Simultaneous space-time finite element methods for parabolic optimal control problems[*]

Ulrich Langer[1] and Andreas Schafelner[2]

[1]*Institute for Computational Mathematics*
Johannes Kepler University Linz
Altenbergerstr. 69, 4040 Linz, Austria
[2]*Doctoral Program "Computational Mathematics"*
Johannes Kepler University Linz
Altenbergerstr. 69, A-4040 Linz, Austria

Abstract

This work presents, analyzes and tests 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.

Keywords: Parabolic optimal control problems; L_2-regularization; Space-time finite element methods.

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 parameter $\varrho > 0$, find the state $y \in Y_0 = \{v \in L^2(0,T;H_0^1(\Omega)) : \partial_t v \in L^2(0,T;H^{-1}(\Omega)), v = 0 \text{ on } \Sigma_0\}$ and the control $u \in U = L_2(0,T;L_2(\Omega)) = L_2(Q)$ minimizing the cost functional

$$J(y,u) = \frac{1}{2} \int_Q |y - y_d|^2 \, \mathrm{d}Q + \frac{\varrho}{2}\|u\|_{L_2(Q)}^2 \tag{1}$$

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

$$\partial_t y - \mathrm{div}_x(\nu \nabla_x y) = u \text{ in } Q, \quad y = 0 \text{ on } \Sigma, \quad y = 0 \text{ on } \Sigma_0, \tag{2}$$

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 serves as control. The spatial domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, is supposed to be bounded and Lipschitz. We assumed that $0 < \nu_1 \leq \nu(x,t) \leq \nu_2$ for almost all $(x,t) \in Q$ with positive constants ν_1 and ν_2.

This standard setting was already investigated in the famous book by J.L. Lions [6]. Since the state equation (2) has a unique solution $y \in Y_0$, one can conclude the existence of a unique control $u \in U$ minimizing the quadratic cost functional $J(S(u), u)$, where S is the solution operator mapping $u \in U$ to the unique solution $y \in Y_0$ of (2); see, e.g., [6] and [9]. There is an huge number of publications devoted to the numerical solution of the optimal control problem (1)–(2) with the standard $L_2(Q)$ regularization; see, e.g., [9]. The overwhelming majority of the publications uses some time-stepping or discontinuous Galerkin method for the time discretization in combination with some space-discretization method like the finite element method; see, e.g., [9]. The unique solvability of the optimal control problem can also be established by showing that the optimality system has a unique solution. In [5], the Banach-Nečas-Babuška theorem was applied to the optimality system to show its well-posedness. Furthermore, the discrete inf-sup condition, which does not follow from the inf-sup condition in the infinite-dimensional setting, was established for continuous space-time finite element discretization on fully unstructured simplicial space-time meshes. The discrete inf-sup condition implies stability of the discretization and a priori discretization error estimates. Distributed controls u from the space $U = L_2(0, T; H^{-1}(\Omega))$ together with energy regularization were investigated in [4], where one can also find a comparison of the energy regularization with the $L_2(Q)$ and the sparse regularizations.

In this paper, we make use of the maximal parabolic regularity of the reduced optimality system in the case of the $L_2(Q)$ regularization and under additional assumptions imposed on the coefficient ν. Then we can derive a stabilized finite element discretization of the reduced optimality system in the same way as it was done for the state equation in our preceding papers [3] The properties of the finite element scheme lead to a priori discretization error estimates that are confirmed by the numerical experiments.

2 Space-Time Finite Element Discretization

Eliminating the control u from the optimality system by means of the gradient equation $p + \varrho u = 0$, we arrive at the reduced optimality system the weak form of which reads as follows: Find the state $y \in Y_0$ and the adjoint state $p \in P_T$ such that, for $v, q \in V = L_2(0, T; H_0^1(\Omega))$, it holds

$$
\begin{aligned}
\varrho \int_Q \left[\partial_t y\, v + \nu\, \nabla_x y \cdot \nabla_x v \right] dQ + \int_Q p\, v\, dQ &= 0, \\
-\int_Q y\, q\, dQ + \int_Q \left[-\partial_t p\, q + \nu\, \nabla_x p \cdot \nabla_x q \right] dQ &= -\int_Q y_d\, q\, dQ,
\end{aligned}
\tag{3}
$$

2

where $P_T := \{p \in L^2(0,T; H_0^1(\Omega)) : \partial_t p \in L^2(0,T; H^{-1}(\Omega)), p = 0 \text{ on } \Sigma_T\}$. The variational reduced optimality system (3) is well-posed; see [5, Theorem 3.3]. Moreover, we additionally assume that the coefficient $\nu(x,t)$ is of bounded variation in t for almost all $x \in \Omega$. Then $\partial_t u$ and $Lu := -\text{div}_x(\nu \nabla_x u)$ as well as $\partial_t p$ and $Lp := -\text{div}_x(\nu \nabla_x p)$ belong to $L_2(Q)$; see [2]. This property is called maximal parabolic regularity. In this case, the parabolic partial differential equations involved in the reduced optimality system (3) hold in $L_2(Q)$. Therefore, the solution of the reduced optimality system (3) is equivalent to the solution of the following system of coupled forward and backward systems 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

$$
\begin{aligned}
\varrho\Big[\partial_t y - \text{div}_x(\nu \nabla_x y)\Big] &= -p \quad \text{in } L_2(Q), \\
-\partial_t p\, q - \text{div}_x(\nu \nabla_x p) &= y - y_d \quad \text{in } L_2(Q)
\end{aligned}
\tag{4}
$$

hold, where $H^{L,1}(Q) = \{v \in H^1(Q) : Lv := -\text{div}_x(\nu \nabla_x v) \in L_2(Q)\}$. The coupled PDE optimality system (4) is now the starting point for the construction of the coercive finite element scheme.

Let $\mathcal{T}_h$ be a regular decomposition of the space-time cylinder Q into simplicial elements, i.e., $\overline{Q} = \bigcup_{K \in \mathcal{T}_h} \overline{K}$, and $K \cap K' = \emptyset$ for all K and K' from $\mathcal{T}_h$ with $K \neq K'$; see, e.g., [1] for more details. On the basis of the triangulation $\mathcal{T}_h$, we define the space-time finite element spaces

$$
\begin{aligned}
Y_{0h} &= \{y_h \in C(\overline{Q}) : y_h(x_K(\cdot)) \in \mathbb{P}_k(\hat{K}), \forall K \in \mathcal{T}_h,\, y_h = 0 \text{ on } \overline{\Sigma} \cap \overline{\Sigma}_0\}, \tag{5} \\
P_{Th} &= \{p_h \in C(\overline{Q}) : p_h(x_K(\cdot)) \in \mathbb{P}_k(\hat{K}), \forall K \in \mathcal{T}_h,\, p_h = 0 \text{ on } \overline{\Sigma} \cap \overline{\Sigma}_T\}, \tag{6}
\end{aligned}
$$

where $x_K(\cdot)$ denotes the map from the reference element $\hat{K}$ to the finite element $K \in \mathcal{T}_h$, and $\mathbb{P}_k(\hat{K})$ is the space of polynomials of the degree k on the reference element $\hat{K}$. For brevity of the presentation, we set ν to 1. The same derivation can be done for ν that fulfill the condition $\text{div}_x(\nu \nabla_x w_h)|_K \in L_2(K)$ for all w_h from Y_{0h} or P_{Th} and for all $K \in \mathcal{T}_h$ (i.e., piecewise smooth) in addition to the conditions imposed above. Multiplying the first PDE in (4) by $v_h + \lambda \partial_t v_h$ with $v_h \in Y_{0h}$, and the second one by $q_h - \lambda \partial_t q_h$ with $q_h \in P_{Th}$, integrating over K, integrating by parts in the elliptic parts where the scaling parameter λ does not appear, and summing over all $K \in \mathcal{T}_h$, we arrive at the variational consistency identity

$$
a_h(y,p;v_h,q_h) = \ell_h(v_h,q_h) \quad \forall (v_h,q_h) \in Y_{0h} \times P_{Th}, \tag{7}
$$

with the combined bilinear and linear forms

$$
\begin{aligned}
a_h(y,p;v,q) &= \sum_{K \in \mathcal{T}_h} \int_K \Big[\varrho\big(\partial_t y\, v + \lambda \partial_t y \partial_t v + \nabla_x y \cdot \nabla_x v - \lambda \Delta_x y\, \partial_t v\big) \\
&\quad + p(v + \lambda \partial_t v) - \partial_t p\, q + \lambda \partial_t p \partial_t q + \nabla_x p \cdot \nabla_x q \\
&\quad + \lambda \Delta_x p\, \partial_t q - u(q - \lambda \partial_t q) \Big]\, \mathrm{d}K \quad \text{and} \tag{8} \\
\ell_h(v,q) &= -\sum_{K \in \mathcal{T}_h} \int_K y_d(q - \lambda \partial_t q)\, \mathrm{d}K, \tag{9}
\end{aligned}
$$

3

respectively. Now, the corresponding consistent finite element scheme reads as follows: Find $(y_h, p_h) \in Y_{0h} \times P_{Th}$ such that

$$a_h(y_h, p_h; v_h, q_h) = \ell_h(v_h, q_h) \quad \forall (v_h, q_h) \in Y_{0h} \times P_{Th}. \tag{10}$$

Subtracting (10) from (7), we immediately get the Galerkin orthogonality relation

$$a_h(y - y_h, p - p_h; v_h, q_h) = 0 \quad \forall (v_h, q_h) \in Y_{0h} \times P_{Th}, \tag{11}$$

which is crucial for deriving discretization error estimates.

3 Discretization Error Estimates

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

$$\begin{aligned}
\|(v, q)\|_h^2 &= \varrho \, \|v\|_{h,T}^2 + \|q\|_{h,0}^2 = \varrho \left(\|v(\cdot, T)\|_{L_2(\Omega)}^2 + \|\nabla_x v\|_{L_2(Q)}^2 + \lambda \|\partial_t v\|_{L_2(Q)}^2 \right) \\
&\quad + \|q(\cdot, 0)\|_{L_2(\Omega)}^2 + \|\nabla_x q\|_{L_2(Q)}^2 + \lambda \|\partial_t q\|_{L_2(Q)}^2.
\end{aligned}$$

Indeed, for all $(v_h, q_h) \in Y_{0h} \times P_{Th}$, we get the estimate

$$\begin{aligned}
a_h(v_h, q_h; v_h, q_h) &= \sum_{K \in \mathcal{T}_h} \int_K \Big[\varrho \big(\partial_t v_h \, v_h + \lambda |\partial_t v_h|^2 + |\nabla_x v_h|^2 - \lambda \Delta_x v_h \, \partial_t v_h \big) \\
&\qquad + q_h(v_h + \lambda \partial_t v_h) - \partial_t q_h \, q_h + \lambda |\partial_t q_h|^2 + |\nabla_x q_h|^2 \\
&\qquad + \lambda \Delta_x q_h \, \partial_t q_h - v_h(q_h - \lambda \partial_t q_h) \Big] \, \mathrm{d}K \\
&\geq \mu_c \, \|(v_h, q_h)\|_h^2, \tag{12}
\end{aligned}$$

with $\mu_c = 1/2$ provided that $\lambda \leq c_{inv}^{-2} h^2$, where c_{inv} denotes the constant in the inverse inequality $\|\mathrm{div}_x(\nabla_x w_h)\|_{L_2(K)} \leq c_{inv} h^{-1} \|\nabla_x w_h\|_{L_2(K)}$ that holds for all $w_h \in Y_{0h}$ or $w_h \in P_{Th}$. For $k = 1$, the terms $\Delta_x v_h$ and $\Delta_x q_h$ are zero, and we do not need the inverse inequality, but λ should be also $O(h^2)$ in order to get an optimal convergence rate estimate. The coercivity of the bilinear form a_h immediately implies uniqueness and existence of the finite element solution $(y_h, p_h) \in Y_{0h} \times P_{Th}$ of (10). In order to prove discretization error estimates, we need the boundedness of the bilinear form

$$|a_h(y, p; v_h, q_h)| \leq \mu_b \|(y, p)\|_{h,*} \|(v_h, q_h)\|_h \quad \forall (v_h, q_h) \in Y_{0h} \times P_{Th}, \tag{13}$$

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

$$\begin{aligned}
\|(y, p)\|_{h,*}^2 &= \|(y, p)\|_h^2 + \varrho \sum_{K \in \mathcal{T}_h} \lambda \|\Delta_x y\|_{L_2(K)}^2 + [(\varrho + 1)\lambda^{-1} + \lambda] \|y\|_{L_2(Q)}^2 \\
&\quad + \sum_{K \in \mathcal{T}_h} \lambda \|\Delta_x p\|_{L_2(K)}^2 + [2\lambda^{-1} + \lambda] \|p\|_{L_2(Q)}^2
\end{aligned}$$

Indeed, using Cauchy's inequalities and the Friedrichs inequality $\|w\|_{L_2(Q)} \leq c_{F\Omega} \|\nabla_x w\|_{L_2(Q)}$ that holds for all $w \in Y_0$ or $w \in P_T$, we can easily prove (13)

4

with $\mu_b = (\max\{4, 1 + \lambda c_{F\Omega}^2, 3 + \varrho^{-1}, 1 + \lambda c_{F\Omega}^2 \varrho^{-1}\})^{1/2}$. Now, (11), (12), and (13) immediately lead to the following Céa-like estimate of the discretization error by some best-approximation error.

Theorem 1. *Let $y_d \in L_2(Q)$ be a given target, and let $\nu \in L_\infty(Q)$ fulfill the assumptions imposed above. Furthermore, we assume that the regularization (cost) parameter $\varrho \in \mathbb{R}_+$ is fixed. Then the Céa-like estimate*

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

holds, where (y, p) and (y_h, p_h) are the solutions of (3) and (10), respectively.

This Céa-like estimate immediately yields convergence rate estimates of the form

$$\|(y - y_h, p - p_h)\|_h \leq c(u, p) h^s \tag{14}$$

with $s = \min\{k, l\}$ provided that $y \in Y_0 \cap H^{L,1}(Q) \cap H^{l+1}(Q)$ and $p \in P_T \cap H^{L,1}(Q) \cap H^{l+1}(Q)$, where l is some positive real number defining the regularity of the solution; see [3] for corresponding convergence rate estimates for the state equation only.

4 Numerical Results

Let $\{\phi^{(j)} : j = 1, \ldots, N_h\}$ be a nodal finite element basis for Y_{0h}, and let $\{\psi^{(m)} : m = 1, \ldots, M_h\}$ be a nodal finite element basis for P_{Th}. Then we can express each finite element function $y_h \in Y_{0h}$ and $p_h \in P_{Th}$ via the finite element basis, i.e. $y_h = \sum_{j=1}^{N_h} y_j \phi^{(j)}$ and $p_h = \sum_{m=1}^{M_h} p_m \psi^{(m)}$, respectively. We insert this ansatz into (10), test with basis functions $\phi^{(i)}$ and $\psi^{(n)}$, and obtain the system

$$\mathbf{K}_h \begin{pmatrix} \mathbf{y}_h \\ \mathbf{p}_h \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{f}_h \end{pmatrix}$$

with $\mathbf{K}_h = (a_h(\phi^{(j)}, \psi^{(m)}; \phi^{(i)}, \psi^{(n)}))_{i,j=1,\ldots,N_h}^{m,n=1,\ldots,M_h}$, $\mathbf{f}_h = (\ell_h(0, \psi^{(n)}))_{n=1,\ldots,M_h}$, $\mathbf{y}_h = (y_j)_{j=1,\ldots,N_h}$ and $\mathbf{p}_h = (p_m)_{m=1,\ldots,M_h}$. The (block)-matrix $\mathbf{K}_h$ is non-symmetric, but positive definite due to (12). Hence the linear system is solved by means of the flexible General Minimal Residual (GMRES) method, preconditioned by a block-diagonal algebraic multigrid (AMG) method, i.e., we apply an AMG preconditioner to each of the diagonal blocks of $\mathbf{K}_h$. Note that we only need to solve once in order to obtain a numerical solution of the space-time tracking optimal control problem (1)–(2), consisting of state and adjoint state. The control can then be recovered from the gradient equation $p + \varrho u = 0$.

The space-time finite element method is implemented by means of the `C++` library MFEM [7]. We use *BoomerAMG*, provided by the linear solver library *hypre*, to realize the preconditioner. The linear solver is stopped once the initial residual is reduced by a factor of 10^{-8}. We are interested in convergence rates with respect to the mesh size h for a fixed regularization (cost) parameter ϱ.

4.1 Smooth Target

For our first example, we consider the space-time cylinder $Q = (0,1)^3$, i.e., $d = 2$, the manufactured state

$$y(x,t) = \sin(x_1\,\pi)\sin(x_2\,\pi)\left(a\,t^2 + b\,t\right),$$

as well as the corresponding adjoint state

$$p(x,t) = -\varrho\sin(x_1\,\pi)\sin(x_2\,\pi)\left(2\,\pi^2\,a\,t^2 + (2\,\pi^2\,b + 2\,a)t + b\right),$$

with $a = \frac{2\,\pi^2+1}{2\,\pi^2+2}$ and $b = 1$. The desired state y_d and the optimal control u are then computed accordingly, and we fix the regularization parameter $\varrho = 0.01$. This problem is very smooth and devoid of any local features or singularities, hence we expect optimal convergence rates. Indeed, as we can observe in Fig. 1, the error in the $\|(\cdot,\cdot)\|_{Y_0\times P_T}$-norm decreases with a rate of $\mathcal{O}(h^k)$, where k is the polynomial degree of the finite element basis functions.
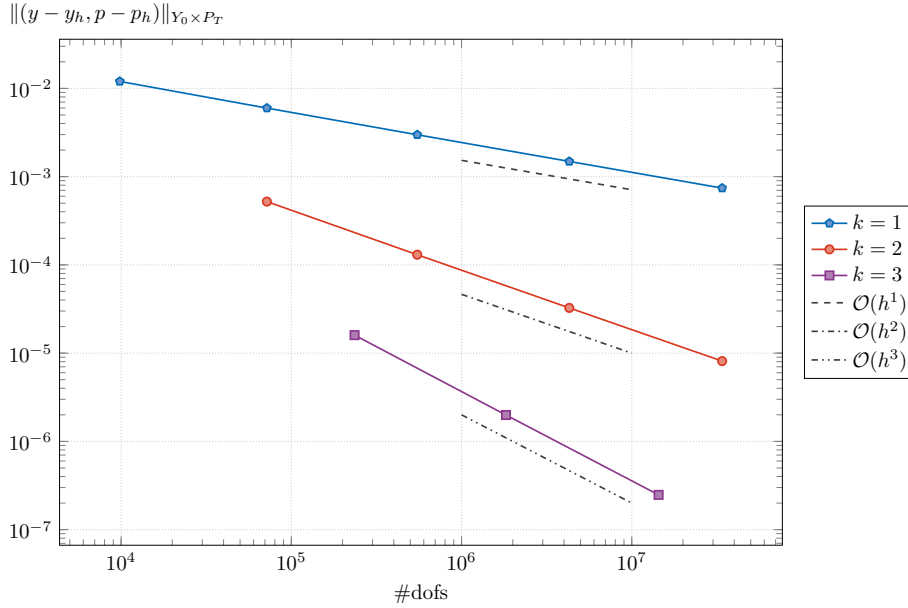

Fig. 1: Convergence rates for different polynomial degrees $k = 1,2,3$.

4.2 Discontinuous Target

For the second example, we consider once more the space-time cylinder $Q = (0,1)^3$, and specify the target state

$$y_d(x,t) = \begin{cases} 1, & \sqrt{(x_1 - 0.5)^2 + (x_2 - 0.5)^2 + (t - 0.5)^2} \leq 0.25, \\ 0, & \text{else,} \end{cases}$$

as an expanding and shrinking circle that is nothing but a *fixed* ball in the space-time cylinder Q. We use the fixed regularization parameter $\varrho = 10^{-6}$. Here, we do not know the exact solutions for the state or the optimal control, thus we cannot consider any convergence rates for the discretization error. However, the discontinuous target state may introduce local features at the (hyper-)surface of discontinuity. Hence it might be beneficial to use adaptive mesh refinements driven by an a posteriori error indicator. In particular, we use the residual based indicator proposed by Steinbach and Yang [8], applied to the residuals of the reduced optimality system (4). The final indicator is then the sum of the squares of both parts.

In Fig. 2, we present the finite element functions y_h, p_h, and u_h, plotted over cuts of the space-time mesh $\mathcal{T}_h$ at different times t. We can observe that the mesh refinements are mostly concentrated in annuli centered at $(0.5, 0.5)$, e.g. for $t = 0.5$, the outer and inner radii are $\sim \frac{7}{36} \pm \frac{1}{36}$, respectively; see Fig. 2 (middle row).

5   Conclusions

We proposed a stable, fully unstructured, space-time simplicial finite element discretization of the reduced optimality system of the standard space-time tracking parabolic optimal control problem with L_2-regularization. We derived a priori discretization error estimates. We presented numerical results for two benchmarks. We observed optimal rates for the example with smooth solutions as predicted by the a priori estimates. In the case of a discontinuous target, we use full space-time adaptivity. In order to get the full space-time solution (y_h, p_h, u_h), one has to solve only *one* system of algebraic equations. In this paper, we used flexible GMRES preconditioned by AMG.

References

[1] CIARLET, P. G. *The Finite Element Method for Elliptic Problems*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978.

[2] DIER, D. Non-autonomous maximal regularity for forms of bounded variation. *J. Math. Anal. Appl. 425* (2015), 33–54.

[3] LANGER, U., NEUMÜLLER, M., AND SCHAFELNER, A. 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*, T. Apel, U. Langer, A. Meyer, and O. Steinbach, Eds., vol. 128 of *Lecture Notes in Computational Science and Engineering (LNCSE)*. Springer, Berlin, Heidelberg, New York, 2019, ch. 13, pp. 247–275.
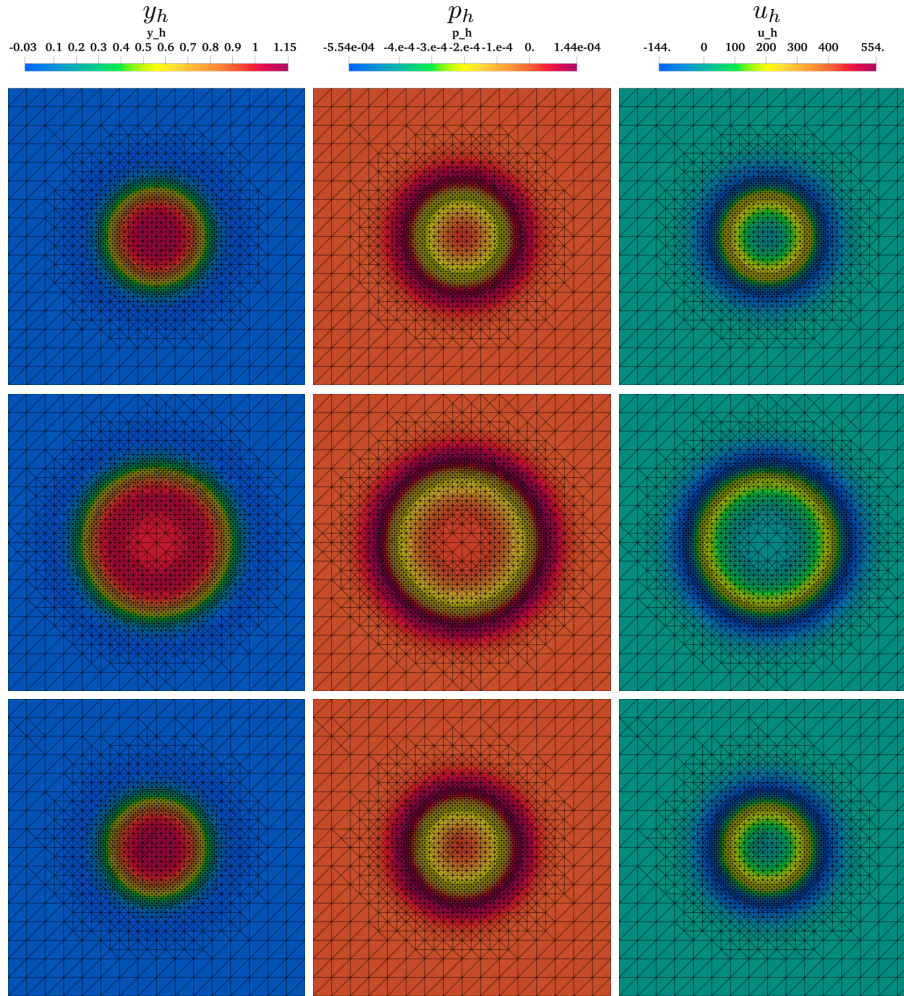
Fig. 2: Finite element solutions, with $J(y_h, u_h) = 3.5095 \times 10^{-3}$, plotted over the space-time mesh, obtained after 20 adaptive refinements, and cut at $t = 0.3125$ (upper row), $t = 0.5$ (middle row), and $t = 0.6875$ (lower row).

[4] LANGER, U., STEINBACH, O., TRÖLTZSCH, F., AND YANG, H. Space-time finite element discretization of parabolic optimal control problems with energy regularization. *SIAM J. Numer. Anal.* (2021). to appear.

[5] LANGER, U., STEINBACH, O., TRÖLTZSCH, F., AND YANG, H. Unstructured space-time finite element methods for optimal control of parabolic equations. *SIAM J. Sci. Comput.* (2021). to appear.

[6] LIONS, J. L. *Optimal control of systems governed by partial differential*

equations, vol. 170. Springer, Berlin, 1971.

[7] MFEM: Modular finite element methods library. mfem.org.

[8] STEINBACH, O., AND YANG, H. 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*, 3 (2018), e2143 nla.2143.

[9] TRÖLTZSCH, F. *Optimal control of partial differential equations: Theory, methods and applications*, vol. 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 2010.