Technische Universität Graz

Unstructured space-time finite element methods for optimal sparse control of parabolic equations

U. Langer, O. Steinbach, F. Tröltzsch, H. Yang

**Berichte aus dem
Institut für Angewandte Mathematik**

Technische Universität Graz

Unstructured space-time finite element methods for optimal sparse control of parabolic equations

U. Langer, O. Steinbach, F. Tröltzsch, H. Yang

Berichte aus dem Institut für Angewandte Mathematik

Technische Universität Graz
Institut für Angewandte Mathematik
Steyrergasse 30
A 8010 Graz

WWW: `http://www.applied.math.tugraz.at`

Unstructured Space-Time Finite Element Methods for Optimal Sparse Control of Parabolic Equations[*]

Ulrich Langer[†]   Olaf Steinbach[‡]   Fredi Tröltzsch[§]   Huidong Yang[¶]

March 31, 2020

Abstract

We consider a space-time finite element method on fully unstructured simplicial meshes for optimal sparse control of semilinear parabolic equations. The objective is a combination of a standard quadratic tracking-type functional including a Tikhonov regularization term and of the L^1-norm of the control that accounts for its spatio-temporal sparsity. We use a space-time Petrov-Galerkin finite element discretization for the first-order necessary optimality system of the associated discrete optimal sparse control problem. The discretization is based on a variational formulation that employs piecewise linear finite elements simultaneously in space and time. Finally, the discrete nonlinear optimality system that consists of coupled forward-backward state and adjoint state equations is solved by a semismooth Newton method.

Keywords: space-time finite element method, optimal sparse control, semilinear parabolic equations
MSC 2010: 49J20, 35K20, 65M60, 65M50, 65M15, 65Y05

1   Introduction

Optimal sparse control with the L^1-norm of the control in the objective functional and with linear elliptic state equations has been analyzed in [30] about

a decade ago. The method was extended to semilinear elliptic optimal control problems in [10] and to problems governed by elliptic equations with uncertain coefficients in [27]. In [15, 16, 29], the authors investigated problems of optimal sparse control for the Schlögl and FitzHugh-Nagumo systems, where traveling wave fronts or spiral waves were controlled. Directionally spatio-temporal optimal sparse control of linear/semilinear elliptic/parabolic equations and its corresponding optimality conditions was studied in [11, 20]. Optimality conditions for directionally sparse parabolic control problems without control constraints were considered in [13]; see also [14]. The optimal sparse controls considered therein exhibit sparsity in space, but not necessarily in time. We also mention another class of optimal sparse control problems for parabolic equations with controls in measure spaces instead of L^1-spaces, see, e.g., [5, 8, 12, 17, 23]. A thorough review of the existing literature on this challenging topic is beyond the scope of this work. Therefore, we refer to the recent survey article [7] on sparse solutions in optimal control of both elliptic and parabolic equations and the references therein.

Moreover, numerical approximations of optimal sparse controls of elliptic and parabolic problems were of great interest. For example, the standard 5-point stencil was used in [30] for the discretization of elliptic optimal sparse control problems. In [10], the authors proved rigorous error estimates for the finite element approximation of semilinear elliptic sparse control problems with with box constraints on the control. In the discrete optimality conditions, they used piecewise linear approximations for the state and adjoint state, while a piecewise constant ansatz was applied to the control and the subdifferential. We also mention the approximation of sparse controls by piecewise linear functions in [9]. Here, the authors adopt special quadrature formulae to discretize the squared L^2-norm and the L^1-norm of the control in the objective functional. This leads to an elementwise representation of the control and of the subdifferential.

Later, error estimates were derived for the space-time finite element approximation of parabolic optimal sparse control problems without control constraints in [13]. The discretization was performed on tensor-structured space-time meshes. In the associated discrete optimal control problem, these authors used a spatio-temporal finite element ansatz for the state that consists of products of continuous, piecewise linear basis functions in space, and piecewise constant basis functions in time. For the control, they utilize space-time elementwise constant basis functions. The same space as for the state was employed for the adjoint state in the discrete optimality system. Improved approximation rates were achieved in [14]. For the control discretization, they employ basis functions that are continuous and piecewise linear in space and piecewise constant in time. These methods can be re-interpreted as an implicit Euler discretization of the spatially discretized optimality system.

For optimal sparse control of the Schlögl and FitzHugh-Nagumo models that were considered in [15], a semi-implicit Euler-method in time and continuous piecewise linear finite elements in space were applied to both the state and adjoint state equations. Recently, in [38], for the optimal control of the convective FitzHugh-Nagumo equation, the state and adjoint state equations were

2

discretized by a symmetric interior penalty Galerkin method in space and the backward Euler method in time.

In contrast to the discretization methods discussed above, we apply continuous space-time finite element approximations on fully unstructured simplicial meshes for parabolic optimal sparse control problems with control constraints. This can be seen as an extension of the Petrov-Galerkin space-time finite element method proposed in [32] for parabolic problems, and in our recent work [25] for parabolic optimal control problems. This kind of unstructured space-time finite element approaches has gained increasing interest; see, e.g., [2, 3, 22, 24, 34, 36, 39], and the survey article [35].

In comparison to the more conventional time-stepping methods or tensor-structured space-time methods [18, 19, 28], this unstructured space-time approach provides us with more flexibility in constructing parallel space-time solvers such as parallel space-time algebraic multigrid preconditioners [24] or space-time balancing domain decomposition by constraints (BDDC) preconditioners [26]. Moreover, it becomes more convenient to realize simultaneous space-time adaptivity on unstructured space-time meshes [24, 25, 33] than the other methods. Here, time is just considered as another spatial coordinate. For more comparisons of our space-time finite element methods with others, we refer to [35].

The remainder of this paper will be structured as follows: Section 2 describes a model optimal sparse control problem that we aim to solve. Some preliminary existing results concerning optimality conditions are given in Section 3. The space-time finite element discretization of the associated discrete optimal control problem, the resulting discrete optimality conditions, as well as the application of the semismooth Newton iteration are discussed in Section 4. The applicability of our proposed method is confirmed by two numerical examples in Section 5. Finally, some conclusions are drawn in Section 6.

2 The optimal sparse control model problem

We consider the optimal sparse control problem

$$\min_{z \in Z_{ad}} \mathcal{J}(z) := \frac{1}{2} \|u_z - u_Q\|_{L^2(Q)}^2 + \frac{\varrho}{2} \|z\|_{L^2(Q)}^2 + \mu \|z\|_{L^1(Q)}, \qquad (1)$$

where the admissible set is

$$Z_{ad} = \{z \in L^\infty(Q) : a \le z(x,t) \le b \text{ for a.a. } (x,t) \in Q\}, \qquad (2)$$

and u_z is the unique solution of the state equation

$$
\begin{aligned}
\partial_t u - \Delta_x u + R(u) &= z && \text{in } Q := \Omega \times (0,T), \\
u &= 0 && \text{on } \Sigma := \partial\Omega \times (0,T), \\
u &= u_0 && \text{on } \Sigma_0 := \Omega \times \{0\}.
\end{aligned} \qquad (3)
$$

Here, the spatial computational domain $\Omega \subset \mathbb{R}^d$, $d \in \{1,2,3\}$, is supposed to be bounded and Lipschitz, $T > 0$ is the fixed terminal time, ∂_t denotes the partial

3

time derivative, $\Delta_x = \sum_{i=1}^d \partial_{x_i}^2$ the spatial Laplacian, and the source term z acts as a distributed control in Q. Moreover, $u_Q \in L^2(Q)$ is a given desired state. We further assume

$$-\infty < a < 0 < b < +\infty, \ \varrho > 0, \ \mu > 0.$$

The nonlinear reaction term R is defined by

$$R(u) = (u - u_1)(u - u_2)(u - u_3)$$

with given real numbers $u_1 \leq u_2 \leq u_3$. The functional $g : L^1(Q) \to \mathbb{R}$ defined by $g(\cdot) = \| \cdot \|_{L^1(Q)}$ is Lipschitz continuous and convex but not Fréchet differentiable. We notice that similar model problems for semilinear equations have been studied, e.g., in [7, 11].

3 Preliminary results

Let us recall some facts that are known from literature, e.g., [7]. For all $z \in L^p(Q)$, $p > d/2 + 1$, the state equation (3) has a unique solution $u_z \in W(0, T) \cap L^\infty(Q)$, where

$$W(0, T) = \left\{ v \in L^2(0, T; H_0^1(\Omega)) : \partial_t v \in L^2(0, T; H^{-1}(\Omega)) \right\}. \tag{4}$$

Here, $H_0^1(\Omega) := \{ v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega \}$. The mapping $z \mapsto u_z$ is continuously Fréchet differentiable in these spaces. The optimal control problem has at least one optimal control that is denoted by $\bar{z}$; the associated optimal state is denoted by $\bar{u}$.

If $\bar{z}$ is a locally optimal control of the model problem, then there exist a unique adjoint state $\bar{p} \in W(0, T)$ and some $\bar{\lambda} \in \partial g(\bar{u}) \subset L^\infty(Q)$ such that $(\bar{u}, \bar{p}, \bar{z}, \bar{\lambda})$ solves the optimality system

$$\partial_t u - \Delta_x u + R(u) = z \text{ in } Q, \ u = 0 \text{ on } \Sigma, \ u = u_0 \text{ on } \Sigma_0, \tag{5a}$$

$$-\partial_t p - \Delta_x p + R'(u)p = u - u_Q \text{ in } Q, \ p = 0 \text{ on } \Sigma, \ p = 0 \text{ on } \Sigma_T, \tag{5b}$$

$$\int_Q (p + \varrho z + \mu \lambda)(v - z) \, dx \, dt \geq 0 \text{ for all } v \in Z_{ad}, \tag{5c}$$

where $\Sigma_T := \Omega \times \{T\}$. A detailed discussion of this optimality system leads to the relations

$$\bar{z}(x, t) = 0 \Leftrightarrow |\bar{p}(x, t)| \leq \mu, \tag{6a}$$

$$\bar{z}(x, t) = \mathbf{Proj}_{[a,b]} \left(-\frac{1}{\varrho} \left(\bar{p}(x, t) + \mu \bar{\lambda}(x, t) \right) \right), \tag{6b}$$

$$\bar{\lambda}(x, t) = \mathbf{Proj}_{[-1,1]} \left(-\frac{1}{\mu} \bar{p}(x, t) \right), \tag{6c}$$

that hold for a.a. $(x, t) \in Q$; see, e.g., [15]. Here, the projection $\mathbf{Proj}_{[\alpha,\beta]} : \mathbb{R} \to [\alpha, \beta]$ is defined by $\mathbf{Proj}_{[\alpha,\beta]}(q) = \max\{\alpha, \min\{q, \beta\}\}$; see, e.g., [37]. The

subdifferential of the L^1-norm of the control is given as follows:

$$\bar{\lambda} \in \partial g(\bar{z}) \Leftrightarrow \begin{cases} \bar{\lambda} = 1 & \text{if } \bar{z}(x,t) > 0, \\ \bar{\lambda} \in [-1,1] & \text{if } \bar{z}(x,t) = 0, \\ \bar{\lambda} = -1 & \text{if } \bar{z}(x,t) < 0 \end{cases}$$

for a.a. $(x,t) \in Q$ and $\bar{z} \in L^\infty(Q)$. By these relations, we obtain the following form of an optimal control:

$$\bar{z} = \begin{cases} a & \text{on } \mathcal{A}_a := \{(x,t) \in Q : -\bar{p}(x,t) + \mu < \varrho a\}, \\ b & \text{on } \mathcal{A}_b := \{(x,t) \in Q : -\bar{p}(x,t) - \mu > \varrho b\}, \\ 0 & \text{on } \mathcal{A}_0 := \{(x,t) \in Q : |\bar{p}(x,t)| \le \mu\}, \\ -\frac{1}{\varrho}(\bar{p} - \mu) & \text{on } \mathcal{I}_- := \{(x,t) \in Q : \varrho a \le -\bar{p}(x,t) + \mu < 0\}, \\ -\frac{1}{\varrho}(\bar{p} + \mu) & \text{on } \mathcal{I}_+ := \{(x,t) \in Q : 0 < -\bar{p}(x,t) - \mu \le \varrho b\}. \end{cases} \tag{7}$$

The set $\mathcal{A}_0$ accounts for the sparsity of the control.

Eliminating the control from the optimality system and using the projection formulae above, we obtain the following system for the state and the adjoint state:

$$\partial_t u - \Delta_x u + R(u) = \mathbf{Proj}_{[a,b]}\left(-\frac{1}{\varrho}\left(p + \mu \mathbf{Proj}_{[-1,1]}\left(-\frac{1}{\mu}p\right)\right)\right) \text{ in } Q, \tag{8a}$$

$u = 0 \text{ on } \Sigma, \ u = u_0 \text{ on } \Sigma_0,$

$$-\partial_t p - \Delta_x p + R'(u)p = u - u_Q \text{ in } Q, \tag{8b}$$

$p = 0 \text{ on } \Sigma, \ p = 0 \text{ on } \Sigma_T.$

Let us define the Bochner spaces for the state and adjoint state variables as follows:

$X_0 := L^2(0,T; H_0^1(\Omega)) \cap H_{0,}^1(0,T; H^{-1}(\Omega)) = \{v \in W(0,T), v = 0 \text{ on } \Sigma_0\},$

$X_T := L^2(0,T; H_0^1(\Omega)) \cap H_{,0}^1(0,T; H^{-1}(\Omega)) = \{v \in W(0,T), v = 0 \text{ on } \Sigma_T\},$

$Y := L^2(0,T; H_0^1(\Omega)).$

Our space-time variational formulation for the coupled system (8) reads: Find $u \in X_0$ and $p \in X_T$ such that the variational equations

$$\int_Q \partial_t u \, v \, dx \, dt + \int_Q \nabla_x u \cdot \nabla_x v \, dx \, dt + \int_Q R(u) \, v \, dx \, dt$$

$$= \int_Q \mathbf{Proj}_{[a,b]}\left(-\frac{1}{\varrho}\left(p + \mu \mathbf{Proj}_{[-1,1]}\left(-\frac{1}{\mu}p\right)\right)\right) v \, dx \, dt \tag{9a}$$

$$-\int_Q u \, q \, dx \, dt - \int_Q \partial_t p \, q \, dx \, dt + \int_Q \nabla_x p \cdot \nabla_x q \, dx \, dt + \int_Q R'(u)p \, q \, dx \, dt$$

$$= -\int_Q u_Q \, q \, dx \, dt \tag{9b}$$

hold for all $v, q \in Y$. This system is solvable, because the optimal control problem has at least one solution. Due to the projection formula on the right hand side of (9a), we need special care to discretize the optimality system. In fact, following the discretization scheme proposed in [1], we go back to the variational inequality (5c) and use piecewise constant approximation for the control in order to derive first order necessary optimality conditions for the associated discrete optimal control problem. We will discuss this in the forthcoming Section.

4 Space-time finite element discretization

For the space-time finite element approximation of the optimal control problem (1), we consider an admissible triangulation $\mathcal{T}_h(Q)$ of the space-time domain Q into shape regular simplicial finite elements τ. Here, the mesh size h is defined by $h = \max_{\tau \in \mathcal{T}_h} h_\tau$ with h_τ being the diameter of the element τ; see, e.g., [6, 31]. For simplicity, we assume Ω to be a polygonal spatial domain. Therefore, the triangulation exactly covers $Q = \Omega \times (0, T)$.

Let $S_h^1(Q)$ be the space of continuous and piecewise linear functions that are defined with respect to the triangulation $\mathcal{T}_h(Q)$. The discrete variational form of the state equation (3) reads as follows: Find $u_h \in X_{0,h} = S_h^1(Q) \cap X_0$ such that

$$\int_Q \partial_t u_h \, v_h \, dx \, dt + \int_Q \nabla_x u_h \cdot \nabla_x v_h \, dx \, dt + \int_Q R(u_h) \, v_h \, dx \, dt = \int_Q z \, v_h \, dx \, dt \quad (10)$$

is satisfied for all $v_h \in X_{0,h}$. For approximating the control z, we define the space

$$Z_h = \Big\{ z_h \in L^\infty(Q) : z_h \text{ is constant in each } \tau \in \mathcal{T}_h \Big\}.$$

An element $z_h \in Z_h$ can be represented in the form

$$z_h = \sum_{\tau \in \mathcal{T}_h} z_\tau \mathcal{X}_\tau,$$

with $\mathcal{X}_\tau$ being the characteristic function of τ. Moreover, the set of discrete admissible controls is defined by

$$Z_{ad,h} = \Big\{ z_h \in Z_h : a \le z_h|_\tau \le b \text{ for all } \tau \in \mathcal{T}_h \Big\}.$$

Now, we consider the discrete optimal control problem

$$\min_{z_h \in Z_{ad,h}} \mathcal{J}_h(z_h) := \frac{1}{2} \|u_{z_h} - u_Q\|_{L^2(Q)}^2 + \frac{\varrho}{2} \|z_h\|_{L^2(Q)}^2 + \mu \|z_h\|_{L^1(Q)},$$

where u_{z_h} denotes the solution of the discrete variational problem (10) with the discrete control z_h. We assume that the discrete optimal control problem has at least one locally optimal control that is denoted by $\bar{z}_h$. The associated state is denoted by $\bar{u}_h$. If $\bar{z}_h = \sum_{\tau \in \mathcal{X}_h} \bar{z}_\tau \mathcal{X}_\tau$ is a locally optimal control of

6

the discrete optimal control problem, then there exists a unique adjoint state $\bar{p}_h \in X_{T,h} = S_h^1(Q) \cap X_T$ and $\bar{\lambda}_h \in \partial g(\bar{z}_h)$ such that $(\bar{u}_h, \bar{p}_h, \bar{z}_h, \bar{\lambda}_h)$ solves the discrete optimality system

$$\int_Q \partial_t u_h \, v_h \, dx \, dt + \int_Q \nabla_x u_h \cdot \nabla_x v_h \, dx \, dt + \int_Q R(u_h) \, v_h \, dx \, dt$$

$$= \int_Q \bar{z}_h \, v_h \, dx \, dt, \text{ for all } v_h \in X_{0,h}, \qquad (11\text{a})$$

$$-\int_Q u_h \, q_h \, dx \, dt - \int_Q \partial_t p_h \, q_h \, dx \, dt + \int_Q \nabla_x p_h \cdot \nabla_x q_h \, dx \, dt$$

$$+ \int_Q R'(u_h) p_h \, q_h \, dx \, dt = -\int_Q u_Q \, q_h \, dx \, dt, \text{ for all } q_h \in X_{T,h}, \qquad (11\text{b})$$

$$\int_Q \Big(p_h + \varrho z_h + \mu \lambda_h \Big) \Big(v_h - z_h \Big) \, dx \, dt \geq 0, \text{ for all } v_h \in Z_{ad,h}. \qquad (11\text{c})$$

The existence of a solution to the discretized optimality system will not be discussed in this paper. We tacitly assume that a locally unique solution exists. Note that $\bar{\lambda}_h \in \partial g(\bar{z}_h)$ is equivalent to the form

$$\bar{\lambda}_h = \sum_{\tau \in \mathcal{T}_h} \bar{\lambda}_\tau \mathcal{X}_\tau \quad \text{with} \quad \begin{cases} \bar{\lambda}_\tau = 1 & \text{if } \bar{z}_\tau > 0, \\ \bar{\lambda}_\tau \in [-1, 1] & \text{if } \bar{z}_\tau = 0, \\ \bar{\lambda}_\tau = -1 & \text{if } \bar{z}_\tau < 0. \end{cases}$$

Then, the inequality (11c) can be represented as follows:

$$\sum_{\tau \in \mathcal{T}_h} \left(\int_\tau \bar{p}_h \, dx \, dt + |\tau| \Big(\varrho \bar{z}_\tau + \mu \bar{\lambda}_\tau \Big) \right) \Big(\bar{v}_\tau - \bar{z}_\tau \Big) \geq 0 \text{ for all } a \leq \bar{v}_\tau \leq b,$$

which is recast in the equivalent form

$$\left(\int_\tau \bar{p}_h \, dx \, dt + |\tau| \Big(\varrho \bar{z}_\tau + \mu \bar{\lambda}_\tau \Big) \right) \Big(\bar{v}_\tau - \bar{z}_\tau \Big) \geq 0, \ a \leq \bar{v}_\tau \leq b,$$

for all $\tau \in \mathcal{T}_h$. Therefore, we have the projection representation formula [37]

$$\bar{z}_\tau = \mathbf{Proj}_{[a,b]} \left(-\frac{1}{\varrho} \left(\frac{1}{|\tau|} \int_\tau \bar{p}_h \, dx \, dt + \mu \bar{\lambda}_\tau \right) \right) \qquad (12)$$

for the optimal control on each element $\tau \in \mathcal{T}_h$. From this, we have the following results:

$$\bar{z}_\tau = 0 \ \Leftrightarrow \ \frac{1}{|\tau|} \left| \int_\tau \bar{p}_h \, dx \, dt \right| \leq \mu,$$

$$\bar{\lambda}_\tau = \mathbf{Proj}_{[-1,1]} \left(-\frac{1}{\mu |\tau|} \int_\tau \bar{p}_h \, dx \, dt \right).$$

The above results are obtained by the discretization approach proposed in [1, 10] for the optimal control of semilinear elliptic equations, and in [15] of the

Schlögl and FitzHugh-Nagumo systems. In fact, by a close look to the projection formula, we have the following form of the discrete optimal control:

$$
\bar{z}_\tau = \begin{cases}
a & \text{on } \mathcal{A}_{a,\mathcal{T}_h} := \{\tau \in \mathcal{T}_h : -\bar{p}_\tau + \mu < \varrho a\}, \\
b & \text{on } \mathcal{A}_{b,\mathcal{T}_h} := \{\tau \in \mathcal{T}_h : -\bar{p}_\tau - \mu > \varrho b\}, \\
0 & \text{on } \mathcal{A}_{0,\mathcal{T}_h} := \{\tau \in \mathcal{T}_h : |\bar{p}_\tau| \le \mu\}, \\
-\frac{1}{\varrho}(\bar{p}_\tau - \mu) & \text{on } \mathcal{I}_{-,\mathcal{T}_h} := \{\tau \in \mathcal{T}_h : \varrho a \le -\bar{p}_\tau + \mu < 0\}, \\
-\frac{1}{\varrho}(\bar{p}_\tau + \mu) & \text{on } \mathcal{I}_{+,\mathcal{T}_h} := \{\tau \in \mathcal{T}_h : 0 < -\bar{p}_\tau - \mu \le \varrho b\},
\end{cases} \tag{13}
$$

where

$$
\bar{p}_\tau = \frac{1}{|\tau|} \int_\tau \bar{p}_h \, dx \, dt.
$$

Inserting (12) into (11a), we obtain an equivalent form of the discrete optimality system that consists of the state and adjoint state equations. Namely, find $\bar{u}_h \in X_{0,h}$ and $\bar{p}_h \in X_{T,h}$ such that $(\bar{u}_h, \bar{p}_h)$ solves the coupled system

$$
\int_Q \partial_t u_h \, v_h \, dx \, dt + \int_Q \nabla_x u_h \cdot \nabla_x v_h \, dx \, dt + \int_Q R(u_h) \, v_h \, dx \, dt
$$
$$
- \sum_{\tau \in \mathcal{T}_h} \int_\tau \mathbf{Proj}_{[a,b]}\left(-\frac{1}{\varrho}\left(\bar{p}_\tau + \mu \, \mathbf{Proj}_{[-1,1]}\left(-\frac{1}{\mu}\bar{p}_\tau\right)\right)\right) v_h \, dx \, dt
$$
$$
= 0, \text{ for all } v_h \in X_{0,h}, \tag{14a}
$$
$$
- \int_Q u_h \, q_h \, dx \, dt - \int_Q \partial_t p_h \, q_h \, dx \, dt + \int_Q \nabla_x p_h \cdot \nabla_x q_h \, dx \, dt
$$
$$
+ \int_Q R'(u_h) p_h \, q_h \, dx \, dt = - \int_Q u_Q \, q_h \, dx \, dt, \text{ for all } q_h \in X_{T,h}. \tag{14b}
$$

The convergence of the solution of the discrete optimal control problem to the solution of its associated continuous optimal control problem as well as the error analysis of our finite element approximation are beyond the scope of this work and will be studied elsewhere.

To solve the above discrete coupled nonlinear optimality system, we apply the semismooth Newton method as discussed in [30], where a generalized derivative needs to be computed at each Newton iteration. In fact, each iteration turns out to be one step of a primal-dual active set strategy [21]: Given (u_h^k, p_h^k), find

$(\delta u_h, \delta p_h)$ such that

$$\int_Q \partial_t \delta u_h\, v_h\, dx\, dt + \int_Q \nabla_x \delta u_h \cdot \nabla_x v_h\, dx\, dt + \int_Q R'(u_h^k)\, \delta u_h\, v_h\, dx\, dt$$

$$+ \frac{1}{\varrho}\left(\mathcal{X}_{\mathcal{I}_-,\mathcal{T}_h} + \mathcal{X}_{\mathcal{I}_+,\mathcal{T}_h}\right)\left(\sum_{\tau \in \mathcal{T}_h} \int_\tau \left(\frac{1}{|\tau|}\int_\tau \delta p_h\, dx\, dt\right) v_h\, dx\, dt\right)$$

$$= -\int_Q \partial_t u_h^k\, v_h\, dx\, dt - \int_Q \nabla_x u_h^k \cdot \nabla_x v_h\, dx\, dt + \int_Q R(u_h^k)\, v_h\, dx\, dt$$

$$+ \mathcal{X}_{\mathcal{A}_a,\mathcal{T}_h}\left(\sum_{\tau \in \mathcal{T}_h} \int_\tau a\, v_h\, dx\, dt\right) + \mathcal{X}_{\mathcal{A}_b,\mathcal{T}_h}\left(\sum_{\tau \in \mathcal{T}_h} \int_\tau b\, v_h\, dx\, dt\right)$$

$$- \frac{1}{\varrho}\mathcal{X}_{\mathcal{I}_-,\mathcal{T}_h}\left(\sum_{\tau \in \mathcal{T}_h} \int_\tau \left(\frac{1}{|\tau|}\int_\tau p_h^k\, dx\, dt - \mu\right) v_h\, dx\, dt\right)$$

$$- \frac{1}{\varrho}\mathcal{X}_{\mathcal{I}_+,\mathcal{T}_h}\left(\sum_{\tau \in \mathcal{T}_h} \int_\tau \left(\frac{1}{|\tau|}\int_\tau p_h^k\, dx\, dt + \mu\right) v_h\, dx\, dt\right)$$

and

$$-\int_Q \delta u_h\, q_h\, dx\, dt - \int_Q \partial_t \delta p_h\, q_h\, dx\, dt + \int_Q \nabla_x \delta p_h \cdot \nabla_x q_h\, dx\, dt$$

$$+ \int_Q R'(u_h^k)\delta p_h\, q_h\, dx\, dt + \int_Q R''(u_h^k)p_h^k \delta u_h\, q_h\, dx\, dt$$

$$= \int_Q u_h^k\, q_h\, dx\, dt + \int_Q \partial_t p_h^k\, q_h\, dx\, dt - \int_Q \nabla_x p_h^k \cdot \nabla_x q_h\, dx\, dt$$

$$- \int_Q R'(u_h^k)p_h^k q_h\, dx\, dt - \int_Q u_Q\, q_h\, dx\, dt$$

are fulfilled, and $u_h^{k+1} = u_h^k + \omega\delta u_h$, $p_h^{k+1} = p_h^k + \omega\delta p_h$ with some damping parameter $\omega \in (0,1]$.

5 Numerical experiments

For the two numerical examples considered in this section, we set $\Omega = (0,1)^2$, $T = 1$, and therefore $Q = (0,1)^3$. Using an octasection-based refinement [4], we uniformly decompose the space-time cylinder Q until the mesh size reaches $h = 1/128$. Therefore, the total number of degrees of freedom for the coupled state and adjoint state equations is $4,194,304$. We will also use an adaptive refinement procedure that is driven by a residual based error indicator for the coupled state and adjoint state system similar to that one that was developed for the state equation in [33]. We perform our numerical tests on a desktop with Intel@ Xeon@ Processor E5-1650 v4 (15 MB Cache, 3.60 GHz), and 64 GB memory. For the nonlinear first order necessary optimality system, we use the
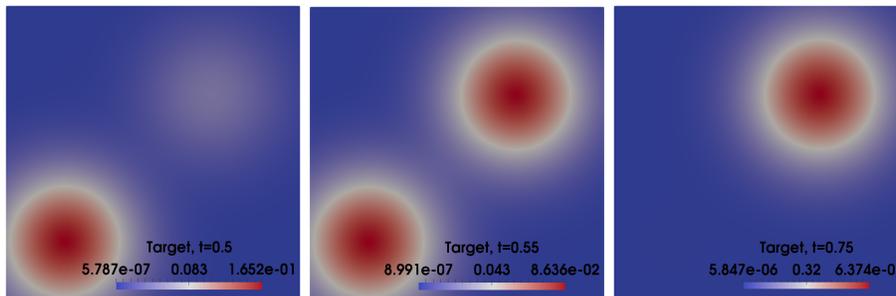
Figure 1: Example 1, plots of the target at time $t = 0.5$, 0.55, 0.75 for the moving target example.

relative residual error 10^{-5} as a stopping criterion in the semismooth Newton iteration, whereas the algebraic multigrid preconditioned GMRES solver for the linearized system at each Newton iteration is stopped after a residual error reduction by 10^{-6}; cf. also [35].

5.1 Moving target (Example 1)

In the first example, the desired state is given by the function

$$u_Q(x, t) = \exp\left(-20(x_1 - 0.2)^2 + (x_2 - 0.2)^2 + (t - 0.2)^2\right) + \exp\left(-20(x_1 - 0.7)^2 + (x_2 - 0.7)^2 + (t - 0.9)^2\right),$$

which is adapted from an example constructed in [11]; see an illustration of the target at time $t = 0.5$, 0.55, and 0.75 in Fig. 1. The same desired state was also used in the numerical test for spatially directional sparse control in [13, 14]. The parameters in the optimal control problem are $\varrho = 10^{-4}$, $\mu = 0.004$, $a = -10$, and $b = 20$. For the nonlinear reaction term in the state equation, we set $R(u) = u(u - 0.25)(u + 1)$. Homogeneous initial and Dirichlet boundary conditions are used for the state equation.

To reach the relative residual error 10^{-5} for the nonlinear first order necessary optimality system, we needed 21 and 37 semismooth Newton iterations for the nonsparse and sparse optimal control, respectively; see Fig.2. We clearly see a superlinear convergence of the semismooth Newton method. The total system assembling and solving time is about 4 and 4.7 hours on the desktop computer, respectively.

Comparisons of sparse and nonsparse controls as well as of associated states at different times are displayed in Fig. 3 and Fig. 4, respectively. A close look to sparse and nonsparse controls along different lines in the space-time domain is illustrated in Fig. 5. In this example, we clearly see that the L^1 cost functional promotes spatial and temporal sparsity, with some precision loss of the associated state to the target, cf. Fig. 4.
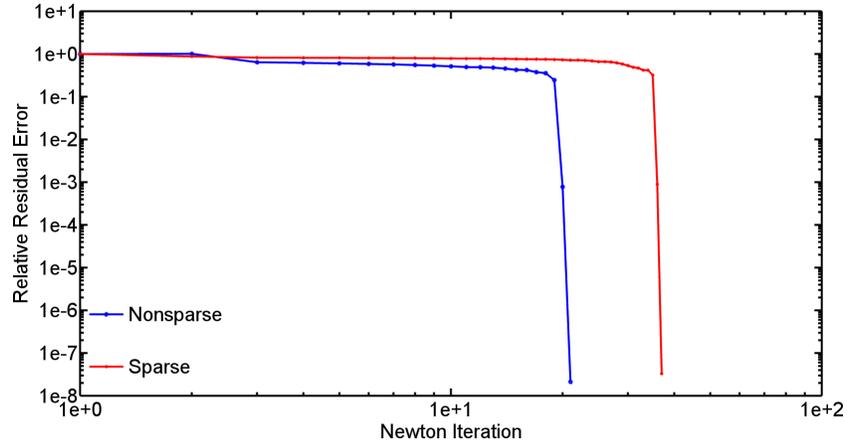
Figure 2: Example 1, relative residual error reduction in the semismooth Newton method for the nonsparse and sparse control in the moving target example.
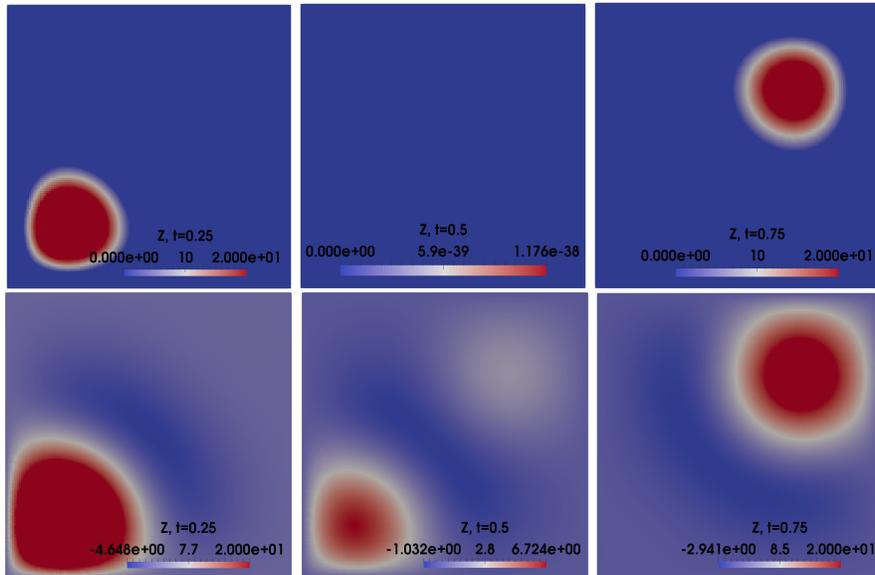
Figure 3: Example 1, Comparisons of sparse (up) and nonsparse controls (down) at time $t = 0.25,\ 0.5,\ 0.75$ for the moving target example.

11
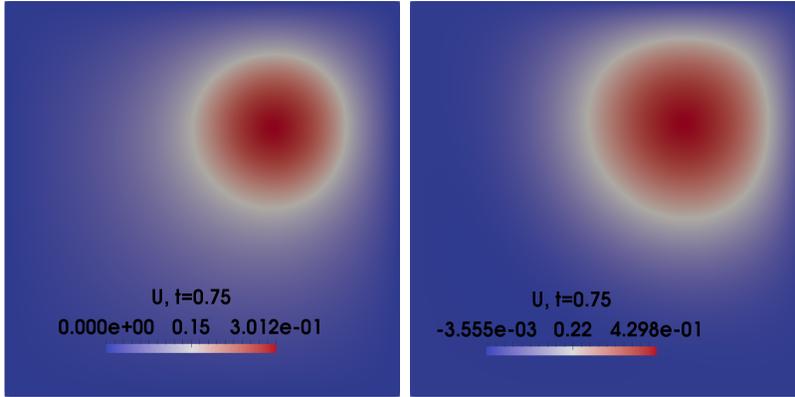
Figure 4: Example 1, Comparisons of states associated to sparse (left) and nonsparse controls (right) at time $t = 0.75$ for the moving target example.
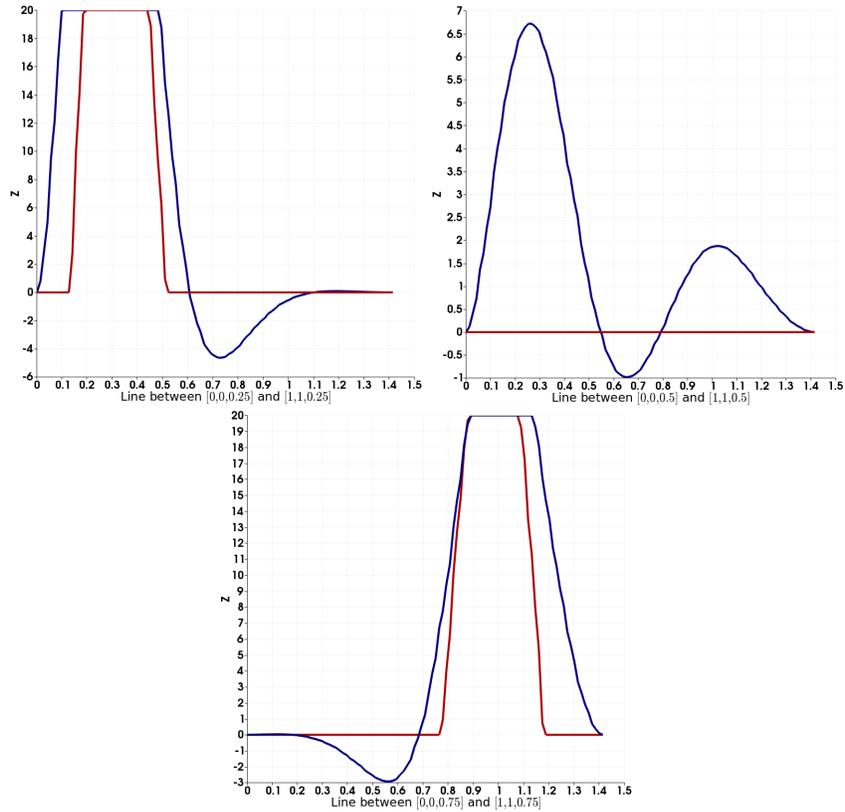
Figure 5: Example 1, comparisons of sparse (red) and nonsparse (blue) controls along the line between $[0, 0, 0.25]$ and $[1, 1, 0.25]$, between $[0, 0, 0.5]$ and $[1, 1, 0.5]$, and between $[0, 0, 0.75]$ and $[1, 1, 0.75]$ for the moving target example.
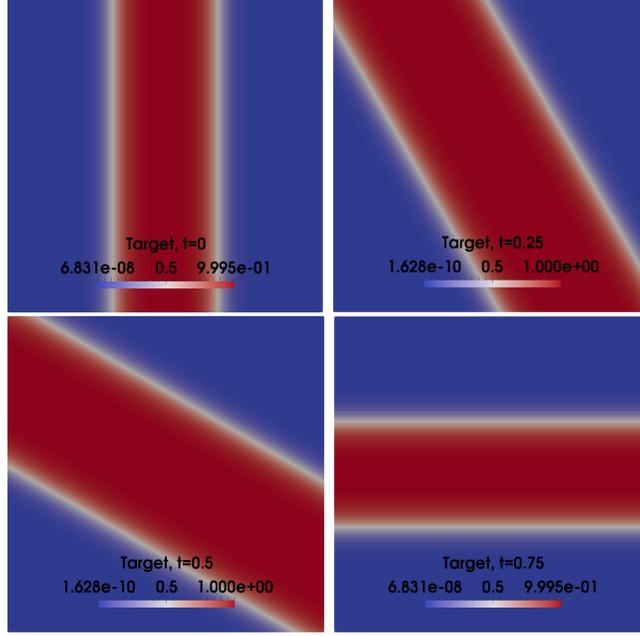
Figure 6: Example 2, plots of the target at time $t = 0,\ 0.25,\ 0.5,\ 0.75$ for the turning wave example.

5.2   Turning wave target (Example 2)

In this example, we consider the target

$$u_Q(x,t) = \left(1.0 + \exp\left(\frac{\cos(g(t))\left(\frac{70}{3} - 70x_1\right) + \sin(g(t))\left(\frac{70}{3} - 70x_2\right)}{\sqrt{2}}\right)\right)^{-1}$$
$$+ \left(1.0 + \exp\left(\frac{\cos(g(t))\left(70x_1 - \frac{140}{3}\right) + \sin(g(t))\left(70x_2 - \frac{140}{3}\right)}{\sqrt{2}}\right)\right)^{-1} - 1,$$

where $g(t) = \frac{2\pi}{3}\min\left\{\frac{3}{4}, t\right\}$. This is an adapted version of the turning wave example considered in [15]. The wave front turns 90 degrees from time $t = 0$ to $t = 0.75$, and remains fixed after $t = 0.75$; see the target at $t = 0$, 0.25, 0.5, and 0.75 as illustrated in Fig. 6.

The nonlinear reaction term is given by $R(u) = u(u - 0.25)(u + 1)$. We use the initial data

$$u_0(x) = \left(1 + \exp\left(\frac{\frac{70}{3} - 70x_1}{\sqrt{2}}\right)\right)^{-1} + \left(1 + \exp\left(\frac{70x_1 - \frac{140}{3}}{\sqrt{2}}\right)\right)^{-1} - 1$$

on Σ_0, and homogeneous Neumann boundary condition on Σ for the state. As parameters, we use $\varrho = 10^{-6}$, $\mu = 10^{-4}$ for the sparse case and $\varrho = 10^{-6}$, $\mu = 0$

13
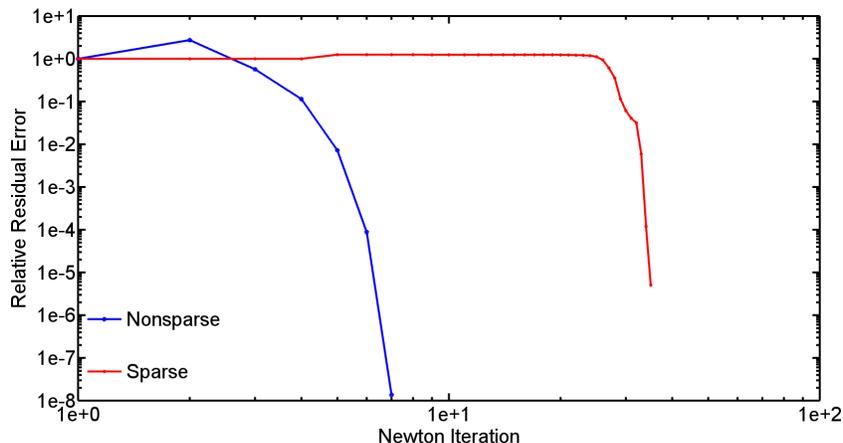
Figure 7: Example 2, relative residual error reduction in the semismooth Newton method for the nonsparse and sparse control in the turning wave example.

for the nonsparse case. The bounds $a = -100$ and $b = 100$ are set for both cases.

To solve the nonlinear first order necessary optimality system, we needed 7 and 35 semismooth Newton iterations for the nonsparse and sparse optimal control, respectively; see Fig. 7. The total system assembling and solving time is about 3 and 12.7 hours on the desktop computer, respectively.

The numerical solutions of sparse and nonsparse controls as well as associated optimal states are illustrated in Fig. 8. We clearly see certain sparsity of our optimal sparse control as compared to pure L^2-regularization, without too much precision loss of the associated state to the target. A closer look to the sparse and nonsparse controls confirms that our optimal sparse control exhibits sparsity with respect to the spatial direction; see Fig. 9.

Instead of an uniform refinement, we may also adopt an adaptive strategy. In this way, local refinements are made in the region where the solution shows a more local character, whereas coarser meshes appear in the other region. For example, in the optimal sparse control case ($\varrho = 10^{-6}$, $\mu = 10^{-4}$), we start from an initial mesh with 729 grid points, 9 in each spatial and the temporal direction. We use a residual based error indicator for the coupled state and adjoint state system to guide our adaptive mesh refinement, similar to the approach [33]. After the 6th adaptive octasection refinement [4], the mesh contains $1,053,443$ grid points; see the adaptive space-time mesh and the meshes on the cutting plans at different times in Fig. 10. As we observe, the adaptive refinements follow the rotation of the wave front of the state.

14
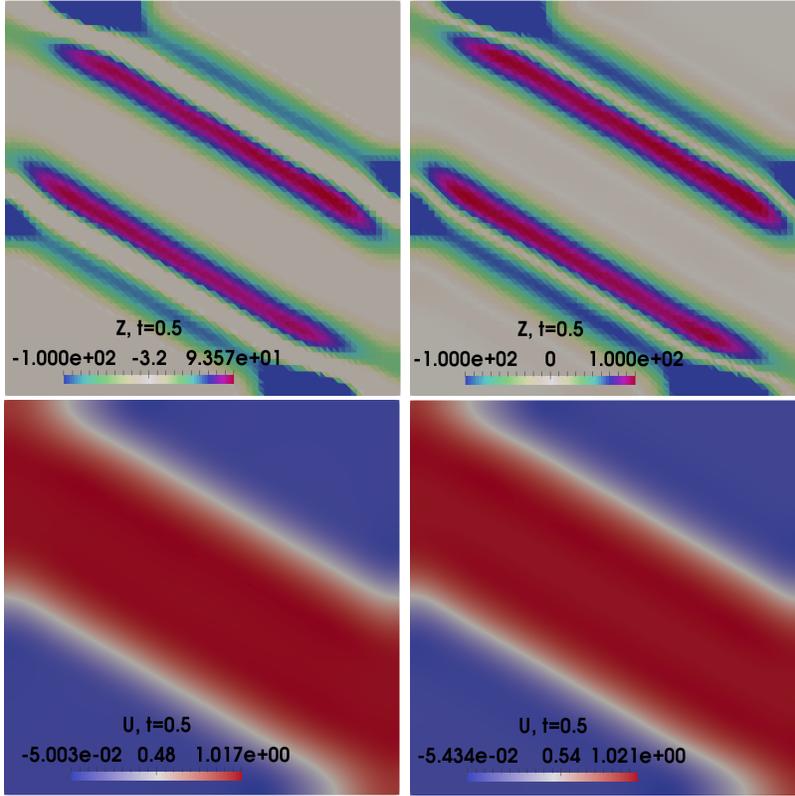
Figure 8: Example 2, plots of the sparse (left) and nonsparse (right) controls (in the first row) and the associated states (in the second row) at time $t = 0.5$ for the turning wave example.

6   Conclusions

In this work, we have considered a space-time Petrov-Galerkin finite element method on fully unstructured simplicial meshes for semilinear parabolic optimal sparse control problems. The objective functional involves the well-known L^1-norm of the control in addition to the standard L^2-regularization term. The proposed method is able to capture spatio-temporal sparsity, which has been confirmed by our numerical experiments. A rigorous convergence and error analysis of our space-time Petrov-Galerkin finite element methods for such optimal sparse control problems is left for future work.
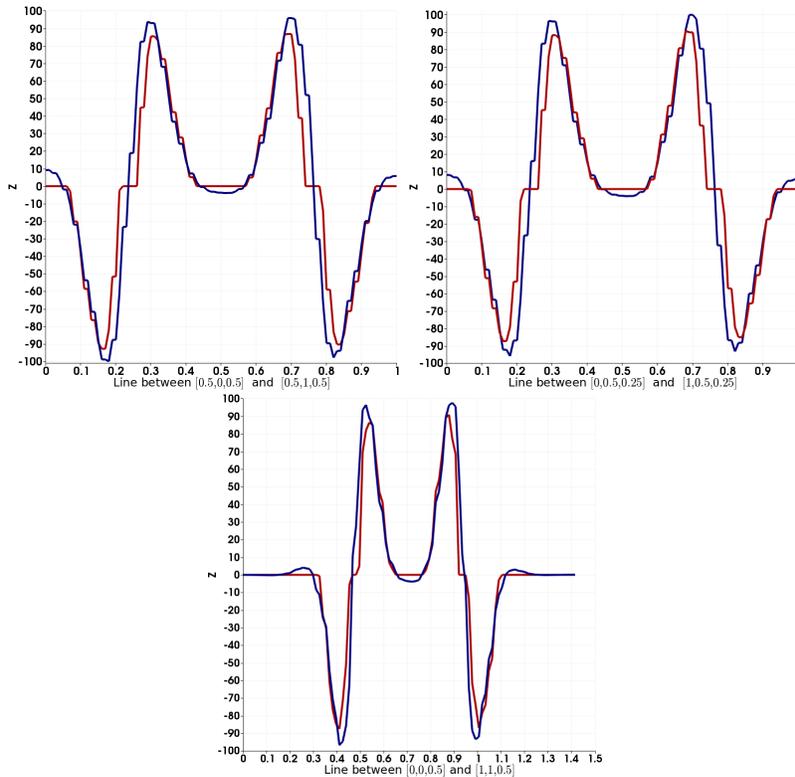
Figure 9: Example 2, comparisons of sparse (red) and nonsparse (blue) controls along the line between $[0.5, 0, 0.5]$ and $[0.5, 1, 0.5]$, between $[0, 0.5, 0.25]$ and $[1, 0.5, 0.25]$, and between $[0, 0, 0.5]$ and $[1, 1, 0.5]$ for the turning wave example.

References

[1] N. Arada, E. Casas, and F. Tröltzsch. Error estimates for the numerical approximation of a semilinear elliptic control problem. *Comput. Optim. Appl.*, 23:201–229, 2002.

[2] R. E. Bank, P. S. Vassilevski, and L. T. Zikatanov. Arbitrary dimension convection-diffusion schemes for space-time discretizations. *J. Comput. Appl. Math.*, 310:19–31, 2017.

[3] M. Behr. Simplex space-time meshes in finite element simulations. *Int. J. Numer. Meth. Fluids*, 57:1421–1434, 2008.

[4] J. Bey. Tetrahedral grid refinement. *Computing*, 55:355–378, 1995.

[5] A. C. Boulanger and P. Trautmann. Sparse optimal control of the KdV-Burgers equation on a bounded domain. *SIAM J. Control Optim.*, 55(6):3673–3706, 2017.
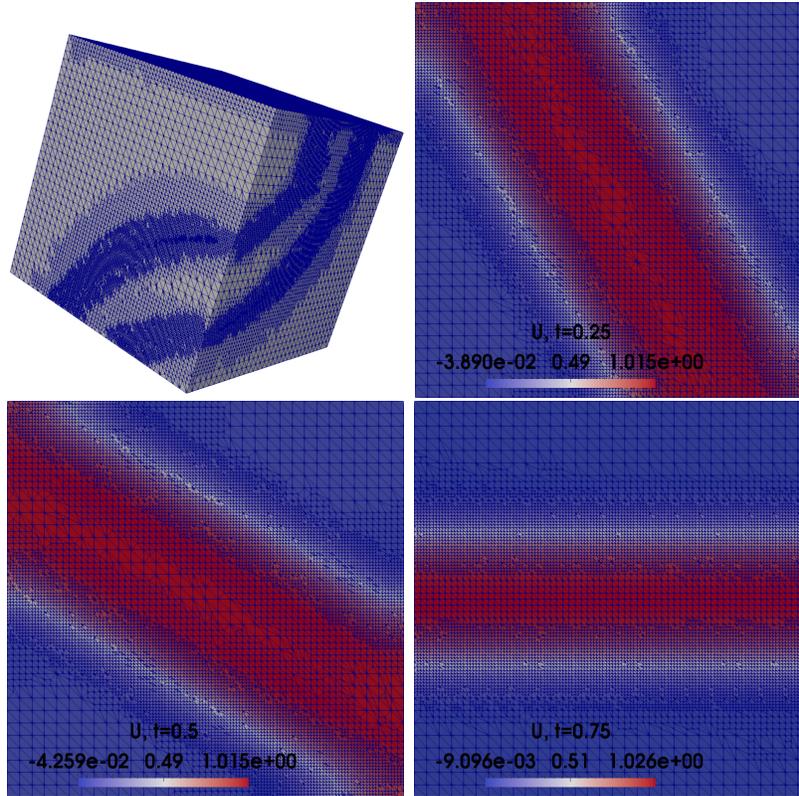
Figure 10: Example 2, plots of the adaptive space-time mesh (top-left) at the 6th step, and the meshes on the cutting plans for times $t = 0.25$, 0.5, and 0.75.

[6] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics.* Cambridge University Press, 2007.

[7] E. Casas. A review on sparse solutions in optimal control of partial differential equations. *SeMA*, 74:319–344, 2017.

[8] E. Casas, C. Clason, and K. Kunisch. Parabolic control problems in measure spaces with sparse solutions. *SIAM J. Control Optim.*, 51(1):28–63, 2013.

[9] E. Casas, R. Herzog, and G. Wachsmuth. Approximation of sparse controls in semilinear equations by piecewise linear functions. *Numer. Math.*, 122:645–669, 2012.

[10] E. Casas, R. Herzog, and G. Wachsmuth. Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional. *SIAM J. Optim.*, 22(3):795–820, 2012.

[11] E. Casas, R. Herzog, and G. Wachsmuth. Analysis of spatio-temporally sparse optimal control problems of semilinear parabolic equations. *ESAIM: COCV*, 23(1):263–295, 2017.

[12] E. Casas and K. Kunisch. Parabolic control problems in space-time measure spaces. *ESAIM: COCV*, 22(2):355–370, 2016.

[13] E. Casas, M. Mateos, and A. Rösch. Finite element approximation of sparse parabolic control problems. *Math. Control Relat. Fields*, 7(3):393–417, 2017.

[14] E. Casas, M. Mateos, and A. Rösch. Improved approximation rates for a parabolic control problem with an objective promoting directional sparsity. *Comput. Optim. Appl.*, 70:239–266, 2018.

[15] E. Casas, C. Ryll, and F. Tröltzsch. Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems. *Comput. Methods Appl. Math.*, 13(4):415–442, 2013.

[16] E. Casas, C. Ryll, and F. Tröltzsch. Second order and stability analysis for optimal sparse control of the FitzHugh-Nagumo equation. *SIAM J. Control and Optim.*, 53(4):2168–2202, 2015.

[17] E. Casas and E. Zuazua. Spike controls for elliptic and parabolic PDEs. *Syst. Control Lett.*, 62(4):311–318, 2013.

[18] M. J. Gander. 50 years of time parallel integration. In Thomas Carraro, Michael Geiger, Stefan Körkel, and Rolf Rannacher, editors, *Multiple Shooting and Time Domain Decomposition*, pages 69–114. Springer Verlag, Heidelberg, Berlin, Cham, 2015.

[19] M. J. Gander and M. Neumüller. Analysis of a new space-time parallel multigrid algorithm for parabolic problems. *SIAM J. Sci. Comput.*, 38(4):A2173–A2208, 2016.

[20] R. Herzog, G. Stadler, and G. Wachsmuth. Directional sparsity in optimal control of partial differential equations. *SIAM J. Control Optim.*, 50(2):943–963, 2012.

[21] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications*, volume 15 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[22] V. Karyofylli, L. Wendling, M. Make, N. Hosters, and M. Behr. Simplex space-time meshes in thermally coupled two-phase flow simulations of mold filling. *Compur. Fluids*, 192:104261, 2019.

[23] K. Kunisch, K. Pieper, and B. Vexler. Measure valued directional sparsity for parabolic optimal control problems. *SIAM J. Control Optim.*, 52(5):3078–3108, 2014.

[24] U. Langer, Neumüller M., and Schafelner A. Space-time finite element methods for parabolic evolution problems with variable coefficients. In T. Apel, U. Langer, A. Meyer, and O. Steinbach, editors, *Advanced Finite Element Methods with Applications: Selected Papers from the 30th Chemnitz Finite Element Symposium 2017*, pages 247–275, Cham, 2019. Springer International Publishing.

[25] U. Langer, O. Steinbach, F. Tröltzsch, and H. Yang. Unstructured space-time finite element methods for optimal control of parabolic equations. In preparation, 2020.

[26] U. Langer and H. Yang. BDDC preconditioners for a space-time finite element discretization of parabolic problems. In *Domain Decomposition Methods in Science and Engineering XXV*, Cham, 2020. Springer International Publishing. To appear.

[27] C. Li and G. Stadler. Sparse solutions in optimal control of PDEs with uncertain parameters: The linear case. *SIAM J. Control Optim.*, 57(1):633–658, 2019.

[28] A. Nägel, D. Logashenko, J. B. Schroder, and U. M. Yang. Aspects of solvers for large-scale coupled problems in porous media. *Transp. Porous. Med.*, 130:363–390, 2019.

[29] C. Ryll, J. Löber, S. Martens, H. Engel, and F. Tröltzsch. Analytical, optimal, and sparse optimal control of traveling wave solutions to reaction-diffusion systems. In E. Schöll, S. H. L. Klapp, and P. Hövel, editors, *Control of Self-Organizing Nonlinear Systems*, pages 189–210. Springer International Publishing, Cham, 2016.

[30] G. Stadler. Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices. *Comput. Optim. Appl.*, 44:159–181, 2009.

[31] O. Steinbach. *Numerical Approximation Methods for Elliptic Boundary Value Problems*. Springer-Verlag New York, 2008.

[32] O. Steinbach. Space-time finite element methods for parabolic problems. *Comput. Methods Appl. Math.*, 15:551–566, 2015.

[33] 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(3):e2143, 2018.

[34] O. Steinbach and H. Yang. A space-time finite element method for the linear bidomain equations. In T. Apel, U. Langer, A. Meyer, and O. Steinbach, editors, *Advanced Finite Element Methods with Applications: Selected Papers from the 30th Chemnitz Finite Element Symposium 2017*, pages 323–339, Cham, 2019. Springer International Publishing.

[35] O. Steinbach and H. Yang. Space-time finite element methods for parabolic evolution equations: discretization, a posteriori error estimation, adaptivity and solution. In O. Steinbach and U. Langer, editors, *Space-Time Methods: Application to Partial Differential Equations*, Radon Series on Computational and Applied Mathematics, pages 207–248, Berlin, 2019. de Gruyter.

[36] I. Toulopoulos. Space-time finite element methods stabilized using bubble function spaces. *Appl. Anal.*, 0(0):1–18, 2018.

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

[38] M. Uzunca, T. Küçükseyhan, H. Yücel, and B. Karasözen. Optimal control of convective Fitzhugh-Nagumo equation. *Comput. Math. Appl.*, 73(9):2151–2169, 2017.

[39] M. von Danwitz, V. Karyofylli, N. Hosters, and M. Behr. Simplex space-time meshes in compressible flow simulations. *Int. J. Numer. Meth. Fluids*, 91(1):29–48, 2019.