

Ulrich Langer* and Andreas Schafelner

Adaptive space–time finite element methods for parabolic optimal control problems

<https://doi.org/10.1515/jnma-2021-0059>

Received May 10, 2021; revised September 17, 2021; accepted October 06, 2021

Abstract: We present, analyze, and test locally stabilized space–time finite element methods on fully unstructured simplicial space–time meshes for the numerical solution of space–time tracking parabolic optimal control problems with the standard L_2 -regularization. We derive a priori discretization error estimates in terms of the local mesh-sizes for shape-regular meshes. The adaptive version is driven by local residual error indicators, or, alternatively, by local error indicators derived from a new functional a posteriori error estimator. The latter provides a guaranteed upper bound of the error, but is more costly than the residual error indicators. We perform numerical tests for benchmark examples having different features. In particular, we consider a discontinuous target in form of a first expanding and then contracting ball in 3d that is fixed in the 4d space–time cylinder.

Keywords: parabolic optimal control problem, L_2 -regularization, space–time finite element methods, discretization error estimates, adaptive versions, parallel solvers

Classification: 49J20, 35K20, 65M60, 65M50, 65M15, 65Y05

Dedicated to Wolfgang L. Wendland on the occasion of his 85th birthday.

1 Introduction

Let us consider the following space–time tracking optimal control problem: For a given target function $y_d \in L_2(Q)$ (desired state) and for some appropriately chosen regularization (cost) parameter $\rho > 0$, find the state $y \in Y$ and the control $u \in U$ minimizing the cost functional

$$J(y, u) = \frac{1}{2} \int_Q |y - y_d|^2 \, dQ + \frac{\rho}{2} \|u\|_U^2 \quad (1.1)$$

subject to the linear parabolic initial-boundary value problem (IBVP):

$$\partial_t y - \operatorname{div}_x(v \nabla_x y) = u \quad \text{in } Q, \quad y = 0 \quad \text{on } \Sigma, \quad y = 0 \quad \text{on } \Sigma_0 \quad (1.2)$$

where $Q := \Omega \times (0, T)$, $\Sigma := \partial\Omega \times (0, T)$, $\Sigma_0 := \Omega \times \{0\}$, $T > 0$ is the final time, ∂_t denotes the partial time derivative, div_x is the spatial divergence operator, ∇_x is the spatial gradient, and the source term u on the right-hand side of the parabolic PDE (1.2) serves as control. The spatial domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, is supposed to be bounded and Lipschitz. The coefficient v is assumed to be uniformly positive and bounded, i.e., there exist positive constants v_1 and v_2 such that

$$0 < v_1 \leq v(x, t) \leq v_2 \quad \text{for almost all } (x, t) \in Q. \quad (1.3)$$

For simplicity, we here consider only scalar coefficients, but it is clear that the scalar coefficient v can be replaced by a symmetric, and uniformly positive definite and bounded $d \times d$ coefficient matrix.

*Corresponding author: Ulrich Langer, Institute for Computational Mathematics, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria. Email: ulanger@numa.uni-linz.ac.at

Andreas Schafelner, Doctoral Program ‘Computational Mathematics’, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria.

In the standard setting, that was already investigated in the famous book by J. L. Lions [38], the distributed control u is taken from $U = L_2(0, T; L_2(\Omega)) = L_2(Q)$, and thus the cost of the control is also measured in the $L_2(Q)$ -norm that mathematically serves as regularization term in (1.1). Since the state equation (1.2) has a unique solution $y \in Y := Y_0 = \{v \in L^2(0, T; H_0^1(\Omega)) : \partial_t v \in L^2(0, T; H^{-1}(\Omega)), v = 0 \text{ on } \Sigma_0\} = \{v \in W(0, T) : v = 0 \text{ on } \Sigma_0\}$, one can conclude the existence of a unique control $u \in U$ minimizing the cost functional $J(Su, u)$, where S is the solution operator mapping $u \in U$ to the unique solution $y \in Y$ of (1.2) (see, e.g., [38, 56]). There is a huge number of publications devoted to the numerical solution of optimal control problems (1.1)–(1.2) with the standard L_2 -regularization (see, e.g., [6, 23, 56]). The overwhelming majority of the publications uses some time-stepping method or discontinuous Galerkin method for the time discretization in combination with some space-discretization method like the finite element method (see, e.g., [40, 41]). The unique solvability of the optimal control problem can also be established by showing that the optimality system (KKT system) has a unique solution, since these problems are equivalent for quadratic cost functionals with linear constraints. In [37], the Banach–Nečas–Babuška (BNB) theorem was applied to the optimality system to show its well-posedness. We refer the reader to [7, Th. 3.6] and [16, Th. 2.6] for the version of the BNB theorem, which was used in [37], and to the original papers [3, 44]. Furthermore, the discrete inf-sup condition, that does not follow from the inf-sup condition in the infinite-dimensional setting, was established for continuous space–time finite element discretizations on fully unstructured simplicial space–time meshes [37, Lem. 3.4]. The discrete inf-sup condition implies stability of the discretization and a priori discretization error estimates (see also [3, 4]). In connection with continuous space–time finite element discretizations for parabolic optimal control problems, we would like to mention the publications [19, 43]. A comprehensive overview of space–time methods for parabolic IBVP can be found in [53].

Besides the L_2 -regularization, other regularization respectively cost terms can be chosen in order to obtain certain desired effects. We here only mention the sparsity techniques where the L_1 term $\mu \|u\|_{L_1(Q)}$ with the sparsity parameter μ is added to the L_2 -regularization term and directional sparsity techniques [11, 21, 49], control in measure spaces also leading to locally concentrated controls [10], and the energy regularization where $U = L^2(0, T; H^{-1}(\Omega))$ (see [36]). The energy regularization is motivated by applications in electrical engineering where controls u from $L^2(0, T; H^{-1}(\Omega))$ are admissible, i.e., controls that are concentrated on spatial hypersurfaces. Furthermore, the observation can be restricted to some subset of the space–time cylinder \bar{Q} including $\Sigma_T := \Omega \times \{0\}$ (observation at the terminal time), or the control can be restricted to some subset of \bar{Q} including the boundary Σ or some parts of Σ (control via Dirichlet, Neumann, and Robin boundary conditions).

In this paper, we consider locally stabilized space–time finite element methods on fully unstructured simplicial space–time meshes for the numerical solution of the space–time tracking parabolic optimal control problem (1.1)–(1.2) with the standard L_2 -regularization as model problem, although the space–time finite element technique presented in this paper can certainly be applied to other optimal control problems as well. In our former works [32, 33], we have successfully applied the locally stabilized space–time finite element methods to the state equation (1.2) with right-hand sides u from $L_2(Q)$ and with special distributional right-hand sides u from $L^2(0, T; H^{-1}(\Omega))$. In particular, we have proposed adaptive space–time finite element schemes based on local error indicators that are derived from the residual error indicator proposed in [52], and Repin’s functional error estimator providing a guaranteed upper bound for any admissible approximation [46]. In our note [34], we report on the first results for globally stabilized space–time finite element methods applied to the optimal control problem (1.1)–(1.2), but on quasi-uniform meshes characterized by a global mesh-size parameter h . There we used global time-upwind finite element test functions $v_h + \partial_t v_h$ and $q_h - \partial_t q_h$ for constructing consistent finite element schemes approximating the reduced optimality system. We mention that upwind test functions were introduced by Hughes and Brooks for constructing stable finite element schemes for stationary convection–diffusion problems in [24]. This stabilization technique is called SUPG, and was later used by Johnson and Saranen [27] for transient problems; see also [25] for the related Galerkin Least-Squares finite element methods, and [5, 31] for more recent papers on these stabilization techniques. In the case of unstructured, but shape-regular meshes naturally produced by adaptive schemes, one should replace the global discretization parameter h by the local mesh-size h_K that can be different for every element K from the triangulation \mathcal{T}_h . This simple replacement of h by h_K works well for the state equation

alone [32, 33], but not for the reduced optimality system. Therefore, in this paper, we introduce a differentiable mesh-density function $\lambda_h(x, t)$ in order to prove coercivity of the mesh-dependent bilinear form $a_h(\cdot, \cdot)$ that corresponds to the finite element scheme for the reduced optimality system. Coercivity together with Galerkin orthogonality and extended boundedness immediately leads to a best-approximation estimate from which one can derive convergence rate estimates by means of interpolation error estimates under additional regularity assumptions. Now adaptivity can be performed simultaneously in space and time on the basis of local error indicators. We use the residual error indicator that was introduced for the state equation in [52], and that was later used for optimal control problems as well [37]. The residual error indicator is computationally cheap, and works very well in all our numerical experiments. However, a rigorous proof of reliability and efficiency is still missing even for the state equation (cf. [52]). Using the inf-sup condition for the reduced optimality system proved in [37], we are able to derive rigorously a functional a posteriori error estimator that provides a guaranteed upper bound of the discretization error and local error indicators for adaptivity. Finally, we have to solve one system of finite element equations for defining all space–time unknowns all at once. The system matrix is non-symmetric, but positive definite. We solve this system by means of a parallel version of the flexible General Minimal Residual (FGMRES) method [47], preconditioned by a block-diagonal algebraic multigrid (AMG) preconditioner. The parallelization is relatively easy since it can be done simultaneously in space and time as known from elliptic problems. We emphasize that our stabilized scheme is coercive with respect to a mesh-dependent norm that is stronger than the norm used in the inf-sup condition in [37].

The remainder of the paper is organized as follows. Section 2 introduces the basic notations and states some preliminary results on the solvability and the space–time discretization of the parabolic initial-boundary value problem (1.2) that serves as state equation in the optimal control problem studied in this paper. Section 3 states the reduced optimality system characterizing the unique solution of the optimal control problem (1.1)–(1.2). The space–time finite element discretization of the reduced optimality system is derived in Section 4. In Section 5, we establish the coercivity of the bilinear form corresponding the finite element scheme. A priori estimates of the discretization error are derived in Section 6. Section 7 provides the derivation of a new functional a posteriori error estimator, and briefly describes the construction of simultaneous space–time adaptive finite element schemes. The algebraic system corresponding to the finite element scheme and the solver are presented in Section 8. Section 9 is devoted to the presentation and discussion of the numerical results. Finally, we draw some conclusion and give an outlook in Section 10.

2 Preliminaries

Throughout the paper, we use the standard notations for Lebesgue spaces $L_p(\cdot)$ and Sobolev spaces $W_p^k(\cdot)$ respectively (resp.) $H^k(\cdot) = W_2^k(\cdot)$ with the corresponding norms resp. scalar products (see, e.g., [1]). Furthermore, we also use Bochner spaces $L_2(0, T; X)$ of square-integrable abstract functions, mapping the time interval $(0, T)$ to the spatial Hilbert space $X = L_2(\Omega), H_0^1(\Omega), H^{-1}(\Omega), \dots$; see, e.g., [38] for the precise definition of the spaces, norms, and scalar products as well as for the main properties.

The standard weak or variational formulation of the state equation (1.2) reads as follows: Find $y \in Y = Y_0 := \{y \in V : \partial_t y \in V^*, v = 0 \text{ on } \Sigma_0\}$ such that

$$b(y, v) = \langle u, v \rangle_Q \quad \forall v \in V := L_2(0, T; H_0^1(\Omega)) \quad (2.1)$$

with the bilinear form $b(\cdot, \cdot) : Y \times V \rightarrow \mathbb{R}$, defined by the identity

$$b(y, v) := \int_Q \left[\partial_t y v + v \nabla_x y \cdot \nabla_x v \right] dQ, \quad \forall (y, v) \in Y \times V \quad (2.2)$$

and the duality product $\langle \cdot, \cdot \rangle_Q : V^* \times V \rightarrow \mathbb{R}$, and for given $u \in V^*$. We mention that the first integral in (2.2) has to be understood as duality pairing as well since $\partial_t y$, in general, belongs to V^* . We also may integrate by parts with respect to time. Then we can include the initial conditions as natural conditions into the variational formulation, and we look for a weak solution y in $H_0^{1,0}(Q) := \{v \in L_2(Q) : \nabla_x v \in (L_2(Q))^d, v = 0 \text{ on } \Sigma\}$

with test functions v from $H_0^{1,1}(Q) := \{v \in H_0^{1,0}(Q) : \partial_t v \in L_2(Q)\}$ vanishing on the top $\Sigma_T := \Omega \times \{T\}$ of the space–time cylinder Q (see, e.g., [28]). The standard methods to show existence and uniqueness are Galerkin’s method in space and a priori estimates. One can also use the BNB theorem as it was done in [50]. Indeed, the bilinear form $b(\cdot, \cdot)$ fulfills the following three conditions:

(BNB1) boundedness: $|b(y, v)| \leq c_b \|y\|_Y \|v\|_V$;

(BNB2) inf-sup condition: $\inf_{y \in Y \setminus \{0\}} \sup_{v \in V \setminus \{0\}} \frac{b(y, v)}{\|y\|_Y \|v\|_V} \geq c_{\text{inf sup}} > 0$;

(BNB3) injectivity of B^* : $\forall v \in V \setminus \{0\} \exists y \in Y : b(y, v) \neq 0$;

that are sufficient and necessary for $B : Y \rightarrow V^*$ being an isomorphism, where the operator B is defined by the bilinear form:

$$\langle By, v \rangle_Q = b(y, v) \quad \forall (y, v) \in Y \times V.$$

Therefore, the solution operator $S = B^{-1}$ is a well-defined bounded linear operator from V^* onto Y . Moreover, if $u \in L_2(Q)$, then the unique solution y belongs to the space

$$H^{L,1}(Q) = \{v \in H^1(Q) : L_x v := -\text{div}_x(v \nabla_x v) \in L_2(Q)\}$$

provided that the coefficient v fulfills some additional conditions (see [14]). This property is called maximal parabolic regularity. We mentioned that already Ladyzhenskaya proved maximal parabolic regularity for the case $v = 1$ in [28]. Until now many papers have been published on this topic (see, e.g., [18]).

3 Optimality system

Eliminating the control u from the optimality system by means of the so-called gradient equation $p + \varrho u = 0$, we get the reduced optimality system. The weak form of the reduced optimality system can be written in the form: Find the state $y \in Y_0$ and the adjoint state $p \in P_T$ such that

$$\varrho \int_Q [\partial_t y v + v \nabla_x y \cdot \nabla_x v] dQ + \int_Q p v dQ = 0 \quad (3.1)$$

$$- \int_Q y q dQ + \int_Q [-\partial_t p q + v \nabla_x p \cdot \nabla_x q] dQ = - \int_Q y_d q dQ \quad (3.2)$$

holds for $v, q \in V$, where $P_T := \{p \in W(0, T) : p = 0 \text{ on } \Sigma_T\}$. It is shown in [37, Th. 3.3] by means of the BNB theorem that the variational reduced optimality system (3.1)–(3.2) is well-posed.

In addition to (1.3), we assume that the coefficient function $v(x, t)$ is of bounded variation in t for almost all $x \in \Omega$. Then we can conclude maximal parabolic regularity for the parabolic equations (3.1) and (3.2) because the corresponding right-hand sides $-\frac{1}{\varrho} p$ and $y - y_d$ belong to $L_2(Q)$ (see [14]). Therefore, $\partial_t y$ and $L_x y := -\text{div}_x(v \nabla_x y)$ as well as $\partial_t p$ and $L_x p$ belong to $L_2(Q)$ too. In this case, the variational reduced optimality system (3.1)–(3.2) can be rewritten in the strong form as coupled forward and backward system of parabolic PDEs: Find $y \in Y_0 \cap H^{L,1}(Q)$ and $p \in P_T \cap H^{L,1}(Q)$ such that the coupled PDE optimality system

$$\varrho [\partial_t y - \text{div}_x(v \nabla_x y)] = -p \quad \text{in } L_2(Q) \quad (3.3)$$

$$-\partial_t p - \text{div}_x(v \nabla_x p) = y - y_d \quad \text{in } L_2(Q) \quad (3.4)$$

holds. Further regularity results for parabolic problems can be found, e.g., in [17, 29]. Later we need such regularity results for convergence rate estimates.

4 Space–time finite element discretization

We now use the PDE optimality system (3.3)–(3.4) as starting point for constructing stable and consistent space–time finite element schemes on fully unstructured simplicial triangulations of the space–time cylinder Q .

Thus, let us first consider a decomposition $\mathcal{T}_h = \{K\}$ of Q into shape-regular simplicial elements, i.e., $\bar{Q} = \bigcup_{K \in \mathcal{T}_h} \bar{K}$, and $K \cap K' = \emptyset$ for all K and K' from \mathcal{T}_h with $K \neq K'$ (see, e.g., [7, 16] for more details). Once a shape-regular triangulation is available, we define the space–time finite element spaces $Y_{0h} = \{y_h \in S_h^k(\bar{Q}) : y_h = 0 \text{ on } \bar{\Sigma} \cup \bar{\Sigma}_0\}$ and $P_{Th} = \{p_h \in S_h^k(\bar{Q}) : p_h = 0 \text{ on } \bar{\Sigma} \cup \bar{\Sigma}_T\}$. The standard finite element space

$$S_h^k(\bar{Q}) = \{y_h \in C(\bar{Q}) : y_h(x_K(\cdot)) \in \mathbb{P}_k(\bar{K}) \quad \forall K \in \mathcal{T}_h\}$$

consists of all continuous and piecewise polynomial functions (provided that $x_K(\cdot)$ is affine-linear), where $x_K(\cdot)$ denotes the map from the reference element \bar{K} (unit simplex) to the finite element $K \in \mathcal{T}_h$, and $\mathbb{P}_k(\bar{K})$ is the space of polynomials of degree $k \in \mathbb{N} := \{1, 2, \dots\}$ on the reference element \bar{K} .

In addition to the assumptions made above, we further assume that v is piecewise smooth in the sense that $\text{div}_x(v \nabla_x w_h)|_K \in L_2(K)$ for all w_h from Y_{0h} or P_{Th} and for all $K \in \mathcal{T}_h$. Now we multiply the first PDE (3.3) by a upwind test function $v_h + \vartheta \lambda_h^2 \partial_t v_h$ with $v_h \in Y_{0h}$, and the second one (3.4) by $q_h - \vartheta \lambda_h^2 \partial_t q_h$ with $q_h \in P_{Th}$, where ϑ is some positive scaling parameter and $\lambda_h \in W_\infty^1(Q)$ is a mesh-density function which we will choose later. Then, integrating over K , integrating by parts in the elliptic terms where the scaling parameter ϑ does not appear, and summing over all $K \in \mathcal{T}_h$, we arrive at the variational consistency identity

$$a_h(y, p; v_h, q_h) = \ell_h(v_h, q_h) \quad \forall (v_h, q_h) \in Y_{0h} \times P_{Th} \quad (4.1)$$

with the combined bilinear and linear forms

$$\begin{aligned} a_h(y, p; v, q) = \sum_{K \in \mathcal{T}_h} \int_K & \left[\varrho (\partial_t y v + \vartheta \lambda_h^2 \partial_t y \partial_t v + v \nabla_x y \cdot \nabla_x v + \vartheta \lambda_h^2 L_x y \partial_t v) \right. \\ & + p(v + \vartheta \lambda_h^2 \partial_t v) - \partial_t p q + \vartheta \lambda_h^2 \partial_t p \partial_t q + v \nabla_x p \cdot \nabla_x q \\ & \left. - \vartheta \lambda_h^2 L_x p \partial_t q - y(q - \vartheta \lambda_h^2 \partial_t q) \right] dK \end{aligned} \quad (4.2)$$

$$\ell_h(v, q) = - \sum_{K \in \mathcal{T}_h} \int_K y_d (q - \vartheta \lambda_h^2 \partial_t q) dK \quad (4.3)$$

respectively. The continuous and piecewise differentiable mesh-density function $\lambda_h(x, t)$ should be chosen in such a way that the inequalities

$$0 < \underline{\lambda}_0 h_K \leq \lambda_h(x, t) \leq \bar{\lambda}_0 h_K, \quad |\partial_t \lambda_h(x, t)| \leq \lambda_1 \quad (4.4)$$

hold for all $(x, t) \in K$, $K \in \mathcal{T}_h$, and for all meshes under consideration, where $h_K = \text{diam}(K)$, and $\underline{\lambda}_0$, $\bar{\lambda}_0$ and λ_1 are some positive generic constants. We can easily define the mesh density function $\lambda_h(x, t)$ by means of the finite element space $S_h^1(\bar{Q}) = \text{span}\{\varphi_i : i = 1, 2, \dots, n_h\}$, spanned by the nodal basis functions $\varphi_i(x, t)$, as follows

$$\lambda_h(x, t) = \sum_{i=1}^{n_h} h_i \varphi_i(x, t)$$

where h_i denotes the average of the length of all edges meeting at the vertex (x^i, t^i) , i.e. the sum of the lengths divided by the number of these edges. We note that $\lambda_h((x^i, t^i)) = h_i$ since $\varphi_j((x^i, t^i)) = \delta_{i,j}$, where $\delta_{i,j}$ denotes Kronecker's symbol. The conformity and the shape regularity of the mesh immediately imply inequalities (4.4). The general discretization parameter h can be chosen as $(n_h)^{-1/(d+1)}$ or $(N_h)^{-1/(d+1)}$ or $(M_h)^{-1/(d+1)}$, where n_h is the number of vertices in \mathcal{T}_h , $N_h = \dim Y_{0h}$, and $M_h = \dim P_{Th}$. We note that $h_K = O(h)$ for all $K \in \mathcal{T}_h$ in the case of a (quasi) uniform mesh.

The consistent finite element scheme corresponding to (4.1) now reads as follows: Find $(y_h, p_h) \in Y_{0h} \times P_{Th}$ such that

$$a_h(y_h, p_h; v_h, q_h) = \ell_h(v_h, q_h) \quad \forall (v_h, q_h) \in Y_{0h} \times P_{Th}. \quad (4.5)$$

Subtracting (4.5) from (4.1), we immediately obtain the Galerkin orthogonality relation

$$a_h(y - y_h, p - p_h; v_h, q_h) = 0 \quad \forall (v_h, q_h) \in Y_{0h} \times P_{Th} \quad (4.6)$$

which is crucial for deriving discretization error estimates.

5 Coercivity and unique solvability

We will now show that the bilinear form a_h is coercive on $Y_{0h} \times P_{Th}$ with respect to the norm

$$\|(v, q)\|_h^2 = \|v\|_{h,T}^2 + \|q\|_{h,0}^2 \quad (5.1)$$

for some suitably chosen scaling parameter ϑ , with

$$\begin{aligned} \|v\|_{h,T}^2 &= \alpha \varrho \|v(\cdot, T)\|_{L_2(\Omega)}^2 + \varrho \vartheta \|\lambda_h \partial_t v\|_{L_2(Q)}^2 + \varrho \|\sqrt{\nu} \nabla_x v\|_{L_2(Q)}^2 \\ \|q\|_{h,0}^2 &= \alpha \|q(\cdot, 0)\|_{L_2(\Omega)}^2 + \vartheta \|\lambda_h \partial_t q\|_{L_2(Q)}^2 + \|\sqrt{\nu} \nabla_x q\|_{L_2(Q)}^2 \end{aligned}$$

where α is some positive parameter. We choose $\alpha = 1$ if not stated otherwise. First, we observe that the identity

$$\begin{aligned} a_h(v_h, q_h; v_h, q_h) &= \sum_{K \in \mathcal{T}_h} \int_K \left[\varrho (\partial_t v_h v_h + \vartheta \lambda_h^2 |\partial_t v_h|^2 + |\sqrt{\nu} \nabla_x v_h|^2 \right. \\ &\quad + \vartheta \lambda_h^2 L_x v_h \partial_t v_h) + q_h (v_h + \vartheta \lambda_h^2 \partial_t v_h) - \partial_t q_h q_h + \vartheta \lambda_h^2 |\partial_t q_h|^2 \\ &\quad \left. + |\sqrt{\nu} \nabla_x q_h|^2 - \vartheta \lambda_h^2 L_x q_h \partial_t q_h - v_h (q_h - \vartheta \lambda_h^2 \partial_t q_h) \right] dK \\ &= \sum_{K \in \mathcal{T}_h} \int_K \varrho (\vartheta \lambda_h^2 |\partial_t v_h|^2 + |\sqrt{\nu} \nabla_x v_h|^2 + \vartheta \lambda_h^2 L_x v_h \partial_t v_h) dK \\ &\quad + \sum_{K \in \mathcal{T}_h} \int_K (\vartheta \lambda_h^2 |\partial_t q_h|^2 + |\sqrt{\nu} \nabla_x q_h|^2 - \vartheta \lambda_h^2 L_x q_h \partial_t q_h) dK \\ &\quad + \frac{\varrho}{2} \|v_h(\cdot, T)\|_{L_2(\Omega)}^2 + \frac{1}{2} \|q_h(\cdot, 0)\|_{L_2(\Omega)}^2 - \int_Q 2\vartheta \lambda_h \partial_t \lambda_h v_h q_h dQ \end{aligned} \quad (5.2)$$

is valid for all $(v_h, q_h) \in Y_{0h} \times P_{Th}$. Here we have used the identities

$$\int_Q \varrho \partial_t v_h v_h dQ = \frac{\varrho}{2} \|v_h(\cdot, T)\|_{L_2(\Omega)}^2, \quad - \int_Q \partial_t q_h q_h dQ = \frac{1}{2} \|q_h(\cdot, 0)\|_{L_2(\Omega)}^2$$

and

$$\begin{aligned} \int_Q \vartheta \lambda_h^2 v_h \partial_t q_h dQ &= - \int_Q \vartheta \partial_t (\lambda_h^2 v_h) q_h dQ + \int_{\partial Q} \vartheta \lambda_h^2 v_h q_h ds \\ &= - \int_Q \vartheta \lambda_h^2 \partial_t v_h q_h dQ - \int_Q 2\vartheta \lambda_h \partial_t \lambda_h v_h q_h dQ. \end{aligned}$$

We note that, in the latter identity, the boundary integral vanishes for all finite element functions (v_h, q_h) from $Y_{0h} \times P_{Th}$.

The special case that

1. $k = 1$ and $v = \text{const}$ on K for all $K \in \mathcal{T}_h$, and
2. $\partial_t \lambda_h = 0$, e.g., uniform mesh $h_i = h$, as considered in [34],

makes the terms with the second-order elliptic operator L_x and the last term in (5.2) zero. Then the identity (5.2) gives

$$a_h(v_h, q_h; v_h, q_h) = \mu_c \|(v, q)\|_h^2 \quad \forall (v_h, q_h) \in Y_{0h} \times P_{Th}$$

with $\mu_c = 1$ provided that α is chosen as 0.5 in the definition of the norm (5.1), i.e., the bilinear form is coercive on $Y_{0h} \times P_{Th}$ with the coercivity constant 1.

In the general case, we have to estimate these terms. Beside Cauchy's and Young's inequalities, we need the generalized Friedrichs' inequality

$$\|v\|_{L_2(Q)} \leq c_F(v) \|\sqrt{\nu} \nabla_x v\|_{L_2(Q)} \quad \forall v \in V = L_2(0, T; H_0^1(\Omega)) \quad (5.3)$$

where $c_F(v)$ can obviously be estimated from above by the Friedrichs constant $c_{F,\Omega}$ of the spatial domain Ω times $(v_1)^{-1/2}$, i.e., $c_F(v) \leq c_{F,\Omega}/v_1^{1/2}$, and the special inverse inequality

$$\lambda_{\max,K} := \max_{\mathbf{w}_h \in \mathbb{V}_x S_h^k(\bar{K})} \frac{(\text{div}_x(v \mathbf{w}_h), \text{div}_x(v \mathbf{w}_h))_{L_2(K)}}{(\sqrt{\nu} \mathbf{w}_h, \sqrt{\nu} \mathbf{w}_h)_{L_2(K)}} \leq c_{\text{inv},K}^2 h_K^{-2} \quad (5.4)$$

where $\lambda_{\max,K}$ can be computed from a small generalized matrix eigenvalue problem for every finite element $K \in \mathcal{T}_h$ (see [32]). The upper bound can be shown by mapping K to the reference element \widehat{K} . Now, using these inequalities, we can proceed to estimate (5.2) from below as follows:

$$\begin{aligned} a_h(v_h, q_h; v_h, q_h) &\geq \frac{\varrho}{2} \|v_h(\cdot, T)\|_{L_2(\Omega)}^2 + \frac{1}{2} \|q_h(\cdot, 0)\|_{L_2(\Omega)}^2 \\ &\quad + \sum_{K \in \mathcal{T}_h} \int_K \varrho \left[\vartheta \lambda_h^2 \left(1 - \frac{1}{2\varepsilon}\right) |\partial_t v_h|^2 + (1 - \varepsilon \vartheta a_K - \vartheta b \varrho^{-1}) |\sqrt{v} \nabla_x v_h|^2 \right] dK \\ &\quad + \sum_{K \in \mathcal{T}_h} \int_K \left[\vartheta \lambda_h^2 \left(1 - \frac{1}{2\bar{\varepsilon}}\right) |\partial_t q_h|^2 + (1 - \bar{\varepsilon} \vartheta a_K - \vartheta b) |\sqrt{v} \nabla_x q_h|^2 \right] dK \\ &\geq \mu_c \|(v_h, q_h)\|_h^2 \quad \forall (v_h, q_h) \in Y_{0h} \times P_{Th} \end{aligned} \quad (5.5)$$

where ε and $\bar{\varepsilon}$ are positive parameters from Young's inequalities, $c_F = c_F(v)$, $a_K = 0.5 \bar{\lambda}_0^2 c_{\text{inv},K}^2$, $b = \bar{\lambda}_0 \lambda_1 h c_F^2$, and $h = \max_{K \in \mathcal{T}_h} h_K$. The coercivity constant μ_c is positive for properly chosen positive parameters ε , $\bar{\varepsilon}$, and ϑ . Indeed, $\mu_c = 1/2$, if $\varepsilon = 1$, $\bar{\varepsilon} = 1$, and

$$0 < \vartheta \leq \frac{1}{2} \min \left\{ \frac{\varrho}{\varrho a + b}, \frac{1}{a + b} \right\} \quad (5.6)$$

with $a = \max_{K \in \mathcal{T}_h} a_K$.

The coercivity (5.5) of the bilinear form a_h on $Y_{0h} \times P_{Th}$ immediately implies that the finite element scheme (4.5) can only have one solution, but, in the finite dimensional case, uniqueness yields existence. Thus, the space–time finite element scheme (4.5) has a unique solution that can be determined via the solution of corresponding algebraic system of finite element equations (see Section 8).

Remark 5.1. In the case of a quasi-uniform mesh, formally defined by the setting $h_K = h$ and $h_i = h$, i.e., $\lambda_h = h$, $\partial_t \lambda_h = 0$, and, therefore, $\bar{\lambda}_0 = \bar{\lambda}_0 = 1$ and $\lambda_1 = 0$, the coercivity estimate (5.5) yields the coercivity estimate presented in [34]. We note that, for $k = 1$, $L_x v_h = L_x q_h = 0$ on $K \in \mathcal{T}_h$ for all $(v_h, q_h) \in Y_{0h} \times P_{Th}$ provided that v is elementwise constant.

6 A priori discretization error estimates

In order to derive a priori discretization error estimates, we first need to establish the extended boundedness of the bilinear form

$$|a_h(y, p; v_h, q_h)| \leq \mu_b \|(y, p)\|_{h,*} \|(v_h, q_h)\|_h \quad \forall (v_h, q_h) \in Y_{0h} \times P_{Th} \quad (6.1)$$

and for all $y \in Y_{0h} + Y_0 \cap H^{L,1}(Q)$ and $p \in P_{Th} + P_T \cap H^{L,1}(Q)$, where

$$\begin{aligned} \|(y, p)\|_{h,*}^2 &= \|(y, p)\|_h^2 + \varrho \vartheta \sum_{K \in \mathcal{T}_h} \|\lambda_h L_x y\|_{L_2(K)}^2 + 3\varrho \vartheta^{-1} \|\lambda_h^{-1} y\|_Q^2 \\ &\quad + \vartheta \sum_{K \in \mathcal{T}_h} \|\lambda_h L_x p\|_{L_2(K)}^2 + 3\vartheta^{-1} \|\lambda_h^{-1} p\|_Q^2. \end{aligned}$$

The boundedness constant μ_b will be defined below. At first, using integration by parts with respect to t , we get the identity

$$\begin{aligned} a_h(y, p; v_h, q_h) &= \int_Q p(v_h + \vartheta \lambda_h^2 \partial_t v_h) dQ + \varrho \int_{\partial Q} y v_h n_t ds - \varrho \int_Q y \partial_t v_h dQ \\ &\quad + \varrho \int_Q (\vartheta \lambda_h^2 \partial_t y \partial_t v_h + v \nabla_x y \cdot \nabla_x v_h) dQ + \varrho \sum_{K \in \mathcal{T}_h} \int_K \vartheta \lambda_h^2 L_x y \partial_t v_h dK \\ &\quad - \int_Q y(q_h - \vartheta \lambda_h^2 \partial_t q_h) dQ - \int_{\partial Q} p q_h n_t ds + \int_Q p \partial_t q_h dQ \\ &\quad + \int_Q (\vartheta \lambda_h^2 \partial_t p \partial_t q_h + v \nabla_x p \cdot \nabla_x q_h) dQ - \sum_{K \in \mathcal{T}_h} \int_K \vartheta \lambda_h^2 L_x p \partial_t q_h dK \end{aligned}$$

that is valid for all $(y, p) \in (Y_{0h} + Y_0 \cap H^{L,1}(Q)) \times (P_{Th} + P_T \cap H^{L,1}(Q))$ and $(v_h, q_h) \in Y_{0h} \times P_{Th}$, and that is the starting point for estimating the right-hand side from above. Indeed, Young's and Cauchy's inequalities, and the generalized Friedrichs inequality (5.3), yield the following estimates:

$$\begin{aligned}
a_h(y, p; v_h, q_h) &\leq \varrho \|y(\cdot, T)\|_{\Omega} \|v_h(\cdot, T)\|_{\Omega} + \varrho \vartheta^{-1/2} \|\lambda_h^{-1} y\|_Q \vartheta^{1/2} \|\lambda_h \partial_t v_h\|_Q \\
&\quad + \varrho \vartheta \|\lambda_h \partial_t y\|_Q \|\lambda_h \partial_t v_h\|_Q + \varrho \|\sqrt{v} \nabla_x y\|_Q \|\sqrt{v} \nabla_x v_h\|_Q \\
&\quad + \left(\sum_{K \in \mathcal{T}_h} \varrho \vartheta \|\lambda_h L_x y\|_K^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}_h} \varrho \vartheta \|\lambda_h \partial_t v_h\|_K^2 \right)^{1/2} \\
&\quad + \vartheta^{-1/2} \|\lambda_h^{-1} p\|_Q \vartheta^{1/2} \|\lambda_h v_h\|_Q + \vartheta^{-1/2} \|\lambda_h^{-1} p\|_Q \vartheta^{3/2} \|\lambda_h^3 \partial_t v_h\|_Q \\
&\quad + \|p(\cdot, 0)\|_{\Omega} \|q_h(\cdot, 0)\|_{\Omega} + \vartheta^{-1/2} \|\lambda_h^{-1} p\|_Q \vartheta^{1/2} \|\lambda_h \partial_t q_h\|_Q \\
&\quad + \vartheta^{1/2} \|\lambda_h \partial_t p\|_Q \vartheta^{1/2} \|\lambda_h \partial_t q_h\|_Q + \|\sqrt{v} \nabla_x p\|_Q \|\sqrt{v} \nabla_x q_h\|_Q \\
&\quad + \left(\sum_{K \in \mathcal{T}_h} \vartheta \|\lambda_h L_x p\|_K^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}_h} \vartheta \|\lambda_h \partial_t q_h\|_K^2 \right)^{1/2} \\
&\quad + \varrho^{1/2} \vartheta^{-1/2} \|\lambda_h^{-1} y\|_Q \varrho^{-1/2} \vartheta^{1/2} \|\lambda_h q_h\|_Q \\
&\quad + \varrho^{1/2} \vartheta^{-1/2} \|\lambda_h^{-1} y\|_Q \varrho^{-1/2} \vartheta^{3/2} \|\lambda_h^3 \partial_t q_h\|_Q \\
&\leq \left[\varrho \|y(\cdot, T)\|_{\Omega}^2 + \varrho \vartheta^{-1} \|\lambda_h^{-1} y\|_Q^2 + \varrho \vartheta \|\lambda_h \partial_t y\|_Q^2 + \varrho \|\sqrt{v} \nabla_x y\|_Q^2 \right. \\
&\quad + \varrho \sum_{K \in \mathcal{T}_h} \vartheta \|\lambda_h L_x y\|_K^2 + \vartheta^{-1} \|\lambda_h^{-1} p\|_Q^2 + \vartheta^{-1} \|\lambda_h^{-1} p\|_Q^2 \\
&\quad + \|p(\cdot, 0)\|_{\Omega}^2 + \vartheta^{-1} \|\lambda_h^{-1} p\|_Q^2 + \vartheta \|\lambda_h \partial_t p\|_Q^2 + \|\sqrt{v} \nabla_x p\|_Q^2 \\
&\quad \left. + \sum_{K \in \mathcal{T}_h} \vartheta \|\lambda_h L_x p\|_K^2 + \varrho \vartheta^{-1} \|\lambda_h^{-1} y\|_Q^2 + \varrho \vartheta^{-1} \|\lambda_h^{-1} y\|_Q^2 \right]^{1/2} \\
&\quad \times \left[\varrho \|v_h(\cdot, T)\|_{\Omega}^2 + \varrho \vartheta \|\lambda_h \partial_t v_h\|_Q^2 + \varrho \vartheta \|\lambda_h \partial_t v_h\|_Q^2 + \varrho \|\sqrt{v} \nabla_x v_h\|_Q^2 \right. \\
&\quad + \varrho \vartheta \|\lambda_h \partial_t v_h\|_Q^2 + \vartheta \|\lambda_h v_h\|_Q^2 + \vartheta^2 \bar{\lambda}_0^4 h^4 \vartheta \|\lambda_h \partial_t v_h\|_Q^2 \\
&\quad + \|q_h(\cdot, 0)\|_{\Omega}^2 + \vartheta \|\lambda_h \partial_t q_h\|_Q^2 + \vartheta \|\lambda_h \partial_t q_h\|_Q^2 + \|\sqrt{v} \nabla_x q_h\|_Q^2 \\
&\quad \left. + \vartheta \|\lambda_h \partial_t q_h\|_Q^2 + \varrho^{-1} \vartheta \|\lambda_h q_h\|_Q^2 + \varrho^{-1} \vartheta^2 \bar{\lambda}_0^4 h^4 \vartheta \|\lambda_h \partial_t q_h\|_Q^2 \right]^{1/2} \\
&\leq \left[\varrho \|y\|_{h,T} + \|p\|_{h,0} + \varrho \sum_{K \in \mathcal{T}_h} \vartheta \|\lambda_h L_x y\|_K^2 + \sum_{K \in \mathcal{T}_h} \vartheta \|\lambda_h L_x p\|_K^2 \right. \\
&\quad \left. + (\varrho + \varrho + \varrho) \vartheta^{-1} \|\lambda_h^{-1} y\|_Q^2 + (2 + 1) \vartheta^{-1} \|\lambda_h^{-1} p\|_Q^2 \right]^{1/2} \\
&\quad \times \left[\varrho \|v_h(\cdot, T)\|_{\Omega}^2 + (3\varrho + \vartheta^2 \bar{\lambda}_0^4 h^4) \vartheta \|\lambda_h \partial_t v_h\|_Q^2 \right. \\
&\quad + (\varrho + \vartheta \bar{\lambda}_0^2 h^2 c_F(v)^2) \|\sqrt{v} \nabla_x v_h\|_Q^2 \\
&\quad + \|q_h(\cdot, 0)\|_{\Omega}^2 + (3 + \varrho^{-1} \vartheta^2 \bar{\lambda}_0^4 h^4) \vartheta \|\lambda_h \partial_t q_h\|_Q^2 \\
&\quad \left. + (1 + \varrho^{-1} \vartheta \bar{\lambda}_0^2 h^2 c_F(v)^2) \|\sqrt{v} \nabla_x q_h\|_Q^2 \right]^{1/2} \\
&\leq \mu_b \|(y, p)\|_{h,*} \|(v_h, q_h)\|_h
\end{aligned}$$

with $\mu_b = (\max \{3 + \varrho^{-1} \vartheta^2 \bar{\lambda}_0^4 h^4, 1 + \vartheta \bar{\lambda}_0^2 h^2 c_F(v)^2 \varrho^{-1}\})^{1/2}$, where we use the shorter notations $\|\cdot\|_D$ for the L_2 -norms $\|\cdot\|_{L_2(D)}$. The following best-approximation error estimate is now an easy consequence of the Galerkin orthogonality (4.6), the coercivity estimate (5.5), the extended boundedness estimate (6.1), and the triangle inequality.

Theorem 6.1. *Let us assume that the coefficient v fulfills the conditions leading to maximal parabolic regularity, i.e., the unique solution (y, p) of the optimality system (3.3)–(3.4) belongs to the space $Y_0 \cap H^{l,1}(Q) \times P_T \cap H^{l,1}(Q)$. Furthermore, let ϱ be some fixed positive regularization parameter, and let ϑ be a fixed scaling parameter such that condition (5.6) is satisfied. Then the discretization error estimate*

$$\|(y, p) - (y_h, p_h)\|_h \leq \inf_{(v_h, q_h) \in Y_{0h} \times P_{Th}} \left(\|(y, p) - (v_h, q_h)\|_h + \frac{\mu_b}{\mu_c} \|(y, p) - (v_h, q_h)\|_{h,*} \right) \quad (6.2)$$

holds, where $(y_h, p_h) \in Y_{0h} \times P_{Th}$ is the unique solution of the finite element optimality system (4.5).

Under additional regularity assumptions imposed on the solution (y, p) , the best-approximation error estimate (6.2) yields discretization error estimates in terms of the mesh-sizes. Indeed, choosing $v_h = I_h y \in Y_{0h}$ and $q_h = I_h p \in P_{Th}$ in the infimum at the right-hand side of (6.2), we can show the following theorem, where I_h is the usual nodal finite element interpolation operator for sufficiently smooth solutions $y, p \in H^{m+1}(Q) \subset C(\bar{Q})$ with $m + 1 > (d + 1)/2$, or a quasi-interpolation operator à la Clément [12] and Scott–Zhang [48] for low-regularity solutions $y, p \in H^{m+1}(Q)$ with $m > 0$ (see also [8]).

Theorem 6.2. *Let the assumptions of Theorem 6.1 hold. We further assume that all transformations $x_K(\cdot)$ from the reference element \bar{K} to the physical finite element $K \in \mathcal{T}_h$ are affine-linear. Then, in the smooth case, when assuming $y, p \in \{v \in H^{m+1}(Q) : v|_K \in H^\ell(K) \forall K \in \mathcal{T}_h\}$ for some $\ell \geq m + 1 > (d + 1)/2$, we get the estimate*

$$\|(y, p) - (y_h, p_h)\|_h \leq \left(1 + \frac{\mu_b}{\mu_c} \right) \left(\sum_{K \in \mathcal{T}_h} h_K^{2(s-1)} (c_1 |y|_{H^s(K)}^2 + c_2 |p|_{H^s(K)}^2) \right)^{1/2} \quad (6.3)$$

with $s = \min\{\ell, k + 1\}$ and some generic positive constants c_1 and c_2 , whereas, in the non-smooth case, when only assuming $y, p \in H^{m+1}(Q)$ for some m such that $(d + 1)/2 \geq m + 1 > 1$, and $v \nabla_x y, v \nabla_x p \in (H^m(K))^d$, we arrive at the estimate

$$\|(y, p) - (y_h, p_h)\|_h \leq \left(1 + \frac{\mu_b}{\mu_c} \right) \left(\sum_{K \in \mathcal{T}_h} h_K^{2(s-1)} (c_3 \mathcal{N}^2(y) + c_4 \mathcal{N}^2(p)) \right)^{1/2} \quad (6.4)$$

with $s = \min\{m + 1, k + 1\} = m + 1 \leq (d + 1)/2$ for $d = 2, 3$, and positive generic constants c_3 and c_4 , where

$$\mathcal{N}^2(v) = |v|_{H^s(S_K)}^2 + |v \nabla_x v|_{H^{s-1}(K)}^2 + h_K^{4-2s} \|\operatorname{div}_x(v \nabla_x v)\|_{L_2(K)}^2$$

and $S_K := \{K' \in \mathcal{T}_h : \bar{K} \cap \bar{K}' \neq \emptyset\}$ denotes the neighborhood of the simplex $K \in \mathcal{T}_h$.

Proof. Using standard interpolation respectively quasi-interpolation error estimates (see, e.g., [8, 12, 22, 48] for space interpolation in the case of semi-norms), we can derive the discretization error estimates (6.3) and (6.4) from the best-approximation error estimate (6.2) along the line of the proofs for the state equations given in [32] and [33] for the smooth and non-smooth cases, respectively. Indeed, in the norms of the right-hand side of (6.2), the same (quasi-) interpolation error terms appear as in the corresponding best-approximation error estimates in [32, 33]. The only difference consists in the inclusions of the positive mesh-density function λ_h or $1/\lambda_h$ as weights into the norms. Using the local lower and upper bounds (4.4) of the mesh-density function λ_h , we can estimate these terms as follows:

$$\|\lambda_h \partial_t e\|_{L_2(Q)}^2 = \sum_{K \in \mathcal{T}_h} \|\lambda_h \partial_t e\|_{L_2(K)}^2 \leq \sum_{K \in \mathcal{T}_h} \bar{\lambda}_0^2 h_K^2 \|\partial_t e\|_{L_2(K)}^2 \quad (6.5)$$

$$\|\lambda_h^{-1} e\|_{L_2(Q)}^2 = \sum_{K \in \mathcal{T}_h} \|\lambda_h^{-1} e\|_{L_2(K)}^2 \leq \sum_{K \in \mathcal{T}_h} \underline{\lambda}_0^{-2} h_K^{-2} \|e\|_{L_2(K)}^2 \quad (6.6)$$

and

$$\sum_{K \in \mathcal{T}_h} \|\lambda_h L_x e\|_{L_2(K)}^2 \leq \sum_{K \in \mathcal{T}_h} \bar{\lambda}_0^2 h_K^2 \|L_x e\|_{L_2(K)}^2 \quad (6.7)$$

where e denotes the (quasi) interpolation errors $y - I_h y$ or $p - I_h p$ for the state or co-state, respectively. For the further estimation of (6.5) and (6.6), we will use standard nodal interpolation or quasi-interpolation error

estimates in dependence of the regularity of the state y or the co-state p . The further estimation of $\|L_x e\|_{L_2(K)}^2$ requires a special treatment. Let us now derive the estimation of $\|L_x e\|_{L_2(K)}^2 = \|\operatorname{div}_x(v\nabla_x(y - I_h y))\|_{L_2(K)}^2$ for the state y in the low-regularity case $y \in H^{m+1}(Q)$ for some $(d+1)/2 \geq m+1 > 1$, and $v\nabla_x y \in (H^m(K))^d$. Inserting and subtracting the average

$$I_h^0 \mathbf{q}(x) = \frac{1}{|K|} \int_K \mathbf{q}(\xi) \, d\xi \in \mathbb{R}^d \quad \text{for } x \in K \quad (6.8)$$

of the flux $\mathbf{q} = -v\nabla_x y$ over $K \in \mathcal{T}_h$ and using the inverse inequality $\|\operatorname{div}_x(c_h + v\nabla_x v_h)\|_{L_2(K)} \leq ch_K^{-1} \|c_h + v\nabla_x v_h\|_{L_2(K)}$ for all $(c_h, v_h) \in \mathbb{R}^d \times V_{0h}$, we can estimate the term $\|\operatorname{div}_x(v\nabla_x(y - I_h y))\|_{L_2(K)}^2$ as follows:

$$\begin{aligned} \|\operatorname{div}_x(v\nabla_x e)\|_{L_2(K)} &= \|\operatorname{div}_x(\mathbf{q} - I_h^0 \mathbf{q} + I_h^0 \mathbf{q} - (-v\nabla_x(I_h y)))\|_{L_2(K)} \\ &\leq \|\operatorname{div}_x \mathbf{q}\|_{L_2(K)} + \|\operatorname{div}_x(I_h^0 \mathbf{q} - (-v\nabla_x(I_h y)))\|_{L_2(K)} \\ &\leq \|\operatorname{div}_x \mathbf{q}\|_{L_2(K)} + ch_K^{-1} \|I_h^0 \mathbf{q} - (-v\nabla_x(I_h y))\|_{L_2(K)} \end{aligned}$$

where I_h is a quasi-interpolation operator like mentioned above. The last term can be estimated by means of the local approximation properties of I_h^0 and I_h . Indeed, inserting and subtracting \mathbf{q} in the last term of the above estimate, we obtain

$$\begin{aligned} \|I_h^0 \mathbf{q} - (-v\nabla_x(I_h y))\|_{L_2(K)}^2 &\leq 2\|I_h^0 \mathbf{q} - \mathbf{q}\|_{L_2(K)}^2 + 2\|\mathbf{q} - (-v\nabla_x(I_h y))\|_{L_2(K)}^2 \\ &\leq c \left(h_K^{2(s-1)} |v\nabla_x(y)|_{H^{s-1}(K)}^2 + h_K^{2(s-1)} |y|_{H^s(S_K)}^2 \right) \end{aligned}$$

with some generic positive constant c , where we have used Poincaré's inequality for estimating the first term. The high-regularity case is straightforward since one can again use the nodal interpolation operator, and one can estimate $\|\operatorname{div}_x(v\nabla_x(y - I_h y))\|_{L_2(K)}^2$ directly. We always assume that the coefficient v restricted to an element K is sufficiently smooth, but v can be discontinuous globally. It is clear that the same estimations can be derived for the co-state p . \square

We mention that the semi-norms in the estimates (6.3) and (6.4) have to be replaced by the corresponding full norms in the case of non-affine linear mappings $x_K(\cdot)$ from \widehat{K} to $K \in \mathcal{T}_h$, e.g., if one uses the popular isoparametric elements ($k \geq 2$) or general non-linear mappings in order to represent (or approximate) curved interfaces or boundaries. Regularity results for the solution (y, p) of the optimality system (3.1)–(3.2) can be derived from known regularity results for parabolic initial-boundary value problems [17, 28, 29] by simple bootstrapping arguments. Here we always assume that the coefficient v is as smooth as necessary. Such kind of bootstrapping arguments can be found in [19] and [36, Rem. 3.3] for L_2 and energy regularizations, respectively. In the case considered in our paper, we always assume maximal parabolic regularity, i.e., $\partial_t y$, $\partial_{x_i} y$, $L_x y$ and $\partial_t p$, $\partial_{x_i} p$, $L_x p$ belong to $L_2(Q)$. The a priori convergence rate estimates of Theorem 6.2 are based on additional (isotropic) regularity information $y, p \in H^{1+m}(Q)$, but not on anisotropic regularity information. Anisotropic regularity information can be the starting point for anisotropic mesh refinement that can be done better for hexahedral elements than simplicial elements; see, e.g., [35] for adaptive hexahedral space–time finite element approximations to parabolic initial boundary value problems, where we consider a parabolic problem on a ‘slit’ spatial domain with a solution that is analytic in t and has low spatial regularity due to the elliptic singularity at the top of the slit.

In the case of low-regularity solutions when singularities in the time or/and space derivatives appear, the convergence rate is affected by these singularities, e.g., we observe $O(h^{s-1}) = O(h^m)$ instead of $O(h)$ for linear shape functions ($k = 1$) if (y, p) only belongs to $H^{1+m}(Q)$, where $s = \min\{1+m, 2\} = 1+m$ for $0 < m \leq 1$. The full rate $O(h)$ can be recovered by mesh grading that was already used for elliptic boundary value problems in early works by Oganessian and Rukhovets [45]; see also [2] for more recent results, and [13] for space–time finite element mesh grading. However, in order to use the mesh grading technique, we must know the strengths of the singularities and their positions in space and time. Thus, in practice, one wants to implement an adaptive finite element method (AFEM) that is based on local error indicators and automatic adaptive mesh refinement. Such an adaptive finite element code should do the same job as an a priori mesh grading; see also [30] for adaptive space–time Isogeometric Analysis of parabolic initial boundary value problems.

7 Error indicators and adaptivity

In many practical applications, such low-regularity solutions can appear due to discontinuous coefficients, re-entrant corners in the computational domain, changing boundary conditions, and missing compatibility of the boundary conditions with the initial conditions. Discontinuous target states y_d will lead to steep gradients, in particular, for small regularization parameters ρ . We note that the target y_d can change discontinuously in space and time (cf. the numerical example from Subsection 9.2). Hence, for such kind of non-smooth problems, it makes sense to use adaptive mesh refinements that are guided by local error indicators based on a posteriori error indicators or even estimators. The unstructured space–time approach considered in this paper allows for adaptivity in space and time simultaneously. This is a huge advantage of fully unstructured space–time techniques over the usual time-stepping methods or tensor-product techniques.

In order to drive the adaptive process, we will first use the error indicator of residual type proposed by Steinbach and Yang in [52] for parabolic PDEs; see also [37] for applications to parabolic optimal control problems. In particular, given the finite element state y_h and co-state p_h , we compute the local indicator

$$\eta_K^2(y_h, p_h) = \eta_{K,y}^2(y_h, p_h) + \eta_{K,p}^2(y_h, p_h)$$

where

$$\eta_{K,y}^2(y_h, p_h) = h_K^2 \|\varrho(\partial_t y_h - \operatorname{div}_x(v \nabla_x y_h)) + p_h\|_K^2 + h_K \|\varrho \llbracket v \nabla_x y_h \rrbracket\|_{\partial K}^2$$

and

$$\eta_{K,p}^2(y_h, p_h) = h_K^2 \|y_h - y_d + \partial_t p_h + \operatorname{div}_x(v \nabla_x y_h)\|_K^2 + h_K \|\llbracket v \nabla_x p_h \rrbracket\|_{\partial K}^2$$

are nothing but the residual and jump terms of the coupled PDE optimality system in the strong form (3.3)–(3.4), respectively.

While the residual error indicator is easy to realize and has a low computational cost, we have no rigorous analysis even in the case of the state equation only (cf. [52]). However, on the basis of the continuous inf-sup condition for the reduced optimality system given in [37], we can rigorously derive a functional a posteriori error estimate that provides a guaranteed upper bound and that can be easily localized. Moreover, this functional a posteriori error estimate is of general importance since it can be used to estimate the error between the exact weak solution $(y, p) \in Y_0 \times P_T$ of the reduced variational optimality system (3.1)–(3.2) and any approximation (\tilde{y}, \tilde{p}) belonging to the space $Y_0 \cap H^1(Q) \times P_T \cap H^1(Q)$. Here we are interested in the finite element approximation $(y_h, p_h) \in Y_{0h} \times P_{Th}$ computed by our finite element scheme (4.5). The variational optimality system (3.1)–(3.2) can be written in the compact form

$$a((y, p), (v, q)) = -(y_d, q) \quad \forall (v, q) \in V \times V \quad (71)$$

obtained by adding (3.1) and (3.2). For simplicity, we assume that $v = 1$. Thus the bilinear form $a(\cdot, \cdot)$ is given by the expression

$$\begin{aligned} a((y, p), (v, q)) := & \varrho \int_Q [\partial_t y v + \nabla_x y \cdot \nabla_x v] dQ + \int_Q p v dQ \\ & - \int_Q y q dQ + \int_Q [-\partial_t p q + \nabla_x p \cdot \nabla_x q] dQ. \end{aligned}$$

Now, using the inf-sup estimate [37, Lem. 3.1], we can estimate the difference between any approximation $(\tilde{y}, \tilde{p}) \in Y_0 \cap H^1(Q) \times P_T \cap H^1(Q)$ and the exact solution $(y, p) \in Y_0 \times P_T$ as follows:

$$\frac{1}{\sqrt{2}} \sqrt{\varrho \|y - \tilde{y}\|_{X_0}^2 + \|p - \tilde{p}\|_{X_T}^2} \leq \sup_{(v, q) \in V \times V} \frac{a((y, p) - (\tilde{y}, \tilde{p}), (v, q))}{\sqrt{\varrho \|\nabla_x v\|_Q^2 + \|\nabla_x q\|_Q^2}} \quad (72)$$

with $\|y\|_{X_0}^2 = \|y\|_V^2 + \|\partial_t y\|_{V^*}^2 = \|\nabla_x y\|_Q^2 + \|\partial_t y\|_{L_2(0, T; H^{-1}(Q))}^2$, and $\|p\|_{X_T}$ defined analogously. We note that the inf-sup constant $1/(2\sqrt{2})$ from [37] can be replaced by the improved constant $1/\sqrt{2}$ (see [51, 54]). Now, adding

and subtracting arbitrary vector functions τ and σ from $H(\operatorname{div}_x, Q) := \{\tau \in [L_2(Q)]^d : \operatorname{div}_x(\tau) \in L_2(Q)\}$, regrouping the terms, integrating by parts, and using Cauchy's as well as Friedrichs's inequalities, we obtain the following estimates of the nominator on the right-hand side of (7.2):

$$\begin{aligned}
a((y, p) - (\bar{y}, \bar{p}), (v, q)) &= (-y_d, q) - a((\bar{y}, \bar{p}), (v, q)) \\
&= \int_Q [(-y_d)q - \varrho \partial_t \bar{y} v - \varrho \nabla_x \bar{y} \cdot \nabla_x v - \bar{p} v + \partial_t \bar{p} q - \nabla_x \bar{p} \cdot \nabla_x q + \bar{y} q] \, dQ \\
&= \int_Q (\varrho (-\partial_t \bar{y})v - \varrho \nabla_x \bar{y} \cdot \nabla_x v - \bar{p} v) \, dQ \pm \varrho \int_Q \tau \cdot \nabla_x v \, dQ \\
&\quad + \int_Q ((-y_d)q + \partial_t \bar{p} q - \nabla_x \bar{p} \cdot \nabla_x q + \bar{y} q) \, dQ \pm \int_Q \sigma \cdot \nabla_x q \, dQ \\
&= \varrho \int_Q \left((-\partial_t \bar{y} - \frac{1}{\varrho} \bar{p})v + \operatorname{div}_x \tau v \right) \, dQ + \varrho \int_Q (\tau - \nabla_x \bar{y}) \cdot \nabla_x v \, dQ \\
&\quad + \int_Q ((-y_d + \partial_t \bar{p} + \bar{y})q + \operatorname{div}_x \sigma q) \, dQ + \int_Q (\sigma - \nabla_x \bar{p}) \cdot \nabla_x q \, dQ \\
&\leq \varrho \|\partial_t \bar{y} - \frac{1}{\varrho} \bar{p} + \operatorname{div}_x \tau\|_Q \|v\|_Q + \varrho \|\tau - \nabla_x \bar{y}\|_Q \|\nabla_x v\|_Q \\
&\quad + \|\partial_t \bar{p} + \bar{y} + \operatorname{div}_x \sigma\|_Q \|q\|_Q + \|\sigma - \nabla_x \bar{p}\|_Q \|\nabla_x q\|_Q \\
&\leq \varrho \left[c_F \|\partial_t \bar{y} - \frac{1}{\varrho} \bar{p} + \operatorname{div}_x \tau\|_Q + \|\tau - \nabla_x \bar{y}\|_Q \right] \|\nabla_x v\|_Q \\
&\quad + [c_F \|\partial_t \bar{p} + \bar{y} + \operatorname{div}_x \sigma\|_Q + \|\sigma - \nabla_x \bar{p}\|_Q] \|\nabla_x q\|_Q \\
&\leq \left(\varrho [c_F \|\partial_t \bar{y} - \frac{1}{\varrho} \bar{p} + \operatorname{div}_x \tau\|_Q + \|\tau - \nabla_x \bar{y}\|_Q]^2 \right. \\
&\quad \left. + [c_F \|\partial_t \bar{p} + \bar{y} + \operatorname{div}_x \sigma\|_Q + \|\sigma - \nabla_x \bar{p}\|_Q]^2 \right)^{1/2} \\
&\quad \times \left(\varrho \|\nabla_x v\|_Q^2 + \|\nabla_x q\|_Q^2 \right)^{1/2}.
\end{aligned}$$

Combining this estimate with the inf-sup condition (7.2), we arrive at the following guaranteed upper bound

$$\sqrt{\varrho \|y - \bar{y}\|_{X_0}^2 + \|p - \bar{p}\|_{X_T}^2} \leq \mathfrak{M}_{\oplus}(\bar{y}, \bar{p}, \tau, \sigma) \quad (7.3)$$

that holds for an arbitrary pair $(\bar{y}, \bar{p}) \in Y_0 \cap H^1(Q) \times P_T \cap H^1(Q)$, where

$$\begin{aligned}
\mathfrak{M}_{\oplus}^2(v, q, \tau, \sigma) &:= 2\varrho [\|\tau - \nabla_x v\|_Q + c_F \|\partial_t v - \frac{1}{\varrho} q + \operatorname{div}_x \tau\|_Q]^2 \\
&\quad + 2[\|\sigma - \nabla_x q\|_Q + c_F \|\partial_t q + v + \operatorname{div}_x \sigma\|_Q]^2.
\end{aligned}$$

In order to estimate the error of our finite element solution, we replace (\bar{y}, \bar{p}) by the finite element solution (y_h, p_h) of (4.5). We still have to choose appropriate fluxes τ and σ from $H(\operatorname{div}_x, Q)$. Unfortunately, we cannot simply take the finite element fluxes $\nabla_x y_h$ and $\nabla_x p_h$ because they do not belong to $H(\operatorname{div}_x, Q)$. Instead, we postprocess the finite element fluxes to produce improved fluxes $\tau_h^{(0)}$ and $\sigma_h^{(0)}$ that belong to $H(\operatorname{div}_x, Q)$. We can then either use the postprocessed fluxes directly and insert them into the majorant \mathfrak{M}_{\oplus} , or we can try to improve the initial fluxes even further. In order to do so, we first estimate the squared majorant once more, and obtain a quadratic functional as an upper bound:

$$\begin{aligned}
\mathfrak{M}_{\oplus}^2(v, q, \tau, \sigma) &\leq 4 \left[\varrho \|\tau - \nabla_x v\|_Q^2 + c_F^2 \|\partial_t v - \frac{1}{\varrho} q + \operatorname{div}_x \tau\|_Q^2 \right. \\
&\quad \left. + \|\sigma - \nabla_x q\|_Q^2 + c_F^2 \|\partial_t q + v + \operatorname{div}_x \sigma\|_Q^2 \right] =: \mathfrak{M}_+^2.
\end{aligned}$$

Using $\tau_h^{(0)}$ and $\sigma_h^{(0)}$ as initial guess, we minimize the functional \mathfrak{M}_+^2 wrt. τ and σ using a few steps of the (preconditioned) CG method, respectively. Hence we obtain improved fluxes $\tau_h^{(1)}$ and $\sigma_h^{(1)}$, which we then

insert into the localized functional \mathfrak{M}_+ to obtain local error indicators, i.e.,

$$\begin{aligned} \eta_K^2(y_h, p_h) = & 2 \left[\varrho \|\tau_h^{(1)} - \nabla_x y_h\|_K^2 + c_F^2 \|\partial_t y_h - \frac{1}{\varrho} p_h + \operatorname{div}_x \tau_h^{(1)}\|_K^2 \right. \\ & \left. + \|\sigma_h^{(1)} - \nabla_x p_h\|_K^2 + c_F^2 \|\gamma_d + \partial_t p_h + y_h + \operatorname{div}_x \sigma_h^{(1)}\|_K^2 \right] \end{aligned}$$

that then drive the adaptive process.

Note that we cannot compute any efficiency indices with respect to (7.3), since the norm in the dual space $L(0, T; H^{-1}(\Omega))$ involved in the left-hand side, is not really computable. Since

$$\varrho \|\nabla_x(y - y_h)\|_Q^2 + \|\nabla_x(p - p_h)\|_Q^2 \leq \varrho \|y - y_h\|_{X_0}^2 + \|p - p_h\|_{X_T}^2$$

we have the estimates

$$\begin{aligned} \varrho \|\nabla_x(y - y_h)\|_Q^2 + \|\nabla_x(p - p_h)\|_Q^2 & \leq \mathfrak{M}_{\oplus}^2(y_h, q_h, \sigma_h^{(1)}, \sigma_h^{(1)}) \\ & \leq \mathfrak{M}_+^2(y_h, q_h, \sigma_h^{(1)}, \sigma_h^{(1)}) \end{aligned}$$

which are used for computing the efficiency indices in Section 9.

Once we have computed the indicators for each $K \in \mathcal{T}_h$, we determine a set \mathcal{M}_h of (almost) minimal cardinality such that

$$\sigma \sum_{K \in \mathcal{T}_h} \eta_K^2(y_h, p_h) \leq \sum_{K \in \mathcal{M}_h} \eta_K^2(y_h, p_h)$$

where $\sigma \in (0, 1)$ is an a priori chosen bulk parameter. This marking strategy is called Dörfler marking [15]. The elements of the set \mathcal{M}_h are then refined, where we might have to refine additional elements in order to maintain the conformity and shape-regularity of the mesh; see, e.g., [39] and [55] for a bisection strategy for simplices that works in any dimension. We use this bisection strategy for our numerical experiments in Section 9 as well.

8 Algebraic system and solvers

Let the finite element spaces $Y_{0h} = \operatorname{span}\{\varphi_j : j = 1, \dots, N_h\}$ and $P_{Th} = \{\psi_m : m = 1, \dots, M_h\}$ be spanned by the standard nodal finite element basis. Then each finite element function $y_h \in Y_{0h}$ and $p_h \in P_{Th}$ can be represented in the form

$$y_h = \sum_{j=1}^{N_h} y_j \varphi_j, \quad p_h = \sum_{m=1}^{M_h} p_m \psi_m. \quad (8.1)$$

Inserting (8.1) into (4.5), and testing with basis functions φ_i and ψ_n , we obtain one big linear system

$$\mathbf{K}_h \begin{pmatrix} \mathbf{y}_h \\ \mathbf{p}_h \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{f}_h \end{pmatrix} \quad (8.2)$$

for determining the unknown coefficient vectors $\mathbf{y}_h = (y_j)_{j=1, \dots, N_h} \in \mathbb{R}^{N_h}$ and $\mathbf{p}_h = (p_m)_{m=1, \dots, M_h} \in \mathbb{R}^{M_h}$ at once for the whole space time cylinder, where $\mathbf{f}_h = (\ell_h(0, \psi_n))_{n=1, \dots, M_h} \in \mathbb{R}^{M_h}$. The system matrix

$$\mathbf{K}_h = (a_h(\varphi_j, \psi_m; \varphi_i, \psi_n))_{i,j=1, \dots, N_h}^{m,n=1, \dots, M_h} = \begin{pmatrix} \mathbf{K}_{yy} & \mathbf{K}_{yp} \\ \mathbf{K}_{py} & \mathbf{K}_{pp} \end{pmatrix} \quad (8.3)$$

obviously has a 2×2 block structure with the blocks \mathbf{K}_{yy} , \mathbf{K}_{yp} , \mathbf{K}_{py} , and \mathbf{K}_{pp} defined by the identities

$$\begin{aligned} (\mathbf{K}_{yy} \mathbf{y}_h, \mathbf{v}_h) &= \varrho \int_Q (\partial_t y v + \vartheta \lambda_h^2 \partial_t y \partial_t v + v \nabla_x y \cdot \nabla_x v + \vartheta \lambda_h^2 L_x y \partial_t v) dQ \\ (\mathbf{K}_{yp} \mathbf{y}_h, \mathbf{q}_h) &= \int_Q y (q - \vartheta \lambda_h^2 \partial_t q) dQ \\ (\mathbf{K}_{py} \mathbf{p}_h, \mathbf{v}_h) &= \int_Q p (v + \vartheta \lambda_h^2 \partial_t v) dQ \\ (\mathbf{K}_{pp} \mathbf{p}_h, \mathbf{q}_h) &= \int_Q (-\partial_t p q + \vartheta \lambda_h^2 \partial_t p \partial_t q + v \nabla_x p \cdot \nabla_x q - \vartheta \lambda_h^2 L_x p \partial_t q) dQ \end{aligned}$$

for all coefficient vectors $\mathbf{y}_h, \mathbf{v}_h \in \mathbb{R}^{N_h}$ and $\mathbf{p}_h, \mathbf{q}_h \in \mathbb{R}^{M_h}$ corresponding to the finite element functions $y = y_h, v = v_h \in Y_{0h}$ and $p = p_h, q = q_h \in P_{Th}$ via the finite element isomorphism; cf. (8.1). Due to (5.5), the system matrix \mathbf{K}_h as well as the diagonal blocks \mathbf{K}_{yy} and \mathbf{K}_{pp} are non-symmetric, but positive definite.

Hence the linear system (8.2) is solved by the FGMRES method [47], preconditioned by a block-diagonal AMG preconditioner $\mathbf{C}_h = \text{blockdiag}(\mathbf{C}_{yy}, \mathbf{C}_{pp})$, where $\mathbf{C}_{yy} = \mathbf{K}_{yy}(\mathbf{I} - \mathbf{M}_{yy})^{-1}$ and $\mathbf{C}_{pp} = \mathbf{K}_{pp}(\mathbf{I} - \mathbf{M}_{pp})^{-1}$ with the AMG iteration matrices \mathbf{M}_{yy} and \mathbf{M}_{pp} resulting from one AMG V-cycles applied to the diagonal blocks of \mathbf{K}_h using zero initial guesses [20]. In the next section, we test the performance of this AMG preconditioned FGMRES as single-grid solver starting with initial guess zero and stopping after reducing the initial residual by the factor 10^{-8} , and in the nested iteration setting where the initial guess for the refinement level $\ell + 1$ is interpolated from the last iterate at the preceding refinement level ℓ . The nested iteration approach aims at stopping the iteration when the corresponding discretization error is reached at level $\ell + 1$ (see, e.g., [20]).

9 Numerical results

We implemented the space–time finite element scheme (4.5) using MFEM [42], a lightweight C++ library. The block-diagonal AMG preconditioner described in Section 8 was realized by means of BoomerAMG, an AMG implementation provided by the linear solver library hypre [26]. Both libraries allow for easy parallelization. We stop the AMG preconditioned FGMRES iterations once the initial residual has been reduced by a factor of 10^{-8} . In the nested iteration approach, we stop the iteration when the residual computed from the interpolated coarse grid approximation is reduced by the factor 10^{-2} for linear elements, and by 10^{-3} for quadratic and cubic elements.

The adaptive bulk parameter is always chosen as $\sigma = 0.25$. In the following, we present the numerical results for two benchmark problems, which were already used in [34, 37] as test problems, but for $d = 2$, i.e., $Q \subset \mathbb{R}^3$. Here we show the corresponding numerical results for the 4-dimensional space–time cylinder Q decomposed into shape-regular four-dimensional simplices (pentatopes).

9.1 Smooth problem with explicitly known solution

As first benchmark, we consider the space–time domain $Q = (0, 1)^4$, i.e., $d = 3$, a constant diffusion coefficient $\nu \equiv 1$, and the manufactured state, co-state, and control

$$\begin{aligned} y(x, t) &= \sum_{i=1}^d \sin(x_i \pi) (a t^2 + b t) \\ p(x, t) &= -\varrho \sum_{i=1}^d \sin(x_i \pi) (d \pi^2 a t^2 + (d \pi^2 b + 2 a)t + b) \\ u(x, t) &= \sum_{i=1}^d \sin(x_i \pi) (d \pi^2 a t^2 + (d \pi^2 b + 2 a)t + b) \end{aligned}$$

respectively, where $a = -(2\pi^2 + 1)/(2\pi^2 + 2)$ and $b = 1$. The desired state y_d is computed accordingly. This solution is highly smooth and has no local features. Hence, we expect to observe the optimal convergence rates predicted by Theorem 6.2. In the left plot of Fig. 1, we present the convergence history of uniform refinements as well as for adaptive refinements using the functional error estimator presented in Section 7, for different polynomial degrees k , and a fixed regularization parameter $\varrho = 0.01$. Indeed, we always obtain a rate of $\mathcal{O}(h^k)$. Note that the ‘bumps’ in the convergence rates are due to the bisection algorithm, where each uniform refinement only doubles the number of elements as opposed to the usual procedure of uniform refinements that subdivides each pentatope into 16 subpentatopes. In the right plot of Fig. 1, we present a comparison between the efficiency indices of the residual indicator, as well as the functional estimator. Here we observe that the efficiency index of the functional estimator is not influenced by the polynomial degree k , whereas the efficiency index of the residual indicator has only a mild dependence on the polynomial degree k . We already

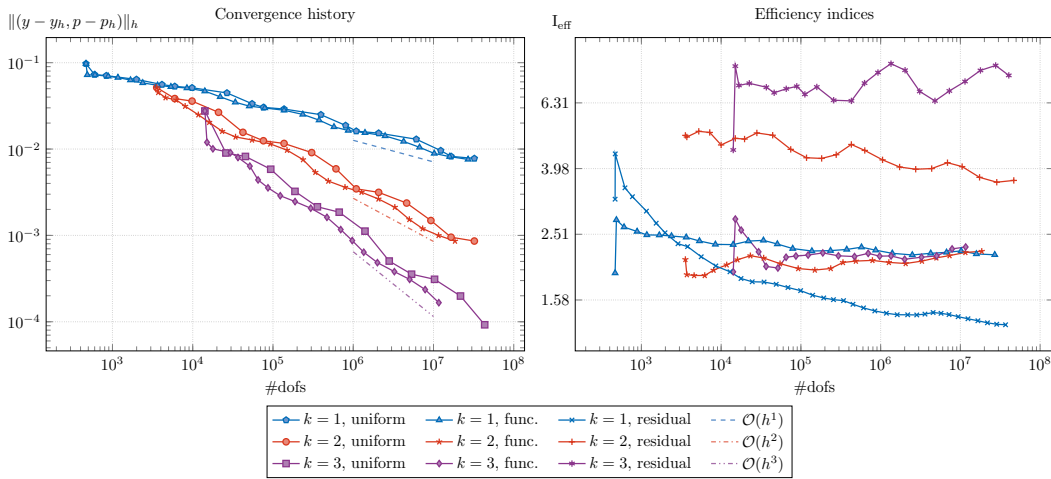


Fig. 1: Convergence rates in the mesh-dependent norm $\|(\cdot, \cdot)\|_h$ (left); Efficiency indices of the residual indicator and the functional estimator wrt. to the error $\sqrt{\varrho\|\nabla_x(y - y_h)\|_Q^2 + \|\nabla_x(p - p_h)\|_Q^2}$ (right); for different polynomial degrees k . Both plots share the same line styles.

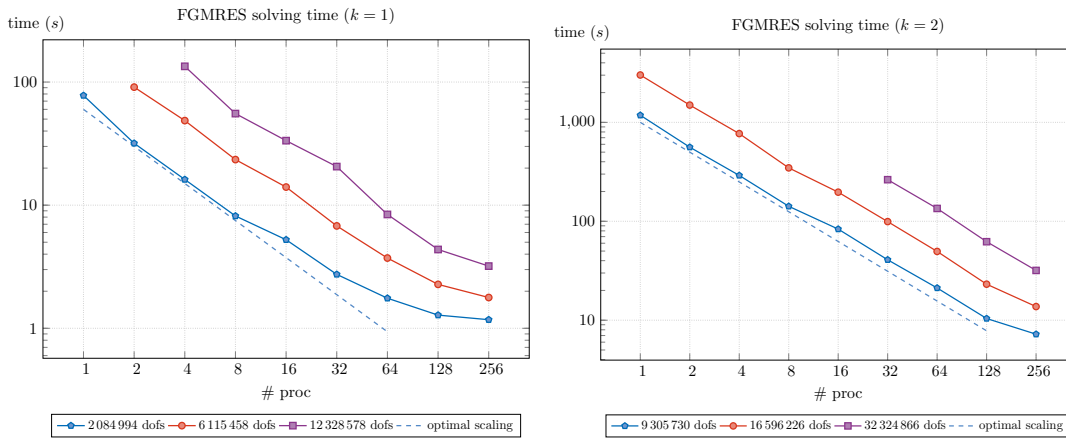


Fig. 2: Strong scaling results of the block-AMG preconditioned FGMRES for fixed problem sizes; using linear elements (left), and quadratic elements (right).

sketched the numerical procedure to obtain improved fluxes $\tau_h^{(1)}$ and $\sigma_h^{(1)}$ in Sect. 7. The postprocessing of the finite element fluxes $\nabla_x y_h$ and $\nabla_x q_h$ is done by means of a nodal averaging procedure, resulting in the initial guesses $\tau_h^{(0)}$ and $\sigma_h^{(0)}$ for the AMG preconditioned CG method. We stop the PCG solver either after the initial residual has been reduced by a factor of 10^{-2} , or after at most 10 CG iterations. The final iterates $\tau_h^{(1)}$ and $\sigma_h^{(1)}$ are then inserted in the localized functional \mathfrak{M}_+ . Additionally, we also performed strong scaling tests for our AMG-preconditioned FGMRES solver (cf. Fig. 2). Here, we plot the solving time (including the setup time of the AMG preconditioner) for three different problem sizes, using linear elements for the left plot, and quadratic elements for the right plot. We can observe that, for linear elements, we obtain almost optimal scaling up to 64 cores, after which the speed up stagnates. This stagnation is due to fact that the problem sizes for each processor are too small, such that the parallel overhead is now the dominating factor. For quadratic elements, the overall speed up is optimal until 128 cores. Then the parallel overhead again affects the speed-up. We note that the size of the problems fitting to a single core can vary greatly depending on the polynomial degree k . Indeed, the biggest problem that we could solve on a single core has 2 084 994 dofs for linear elements, compared to 16 596 226 dofs for quadratic elements.

Finally, we also tested the linear solver in a nested iterations settings. In Table 1, we present the solving times and number of iterations, comparing non-nested and nested iterations, for different polynomial

Tab. 1: Solving times and number of iterations for non-nested and nested iterations using 64 cores, where ℓ indicates the number of uniform refinements. The largest problem sizes are 32 278 018, 72 386 050, and 43 427 330 dofs, for the polynomial degrees $k = 1, 2$, and 3 , respectively.

ℓ	$k = 1$		$k = 2$		$k = 3$	
	non-nested	nested	non-nested	nested	non-nested	nested
0	11 (0.02 s)	13 (0.02 s)	21 (0.05 s)	26 (0.06 s)	44 (0.16 s)	55 (0.20 s)
4	13 (0.03 s)	3 (0.02 s)	28 (0.17 s)	10 (0.09 s)	47 (1.11 s)	17 (0.49 s)
8	16 (0.09 s)	4 (0.05 s)	37 (1.32 s)	14 (0.71 s)	62 (15.84 s)	20 (5.63 s)
12	21 (0.55 s)	4 (0.27 s)	49 (25.03 s)	13 (8.76 s)	83 (390.99 s)	20 (111.00 s)
15	30 (4.57 s)	4 (1.56 s)	70 (380.27 s)	10 (76.67 s)	—	—
16	28 (10.39 s)	4 (3.94 s)	—	—	—	—
18	40 (55.35 s)	2 (11.35 s)	—	—	—	—

degrees k . Here, we can deduce that the AMG preconditioner works very well in the nested iterations framework, reducing the number of iterations as well as the overall solving time drastically.

9.2 Discontinuous target function

In order to showcase the adaptive capabilities of a space–time scheme, we once more consider the four-dimensional space–time domain, $Q = (0, 1)^4$, and a constant diffusion coefficient $\nu \equiv 1$. However, we now want to approximate the discontinuous desired state

$$y_d(x, t) = \begin{cases} 1, & \sqrt{|x - 0.5|^2 + (t - 0.5)^2} \leq \frac{1}{4} \\ 0, & \text{otherwise} \end{cases}$$

that is nothing but a first expanding and then contracting ball, which is however fixed in space–time. Moreover, we use a smaller, fixed regularization parameter $\rho = 10^{-6}$. The discontinuity at the interface between the ball and its surrounding introduces steep gradients for the state and co-state, which require a rather fine local mesh-size in order to resolve the rapid changes properly. Hence, this problem is well suited for using adaptive refinements, driven by the residual error indicator introduced in Section 7. Indeed, the residual indicator leads to mesh refinements concentrated along the hypersurface of discontinuity (see Fig. 4), where we plot the finite element state y_h , co-state p_h , and control u_h on cuts through the space–time mesh \mathcal{T}_h . These cuts are nothing but unit cubes, which we then further cut by three planes along the three spatial coordinates centered at $(0.5, 0.5, 0.5)$. While we do not have explicit knowledge about the exact solutions, we can take a look at the convergence behavior of the sum of local indicators $\eta(y_h, p_h)^2 = \sum_{K \in \mathcal{T}_h} \eta_K(y_h, p_h)^2$ (cf. Fig. 3). Here we can indeed observe almost optimal rates of $\mathcal{O}(h^k)$ in terms of $h = \#\text{dofs}^{-1/(d+1)}$. However, in contrast to residual estimators for elliptic problems, where the sum of indicators is indeed proportional to the discretization error in the energy norm (see, e.g., [9]), no such result is available for residual indicators for parabolic problems on totally unstructured space–time decompositions.

10 Conclusions and outlook

We presented and analyzed new locally stabilized space–time finite element schemes on fully unstructured, but shape regular simplicial decompositions of the space–time cylinder Q for the numerical solution of parabolic optimal control problems. Such meshes typically arise from adaptive mesh refinement driven by local error indicators. The meshes are described by a continuous, elementwise differentiable mesh density function $\lambda_h(x, t)$ that also provides the right scaling of the time-upwind test functions. The resulting space–time finite element scheme is consistent, coercive, and bounded. The boundedness of the corresponding

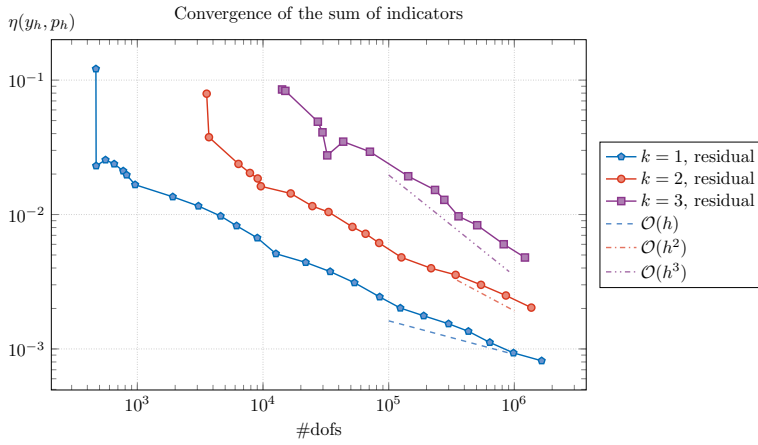


Fig. 3: Convergence plot of the sum of local error indicators $\eta^2(y_h, p_h) = \sum_{K \in \mathcal{T}_h} \eta_K^2(y_h, p_h)$ for different polynomial degrees k .

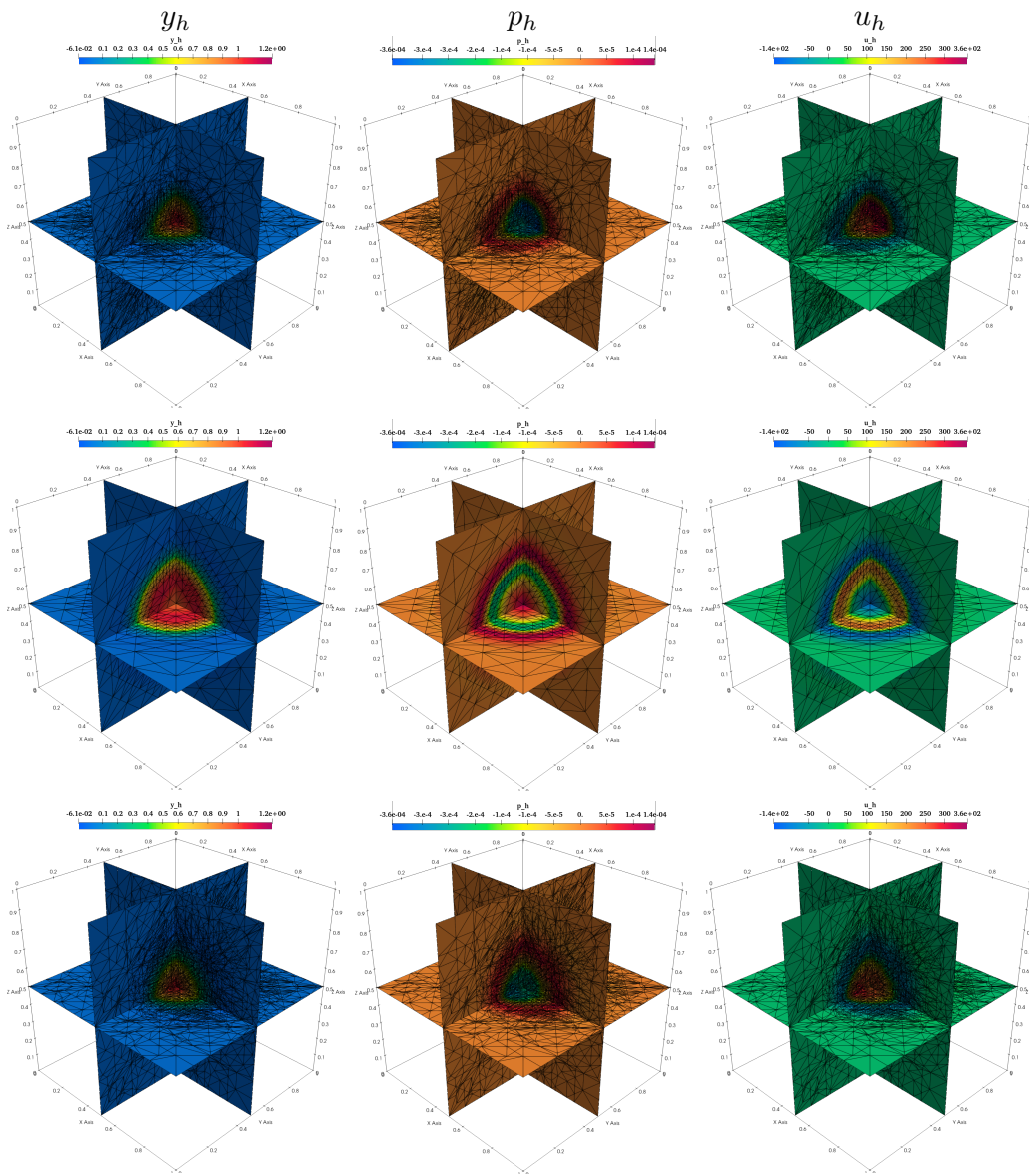


Fig. 4: Plots of the finite element functions y_h , p_h , and u_h over the mesh, cut at $t = 0.3$ (upper row); at $t = 0.5$ (middle row); and at $t = 0.7$ (bottom row).

bilinear form $a_h(\cdot; \cdot)$ holds for an extended space with respect to the first couple of functions. Beside the finite element functions, this space also contains the solution $(y, p) \in (Y_0 \cap H^{L,1}(Q)) \times (P_T \cap H^{L,1}(Q))$ in the maximal parabolic regularity setting. Coercivity, extended boundedness, and Galerkin orthogonality, which results from consistency, immediately yield a best-approximation estimate from which convergence rate estimates follow under additional regularity assumptions. The first adaptive version of our space–time finite element method is based on residual error indicators, which work well in our numerical experiment. However, there is no theory concerning reliability, efficiency, convergence, and optimality in sense of the paper [9]. Therefore, we proposed a second version that is based on a new functional error estimator providing a guaranteed upper bound of the error. However, the computation of the local error indicators are more costly than the residual ones. Finally, we have to solve one big system of finite element equations providing the finite element solution of the reduced optimality system all at once. We used a parallel version of an AMG preconditioned FGMRES that shows an excellent parallel performance in our numerical experiments. The construction and analysis of parallel solvers that are not only robust with respect to h and k , but also with respect to small regularization parameters ϱ is certainly a challenging task for future work. Simultaneous adaptivity and parallelization in space and time are big advantages of our really unstructured space–time finite element solver for the reduced optimality system that is forward and backward in time anyway. It is clear that this space–time technique can be applied to other optimal control problems and to other state equations including convection–diffusion problems and PDE system like the Navier–Stokes equations.

Funding: The authors would like to thank the Austrian Science Fund (FWF) for the financial support under the grant DK W1214-04.

Acknowledgment: The authors would like to express their thanks to the anonymous referees for their helpful comments.

References

- [1] R.A. Adams and J.J.F. Fournier, *Sobolev Spaces*, Academic Press & Elsevier, 2003.
- [2] T. Apel, A.-M. Sändig, and J. R. Whiteman, Graded mesh refinement and error estimates for finite element solutions of elliptic boundary value problems in non-smooth domains, *Math. Methods Appl. Sci.*, **19** (1996), No. 30, 63–85.
- [3] I. Babuška, Error-bounds for finite element method, *Numer. Math.*, **16** (1971), No. 4, 322–333.
- [4] I. Babuška and A.K. Aziz, Survey lectures on the mathematical foundation of the finite element method, In: *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, Academic Press, New York, 1972, pp. 1–359.
- [5] R. E. Bank, P. S. Vassilevski, and L. T. Zikatanov, Arbitrary dimension convection–diffusion schemes for space–time discretizations., *J. Comput. Appl. Math.*, **310** (2017), 19–31.
- [6] A. Borzi and V. Schulz, *Computational Optimization of Systems Governed By Partial Differential Equations*, Computational Science & Engineering 8, Society for Industrial and Applied Mathematics (SIAM), 2011.
- [7] D. Braess, *Finite Elements. Theory, Fast Solvers and Applications in Solid Mechanics*, Cambridge University Press, 2007.
- [8] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, third ed, Texts in Applied Mathematics, Vol. 15, Springer, New York, 2008.
- [9] C. Carstensen, M. Feischl, M. Page, and C. Praetorius, Axioms of adaptivity, *Comput. Methods Appl. Math.*, **67** (2014), 1195–1253.
- [10] E. Casas, C. Clason, and K. Kunisch, Parabolic control problems in measure spaces with sparse solutions, *SIAM J. Control Optim.*, **51** (2013), No. 1, 28–63.
- [11] E. Casas, M. Mateos, and A. Rösch, Improved approximation rates for a parabolic control problem with an objective promoting directional sparsity, *Comput. Optim. Appl.*, **70** (2018), 239–266.
- [12] P. Clément, Approximation by finite element functions using local regularization, *Rev. Française Automat. Informat. Recherche Opérationnelle Sér.*, **9** (1975), No. R-2, 77–84.
- [13] D. Devaud and C. Schwab, Space–time hp -approximation of parabolic equations, *Calcolo*, **55** (2018), No. 3, 1–23.
- [14] D. Dier, Non-autonomous maximal regularity for forms of bounded variation, *J. Math. Anal. Appl.*, **425** (2015), 33–54.
- [15] W. Dörfler, A convergent adaptive algorithm for Poisson’s equation, *SIAM J. Numer. Anal.*, **33** (1996), No. 3, 1106–1124.
- [16] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Springer, New York, 2004.

- [17] L.C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics 19, American Mathematical Society, Providence, Rhode Island, 2010.
- [18] S. Fackler, J.-L. Lions' problem concerning maximal regularity of equations governed by non-autonomous forms, *Ann. I. H. Poincaré*, **34** (2017), 699–709.
- [19] W. Gong, M. Hinze, and Z.J. Zhou, Space–time finite element approximation of parabolic optimal control problems, *J. Numer. Math.*, **20** (2012), No. 2, 111–145.
- [20] G. Haase and U. Langer, *Multigrid Methods: From Geometrical to Algebraic Versions*, Modern Methods in Scientific Computing and Applications (Eds. A. Bourlioux, M.J. Gander, and G. Sabidussi), Kluwer Academic Press, Dordrecht, 2002, pp. 103–154.
- [21] R. Herzog, G. Stadler, and G. Wachsmuth, Directional sparsity in optimal control of partial differential equations, *SIAM J. Control Optim.*, **50** (2012), No. 2, 943–963.
- [22] N. Heuer, On the equivalence of fractional-order Sobolev semi-norms, *J. Math. Anal. Appl.*, **417** (2014), No. 2, 505–518.
- [23] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, Mathematical Modelling: Theory and Applications, Vol. 23, Springer-Verlag, Berlin, 2008.
- [24] T.J. Hughes and A. Brooks, A multidimensional upwind scheme with no crosswind diffusion, In: *Finite Element Methods for Convection Dominated Flows* (Ed. T. J. Hughes), AMD 34, ASME, New York, 1979.
- [25] T. J. Hughes, L. P. Franca, and G. M. Hulbert, A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least-squares method for advection–diffusive equations, *Comput. Methods Appl. Mech. Engrg.*, **73** (1989), 173–189.
- [26] HYPRE: Scalable Linear Solvers and Multigrid Methods, <https://computing.llnl.gov/projects/hypre-scalable-linear-solvers-multigrid-methods>
- [27] C. Johnson and J. Saranen, Streamline diffusion methods for the incompressible Euler and Navier–Stokes equations, *Math. Comp.*, **47** (1986), No. 175, 1–18.
- [28] O. A. Ladyžhenskaya, *The Boundary Value Problems of Mathematical Physics*, Applied Mathematical Sciences, Vol. 49, Springer-Verlag, New York, 1985, Translated from the Russian edition, Nauka, Moscow, 1973.
- [29] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Uraltseva, *Linear and Quasilinear Equations of Parabolic Type*, Nauka, Moscow, 1967 (in Russian). Translated in Vol. 23, American Mathematical Soc., 1988.
- [30] U. Langer, S. Matculevich, and S. Repin, Guaranteed error bounds and local indicators for adaptive solvers using stabilised space–time IgA approximations to parabolic problems, *Comput. Math. with Appl.*, **78** (2019), 2641–2671.
- [31] U. Langer, S. Moore, and M. Neumüller, Space–time isogeometric analysis of parabolic evolution equations, *Comput. Methods Appl. Mech. Engrg.*, **306** (2016), 342–363.
- [32] U. Langer, M. Neumüller, and A. Schafelner, Space–time finite element methods for parabolic evolution problems with variable coefficients. In: *Advanced Finite Element Methods with Applications – Selected Papers from the 30th Chemnitz Finite Element Symposium 2017* (Eds. T. Apel, U. Langer, A. Meyer, and O. Steinbach), LNCSE 128, Springer, Berlin, Heidelberg, New York, 2019, pp. 229–256.
- [33] U. Langer and A. Schafelner, Adaptive space–time finite element methods for non-autonomous parabolic problems with distributional sources. *Comput. Methods Appl. Math.*, **20** (2020), No. 4, 677–693.
- [34] U. Langer and A. Schafelner, Simultaneous space–time finite element methods for parabolic optimal control problems, In: *Large-Scale Scientific Computing, 2021* (Eds. I. Lirkov and S. Margenov), LNCS, Vol. 13127, Springer, Cham, 2022, pp. 314–321.
- [35] U. Langer and A. Schafelner, Space–time hexahedral finite element methods for parabolic evolution problems, In: *Domain Decomposition Methods in Science and Engineering* (Eds. S. C. Brenner, E. T. S. Chung, A. Klawonn, P. Kwok, J. Xu, and J. Zou), LNCSE, Vol. 145, Springer, Cham, 2023.
- [36] U. Langer, O. Steinbach, F. Tröltzsch, and H. Yang, Space–time finite element discretization of parabolic optimal control problems with energy regularization, *SIAM J. Numer. Anal.*, **59** (2021), No. 2, 660–674.
- [37] U. Langer, O. Steinbach, F. Tröltzsch, and H. Yang, Unstructured space–time finite element methods for optimal control of parabolic equation, *SIAM J. Sci. Comput.*, **43** (2021), No. 2, A744–A771.
- [38] J. L. Lions, *Contrôle Optimal De Systèmes Gouvernés Par Des Équations Aux Dérivées Partielles*, Dunod Gauthier-Villars, Paris, 1968.
- [39] J. M. Maubach, Local bisection refinement for n -simplicial grids generated by reflection, *SIAM J. Sci. Comput.*, **16** (1995), No. 1, 210–227.
- [40] D. Meidner and B. Vexler, A priori error estimates for space–time finite element discretization of parabolic optimal control problems. Part I: Problems without control constraints, *SIAM J. Control Optim.*, **47** (2008), No. 3, 1150–1177.
- [41] D. Meidner and B. Vexler, A priori error estimates for space–time finite element discretization of parabolic optimal control problems. Part II: Problems with control constraints, *SIAM J. Control Optim.*, **47** (2008), No. 3, 1301–1329.
- [42] MFEM: Modular Finite Element Methods Library, <https://mfem.org>.
- [43] I. Neitzel, U. Prüfert, and T. Slawig, A smooth regularization of the projection formula for constrained parabolic optimal control problems, *Numer. Funct. Anal. Optim.*, **33** (2011), No. 12, 1283–1315.
- [44] J. Nečas, Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la variationnelle, *Ann. Scuola Norm. Sup. Pisa*, **16** (1962), No. 4, 305–326.

- [45] L.A. Oganjesjan and L.A. Ruchovetz, *Variational Difference Methods for the Solution of Elliptic Equations*, Izdatelstvo Akademi Nank Armjanskoj SSR, Erevan, 1979, (in Russian).
- [46] S. Repin, *A Posteriori Estimates for Partial Differential Equations*, de Gruyter, Berlin, 2008.
- [47] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
- [48] L.R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, *Math. Comput.*, **54** (1990), No. 190, 483–493.
- [49] G. Stadler, Elliptic optimal control problems with L^1 -control cost and applications for the placement of control devices, *Comput. Optim. Appl.*, **44** (2009), 159–181.
- [50] O. Steinbach, Space–time finite element methods for parabolic problems, *Comput. Methods Appl. Math.*, **15** (2015), No. 4, 551–566.
- [51] O. Steinbach, *Improved Error Estimates for a Space–Time Finite Element Method for Parabolic Problems*, Berichte aus dem Institut für Angewandte Mathematik, Bericht 2022/7, TU Graz, 2022.
- [52] O. Steinbach and H. Yang, 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** (2018), No. 3, e2143.
- [53] O. Steinbach and H. Yang, Space–time finite element methods for parabolic evolution equations: discretization, a posteriori error estimation, adaptivity and solution, In: *Space–Time Methods: Application to Partial Differential Equations* (Eds. O. Steinbach and U. Langer), de Gruyter, Berlin, 2019, pp. 207–248.
- [54] O. Steinbach and M. Zank, Coercive space–time finite element methods for initial boundary value problems, *Electron. Trans. Numer. Anal.*, **52** (2020), 154–194.
- [55] R. Stevenson, The completion of locally refined simplicial partitions created by bisection, *Math. Comp.*, **77** (2008), No. 261, 227–241.
- [56] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, Graduate Studies in Mathematics, Vol. 112, American Mathematical Society, Providence, Rhode Island, 2010.