR}}) \rangle \\ & - \langle \mathbf{v}_{\tau,h}(t_{n,\mu}^{\text{GR}}), \boldsymbol{\phi}_{\tau,h}(t_{n,\mu}^{\text{GR}}) \rangle) \\ & + \langle \mathbf{u}_{\tau,h}^+(t_{n-1}), \boldsymbol{\phi}_{\tau,h}^+(t_{n-1}) \rangle \\ & = \frac{\tau_n}{2} \hat{\omega}_l^{\text{GR}} (\langle \partial_t \mathbf{u}_{\tau,h}(t_{n,l}^{\text{GR}}) - \mathbf{v}_{\tau,h}(t_{n,l}^{\text{GR}}), \boldsymbol{\phi}_h \rangle \\ & + \langle \mathbf{u}_{\tau,h}^+(t_{n-1}), \xi_{n,l}^+(t_{n-1}) \boldsymbol{\phi}_h \rangle). \end{aligned}$$

Thus, we have that ( $l = 1, \dots, k+1$ )

$$\begin{aligned} \mathbf{v}_{\tau,h}(t_{n,l}^{\text{GR}}) - \partial_t \mathbf{u}_{\tau,h}(t_{n,l}^{\text{GR}}) & = c_{n,l} \mathbf{u}_{\tau,h}^+(t_{n-1}) \\ \text{with } c_{n,l} & = 2\tau_n^{-1} (\hat{\omega}_l^{\text{GR}})^{-1} \xi_{n,l}^+(t_{n-1}). \end{aligned} \tag{B.3}$$

Substituting (B.3) into the third term on the left-hand side of (B.2), we get that

$$\begin{aligned} & Q_n(C_\gamma(\mathbf{v}_{\tau,h} - \partial_t \mathbf{u}_{\tau,h}, p_{\tau,h})) \\ & = \sum_{\mu=1}^{k+1} C_\gamma(\mathbf{u}_{\tau,h}^+(t_{n-1}), \xi_{n,\mu}^+(t_{n-1}) p_{\tau,h}(t_{n,\mu}^{\text{GR}})) \\ & = C_\gamma(\mathbf{u}_{\tau,h}^+(t_{n-1}), p_{\tau,h}^+(t_{n-1})) \\ & = -\alpha \langle \nabla \cdot \mathbf{u}_{\tau,h}^+(t_{n-1}), p_{\tau,h}^+(t_{n-1}) \rangle \\ & + \alpha \langle \mathbf{u}_{\tau,h}^+(t_{n-1}) \cdot \mathbf{n}, p_{\tau,h}^+(t_{n-1}) \rangle_{\Gamma_\mu^D}. \end{aligned} \tag{B.4}$$

Combining (B.2) with (B.4) then implies that

$$\begin{aligned} & A_\gamma(\mathbf{u}_{\tau,h}(t_n), \mathbf{u}_{\tau,h}(t_n)) + \langle \rho \mathbf{v}_{\tau,h}(t_n), \mathbf{v}_{\tau,h}(t_n) \rangle \\ & + \langle c_0 p_{\tau,h}(t_n), p_{\tau,h}(t_n) \rangle + 2Q_n(B_\gamma(p_{\tau,h}, p_{\tau,h})) \\ & + A_\gamma(\mathbf{u}_{\tau,h}^+(t_{n-1}), \mathbf{u}_{\tau,h}^+(t_{n-1})) + \langle \rho \mathbf{v}_{\tau,h}^+(t_{n-1}), \boldsymbol{\chi}_{\tau,h}^+(t_{n-1}) \rangle \\ & + \langle c_0 p_{\tau,h}^+(t_{n-1}), p_{\tau,h}^+(t_{n-1}) \rangle = 0. \end{aligned} \tag{B.5}$$

From (B.5) along with the discrete coercivity properties (A.3) and (A.5) we deduce that

$$\begin{aligned} \mathbf{u}_{\tau,h}(t_n) & = \mathbf{u}_{\tau,h}^+(t_{n-1}) = \mathbf{0}, \\ \mathbf{v}_{\tau,h}(t_n) & = \mathbf{v}_{\tau,h}^+(t_{n-1}) = \mathbf{0}, \\ p_{\tau,h}(t_n) & = p_{\tau,h}^+(t_{n-1}) = 0 \end{aligned}$$

as well as

$$p_{\tau,h}(t_{n,\mu}^{\text{GR}}) = 0, \quad \text{for } \mu = 1, \dots, k+1.$$

The rest then follows as in the proof of Lem. 3.2.

### C Additional numerical experiments

Here, we present some additional results of our numerical experiments. Table 9 shows the results of the convergence study introduced in Subsec. 5.1 for the solution (5.1) of (1.1). For the computations of Table 9 the Taylor-Hood pair of finite element spaces  $\mathbb{Q}_r^2/\mathbb{Q}_{r-1}$ , with  $r = 4$ , is used instead of the pair  $\mathbb{Q}_r^2/\mathbb{P}_{r-1}^{\text{disc}}$  chosen for the results of Table 2. For the pressure approximation marginally smaller errors are observed. In Table 10 the computed errors of the approximation of (5.1) for larger values of the Lamé parameters  $\lambda$  and  $\mu$  in the elasticity tensor  $\mathbf{C}$  are summarized. No significant increase of the errors is observed, indicating the independence of the error constant on the magnitude of  $\mathbf{C}$ .

**Table 9**  $L^2(L^2)$  and  $L^\infty(L^2)$  errors and experimental orders of convergence (EOC) with temporal polynomial degree  $k = 3$  and spatial degree  $r = 4$  for local spaces  $\mathbb{Q}_r^2/\mathbb{Q}_{r-1}$

$\tau$	$h$	$\ \nabla(\mathbf{u} - \mathbf{u}_{\tau,h})\ _{L^2(L^2)}$	EOC	$\ \mathbf{v} - \mathbf{v}_{\tau,h}\ _{L^2(L^2)}$	EOC	$\ p - p_{\tau,h}\ _{L^2(L^2)}$	EOC
$\tau_0/2^0$	$h_0/2^0$	2.3958497242e-03	–	1.7185669625e-02	–	7.6634974254e-04	–
$\tau_0/2^1$	$h_0/2^1$	1.0529091363e-04	4.51	5.4558642696e-04	4.98	3.8993048591e-05	4.30
$\tau_0/2^2$	$h_0/2^2$	5.9749805593e-06	4.14	1.9143193940e-05	4.83	1.6140682324e-06	4.59
$\tau_0/2^3$	$h_0/2^3$	3.6858181764e-07	4.02	9.6894400439e-07	4.30	9.0805047399e-08	4.15
$\tau_0/2^4$	$h_0/2^4$	2.2976509245e-08	4.00	5.6982904446e-08	4.09	5.5524640600e-09	4.03
$\tau$	$h$	$\ \nabla(\mathbf{u} - \mathbf{u}_{\tau,h})\ _{L^\infty(L^2)}$	EOC	$\ \mathbf{v} - \mathbf{v}_{\tau,h}\ _{L^\infty(L^2)}$	EOC	$\ p - p_{\tau,h}\ _{L^\infty(L^2)}$	EOC
$\tau_0/2^0$	$h_0/2^0$	1.3014883491e-02	–	1.1587851233e-01	–	4.9980604461e-03	–
$\tau_0/2^1$	$h_0/2^1$	8.7944378641e-04	3.89	4.1645287352e-03	4.80	2.6289118267e-04	4.25
$\tau_0/2^2$	$h_0/2^2$	5.6086076018e-05	3.97	2.0121086363e-04	4.37	9.4136464267e-06	4.80
$\tau_0/2^3$	$h_0/2^3$	3.5171476997e-06	4.00	1.1873567381e-05	4.08	5.0122396011e-07	4.23
$\tau_0/2^4$	$h_0/2^4$	2.1959938532e-07	4.00	7.3034502924e-07	4.02	2.8669553969e-08	4.13

**Table 10**  $L^2(L^2)$  and  $L^\infty(L^2)$  errors and experimental orders of convergence (EOC) with temporal polynomial degree  $k = 2$  and spatial degree  $r = 3$  for local spaces  $\mathbb{Q}_r^2/\mathbb{P}_{r-1}^{\text{disc}}$  for Young’s modulus

$E = 10000$  and Poisson’s ratio  $\nu = 0.35$ , corresponding to the Lamé parameters  $\lambda = 8642$  and  $\mu = 3704$

$\tau$	$h$	$\ \nabla(\mathbf{u} - \mathbf{u}_{\tau,h})\ _{L^2(L^2)}$	EOC	$\ \mathbf{v} - \mathbf{v}_{\tau,h}\ _{L^2(L^2)}$	EOC	$\ p - p_{\tau,h}\ _{L^2(L^2)}$	EOC
$\tau_0/2^0$	$h_0/2^0$	1.1835824122e-02	–	2.8040747896e-02	–	2.9610298753e-03	–
$\tau_0/2^1$	$h_0/2^1$	1.5714797742e-03	2.91	5.7609393975e-03	2.28	3.9787786632e-04	2.90
$\tau_0/2^2$	$h_0/2^2$	1.8935735637e-04	3.05	5.1606189537e-04	3.48	3.7075698634e-05	3.42
$\tau_0/2^3$	$h_0/2^3$	2.3548556912e-05	3.01	6.0883693658e-05	3.08	4.5049814701e-06	3.04
$\tau_0/2^4$	$h_0/2^4$	2.9374871960e-06	3.00	7.5088512252e-06	3.02	5.6225700500e-07	3.00
$\tau$	$h$	$\ \nabla(\mathbf{u} - \mathbf{u}_{\tau,h})\ _{L^\infty(L^2)}$	EOC	$\ \mathbf{v} - \mathbf{v}_{\tau,h}\ _{L^\infty(L^2)}$	EOC	$\ p - p_{\tau,h}\ _{L^\infty(L^2)}$	EOC
$\tau_0/2^0$	$h_0/2^0$	6.5064235342e-02	–	1.3580389516e-01	–	1.4838995481e-02	–
$\tau_0/2^1$	$h_0/2^1$	1.1842822460e-02	2.46	1.8023863764e-02	2.91	2.1492360068e-03	2.79
$\tau_0/2^2$	$h_0/2^2$	1.5191266076e-03	2.96	3.4010640997e-03	2.41	2.3825320454e-04	3.17
$\tau_0/2^3$	$h_0/2^3$	1.9340613477e-04	2.97	4.6154471881e-04	2.88	3.0597316621e-05	2.96
$\tau_0/2^4$	$h_0/2^4$	2.4393096482e-05	2.99	5.8122744877e-05	2.99	3.8749301950e-06	2.98

**References**

- Ahmed N, Matthies G (2016) Numerical study of SUPG and LPS methods combined with higher order variational time discretization applied to time-dependent linear convection-diffusion-reaction equations. *J Sci Comput* 67:988–1018
- Amdahl GM (1967) Validity of the single processor approach to achieving large scale computing capabilities. In: Proceedings of the April 18–20, 1967, Spring joint computer conference on—AFIPS '67 (Spring), ACM Press, Atlantic City, New Jersey, p 483
- Anselmann M, Bause M (2023) A geometric multigrid method for space-time finite element discretizations of the Navier-Stokes equations and its application to 3d flow simulation. *ACM Trans Math Softw* 49:5
- Anselmann M, Bause M (2022) Efficiency of local Vanka smoother geometric multigrid preconditioning for space-time finite element methods to the Navier-Stokes equations. *PAMM Proc Appl Math Mech* 22:e202200088
- Anselmann M, Bause M (2022) CutFEM and ghost stabilization techniques for higher order space-time discretizations of the Navier-Stokes equations. *Int J Numer Meth Fluids* 94:775–802
- Anselmann M, Bause M, Becher S, Matthies G (2020) Galerkin-collocation approximation in time for the wave equation and its post-processing. *ESAIM Math Model Numer Anal* 54:2099–2123
- Arndt D, Bangerth W, Feder M, Fehling M, Gassmüller R, Heister T, Heltai L, Kronbichler M, Maier M, Munch P, Pelteret J-P, Stiecko S, Turcksin B, Wells D (2023) The deal. II Library, Version 9.4. *J Numer Math* 30:231–246
- Arnold DN, Boffi D, Falk RS (2002) Approximation by quadrilateral finite elements. *Math Comput* 71:909–922
- Bangerth W, Geiger M, Rannacher R (2010) Adaptive Galerkin finite element methods for the wave equation. *Comput Meth Appl Math* 10:3–48
- Bangerth W, Rannacher R (2003) Adaptive finite element methods for differential equations. Birkhäuser, Basel

11. Bause M, Anselmann M, Köcher U, Radu FA (2021) Convergence of a continuous Galerkin method for hyperbolic-parabolic systems. *Comput. Math. with Appl.* 158:118–138
12. Bause M, Franz S (2023) Structure preserving discontinuous Galerkin approximation of a hyperbolic-parabolic system. *Electron Trans Numer Anal* 1–24. [arXiv:2311.01264](https://arxiv.org/abs/2311.01264) (in review)
13. Bause M, Radu R, Köcher U (2017) Space-time finite element approximation of the Biot poroelasticity system with iterative coupling. *Comput Methods Appl Mech Eng* 320:745–768
14. Benzi M, Golub GH, Liesen J (2005) Numerical solution of saddle point problems. *Acta Numer* 14:1–137
15. Becker R (2002) Mesh adaptation for Dirichlet flow control via Nitsche's method. *Commun Numer Meth Eng* 18:669–680
16. Biot M (1941) General theory of three-dimensional consolidation. *J Appl Phys* 12:155–164
17. Biot M (1955) Theory of elasticity and consolidation for a porous anisotropic solid. *J Appl Phys* 26:182–185
18. Biot M (1972) Theory of finite deformations of porous solids. *Indiana Univ Math J* 21:597–620
19. Both JW, Barnafi NA, Radu FA, Zunino P, Quarteroni A (2022) Iterative splitting schemes for a soft material poromechanics model. *Comput Methods Appl Mech Eng* 388:114183
20. Brenner SC (2003) Korn's inequalities for piecewise  $H^1$  vector fields. *Math Comput* 73:1067–1087
21. Bruchhäuser MP, Köcher U, Bause M (2022) On the implementation of an adaptive multirate framework for coupled transport and flow. *J Sci Comput* 93:59
22. Carlson DE (1972) Linear thermoelasticity. *Handbuch der Physik V Ia/2*. Springer, Berlin
23. Chapelle D, Moireau P (2014) General coupling of porous flows and hyperelastic formulations—from thermodynamics principles to energy balance and compatible time schemes. *Eur J Mech B Fluids* 46:82–96
24. Cooper WW, Seiford LMK (2000) *Tone data envelopment analysis*. Kluwer Academic Publishers, Dordrecht
25. Corti M, Antonietti PF, Luca Dedé, Quarteroni AM (2023) Numerical modeling of the brain poromechanics by high-order discontinuous Galerkin methods. *Math Models Methods Appl Sci* 33:1577–1609
26. Di Pietro DA, Ern A (2012) *Mathematical aspects of discontinuous Galerkin methods*. Springer, Berlin
27. Dörfler W, Findeisen S, Wieners C, Ziegler D (2019) Parallel adaptive discontinuous Galerkin discretizations in space and time for linear elastic and acoustic waves. In: Langer U, Steinbach O (eds) *Space-time methods. Applications to partial differential equations, radon series on computational and applied mathematics*, de Gruyter, Berlin, vol 25, pp 61–88
28. Drzisga D, John L, Råde U, Wohlmuth B, Zulehner W (2018) On the analysis of block smoothers for saddle point problems. *SIAM J Matrix Anal Appl* 39:932–960
29. Ernesti J, Wieners C (2019) A space-time discontinuous Petrov–Galerkin method for acoustic waves. In: Langer U, Steinbach O (eds) *Space-time methods. Applications to partial differential equations, radon series on computational and applied mathematics*, de Gruyter, Berlin, vol 25, pp 89–115
30. Fairweather G (1978) *Finite element Galerkin methods for differential equations*. Lecture notes in pure and applied mathematics. Marcel Dekker Inc., New York, vol 34
31. Franco SR, Francisco FG, Pinto MAV, Rodrigo C (2018) Multigrid method based on a space-time approach with standard coarsening for parabolic problems. *Appl Math Comput* 317:25–34
32. Gander MJ (2015) 50 years of time parallel integration. In: Carraro T et al (eds) *Multiple shooting and time domain decomposition*. Springer, Heidelberg, pp 69–114
33. Gander MJ, Neumüller M (2016) Analysis of a new space-time parallel multigrid algorithm for parabolic problems. *SIAM J Sci Comput* 38:A2173–A2208
34. Gmeiner B, Huber M, John L, Råde U, Wohlmuth B (2016) A quantitative performance study for Stokes solvers at the extreme scale. *J Comput Sci* 17:509–521
35. Gmeiner B, Råde U, Stengel H, Waluga C, Wohlmuth B (2015) Performance and scalability of hierarchical hybrid multigrid solvers for Stokes systems. *SIAM J Sci Comput* 37:C143–C168
36. Güdücü C, Liesen J, Mehrmann V, Szyld DB (2022) On non-Hermitian positive (semi)definite linear algebraic systems arising from dissipative Hamiltonian DAEs. *SIAM J Sci Comput* 44:A2871–A2894
37. Hong Q, Kraus J (2018) Parameter-robust stability of classical three-field formulation of Biot's consolidation model. *Electron Trans Numer Anal* 48:202–226
38. Hong Q, Kraus J, Xu J, Zikatanov L (2016) A robust multigrid method for discontinuous Galerkin discretizations of Stokes and linear elasticity equations. *Numer Math* 132:23–49
39. Horton G, Vandewalle S (1995) A space-time multigrid method for parabolic partial differential equations. *SIAM J Sci Comput* 16:848–864
40. Hussain S, Schieweck F, Turek S (2014) Efficient Newton-multigrid solution techniques for higher order space-time Galerkin discretizations of incompressible flow. *Appl Numer Math* 83:51–71
41. Hussain S, Schieweck F, Turek S (2013) An efficient and stable finite element solver of higher order in space and time for nonstationary incompressible flow. *Int J Numer Methods Fluids* 73:927–952
42. Hussain S, Schieweck F, Turek S (2011) Higher order Galerkin time discretizations and fast multigrid solvers for the heat equation. *J Numer Math* 19:41–61
43. Jiang S, Racke R (2018) *Evolution equations in thermoelasticity*. CRC Press, Boca Raton
44. John V (2016) *Finite element methods for incompressible flow problems*. Springer, Cham
45. John V (2002) Higher order finite element methods and multigrid solvers in a benchmark problem for the 3D Navier-Stokes equations. *Int J Numer Meth Fluids* 40:775–798
46. John V, Matthies G (2001) Higher-order finite element discretizations in a benchmark problem for incompressible flows. *Int J Numer Meth Fluids* 37:885–903
47. John V, Tobiska L (2000) Numerical performance of smoothers in coupled multigrid methods for the parallel solution of the incompressible Navier-Stokes equations. *Int J Numer Meth Fluids* 33:453–473
48. Kanschat G, Riviere B (2018) A finite element method with strong mass conservation for Biot's linear consolidation model. *J Sci Comput* 77:1762–1779
49. Karakashian O, Makridakis C (1999) A space-time finite element method for the nonlinear Schrödinger equation: the continuous Galerkin method. *SIAM J Numer Anal* 36:1779–1807
50. List F, Radu FA (2016) A study on iterative methods for solving Richards' equation. *Comput Geosci* 20:341–353
51. Langer U, Neumüller M, Schafelner A (2019) Space-time finite element methods for parabolic evolution problems with variable coefficients. In: Apel T et al (eds) *Advanced finite element methods with applications*. Springer, Cham, pp 247–275
52. Leis R (1986) *Initial boundary value problems in mathematical physics*. Teubner, Stuttgart. Wiley, Chichester
53. Li XS, Demmel JW (2003) SuperLU\_DIST: a scalable distributed-memory sparse direct solver for unsymmetric linear systems. *ACM Trans Math Softw* 29:110–140

54. Linden J, Lonsdale G, Ritzdorf H, Schüller A (1994) Scalability aspects of parallel multigrid. *Future Gener Comput Syst* 10:429–439
55. Manservigi S (2006) Numerical analysis of Vanka-type solvers for steady Stokes and Navier-Stokes flows. *SIAM J Numer Anal* 44:2025–2056
56. Matthies G (2001) Mapped finite elements on hexahedra. Necessary and sufficient conditions for optimal interpolation errors. *Numer Algorithms* 27:317–327
57. Matthies G, Tobiska L (2002) The inf-sup condition for the mapped  $Q_k^d/P_{k-1}^{\text{disc}}$  element in arbitrary space dimensions. *Computing* 69:119–139
58. Mikelic A, Wheeler MF (2012) Theory of the dynamic Biot-Allard equations and their link to the quasi-static Biot system. *J Math Phys* 53(123702):1–15
59. Neumüller M (2013) Space-time methods: fast solvers and applications, PhD Thesis, TU Graz
60. Nitsche J (1971) über ein Variationsprinzip zur Lösung von Dirichlet Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. In: *Abh Math Sem Univ Hamburg*, vol 36pp 9–15 (in German)
61. Rodrigo C, Hu X, Ohm P, Adler JH, Gaspar FJ, Zikatanov LT (2018) New stabilized discretizations for poroelasticity and the Stokes' equations. *Comput Methods Appl Mech Eng* 341:467–484
62. Rüdiger U (2017) Algorithmic efficiency and the energy wall. In: 2nd workshop on power-aware computing (PACO), ResearchGate. <https://doi.org/10.13140/RG.2.2.33914.18881>
63. Schafelner A (2021) Space-time finite element methods, PhD Thesis, Johannes Kepler University Linz
64. Seifert C, Trostorff S, Waurick M (2022) Evolutionary equations: Picard's theorem for partial differential equations, and applications. Birkhäuser, Cham
65. Saad Y (2003) Iterative methods for sparse linear systems, 2nd edn. SIAM, Philadelphia
66. Showalter R (2000) Diffusion in poro-elastic media. *J Math Anal Appl* 251:310–340
67. Slodička M (1989) Application of Rothe's method to integrodifferential equation. *Comment Math Univ Carolinae* 30:57–70
68. Slurm Workload Manager (2023) Version 22.05. [https://github.com/SchedMD/slurm/tree/master/src/plugins/acct\\_gather\\_energy/rapl](https://github.com/SchedMD/slurm/tree/master/src/plugins/acct_gather_energy/rapl)
69. Steinbach O (2015) Space-time finite element methods for parabolic problems. *Comput Methods Appl Math* 15:551–566
70. Steinbach O, Yang H (2019) Space-time finite element methods for parabolic evolution equations: Discretization, a posteriori error estimation, adaptivity and solution, In: Langer U, Steinbach O (eds) Space-time methods. Applications to partial differential equations, radon series on computational and applied mathematics, de Gruyter, Berlin, vol 25, pp 207–248
71. Steinbach O, Yang H (2018) An algebraic multigrid method for an adaptive space-time finite element discretization. In: *Lecture notes in computer science*. Springer, Cham, vol 10665, pp 63–73
72. Steinbach O, Yang H (2018) Comparison of algebraic multigrid methods for an adaptive space-time finite-element discretization of the heat equation in 3D and 4D. *Numer Linear Algebra Appl* 25:e2143
73. Steinbach O, Yang H (2018) An algebraic multigrid method for an adaptive space-time finite element discretization. In: Lirkov I, Margenov SD (eds) Large-scale scientific computing: 11th international conference, LSSC 2017, Springer, Sozopol, Bulgaria, pp 66–73
74. Steinbach O, Zank M (2022) A generalized inf-sup stable variational formulation for the wave equation. *J Math Anal Appl* 505:125457
75. Thomeé V (2006) Galerkin finite element methods for parabolic problems. Springer, Berlin
76. Trottenberg U, Oosterlee CW, Schüller A (2001) Multigrid. Academic Press, San Diego
77. Turek S (1999) Efficient solvers for incompressible flow problems. Springer, Berlin
78. Turek S, Becker C, Kilian D (2006) Hardware-oriented numerics and concepts for PDE software. *Future Gener Comput Syst* 22:217–238
79. Turek S, Göttsche D, Becker C, Buijssen S, Wobker H (2008) FEAST—realisation of hardware-oriented numerics for HPC simulations with finite elements, Concurrency and computation: practice and experience 6(May), 2247–2265. <https://doi.org/10.1002/cpe.1584>. (Special Issue Proceedings of ISC)
80. Vanka S (1986) Block-implicit multigrid solution of Navier-Stokes equations in primitive variables. *J Comput Phys* 65:138–158
81. Wirtz DC, Schiffers N, Pandorf T, Rademacher K, Weichert D, Forst R (2000) Critical evaluation of known bone material properties to realize anisotropic FE-simulation of the proximal femur. *J Biomech* 33:1325–1330
82. Zulehner W (2022) A short note on inf-sup conditions for the Taylor-Hood family  $Q_k-Q_{k-1}$ , pp 1–15. Preprint [arXiv:2205.14223](https://arxiv.org/abs/2205.14223)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.