arXiv:1804.10549v2 [math.OC] 20 Jun 2018

Maximal discrete sparsity in parabolic optimal control with measures

Evelyn Herberg, Michael Hinze, Henrik Schumacher

June 20, 2018

Abstract. We consider a parabolic optimal control problem governed by space-time measure controls. Two approaches to discretize this problem will be compared. The first approach has been considered in [4] and employs a discontinuous Galerkin method for the state discretization where controls are discretized piecewise constant in time and by Dirac measures concentrated in the finite element nodes in space. In the second approach we use variational discretization for the control problem utilizing a Petrov-Galerkin approximation of the state which induces controls that are composed of Dirac measures in space and time, i.e. Dirac measures concentrated in finite element nodes with respect to space, and concentrated on the grid points of the time integration scheme with respect to time. The latter approach then yields *maximal* sparsity in space-time on the discrete level. Numerical experiments show the differences of the two approaches.

1 Introduction

We follow [4] and consider the continuous minimization problem

$$\min_{(u_0,u)\in\mathcal{M}(\Omega)\times\mathcal{M}(Q)} J(u_0,u) := \tfrac{1}{q}\|y-y_d\|^q_{L^q(Q)} + \alpha\|u\|_{\mathcal{M}(Q)} + \beta\|u_0\|_{\mathcal{M}(\Omega)}, \tag{P}$$

where the *state* $y \in L^q(Q)$ solves the following parabolic *state equation*

$$\begin{cases} \frac{\partial y}{\partial t} - \Delta y, & = u \qquad \text{in } Q, \\ y(x,0) & = u_0, \qquad \text{in } \Omega, \\ y(x,t) & = 0, \qquad \text{on } \Sigma = \Gamma \times (0,T) \end{cases} \tag{1}$$

in a very weak sense specified in the next definition (see [4, Definition 2.1.]).

Definition 1.
A function $y \in L^1(Q)$ is a solution to (1)*, if the following identity holds for all $w \in W$:*

$$\langle S\,y, w\rangle_{\mathcal{M}(\Omega)\times\mathcal{M}(Q),W} := \int_Q -\left(\frac{\partial w}{\partial t} + \triangle w\right) y \, dx\, dt = \int_Q w \, du + \int_\Omega w(0)\, du_0 =: \langle \tilde{u}, w\rangle_{\mathcal{M}(\Omega)\times\mathcal{M}(Q),W}, \tag{2}$$

where $\tilde{u} := (u_0, u) \in \mathcal{M}(\Omega) \times \mathcal{M}(Q)$ and

$$W = \left\{ w \in L^2(0,T; H_0^1(\Omega)) : \left(\frac{\partial w}{\partial t} + \triangle w\right) \in L^\infty(Q) \text{ and } w(x,T) = 0 \text{ in } \Omega \right\}.$$

Here, Ω is an open bounded domain in $\mathbb{R}^d$, $d \in \{1,2,3\}$ with Lipschitz boundary $\Gamma := \partial\Omega$. The time interval $(0,T)$ is denoted by I and $Q := \Omega \times I$ is the space-time domain. Let $\alpha > 0, \beta > 0$ be suitable penalty parameters and let $q \in [1, \min\{2, (d+2)/d\})$. The *controls* (u_0, u) reside in the *control space* $\mathcal{M}(\Omega) \times \mathcal{M}(Q)$, the product of the topological duals of the spaces of continuous functions $C_0(\bar{\Omega})$ and $C_0(\bar{Q})$ with compact support in Ω and Q, respectively. Consequently, the measure norms in the objective functional (P) are defined by:

$$\|u_0\|_{\mathcal{M}(\Omega)} := \sup_{\|f\|_{C_0(\bar{\Omega})}\le 1} \int_\Omega f(x)\, \mathrm{d}u_0 \qquad \text{and} \qquad \|u\|_{\mathcal{M}(Q)} := \sup_{\|f\|_{C_0(\bar{Q})}\le 1} \int_Q f(x,t)\, \mathrm{d}u.$$

In this setting, (P) has a unique solution (see [4, Theorem 2.7]). The discrete concept introduced in [4] proposes a control space consisting of piecewise constant controls in time and Dirac measures in space. As a consequence the *maximal* possible sparsity in this setting yields controls which are constant on a time interval. In source identification applications it may be desirable that the optimal control can be represented even more sparsely e.g. by a combination of Dirac measures in space and time. Here we propose variational discretization from [9], which allows to *control* the discrete structure of the controls through the choice of Petrov-Galerkin-Ansatz and -Test spaces in the discretization of the state equation and so to achieve the desired sparsity. For

this approach we obtain analogous convergence results as reported in [4, Theorem 4.3.]. More precisely, we will prove Theorem 2 (see Section 4 for the notation):

Theorem 2. *Let* $(\bar{u}_{0,h}, \bar{u}_\sigma)$ *be the unique solution of the problem* (P_{vd}) *belonging to* $U_h \times \mathcal{U}_{\mathrm{vd}}$ *and* Γ *be of class* $C^{1,1}$ *and* $1 < q < \min\{2, \frac{d+2}{d}\}$. *If* $\{(\bar{u}_{0,h}, \bar{u}_\sigma)\}_\sigma$ *is a sequence of such solutions with associated states* $\{\bar{y}_\sigma\}_\sigma$ *the following convergence properties hold:*

$$\lim_{|\sigma|\to 0} \|\bar{y} - \bar{y}_\sigma\|_{L^q(Q)} = 0, \tag{3}$$

$$(\bar{u}_{0,h}, \bar{u}_\sigma) \overset{*}{\rightharpoonup} (\bar{u}_0, \bar{u}) \ \text{ as } |\sigma| \to 0 \ \text{ in } \mathcal{M}(\Omega) \times \mathcal{M}(Q), \tag{4}$$

$$\lim_{|\sigma|\to 0} (\|\bar{u}_{0h}\|_{\mathcal{M}(\Omega)}, \|\bar{u}_\sigma\|_{\mathcal{M}(Q)}) = (\|\bar{u}_0\|_{\mathcal{M}(\Omega)}, \|\bar{u}\|_{\mathcal{M}(Q)}), \tag{5}$$

where $(\bar{u}_0, \bar{u})$ *is the unique solution of* (P) *and* $\bar{y}$ *associated state.*

The paper is organized as follows: In Section 2 we analyze the continuous problem (P) and its sparsity structure. Similar control problems in measure spaces are analyzed and discretized, e.g. in [2, 6, 12] for the elliptic case and in [3, 4, 5, 11] for the parabolic case. In [5] the special case of initial data is covered and in [12, 11] a priori error estimates are derived. We will take the approach to set up the predual problem with respect to Fenchel duality as is done in [6]. The resulting predual problem will be discretized with two different strategies. Our first discretization strategy in Section 3 is also presented in [4], which allows us to use their results. Here we add the derivation of a semi-smooth Newton method to solve the discrete problem (P_σ) from [4]. The second strategy in Section 4 uses variational discretization from [9] with a time-discrete scheme similar to the one in [7]. The emerging discrete problem (P_{vd}) is solved analogously to (P_σ). Computational results of both approaches are compared in Section 5.

2 Continuous optimality system

In this section we take a closer look at the structure of (P). As we want to numerically solve our problem with derivative based methods we have to cope with the L^q norm for $q < 2$ in our cost functional. We take an approach which is closely related to the one proposed in [6, Chapter 2.1.], where a Fenchel duality approach is used for the solution of the optimal control problem. We present the problem (P^*), show that this problem has a unique solution and prove that it is the Fenchel predual problem of (P).

To begin with, we recall the optimality conditions from [4, Theorem 3.1.] for the solution of (P) and the resulting sparsity structure [4, Corollary 3.2.] of the optimal controls $(\bar{u}_0, \bar{u})$.

Lemma 3. *Let $(\bar{u}_0, \bar{u})$ denote a solution to (P) with associated state $\bar{y}$. Then there exists an element $\bar{w} \in L^2(0,T;H_0^1(\Omega)) \cap C(\bar{Q})$ satisfying*

$$\begin{cases} -\frac{\partial \bar{w}}{\partial t} - \Delta \bar{w} &= \bar{z} \quad in\ Q, \\ \bar{w}(x,T) &= 0 \quad in\ \Omega, \\ \bar{w}(x,t) &= 0 \quad on\ \Sigma, \end{cases} \qquad \begin{cases} \int_Q \bar{w}\, d\bar{u} + \alpha \|\bar{u}\|_{\mathcal{M}(Q)} = 0, \\ \|\bar{w}\|_{C(\bar{Q})} \begin{cases} = \alpha, & if\ \bar{u} \neq 0, \\ \leq \alpha, & if\ \bar{u} = 0, \end{cases} \end{cases} \qquad \begin{cases} \int_\Omega \bar{w}(0)\, d\bar{u}_0 + \beta \|\bar{u}_0\|_{\mathcal{M}(\Omega)} = 0, \\ \|\bar{w}(0)\|_{C(\bar{\Omega})} \begin{cases} = \beta, & if\ \bar{u}_0 \neq 0, \\ \leq \beta, & if\ \bar{u}_0 = 0, \end{cases} \end{cases}$$

where

$$\bar{z} \begin{cases} = \|\bar{y} - y_d\|^{q-2}(\bar{y} - y_d), & if\ 1 < q < \min\{2, \frac{d+2}{d}\}, \\ \in \operatorname{sign}(\bar{y} - y_d), & if\ q = 1. \end{cases}$$

Furthermore, $\bar{w}$ is unique if $q > 1$.

Remark 4. *Under the assumptions of the previous Theorem we have the following sparsity structure:*

$$\operatorname{supp}(\bar{u}_0^+) \subset \{x \in \bar{\Omega} : \bar{w}(x,0) = -\beta\} \qquad \operatorname{supp}(\bar{u}_0^-) \subset \{x \in \bar{\Omega} : \bar{w}(x,0) = +\beta\}$$
$$\operatorname{supp}(\bar{u}^+) \subset \{(x,t) \in \bar{Q} : \bar{w}(x,t) = -\alpha\} \qquad \operatorname{supp}(\bar{u}^-) \subset \{(x,t) \in \bar{Q} : \bar{w}(x,t) = +\alpha\}$$

where $\bar{u}_0 = \bar{u}_0^+ - \bar{u}_0^-$ and $\bar{u} = \bar{u}^+ - \bar{u}^-$ are the Jordan decompositions.

If one considers it as the generic case that the function $\bar{w}$ is not constant on sets of measure greater than zero the controls have support sets of measure zero. This is our motivation to propose a discretization strategy which reflects this behavior on the discrete level in space and time. Before we go on to the discretization we recall the Fenchel duality theorem from [6, Chapter 1.1.3.].

Let U and Y be Banach spaces with topological duals U^* and Y^* and let $\Lambda : U \to Y$ be a continuous linear operator. Let $F : U \to \bar{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ and $G : Y \to \bar{\mathbb{R}}$ be convex lower semicontinuous functionals, such that F and G are not identically equal to ∞. Furthermore, let the regular point condition be fulfilled, i.e., there exists $v_0 \in U$, such that $F(v_0) < \infty$, $G(\Lambda v_0) < \infty$ and G is continuous at Λv_0. Denote by

$$F^* : U^* \to \bar{\mathbb{R}}, \quad F^*(w) = \sup_{u \in U} \langle w, u \rangle_{U^*,U} - F(u). \tag{6}$$

the Fenchel conjugate of F. In order to calculate F^*, we use the following equivalence

$$\sup_{u\in U} \langle w, u\rangle_{U^*,U} = \langle \bar{w}, u\rangle_{U^*,U} \quad \text{iff} \quad \bar{w} \in \partial F(u), \tag{7}$$

where $\partial F(u)$ denotes the subdifferential of the convex functional F, which reduces to the Gâteaux-derivative if it exists. Under the given assumptions the Fenchel duality theorem states that

$$\inf_{u\in U} F(u) + G(\Lambda u) = \sup_{z\in Y^*} -F^*(\Lambda^* z) - G^*(-z), \tag{8}$$

holds, and that the right side of (8) has at least one solution. Furthermore, we know that the following two statements are equivalent:

$$F(\bar{u}) + G(\Lambda\bar{u}) = -F^*(\Lambda^*\bar{z}) - G^*(-\bar{z}) \tag{9}$$

$$\Lambda^*\bar{z} \in \partial F(\bar{u}) \quad \wedge \quad -\bar{z} \in \partial G(\Lambda\bar{u}) \tag{10}$$

We want to define the problem (P^*). The argument hereof will be the adjoint variable w. From [15, Theorem 27.7.] we know that for every $\psi \in L^\infty(Q) \subset L^2(Q) \subset L^2(0,T;H^{-1}(\Omega))$ the adjoint problem

$$\begin{cases} -\left(\frac{\partial w}{\partial t} + \triangle w\right) &= \psi \qquad \text{in } Q \\ w(x,T) &= 0 \qquad \text{in } \Omega \\ w(x,t) &= 0 \qquad \text{on } \Sigma \end{cases} \tag{11}$$

has a unique solution $w \in W(0,T)$. With $W(0,T) \hookrightarrow C([0,T];L^2(\Omega))$ from [10, Theorem 1.32.] we see that $w \in L^2(0,T;H_0^1(\Omega)) \cap C([0,T];L^2(\Omega))$. As (11) fulfills the requirements of [1, Theorem 5.1.] we even get the regularity $w \in C(\bar{Q})$.

Based on the characterization of solutions in (2) we know from [4, Theorem 2.2.] that there exists a unique solution $y \in L^1(Q)$ to (1), where additionally $y \in L^q(0,T;W_0^{1,p}(\Omega))$ for all $p,q \in [1,2)$ with $\frac{2}{q} + \frac{d}{p} > d+1$, and the following estimate holds:

$$\|y\|_{L^q(0,T;W_0^{1,p}(\Omega))} \leq C_{p,q}\left(\|u\|_{\mathcal{M}(Q)} + \|u_0\|_{\mathcal{M}(\Omega)}\right). \tag{12}$$

To justify the choice of q in the objective functional we recall [4, Remark 2.4.]: There exists a parameter p fulfilling the assumptions formulated above and $W_0^{1,p}(\Omega) \subset L^q(\Omega)$ compactly. Thus $y \in L^q(0,T;W_0^{1,p}(\Omega)) \subset L^q(Q)$. The density of $L^\infty(Q)$ in $L^p(Q)$, where $\frac{1}{q} + \frac{1}{p} = 1$, now implies

that the identity (2) is valid for every w in the space

$$W_q = \left\{ w \in L^2(0,T;H_0^1(\Omega)) : \left(\frac{\partial w}{\partial t} + \triangle w \right) \in L^p(Q) \text{ and } w(x,T) = 0 \text{ in } \Omega \right\}.$$

Our assumptions on Γ ensure that Assumption (A) from [4] is fulfilled, so that given $1 \leq q < \min\left\{2, \frac{d+2}{d}\right\}$ and p with $\frac{1}{q} + \frac{1}{p} = 1$, problem (11) for every $\psi \in L^p(Q)$ admits a unique solution w belonging to $L^p(0,T;W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega))$. The requirements of [1, Theorem 5.1.] are still fulfilled in this case, so we get the regularity $w \in C(\bar{Q})$. Note that due to the compactness of $\bar{Q}$ we have $C(\bar{Q}) = C_0(\bar{Q})$.

We write (2) as $Sy = \tilde{u} \in \mathcal{M}(\Omega) \times \mathcal{M}(Q)$, i.e. $\langle Sy, w \rangle_{\mathcal{M}(\Omega)\times\mathcal{M}(Q),C_0(\bar{Q})} = \langle \tilde{u}, w \rangle_{\mathcal{M}(\Omega)\times\mathcal{M}(Q),C_0(\bar{Q})}$ for all $w \in W$. We know that this identity is true for all $w \in W_q \hookrightarrow C_0(\bar{Q})$ and then by definition of W_q we have $\left(-\frac{\partial w}{\partial t} - \triangle w\right) = S^* w \in L^p(Q)$. In this sense $\langle Sy, w \rangle_{\mathcal{M}(\Omega)\times\mathcal{M}(Q),C_0(\bar{Q}))} = \langle S^* w, y \rangle_{L^p(Q),L^q(Q)}$ holds for any $y \in L^q(0,T;W_0^{1,p}(\Omega)) \subset L^q(Q)$ and $w \in W_q$, i.e. S^* is the adjoint operator of S in the described sense. To abbreviate notation we from here onwards write $\langle \cdot, \cdot \rangle_{\mathcal{M},C_0}$ instead of $\langle \cdot, \cdot \rangle_{\mathcal{M}(\Omega)\times\mathcal{M}(Q),C_0(\bar{Q})}$ and also assume $q > 1$ so that $p < \infty$ holds. Now we are in the position to define problem (P^*):

$$\min_{w \in W_q} K(w) := \frac{1}{p} \|S^* w\|_{L^p(Q)}^p + \langle S^* w, y_d \rangle_{L^p(Q),L^q(Q)} + \ell_{\alpha,\beta}(w). \tag{P^*}$$

Here the indicator function $\ell_{\alpha,\beta}(w)$ is given by

$$\ell_{\alpha,\beta}(w) = \begin{cases} 0, & \text{if } \|w\|_{C_0(\bar{Q})} \leq \alpha \text{ and } \|w(0)\|_{C_0(\bar{\Omega})} \leq \beta, \\ \infty, & \text{else.} \end{cases}$$

As $W_q \hookrightarrow C_0(\bar{Q})$, all appearing norms are well defined. We have

Theorem 5. *Let* $1 < q < \min\{2, \frac{d}{d-1}\}$. *Then problem* ($P^*$) *has a unique solution* $\bar{w} \in W_q$.

Proof. We proceed analogously to [6, Theorem 2.3.]. Let $\{w_k\}_k \subset W_q$ be a minimizing sequence, such that

$$\lim_{k\to\infty} K(w_k) = \inf_{w \in W_q} K(w) =: \underline{K}.$$

As $W_q \neq \emptyset$ we know that $\underline{K} < \infty$. With Cauchy-Schwarz and Young's inequality we see for all k

$$\begin{aligned} K(w_k) &= \tfrac{1}{p}\|S^*w_k\|^p_{L^p(Q)} + \langle S^*w_k, y_d\rangle_{L^p(Q),L^q(Q)} + \underbrace{\ell_{\alpha,\beta}(w_k)}_{\geq 0} \\ &\geq \tfrac{1}{p}\|S^*w_k\|^p_{L^p(Q)} - \|S^*w_k\|_{L^p(Q)}\|y_d\|_{L^q(Q)} \\ &\geq -\tfrac{1}{q}\|y_d\|^q_{L^q(Q)} \\ &> -\infty \end{aligned}$$

and hence $\underline{K} > -\infty$. As $\underline{K} \in \mathbb{R}$ the indicator function $\ell_{\alpha,\beta}(w)$ yields $\|w_k\|_{C_0(\bar{Q})} \leq \alpha$ and $\|w_k(0)\|_{C_0(\bar{\Omega})} \leq \beta$ for all $k \geq M$ with $M \in \mathbb{N}$ large enough. We deduce that there exist weak-* convergent subsequences such that $w_{k'} \overset{*}{\rightharpoonup} \bar{w} \in C_0(\bar{Q})$ and $w_{k'}(0) \overset{*}{\rightharpoonup} \bar{w}(0) \in C_0(\bar{\Omega})$. Due to $p > 2$ the term $\frac{1}{p}\|S^*w_k\|^p_{L^p(Q)}$ is coercive and we deduce the boundedness of $\|S^*w_k\|_{L^p(Q)}$ for all k. Together with the previous finding this yields the existence of a weakly converging subsequence such that $S^*w_{k'} \rightharpoonup z \in L^p(Q)$, i.e. $\langle S^*w_{k'}, y\rangle_{L^p(Q),L^q(Q)} \to \langle z, y\rangle_{L^p(Q),L^q(Q)}$ for all $y \in L^q(Q)$. We also have

$$\langle S^*w_{k'}, y\rangle_{L^p(Q),L^q(Q)} = \langle Sy, w_{k'}\rangle_{\mathcal{M},C_0} \to \langle Sy, \bar{w}\rangle_{\mathcal{M},C_0} = \langle S^*\bar{w}, y\rangle_{L^p(Q),L^q(Q)} \qquad \forall\, y \in L^q(Q)$$

This gives $z = S^*\bar{w} \in L^p(Q)$. From $\left(-\frac{\partial w}{\partial t} - \triangle w\right) = S^*\bar{w} \in L^p(Q)$ and $p > 2$ we can deduce $\bar{w} \in L^2(0,T;H^1_0(\Omega))$. Furthermore from $\langle u, w_{k'}\rangle_{\mathcal{M},C_0} \to \langle \tilde{u}, \bar{w}\rangle_{\mathcal{M},C_0}$ for all $\tilde{u} \in \mathcal{M}(\Omega)\times\mathcal{M}(Q)$ we get $\bar{w}(x,T) = 0$ if we consider the dirac measure located in T. Altogether this yields $\bar{w} \in W_q$.
We know that all continuous norms are weakly lower semicontinuous and because of the weak-* convergence of $w_{k'}$ and $w_{k'}(0)$ we get $\ell_{\alpha,\beta}(\bar{w}) = 0$. Furthermore from $S^*w_{k'} \rightharpoonup S^*\bar{w} \in L^p(Q)$ follows directly $\langle S^*w_{k'}, y_d\rangle_{L^p(Q),L^q(Q)} \to \langle S^*\bar{w}, y_d\rangle_{L^p(Q),L^q(Q)}$. Altogether we get

$$\begin{aligned} \underline{K} \leq K(\bar{w}) &= \tfrac{1}{p}\|S^*\bar{w}\|^p_{L^p(Q)} + \langle S^*\bar{w}, y_d\rangle_{L^p(Q),L^q(Q)} + \ell_{\alpha,\beta}(\bar{w}) \\ &\leq \liminf_{k'\to\infty} \tfrac{1}{p}\|S^*w_{k'}\|^p_{L^p(Q)} + \lim_{k''\to\infty} \langle S^*w_{k'}, y_d\rangle_{L^p(Q),L^q(Q)} \\ &= \lim_{k\to\infty} \left(\tfrac{1}{p}\|S^*w_k\|^p_{L^p(Q)} + \langle S^*w_k, y_d\rangle_{L^p(Q),L^q(Q)}\right) \\ &\leq \lim_{k\to\infty} \left(\tfrac{1}{p}\|S^*w_k\|^p_{L^p(Q)} + \langle S^*w_k, y_d\rangle_{L^p(Q),L^q(Q)} + \ell_{\alpha,\beta}(w_k)\right) \\ &= \underline{K}. \end{aligned}$$

thus $\bar{w}$ is a solution of (P^*). The uniqueness follows from $q > 1$ and thus $p < \infty$.

□

Theorem 6. *The Fenchel dual of* (P^*) *is* (P).

Proof. To apply Fenchel duality we split the objective functional of (P^*) into two parts:

$$\mathcal{F} : W_q \to \bar{\mathbb{R}}, \qquad \mathcal{F}(w) = \frac{1}{p}\|S^*w\|^p_{L^p(Q)} + \langle S^*w, y_d\rangle_{L^p(Q),L^q(Q)} \tag{13}$$

$$\mathcal{G} : C_0(\bar{Q}) \to \bar{\mathbb{R}}, \qquad \mathcal{G}(w) = \ell_{\alpha,\beta}(w) \tag{14}$$

and as $\Lambda : W_q \to C_0(\bar{Q})$ we choose the injection given by the continuous embedding. The functional $\mathcal{F}$ is convex and lower semicontinuous as a sum of functionals with these properties. Moreover, $\mathcal{G}$ is an indicator function of a closed and convex set in $C_0(\bar{Q})$ and thus fulfills the requirements as well. For $\tilde{v} = 0$ it holds that $S^*\tilde{v} = (0,0)$ and $\ell_{\alpha,\beta}(\tilde{v}) = 0$. This directly shows that $\mathcal{F}(\tilde{v}) < \infty$ and $\mathcal{G}(\Lambda\tilde{v}) < \infty$. We also know that $\mathcal{G}(\Lambda\tilde{v}) = \ell_{\alpha,\beta}(\Lambda(0))$ is continuous due to $\alpha,\beta > 0$, consequently the regular point condition holds. Thus we can apply the Fenchel duality theorem. Now the Fenchel conjugates have to be derived. From Fenchel Duality we know that $\mathcal{F}^*$ has the following form, if $\tilde{u} \in \partial\mathcal{F}(w)$:

$$\mathcal{F}^* : W_q^* \to \bar{\mathbb{R}}, \qquad \mathcal{F}^*(\tilde{u}) = \langle \tilde{u}, w\rangle_{\mathcal{M},C_0} - \mathcal{F}(w) \tag{15}$$

We can derive a representation of $\tilde{u} \in \partial\mathcal{F}(w)$ and reformulate:

$$\tilde{u} = S\left(|S^*w|^{p-2}\, S^*w + y_d\right) \quad \Leftrightarrow \quad S^*w = \operatorname{sign}(S^{-1}\tilde{u} - y_d)\,|S^{-1}\tilde{u} - y_d|^{\frac{1}{p-1}}, \tag{16}$$

where $\Lambda^* : (\mathcal{M}(\Omega) \times \mathcal{M}(Q)) \to W_q^*$ is the injection from the dual of $C_0(\bar{Q})$ in W_q^*. Inserting this into (15) we derive $\mathcal{F}^*(\tilde{u}) = {}^1/_q\,|S^{-1}\tilde{u} - y_d|^q_{L^q(Q)}$. To derive $\mathcal{G}^*(\tilde{u})$ we can deconstruct $\mathcal{G}(w)$, such that it consists of two summands, which represent an indicator function with only one constraint respectively, such that $(w(0), w) \in C_0(\bar{\Omega}) \times C_0(\bar{Q})$ and $(C_0(\bar{\Omega}) \times C_0(\bar{Q}))^* = C_0(\bar{\Omega})^* \times C_0(\bar{Q})^* = \mathcal{M}(\Omega) \times \mathcal{M}(Q)$. From [13, Theorem 2.2.8] we know that for $\tilde{u} = (0,u) + (u_0, 0) \in \mathcal{M}(\Omega) \times \mathcal{M}(Q)$ it holds that :

$$\mathcal{G}^*(\tilde{u}) = (\ell_\alpha + \ell_\beta)^*(u_0, u) = \ell^*_\alpha(0,u) + \ell^*_\beta(u_0, 0).$$

Looking at both conjugates separately, we derive:

$$\ell^*_\alpha(0,u) = \sup_{w\in W_q} \langle u, w\rangle_{\mathcal{M},C_0} - \ell_\alpha(0,w) = \sup_{w\in W_q,\, \|w\|_{C_0(\bar{Q})}\le\alpha} \langle u, w\rangle_{\mathcal{M},C_0} = \alpha\|u\|_{\mathcal{M}(Q)}$$

Analogous we can see that $\ell^*_\beta(u_0, 0) = \beta\|u_0\|_{\mathcal{M}(\Omega)}$ holds. Assembling the information we obtain $\mathcal{G}^*(\tilde{u}) = \alpha\|u\|_{\mathcal{M}(Q)} + \beta\|u_0\|_{\mathcal{M}(\Omega)}$. Finally we can use the Fenchel duality theorem and deduce:

$$\begin{aligned} & \min_{w \in W_q} \quad \frac{1}{p}\|S^* w\|^p_{L^p(Q)} + \langle S^* w, y_d\rangle_{L^p(Q),L^q(Q)} + \ell_{\alpha,\beta}(w) \\ = & \min_{\tilde{u} \in \mathcal{M}(\Omega)\times\mathcal{M}(Q)} \frac{1}{q}\|S^{-1}\tilde{u} - y_d\|^q_{L^q(Q)} + \alpha\|{-u}\|_{\mathcal{M}(Q)} + \beta\|{-u_0}\|_{\mathcal{M}(\Omega)}. \end{aligned}$$

As $S^{-1}\tilde{u} = \tilde{y}$ and $\tilde{u} = (u_0, u)$ we recover the problem (P). □

This means that we can solve (P^*) and use (16) and the determined unique optimal solution $\bar{w} \in W_q$ to calculate the optimal control $\hat{u} := (\bar{u}_0, \bar{u})$ for (P).

3 Discontinuous Galerkin Discretization

This section deals with the discontinuous Galerkin discretization of (P), which was introduced in [4]. The mentioned paper provides a convergence result for the discrete problem (P_σ) analogous to Theorem 2. After recapitulating the discrete setting from [4, Section 4] we will additionally discretize (P^*), reformulate the problem equivalently and derive an optimality system by a Lagrange approach. If the necessary conditions are fulfilled we can then apply a semismooth Newton method to solve the optimality system.

In [3] we can find the numerical solution of a problem with a similar parabolic state equation. More details on how to deal with the norm $\|\cdot\|_{\mathcal{M}(\Omega)}$ can be found, e.g. in [2] and [6]. Here, we merely treat $\|\cdot\|_{\mathcal{M}(Q)}$ as the dual norm of $\|\cdot\|_{C_0(\bar{Q})}$. The remaining difficulties are to transfer the results from a space norm to a space-time norm for the control and to deal with the occurring $L^q(Q)$ norm of the state. Additionally the appearance of the initial control u_0 in the objective functional will be discussed.

As a first step to characterizing the discrete spaces, we have to set up the space-time grid. Define the partition $0 = t_0 < t_1 < \ldots < t_{N_\tau} = T$. For the temporal grid the interval I is split in subintervals $I_k = (t_{k-1}, t_k]$ for $k = 1, \ldots, N_{\tau-1}$ and $I_{N_\tau} = (t_{N_\tau-1}, t_{N_\tau})$. The temporal gridsize is denoted by $\tau = \max_{0\le k\le N_\tau} \tau_k$, where $\tau_k := t_k - t_{k-1}$. Let $\mathcal{K}_h$ be a triangulation of Ω for a fixed $h > 0$. Then we define $\rho(K)$ to be the diameter of K and the gridsize of $\mathcal{K}_h$ as $h = \max_{K\in\mathcal{K}_h} \rho(K)$. We set $\bar{\Omega}_h = \bigcup_{K\in\mathcal{K}_h} K$ and denote by Ω_h the interior and by Γ_h the boundary of $\bar{\Omega}_h$. We assume that vertices on Γ_h are points on Γ. We can set up the space-time grid as $Q_h := \Omega_h \times (0, T)$ and define $Q_k := \Omega_h \times I_k$. To avoid having to use two indices, we define the discretization parameter

$\sigma = (\tau, h)$. The space discrete spaces U_h, Y_h and the space-time discrete spaces $\mathcal{U}_\sigma, \mathcal{Y}_\sigma$ can be found in [4, Chapter 4.1.]. We will only recall the representation of contained elements. Here $(\delta_{x_j})_{j=1}^{N_h}$ denote the Dirac-measures and $(e_{x_j})_{j=1}^{N_h}$ is the nodal basis formed by continuous piecewise linear functions. Thus $e_{x_j}(x_i) = \delta_{ij}$ holds. Let $u_{k,h} \in U_h$ and $y_{k,h} \in Y_h$. The elements u_σ and y_σ can be represented using an indicator function χ_k of I_k:

$$u_\sigma = \sum_{k=1}^{N_\tau} u_{k,h} \otimes \chi_k \quad \text{and} \quad y_\sigma = \sum_{k=1}^{N_\tau} y_{k,h} \otimes \chi_k. \tag{17}$$

Now inserting the definition of space-discrete elements $u_{k,h}$ and $y_{k,h}$ we obtain:

$$u_\sigma = \sum_{k,j=1}^{N_\tau, N_h} u_{k,j} \chi_k \, \delta_{x_j} \quad \text{and} \quad y_\sigma = \sum_{k,j=1}^{N_\tau, N_h} y_{k,j} \chi_k \, e_{x_j}. \tag{18}$$

Consequently $\mathcal{U}_\sigma$ and $\mathcal{Y}_\sigma$ are spaces of finite dimension $N_\sigma = N_\tau \times N_h$ with bases given by $\{\chi_k \, \delta_{x_j}\}_{k,j}$ and $\{\chi_k \, e_{x_j}\}_{k,j}$. In [4] an implicit Euler time stepping scheme is used to write down the discrete state equation. We know that $y_{k,h} = y_\sigma|_{I_k}$ for every $k \in \{1, \dots, N_\tau\}$. Let $(u_{0,h}, u_\sigma) \in U_h \times \mathcal{U}_\sigma$ be given and $z_h \in Y_h$ arbitrary. Then for $k \in \{1, \dots, N_\tau\}$ the following equations form the discrete state equation:

$$\begin{cases} \tau_k(\frac{y_{k,h} - y_{k-1,h}}{\tau_k}, z_h) + \tau_k \int_\Omega \nabla y_{k,h} \nabla z_h \, dx = \int_{Q_k} z_h \, \mathrm{d}u_\sigma, \\ y_{0,h} = y_{0h}, \end{cases} \tag{19}$$

where $y_{0h} \in Y_h$ is the unique element satisfying:

$$(y_{0h}, z_h) = \int_\Omega z_h \, \mathrm{d}u_{0,h} \qquad \forall \, z_h \in Y_h \tag{20}$$

Here $(\cdot, \cdot)$ denotes the scalar product in $L^2(\Omega)$. Now we can formulate (P_σ) as follows:

$$\min_{(u_{0h}, u_\sigma) \in \mathcal{U}_\sigma \times U_h} J_\sigma(u_{0h}, u_\sigma) = \frac{1}{q} \| y_\sigma(u_{0h}, u_\sigma) - y_d \|^q_{L^q(Q_h)} + \alpha \|u_\sigma\|_{\mathcal{M}(Q)} + \beta \|u_{0h}\|_{\mathcal{M}(\Omega)} \tag{P_σ}$$

where $y_\sigma(u_{0h}, u_\sigma)$ solves the discrete state equation (19). A central result we have for the discrete setting explained above is [4, Theorem 4.3.], which is the analogon to Theorem 2.

We move on to discretizing (P^*), which we can equivalently reformulate in the following way:

$$\begin{aligned} \min_{w \in W_q} \quad & \hat{K}(w) := \tfrac{1}{p}\|S^* w\|^p_{L^p(Q)} + \langle S^* w, y_d \rangle_{L^p(Q), L^q(Q)} \\ \text{s.t.} \quad & \|w\|_{C_0(\bar{Q})} - \alpha \leq 0 \quad \text{and} \quad \|w(0)\|_{C_0(\bar{\Omega})} - \beta \leq 0. \end{aligned}$$

The discrete representatives of $\|\cdot\|_{C_0(\bar{\Omega})}$ and $\|\cdot\|_{C_0(\bar{Q})}$ are the $\|\cdot\|_\infty$ norms in the respective domains. We set $\tilde{Y} = Y_h \times \mathcal{Y}_\sigma$ and $\tilde{U} = U_h \times \mathcal{U}_\sigma$. For $(y_{0,h}, y_\sigma)$ and y_d to be from the same discrete space, we define $y_{d,\sigma} := \begin{pmatrix} (y_d(x,0))_h & (y_d)_\sigma \end{pmatrix}^\top$. Here $(y_d(x,0))_h$ and $(y_d)_\sigma$ denote the evaluations of $y_d(x,0)$ and y_d on the inner nodes of Ω_h and Q_h respectively. We know the discrete problem (P^*) for $w = (w_{0,h}, w_\sigma)$:

$$\begin{aligned} \min_{w \in U_h^* \times \mathcal{U}_\sigma^*} \quad & K_\sigma(w) := \tfrac{1}{p}\|S_\sigma^* w\|^p_{L^p(Q_h)} + \langle S_\sigma^* w, y_{d,\sigma} \rangle_{L^p(Q_h), L^q(Q_h)} \\ \text{s.t.} \quad & \|w_\sigma\|_\infty - \alpha \leq 0 \quad \text{and} \quad \|w_{0,h}\|_\infty - \beta \leq 0. \end{aligned} \tag{P^*_σ}$$

In order to specify S_σ, we transform the discrete state equation from (19) into a matrix-vector-multiplication. For this purpose we will from now on identify elements from $\mathcal{U}_\sigma$ and $\mathcal{Y}_\sigma$ with vectors in $\mathbb{R}^{N_\sigma}$. From (18) we know that our discrete elements u_σ and y_σ can be expressed via their expansion coefficients $u_{k,j}$ and $y_{k,j}$ respectively. To simplify the notation we will define $u_k := (u_{k,1}, \dots, u_{k,N_h})^\top \in \mathbb{R}^{N_h}$ and can then write $u_\sigma = (u_1, \dots, u_{N_\tau})^\top \in \mathbb{R}^{N_\sigma}$. Analogous we get the vectors y_k and y_σ. The elements from U_h and Y_h are identified in the same way. To avoid complicating the notation, we will not add arrows above the vectors. Similar to [3] we can set up a solution matrix for the discrete state equation. One difference we need to consider is $u_{0,h} \neq 0$. This leads to an additional column and an additional row. We calculate the right hand side by inserting the discrete representations $u_\sigma = \sum_{i=1}^{N_h} \sum_{j=1}^{N_\tau} u_{j,i}\, \delta_{x_i} \otimes \chi_j$ and $z_h = \sum_{l=1}^{N_h} z_l\, e_{x_l}$:

$$\int_{Q_k} z_h \, \mathrm{d}u_\sigma = \tau_k \sum_{j=1}^{N_h} \sum_{l=1}^{N_h} \int_{\Omega_h} \langle u_{k,j}\, \delta_{x_j}, z_l\, e_{x_l} \rangle \, \mathrm{d}x = \tau_k \sum_{j=1}^{N_h} u_{k,j}\, z_j = \tau_k\, u_{k,h}^\top\, z_h.$$

We define $M_h := (\langle e_{x_j}, e_{x_k} \rangle)_{j,k=1}^{N_h}$ as the mass matrix and $A_h := (\int_\Omega \nabla e_{x_j} \nabla e_{x_k} \, \mathrm{d}x)_{j,k=1}^{N_h}$ as the stiffness matrix corresponding to Y_h. Also we notice that the "mass matrix" $(\langle \delta_{x_j}, e_{x_l} \rangle)_{j,l=1}^{N_h}$ is the identity in $\mathbb{R}^{N_h \times N_h}$. This delivers S_σ except for the first row, where we insert the relation $M_h\, y_{0,h} = u_{0,h}$. The discrete solution operator of the state equation $S_\sigma : \mathbb{R}^{N_\sigma + N_h} \to \mathbb{R}^{N_\sigma + N_h}$ is

represented by:

$$\begin{pmatrix} M_h & 0 & \dots & \dots & 0 \\ -M_h & M_h + \tau_1 A_h & & & \vdots \\ 0 & -M_h & M_h + \tau_2 A_h & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -M_h & M_h + \tau_{N_\tau} A_h \end{pmatrix} \begin{pmatrix} y_{0,h} \\ y_1 \\ y_2 \\ \vdots \\ y_{N_\tau} \end{pmatrix} = \begin{pmatrix} u_{0,h} \\ \tau_1 u_1 \\ \tau_2 u_2 \\ \vdots \\ \tau_{N_\tau} u_{N_\tau} \end{pmatrix}. \tag{21}$$

A representation of the adjoint is $S_\sigma^* = S_\sigma^\top$. We reformulate the $\|\cdot\|_\infty$ norms in (P^*) by:

$$\|w_\sigma\|_\infty - \alpha \leq 0 \quad \Leftrightarrow \quad \max_{k=1,\dots,N_\tau, j=1,\dots,N_h} \{w_{k,j}, -w_{k,j}\} - \alpha \leq 0,$$

and can now state $2 \cdot N_\sigma$ linear inequality constraints that are equal to the first inequality in (P^*). Analogously we can equivalently reformulate the second part by $2 \cdot N_h$ linear constraints. Setting $w_{k,j} = 0$ for $k = 0, \dots, N_\tau$ and $j = 1, \dots, N_h$ all inequalities are strictly fulfilled and the Slater condition is satisfied, as may be confirmed in [8, Definition 2.44]. We can proceed by setting up the corresponding Lagrangian L with multipliers $\lambda^1, \lambda^2 \in \mathbb{R}^{N_\sigma}$ and $\lambda^3, \lambda^4 \in \mathbb{R}^{N_h}$:

$$\begin{aligned} L(w, \lambda^1, \lambda^2, \lambda^3, \lambda^4) = {} & \tfrac{1}{p} \|S_\sigma^T w\|_{L^p(Q_h)}^p + \langle S_\sigma^\top w, y_{d,\sigma} \rangle_{L^p(Q_h), L^q(Q_h)} + \textstyle\sum_k^{N_\tau} \sum_{j=1}^{N_h} \lambda_{k,j}^1 (w_{k,j} - \alpha) \\ & + \textstyle\sum_k^{N_\tau} \sum_{j=1}^{N_h} \lambda_{k,j}^2 (-w_{k,j} - \alpha) + \sum_{j=1}^{N_h} \lambda_j^3 (w_{0,j} - \beta) + \sum_{j=1}^{N_h} \lambda_j^4 (-w_{0,j} - \beta). \end{aligned}$$

We can now form the optimality system using the Karush-Kuhn-Tucker conditions. These state that the partial differential of L by the main variable w has to be zero:

$$\frac{\partial L}{\partial w} = \frac{\partial K_\sigma(w)}{\partial w} + \begin{pmatrix} \lambda^3 \\ \lambda^1 \end{pmatrix} - \begin{pmatrix} \lambda^4 \\ \lambda^2 \end{pmatrix} = 0 \tag{22}$$

and for the inequality constraints we have the following complementary conditions:

$$\lambda_{k,j}^i ((-1)^{(i-1)} w_{k,j} - \alpha) = 0 \quad \wedge \quad \lambda_{k,j}^i \geq 0 \quad \wedge \quad ((-1)^{(i-1)} w_{k,j} - \alpha) \leq 0 \quad \forall k, \forall j, i \in \{1, 2\}, \tag{23}$$

$$\lambda_j^i ((-1)^{(i-1)} w_{0,j} - \beta) = 0 \quad \wedge \quad \lambda_j^i \geq 0 \quad \wedge \quad ((-1)^{(i-1)} w_{0,j} - \beta) \leq 0 \quad \forall j, i \in \{3, 4\}. \tag{24}$$

From Fenchel duality we know that $\tilde{u} \in \frac{\partial K_\sigma(w)}{\partial w} = \partial \mathcal{F}_\sigma(w)$, because this is the discrete representative of $\mathcal{F}$. In [3] this property is used to recover a problem with the variable $\tilde{u}$. Consequently the constraint $\tilde{u} = \partial \mathcal{F}_\sigma(w)$ is added. Theoretically we could do the same, but as we want to use

derivative based methods, we will have to differentiate all constraints again. As $p > 2$ we can see that in (16) the exponent $\frac{1}{p-1}$ is strictly smaller than 1, which is problematic. Instead we will solve for w and recover the optimal control $\tilde{u}$ afterwards.
We would like to apply a semismooth Newton method ([10, Algorithm 2.11]), which is a way to find x^* solving $F(x^*) = 0$, to the optimality system (22)-(24), thus we reformulate (23) - (24) equivalently for all $k = 1, \dots, N_\tau$ and all $j = 1, \dots, N_h$:

$$N^1_{k,j} := \max\left\{0,\ \lambda^1_{k,j} + \kappa(w_{k,j} - \alpha)\right\} - \lambda^1_{k,j} = 0, \quad N^2_{k,j} := \max\left\{0,\ \lambda^2_{k,j} + \kappa(-w_{k,j} - \alpha)\right\} - \lambda^2_{k,j} = 0,$$
$$N^3_j := \max\left\{0,\ \lambda^3_j + \kappa(w_{0,j} - \beta)\right\} - \lambda^3_j = 0, \quad N^4_j := \max\left\{0,\ \lambda^4_j + \kappa(-w_{0,j} - \beta)\right\} - \lambda^4_j = 0.$$

This allows us to define $F(w, \lambda^1, \lambda^2, \lambda^3, \lambda^4) = \begin{pmatrix} \partial L/\partial w & N^1 & N^2 & N^3 & N^4 \end{pmatrix}^\top \in \mathbb{R}^{3\cdot(N_\sigma + N_h)}$ containing the left sides of our optimality system and solve the equation $F(w, \lambda^1, \lambda^2, \lambda^3, \lambda^4) = 0$, which will deliver the optimal solution of the problem (P^*_σ). Hereafter we denote Clarke's generalized Jacobian as in [10, Example 2.4] by ∂F:

$$\partial F(x) := \operatorname{conv}\left\{M \ :\ x_k \overset{k\to\infty}{\longrightarrow} x,\ F'(x_k) \longrightarrow M,\ F \text{ differentiable at } x_k\right\}$$

The next step is to write down $DF(w, \lambda^1, \lambda^2, \lambda^3, \lambda^4)$, which has the following structure:

$$DF(w, \lambda^1, \lambda^2, \lambda^3, \lambda^4) = \begin{pmatrix} \frac{\partial^2 L}{\partial^2 w^2} & \frac{\partial^2 L}{\partial w \partial \lambda^1} & \frac{\partial^2 L}{\partial w \partial \lambda^2} & \frac{\partial^2 L}{\partial w \partial \lambda^3} & \frac{\partial^2 L}{\partial w \partial \lambda^4} \\ \frac{N^1}{\partial w} & \frac{N^1}{\partial \lambda^1} & 0 & 0 & 0 \\ \frac{N^2}{\partial w} & 0 & \frac{N^2}{\partial \lambda^2} & 0 & 0 \\ \frac{N^3}{\partial w} & 0 & 0 & \frac{N^3}{\partial \lambda^3} & 0 \\ \frac{N^4}{\partial w} & 0 & 0 & 0 & \frac{N^4}{\partial \lambda^4}. \end{pmatrix} \tag{25}$$

The first row of the matrix can be calculated by derivation rules. Due to the max-norms the differentials of N^i for $i \in \{1, 2, 3, 4\}$ are not distinct. For an arbitrary max-norm the generalized Jacobian is:

$$\frac{\partial}{\partial x}(\max\{0, g(x)\}) = \begin{cases} 0, & \text{if } g(x) < 0, \\ [0, \frac{\partial g(x)}{\partial x}], & \text{if } g(x) = 0, \\ \frac{\partial g(x)}{\partial x}, & \text{if } g(x) > 0. \end{cases}$$

We decide to always choose $\frac{\partial g(x)}{\partial x}$, if $g(x) = 0$. Using this the remaining blocks can be calculated. An interesting observation is that for $\kappa = 1$ we have a symmetric matrix DF on the active sets. As we will solve for the optimal adjoint $\bar{w}$ we need to recover the optimal control $\bar{\tilde{u}}$ through the

discrete version of (16):

$$\bar{\bar{u}} = |S_\sigma^\top \bar{w}|^{p-2}(S_\sigma^\top \bar{w}) + y_{d,\sigma} \tag{26}$$

4 Variational Discretization

Here we want to achieve the desired maximal discrete sparsity, i.e. dirac-measures in space-time, by choosing the Petrov-Galerkin-Ansatz and -Test space that will induce this structure. The variational discretization concept was introduced in [9] and its key feature is to not discretize the control space. Instead, via the discretization of the test space and the optimality conditions, an implicit discretization of the control is achieved. This is how we *control* the discrete structure of the controls, as mentioned in the introduction. Looking at the relation (16) between the adjoint state and the control it becomes even more obvious, that the discrete structure of the test space affects the structure of the control. We set up the discrete spaces, but differently to [4] we will define a test space $\mathcal{V}_\sigma$ consisting of continuous and piecewise linear functions in space and time. This is motivated by the fact that the controls $(u_0, u) \in \mathcal{M}(\Omega) \times \mathcal{M}(Q)$ will be affected by the structure of these test functions. Whether the choice of continuous and piecewise quadratic functions leads to even better results will be part of further research. Afterwards we will set up the discrete state equation and the discrete problem. The main result will be a convergence result similar to the main result of [4].
The spaces Y_h and $\mathcal{Y}_\sigma$ remain the same. We define the test space:

$$\mathcal{V}_\sigma := \{v_\sigma \in C(I; Y_h) : v_\sigma|_{I_k} \in \mathcal{P}_1(I_k, Y_h), 1 \le k \le N_\tau \text{ and } v_\sigma(T) = 0\} \subset W_q. \tag{27}$$

Any element from $\mathcal{V}_\sigma$ can be written as $v_\sigma = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} v_{k,j}\, e_{t_k} \otimes e_{x_j}$ with $(e_{t_k} \otimes e_{x_j})(x,t) := e_{x_j}(x)\, e_{t_k}(t)$, where $(e_{t_k})_{k=1}^{N_\tau}$ is the nodal basis formed by continuous piecewise linear functions on the time grid. To analyze how the controls (u_0, u) are implicitly discretized, we look at the optimality conditions (Lemma 3) and the sparsity structure (Remark 4) in the continuous setting. We can now variationally discretize these conditions with $y \in \mathcal{Y}_\sigma$ and $w = (w_{0,h}, w_\sigma) \in Y_h \times \mathcal{V}_\sigma$. We know:

$$\|\bar{w}_\sigma\|_\infty \le \alpha \qquad \text{and} \qquad \|\bar{w}_{0,h}\|_\infty \le \beta.$$

Additionally in the discrete setting we deduce:

$$\operatorname{supp}(\bar{u}_0^+) \subset \{x \in \bar{\Omega} : \bar{w}_{0,h}(x) = -\beta\}, \qquad \operatorname{supp}(\bar{u}_0^-) \subset \{x \in \bar{\Omega} : \bar{w}_{0,h}(x) = +\beta\},$$
$$\operatorname{supp}(\bar{u}^+) \subset \{(x,t) \in \bar{Q} : \bar{w}_\sigma(x,t) = -\alpha\}, \qquad \operatorname{supp}(\bar{u}^-) \subset \{(x,t) \in \bar{Q} : \bar{w}_\sigma(x,t) = +\alpha\}.$$

Here we see that the variational discretization concept delivers that the structure of the test space $\mathcal{V}_\sigma$ has an affect on the structure of the controls (u_0, u). In the chosen case for the discrete adjoint state w the maximal values $\pm\alpha$ and $\pm\beta$, in the generic case, can only be attained at grid points. Consequently we know:

$$\operatorname{supp}(\bar{u}) \subset \{(x_j, t_k)\}_{j=1,k=1}^{N_h,N_\tau} \qquad \text{and} \qquad \operatorname{supp}(\bar{u}_0) \subset \{(x_j)\}_{j=1}^{N_h}.$$

Hence we define sets, whose elements are sums of dirac measures on the grid points:

$$U_h = \left\{u_h \in \mathcal{M}(\Omega) : u_h = \textstyle\sum_{j=1}^{N_h} u_j\, \delta_{x_j}, \text{ with } u_j \in \mathbb{R}\right\},$$
$$\mathcal{U}_{\mathrm{vd}} = \left\{u_\sigma \in \mathcal{M}(Q) : u_\sigma = \textstyle\sum_{k,j=1}^{N_\tau,N_h} u_{k,j}\, \delta_{x_j} \otimes \delta_{t_k}, \text{ with } u_{k,j} \in \mathbb{R}\right\}.$$

The space U_h is the same as before, but $\mathcal{U}_{\mathrm{vd}}$ is not the same as $\mathcal{U}_\sigma$. We will now cite [4, Proposition 4.1.], which only depends on the space U_h, without proof.

Lemma 7. *Let the linear operators Λ_h and Π_h be defined as below:*

$$\Lambda_h : \mathcal{M}(\Omega) \to U_h \subset \mathcal{M}(\Omega), \qquad \Lambda_h u_0 = \textstyle\sum_j \langle u_0, e_{x_j}\rangle\, \delta_{x_j}$$
$$\Pi_h : C(\bar{\Omega}) \to Y_h \subset C(\bar{\Omega}), \qquad \Pi_h y = \textstyle\sum_j y(x_j)\, e_{x_j}$$

Then for every $u_0 \in \mathcal{M}(\Omega)$ and every $y \in C(\bar{\Omega})$ and $y_h \in Y_h$ the following properties hold.

$$\langle u_0, y_h\rangle = \langle \Lambda_h u_0, y_h\rangle, \tag{28}$$
$$\langle u_0, \Pi_h y\rangle = \langle \Lambda_h u_0, y\rangle, \tag{29}$$
$$\|\Lambda_h u_0\|_{\mathcal{M}(\Omega)} \le \|u_0\|_{\mathcal{M}(\Omega)}, \tag{30}$$
$$\Lambda_h u_0 \overset{*}{\rightharpoonup} u_0 \in \mathcal{M}(\Omega) \quad \textit{and} \quad \|\Lambda_h u_0\|_{\mathcal{M}(\Omega)} \overset{h\to 0}{\longrightarrow} \|u_0\|_{\mathcal{M}(\Omega)}. \tag{31}$$

We derive an analogous result for the space-time discrete spaces $\mathcal{U}_{\mathrm{vd}}$ and $\mathcal{Y}_\sigma$. Due to the different test space, we can not simply copy [4, Proposition 4.2.]. Hence we will state a theorem adjusted to the changes. The structure of the proof remains the same, only the technical calcula-

tions are different.

Lemma 8. *Let the linear operators Υ_{vd} and Ψ_{vd} be defined as below:*

$$\Upsilon_{\mathrm{vd}} : \mathcal{M}(Q) \to \mathcal{U}_{\mathrm{vd}} \subset \mathcal{M}(Q), \qquad \Upsilon_{\mathrm{vd}}\, u = \textstyle\sum_{k,j} \delta_{x_j} \otimes \delta_{t_k} \int_{Q_k} e_{t_j}\, \mathrm{d}u$$
$$\Psi_{\mathrm{vd}} : C(\bar{Q}) \ \to \mathcal{Y}_\sigma, \qquad \Psi_{\mathrm{vd}}\, y = \textstyle\sum_{k,j} y(x_j, t_k)\, e_{x_j} \otimes \chi_k$$

Then for every $u \in \mathcal{M}(Q), y \in C(\bar{Q})$ and arbitrary $y_\sigma \in \mathcal{Y}_\sigma$ the following properties hold.

$$\langle u, y_\sigma \rangle \qquad = \langle \Upsilon_{\mathrm{vd}}\, u, y_\sigma \rangle, \tag{32}$$
$$\langle u, \Psi_{\mathrm{vd}}\, y \rangle \quad = \langle \Upsilon_{\mathrm{vd}}\, u, y \rangle, \tag{33}$$
$$\|\Upsilon_{\mathrm{vd}}\, u\|_{\mathcal{M}(Q)} \le \|u\|_{\mathcal{M}(Q)}, \tag{34}$$
$$\Upsilon_{\mathrm{vd}}\, u \overset{*}{\rightharpoonup} u \ \in \ \mathcal{M}(Q) \quad \textit{and} \quad \|\Upsilon_{\mathrm{vd}}\, u\|_{\mathcal{M}(Q)} \overset{|\sigma|\to 0}{\longrightarrow} \|u\|_{\mathcal{M}(Q)} \tag{35}$$

The next step is to set up the new discrete state equation. To this end we start by deriving a very weak formulation of (1) which will be discretized afterwards. By multiplication with $z \in W_q$, integration over the domain Q, and utilizing $z(x, T) = 0$ and $y(x, 0) = 0$, we arrive at

$$A(y, z) \coloneqq \int_Q \left(- y\, \tfrac{\partial z}{\partial t} + \nabla y \nabla z \right) \mathrm{d}x\, \mathrm{d}t = \int_\Omega z(\cdot, 0)\, \mathrm{d}u_0 + \int_Q z\, \mathrm{d}u. \tag{36}$$

We can now discretize by inserting $y_\sigma \in \mathcal{Y}_\sigma$ and testing for $z_\sigma \in \mathcal{V}_\sigma$. This delivers the following discrete representation of the state equation: Find $y_\sigma \in \mathcal{Y}_\sigma$, such that

$$A(y_\sigma, z_\sigma) = \int_\Omega z_\sigma(\cdot, 0)\, \mathrm{d}u_0 + \int_Q z_\sigma\, \mathrm{d}u \qquad \forall z_\sigma \in \mathcal{V}_\sigma. \tag{37}$$

We can formulate the discrete problem (P_{vd}):

$$\min_{(u_0,u) \in \mathcal{M}(\Omega) \times \mathcal{M}(Q)} J_{\mathrm{vd}}(u_0, u) = \tfrac{1}{q} \|y_\sigma(u_0, u) - y_d\|^q_{L^q(Q_h)} + \alpha \|u\|_{\mathcal{M}(Q)} + \beta \|u_0\|_{\mathcal{M}(\Omega)} \tag{P_{vd}}$$

where $y_\sigma(u_0, u)$ solves the discrete state equation (37).
We observe, that J_{vd} is convex, but not strictly convex like J. In the continuous setting the strict convexity came from the norm $\|\cdot\|_{L^q(Q_h)}$, but the mapping from the control to the discrete state is not injective. Consequently the uniqueness of the solution cannot be concluded. In the following Theorem 9 we prove the existence of solutions and discuss uniqueness in the discrete setting as done in [3, Section 4.3.].

Theorem 9. *The problem* (P_{vd}) *has at least one solution in* $\mathcal{M}(\Omega) \times \mathcal{M}(Q)$ *and there exists a unique solution* $(\bar{u}_{0,h}, \bar{u}_\sigma) \in U_h \times \mathcal{U}_{\mathrm{vd}}$. *Furthermore we know for any solution* $(\hat{u}_0, \hat{u}) \in \mathcal{M}(\Omega) \times \mathcal{M}(Q)$ *to* (P_{vd}) *it holds* $(\Lambda_h \hat{u}_0, \Upsilon_{\mathrm{vd}} \hat{u}) = (\bar{u}_{0,h}, \bar{u}_\sigma)$.

Proof. The existence of a solution can be derived as in the proof of [4, Theorem 2.7.], because the control domain is still continuous. Let $(\hat{u}_0, \hat{u}) \in \mathcal{M}(\Omega) \times \mathcal{M}(Q)$ be such a solution to (P_{vd}) and define $(\bar{u}_{0,h}, \bar{u}_\sigma) := (\Lambda_h \hat{u}_0, \Upsilon_{\mathrm{vd}} \hat{u}) \in U_h \times \mathcal{U}_{\mathrm{vd}}$. We can deduce from (28) and (32) that

$$y_\sigma(u_0, u) = y_\sigma(\Lambda_h u_0, \Upsilon_{\mathrm{vd}} u) \quad \forall\, (u_0, u) \in \mathcal{M}(\Omega) \times \mathcal{M}(Q). \tag{38}$$

Additionally (30) and (34) deliver $\|\bar{u}_{0,h}\|_{\mathcal{M}(\Omega)} \leq \|\hat{u}_0\|_{\mathcal{M}(\Omega)}$ and $\|\bar{u}_\sigma\|_{\mathcal{M}(Q)} \leq \|\hat{u}\|_{\mathcal{M}(Q)}$. Combining these properties we can deduce $J_{\mathrm{vd}}(\bar{u}_{0,h}, \bar{u}_\sigma) \leq J_{\mathrm{vd}}(\hat{u}_0, \hat{u})$. Therefore $(\bar{u}_{0,h}, \bar{u}_\sigma) \in U_h \times \mathcal{U}_{\mathrm{vd}}$ is a solution of (P_{vd}) and we proved the existence of solutions in the discrete space $U_h \times \mathcal{U}_{\mathrm{vd}}$. The mapping $(\bar{u}_{0,h}, \bar{u}_\sigma) \mapsto y_\sigma(\bar{u}_{0,h}, \bar{u}_\sigma)$, where $y_\sigma(\bar{u}_{0,h}, \bar{u}_\sigma)$ solves (37) for $(u_0, u) = (\bar{u}_{0,h}, \bar{u}_\sigma) \in U_h \times \mathcal{U}_{\mathrm{vd}}$, is linear, injective and we know that $\dim U_h = \dim Y_h$ and $\dim \mathcal{U}_{\mathrm{vd}} = \dim \mathcal{Y}_\sigma$. Hence this mapping is bijective. Therefore the functional J_{vd} is strictly convex on $U_h \times \mathcal{U}_{\mathrm{vd}}$ and consequently (P_{vd}) has a unique solution $(\bar{u}_{0,h}, \bar{u}_\sigma) \in U_h \times \mathcal{U}_{\mathrm{vd}}$.
From the uniqueness in the discrete space and the fact that any projection of a continuous solution $(\Lambda_h \hat{u}_0, \Upsilon_{\mathrm{vd}} \hat{u}) \in U_h \times \mathcal{U}_{\mathrm{vd}}$ is a solution in the discrete space, we deduce that all projections must be equal, i.e. $(\Lambda_h \hat{u}_0, \Upsilon_{\mathrm{vd}} \hat{u}) = (\bar{u}_{0,h}, \bar{u}_\sigma)$ for any solution $(\hat{u}_0, \hat{u}) \in \mathcal{M}(\Omega) \times \mathcal{M}(Q)$. □

Since all projections of solutions yield the unique discrete solution $(\bar{u}_{0,h}, \bar{u}_\sigma) \in U_h \times \mathcal{U}_{\mathrm{vd}}$ it suffices to analyze the convergence properties for this discrete solution. Furthermore for the computational results we can use the representations $\bar{u}_{0,h} = \sum_{j=1}^{N_h} \bar{u}_{0,j}\, \delta_{x_j}$ and $\bar{u}_\sigma = \sum_{k,j=1}^{N_\tau, N_h} \bar{u}_{k,j}\, \delta_{x_j} \otimes \delta_{t_k}$ and uniquely determine the discrete optimal control by calculating the coefficients $\bar{u}_{k,j}$ for $k = 0, \ldots, N_\tau$ and $j = 1, \ldots, N_h$.
We can now prove the convergence result formulated in Theorem 2 along the lines of the proof of [4, Theorem 4.3.].

Proof. By the coercivity of J_{vd} we know that $\{(\bar{u}_{0,h}, \bar{u}_\sigma)\}_\sigma$ is bounded in $\mathcal{M}(\Omega) \times \mathcal{M}(Q)$ and consequently [4, Theorem 2.2.] delivers the boundedness of $\{\bar{y}_\sigma\}_\sigma$ in $L^q(Q)$. Therefore, there exist subsequences, such that for $|\sigma| \to 0$ the following holds true

$$(\bar{u}_{0,h}, \bar{u}_\sigma) \overset{*}{\rightharpoonup} (\tilde{u}_0, \tilde{u}) \in \mathcal{M}(\Omega) \times \mathcal{M}(Q) \quad \text{and} \quad \bar{y}_\sigma \rightharpoonup \tilde{y} \in L^q(Q). \tag{39}$$

As in [4, Theorem 4.3.] we will split the proof into several steps.

I - $\tilde{y}$ *is the solution of* (37) *corresponding to* $(\tilde{u}_0, \tilde{u})$
By the denseness of $\{\xi \in C^1(0,T) : \xi(T) = 0\} \otimes (W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega))$ in $\mathcal{V}$, it is sufficient to test (36) against $z = \varphi \otimes \xi$ with $\xi \in C^1(0,T)$ satisfying $\xi(T) = 0$ and $\varphi \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$. Let φ be approximated by $\varphi_h \in Y_h$, such that

$$\int_\Omega \langle \nabla \varphi_h, \nabla z_h \rangle_{\mathbb{R}^d} \,\mathrm{d}x = \int_\Omega \langle \nabla \varphi, \nabla z_h \rangle_{\mathbb{R}^d} \,\mathrm{d}x \quad \text{for all } z_h \in Y_h \quad \text{and} \quad \|\varphi - \varphi_h\|_{C(\bar{\Omega})} \overset{h\to 0}{\longrightarrow} 0. \tag{40}$$

Moreover, let $\xi_\tau = \sum_k \xi(t_k)\, e_{t_k}$ be the piecewise linear interpolation of ξ so that $\xi_\tau \to \xi$ in $C(\bar{Q})$ and $\xi_\tau' \to \xi'$ in L^∞. Testing (37) against $z_\sigma = \varphi_h \otimes \xi_\tau$, we obtain

$$A(\bar{y}_\sigma, z_\sigma) = \int_\Omega z_\sigma(\cdot, 0) \,\mathrm{d}\bar{u}_{0,h} + \int_Q z_\sigma \,\mathrm{d}\bar{u}_\sigma. \tag{41}$$

On the right hand side, we can perform the limit directly:

$$\int_\Omega z_\sigma \,\mathrm{d}\bar{u}_{0,h} + \int_Q z_\sigma \,\mathrm{d}\bar{u}_\sigma \overset{|\sigma|\to 0}{\longrightarrow} \int_\Omega z \,\mathrm{d}\tilde{u}_0 + \int_Q z \,\mathrm{d}\tilde{u}.$$

The left hand side of (41) can be expanded to

$$A(\bar{y}_\sigma, z_\sigma) = -\int_Q \bar{y}_\sigma \,(\varphi_h \otimes \xi_\tau')\,\mathrm{d}x\,\mathrm{d}t + \int_Q \nabla \bar{y}_\sigma(x,t)\, \nabla \varphi_h(x)\, \xi_\tau(t)\,\mathrm{d}x\,\mathrm{d}t. \tag{42}$$

Applying the very definition of φ_h and integration by parts, we observe that

$$\int_Q \nabla \bar{y}_\sigma(x,t)\, \nabla \varphi_h(x)\, \xi_\tau(t)\,\mathrm{d}x\,\mathrm{d}t = -\int_Q \bar{y}_\sigma \,(\Delta\, \varphi \otimes \xi_\tau)\,\mathrm{d}x\,\mathrm{d}t \overset{|\sigma|\to 0}{\longrightarrow} -\int_Q \tilde{y}\ \Delta\, z \,\mathrm{d}x\,\mathrm{d}t.$$

Along with $-\int_Q \bar{y}_\sigma \,(\varphi_h \otimes \xi')\,\mathrm{d}x\,\mathrm{d}t \overset{|\sigma|\to 0}{\longrightarrow} -\int_Q \tilde{y}\,(\varphi \otimes \xi')\,\mathrm{d}x\,\mathrm{d}t = -\int_Q \tilde{y}\, \frac{\partial z}{\partial t}\,\mathrm{d}x\,\mathrm{d}t$, this implies that $A(\bar{y}_\sigma, z_\sigma) \overset{|\sigma|\to 0}{\longrightarrow} A(\tilde{y}, z)$ and thus $A(\tilde{y}, z) = \int_\Omega z \,\mathrm{d}\tilde{u}_0 + \int_Q z \,\mathrm{d}\tilde{u}$ for all tensor products $z = \varphi \otimes \xi$.
II - $J(\tilde{u}_0, \tilde{u}) \le J(u_0, u) \ \ \forall\, (u_0, u) \in C(\bar{\Omega}) \times C(\bar{Q})$
From [4] we know that an associated solution y to (1) for regular controls (u_0, u) belongs to $L^2(0,T; H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(Q)$. Additionally from [4, page 10] we know in the case of regular controls

$$y_\sigma \to y \ \in L^2(0,T; H_0^1(\Omega)) \overset{q<2}{\subset} L^q(Q). \tag{43}$$

Now we set $(u_{0,h}, u_\sigma) = (\Lambda_h u_0, \Upsilon_{\mathrm{vd}} u)$ for this step. Using (38) and the convergence properties,

we can also see for $|\sigma| \to 0$

$$J_{\mathrm{vd}}(u_{0,h}, u_\sigma) = \underbrace{\frac{1}{q}\|y_\sigma(u_{0,h}, u_\sigma) - y_d\|^q_{L^q(Q)}}_{\overset{(43)}{\to} \frac{1}{q}\|y - y_d\|^q_{L^q(Q)}} + \underbrace{\alpha\|u_\sigma\|_{\mathcal{M}(Q)}}_{\overset{(35)}{\to} \alpha\|u\|_{\mathcal{M}(Q)}} + \underbrace{\beta\|u_{0,h}\|_{\mathcal{M}(\Omega)}}_{\overset{(31)}{\to} \beta\|u_0\|_{\mathcal{M}(\Omega)}} \to J(u_0, u) \tag{44}$$

For the final estimation we will use that $(\bar{u}_{0,h}, \bar{u}_\sigma)$ solves (P_{vd}).

$$J(\tilde{u}_0, \tilde{u}) \overset{(39)}{\leq} \liminf_{|\sigma|\to 0} J_{\mathrm{vd}}(\bar{u}_{0,h}, \bar{u}_\sigma) \leq \liminf_{|\sigma|\to 0} J_{\mathrm{vd}}(u_{0,h}, u_\sigma) \overset{(44)}{=} J(u_0, u) \tag{45}$$

III - $(\tilde{u}_0, \tilde{u}) = (\bar{u}_0, \bar{u})$

We know that the solution to (P) is unique for $q > 1$. Thus, it suffices to show that $(\tilde{u}_0, \tilde{u})$ solves the problem. Choose a sequence $\{(u_{0k}, u_k)\}_k \in C(\bar{\Omega}) \times C(\bar{Q})$, such that

$$(u_{0k}, u_k) \overset{*}{\rightharpoonup} (\bar{u}_0, \bar{u}) \quad \in \mathcal{M}(\Omega) \times \mathcal{M}(Q), \tag{46}$$

$$\|u_{0k}\|_{L^1(\Omega)} = \|u_{0k}\|_{\mathcal{M}(\Omega)} \leq \|\bar{u}_0\|_{\mathcal{M}(\Omega)} \quad \forall k, \tag{47}$$

$$\|u_k\|_{L^1(Q)} = \|u_k\|_{\mathcal{M}(\Omega)} \leq \|\bar{u}\|_{\mathcal{M}(Q)} \quad \forall k. \tag{48}$$

From [4, Lemma 2.6.] we know that in this setting the sequence $\{y_k\}_k$ converges strongly to $\bar{y}(\bar{u}_0, \bar{u})$. The weak* convergence property (46) delivers the following estimates:

$$\|\bar{u}_0\|_{\mathcal{M}(\Omega)} \leq \liminf_{k\to\infty} \|u_{0k}\|_{\mathcal{M}(\Omega)} \overset{(47)}{\leq} \|\bar{u}_0\|_{\mathcal{M}(\Omega)} \quad \text{and} \quad \|\bar{u}\|_{\mathcal{M}(Q)} \leq \liminf_{k\to\infty} \|u_k\|_{\mathcal{M}(Q)} \overset{(48)}{\leq} \|\bar{u}\|_{\mathcal{M}(Q)}.$$

Hence, $\|u_{0k}\|_{\mathcal{M}(\Omega)} \to \|\bar{u}_0\|_{\mathcal{M}(\Omega)}$ and $\|u_k\|_{\mathcal{M}(Q)} \to \|\bar{u}\|_{\mathcal{M}(Q)}$. Analogously as in step **II**, we can now deduce that

$$J(u_{0k}, u_k) \overset{k\to\infty}{\longrightarrow} J(\bar{u}_0, \bar{u}). \tag{49}$$

As this sequence consists of regular controls we know from (45) that $J(\tilde{u}_0, \tilde{u}) \leq J(\bar{u}_0, \bar{u})$. Due to the uniqueness of the solution it is evident that $(\tilde{u}_0, \tilde{u}) = (\bar{u}_0, \bar{u})$ and we can deduce:

$$J(\tilde{u}_0, \tilde{u}) \ = J(\bar{u}_0, \bar{u}) \overset{(39)}{\leq} \liminf_{|\sigma|\to 0} J_{\mathrm{vd}}(\bar{u}_{0,h}, \bar{u}_\sigma) \overset{(45)}{\leq} \liminf_{k\to\infty} J(u_{0k}, u_k) \overset{(49)}{=} J(\bar{u}_0, \bar{u}).$$

This shows $\lim_{|\sigma|\to 0} J_{\mathrm{vd}}(\bar{u}_{0,h}, \bar{u}_\sigma) = J(\bar{u}_0, \bar{u})$ and from (39) we also know $\bar{y}_\sigma \rightharpoonup \bar{y} \in L^q(Q)$.

IV - *proof of* (3)*,* (4) *and* (5)

The convergence $\lim_{|\sigma|\to 0} J_{\mathrm{vd}}(\bar{u}_{0,h}, \bar{u}_\sigma) = J(\bar{u}_0, \bar{u})$ gives (4). We can calculate

$$\begin{aligned}
\tfrac{1}{q}\|\bar{y} - y_d\|^q_{L^q(Q)} &\le \liminf_{|\sigma|\to 0} \tfrac{1}{q}\|\bar{y}_\sigma - y_d\|^q_{L^q(Q)} \le \limsup_{|\sigma|\to 0} \left(J_{\mathrm{vd}}(\bar{u}_{0,h}, \bar{u}_\sigma) - \alpha\|\bar{u}_\sigma\|_{\mathcal{M}(Q)} - \beta\|\bar{u}_{0,h}\|_{\mathcal{M}(\Omega)}\right) \\
&\le \limsup_{|\sigma|\to 0} J_{\mathrm{vd}}(\bar{u}_{0,h}, \bar{u}_\sigma) - \liminf_{|\sigma|\to 0} \left(\alpha\|\bar{u}_\sigma\|_{\mathcal{M}(Q)} + \beta\|\bar{u}_{0,h}\|_{\mathcal{M}(\Omega)}\right) \\
&= J(\bar{u}_0, \bar{u}) - \liminf_{|\sigma|\to 0} \left(\alpha\|\bar{u}_\sigma\|_{\mathcal{M}(Q)} + \beta\|\bar{u}_{0,h}\|_{\mathcal{M}(\Omega)}\right) \\
&\overset{(39)}{\le} J(\bar{u}_0, \bar{u}) - \left(\alpha\|\bar{u}\|_{\mathcal{M}(Q)} + \beta\|\bar{u}\|_{\mathcal{M}(\Omega)}\right) = \tfrac{1}{q}\|\bar{y} - y_d\|^q_{L^q(Q)}.
\end{aligned}$$

Combined with the weak convergence in $L^q(Q)$ this shows the strong convergence (3). In a similar way we can prove the first part of (5).

$$\begin{aligned}
\alpha\|\bar{u}\|_{\mathcal{M}(Q)} &\overset{(39)}{\le} \liminf_{|\sigma|\to 0} \alpha\|\bar{u}_\sigma\|_{\mathcal{M}(Q)} \le \limsup_{|\sigma|\to 0} \left(J_{\mathrm{vd}}(\bar{u}_{0,h}, \bar{u}_\sigma) - \tfrac{1}{q}\|\bar{y}_\sigma - y_d\|^q_{L^q(Q)} - \beta\|\bar{u}_{0,h}\|_{\mathcal{M}(\Omega)}\right) \\
&\overset{(3)}{=} J(\bar{u}_0, \bar{u}) - \tfrac{1}{q}\|\bar{y} - y_d\|^q_{L^q(Q)} - \liminf_{|\sigma|\to 0} \left(\beta\|\bar{u}_{0,h}\|_{\mathcal{M}(\Omega)}\right) \\
&\overset{(39)}{\le} J(\bar{u}_0, \bar{u}) - \tfrac{1}{q}\|\bar{y} - y_d\|^q_{L^q(Q)} - \beta\|\bar{u}_0\|_{\mathcal{M}(\Omega)} = \alpha\|\bar{u}\|_{\mathcal{M}(Q)}.
\end{aligned}$$

Finally, the remaining part of (5) follows directly from $\lim_{|\sigma|\to 0} J_{\mathrm{vd}}(\bar{u}_{0,h}, \bar{u}_\sigma) = J(\bar{u}_0, \bar{u})$ and the fact that we already showed the convergence of the other two terms. □

In order to solve (P^*_{vd}) numerically we want to represent (37) by a matrix vector multiplication. From [7, Section 4] we know that this will deliver a Crank-Nicholson scheme with a smoothing step. Setting $z_{k,h} \coloneqq z_\sigma(\cdot, t_k) \in Y_h$ and $z_k \coloneqq z_{k,h} \otimes e_{t_k} \in \mathcal{V}_\sigma$, we obtain the left hand side of (37):

$$A(y_\sigma, z_k) = (y_{k+1,h} - y_{k,h})^\top M_h\, z_{k,h} + \left(\tfrac{\tau_k}{2}\, y_{k,h} + \tfrac{\tau_{k+1}}{2}\, y_{k+1,h}\right)^\top A_h\, z_{k,h} \qquad \forall\, k \in \{1, \dots, N_\tau - 1\}.$$

The next step is to calculate $r(z_\sigma) \coloneqq \int_\Omega z_\sigma(x, 0)\, \mathrm{d}u_0 + \int_Q z_\sigma\, \mathrm{d}u$ for the basis functions z_k. Keeping the implicit discrete structure of the controls (u_0, u) in mind, we identify u_0 with $u_{0,h} \in U_h$ and u with $\sum_k u_{k,h} \otimes \delta_{t_k} \in \mathcal{U}_{\mathrm{vd}}$. Additionally we know that $\langle \delta_{x_j} \otimes \delta_{t_k}, e_i \otimes e_{t_l} \rangle^{N_\tau, N_h}_{k,l=1,j,i=1} = I_{N_\sigma}$. Thus we obtain $r(z_k) = u^\top_{k,h} z_{k,h}$ for $k \in \{1, \dots, N_\tau - 1\}$. For the initial control we have the relation $M_h Y_{0,h} = u_{0,h}$. Transferring the equations into a matrix vector multiplication with S_{vd} :

$Y_h \times \mathcal{Y}_\sigma \to U_h \times \mathcal{U}_{\mathrm{vd}}$, we obtain

$$\begin{pmatrix} M_h & 0 & \dots & \dots & 0 \\ (-M_h + \frac{\tau_1}{2}A_h) & (M_h + \frac{\tau_2}{2}A_h) & & & \vdots \\ 0 & \ddots & \ddots & & \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & (-M_h + \frac{\tau_{N_\tau-1}}{2}A_h) & (M_h + \frac{\tau_{N_\tau}}{2}A_h) \end{pmatrix} \begin{pmatrix} y_{0,h} \\ y_{1,h} \\ y_{2,h} \\ \vdots \\ y_{N_\tau,h} \end{pmatrix} = \begin{pmatrix} u_{0,h} \\ u_{1,h} \\ u_{2,h} \\ \vdots \\ u_{N_\tau-1,h} \end{pmatrix}. \tag{50}$$

As before it holds $S^*_{\mathrm{vd}} = S^\top_{\mathrm{vd}}$. Analogous to the previous chapter we want to look at the discrete version of the dual problem (P^*). In this case (P^*_{vd}) can be formulated as:

$$\begin{aligned} \min_{w \in U_h^* \times \mathcal{U}^*_{\mathrm{vd}}} \; & K_\sigma(w) \coloneqq \tfrac{1}{p}\|S^*_{\mathrm{vd}} w\|^p_{L^p(Q_h)} + \langle S^*_{\mathrm{vd}} w, y_{d,\sigma}\rangle_{L^p(Q_h),L^q(Q_h)} \\ \text{s.t.} \quad & \|w_\sigma\|_\infty - \alpha \le 0 \quad \text{and} \quad \|w_{0,h}\|_\infty - \beta \le 0. \end{aligned} \tag{P^*_{vd}}$$

Obviously the above problem is similar to (P^*_σ), except for $S_{\mathrm{vd}} \neq S_\sigma$. Consequently the derivation of the optimality system is almost coincident with the procedure in Section 3. Substituting S_{vd} for S_σ the setup is elementary and will not be explained here.

5 Computational Results

We will numerically solve (P^*_σ) and (P^*_{vd}) by a semismooth Newton's method, using the respective optimality systems. To simplify, we fix $u_0 = 0$. This leads to simplifications in the previous results from Section 3 and Section 4. The first row and column of S_σ and S_{vd} can be eliminated. Consequently the constraint $\|w_{0,h}\|_\infty - \beta \le 0$ in the problems (P^*_σ) and (P^*_{vd}) disappears. Furthermore the dimension shrinks from $N_\sigma + N_h$ to N_σ and the variables λ^3 and λ^4 do not appear in the Lagrangians. The dimensions in the optimality system are reduced accordingly, as we only look at $k = 1, \dots, N_\tau$ and (24) does not have to be considered. In this section all variables are specified as their discrete representatives, hence we omit the indices. As our domain we choose $\Omega = [0, 1] \subset \mathbb{R}$ and $I = [0, \frac{3}{2}]$. We assume that our mesh is equidistant, consequently every cell is of size $\tau \cdot h$. We set $\kappa = 1$ and $q = \frac{4}{3}$ and can directly calculate $p = 4$. Using quadrature formulas, we can calculate the representation of (22) and the first block of (25) .

We generate a target by calculating the associated state y_{true} for a known u_{true}. Our example

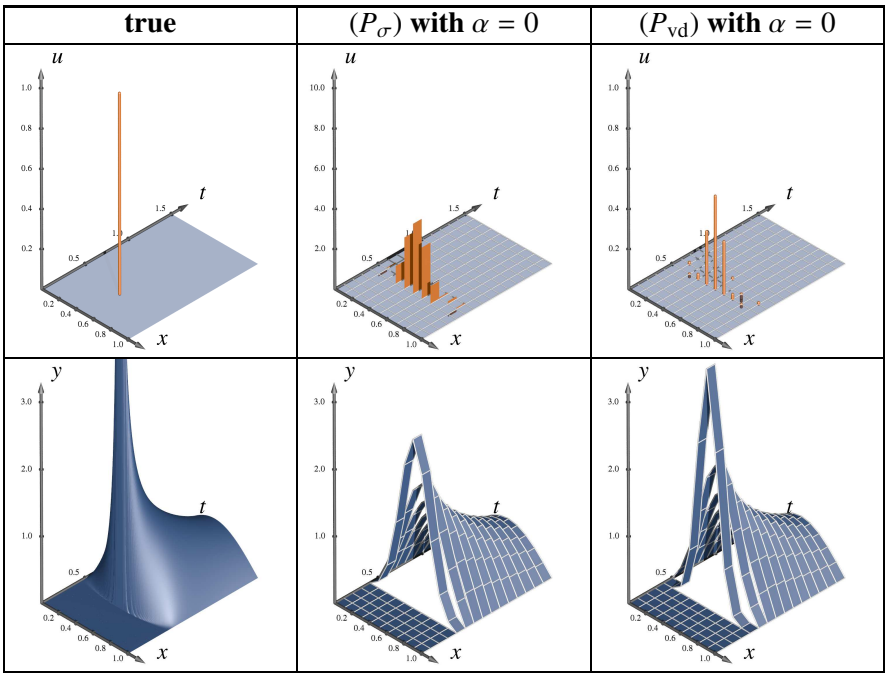

Figure 1: Numerical setup on 10×15 space-time grid with $q = \frac{4}{3}$. **Top row:** True source u_{true} and calculated controls $u_{\sigma,\mathrm{d}}$ and $u_{\mathrm{vd,d}}$ for $\alpha = 0$. **Bottom row:** The true state y_{true} (sampled from the analytic solution with spacial Fourier modes) and the corresponding calculated states $y_{\sigma,\mathrm{d}}$ and $y_{\mathrm{vd,d}}$ for $\alpha = 0$.

problem is a source identification that inherits an obvious sparsity. If the penalty parameter α equals zero, the only admissible point for the problems (P^*_σ) and (P^*_{vd}) is $w \equiv 0$ and (26) shows that this leads to $u_{\sigma,\mathrm{d}} = S_\sigma y_{\sigma,\mathrm{d}}$ and $u_{\mathrm{vd,d}} = S_{\mathrm{vd}} y_{\mathrm{vd,d}}$ respectively. The corresponding visualizations are displayed in Figure 1.

Due to discretization errors the controls for $\alpha = 0$ are not very sparse. We will raise the penalty parameter α, because this will lead to a decrease in the norm of the control and we expect a smaller support. The influence of α can be observed by plotting the norm of u_σ and u_{vd} respectively for a range of α. There exists a value $\bar{\alpha}_i$, such that for all $\alpha_i \geq \bar{\alpha}_i$ the optimal control corresponding to y_{d} is $u_i \equiv 0$ with $i \in \{\sigma, \mathrm{vd}\}$. Additionally it is interesting to look at the values of $\|y_i - y_{\mathrm{true}}\|_{L^{4/3}}$, $i \in \{\sigma, \mathrm{vd}\}$ for changing α. We plotted the dependences in Figure 2.

True to our expectations, the control norms are monotonically decreasing in α and eventually go to zero, while the errors $\|y_\sigma - y_{\mathrm{true}}\|_{L^{4/3}}$, $i \in \{\sigma, \mathrm{vd}\}$ and $\|y_{\mathrm{vd}} - y_{\mathrm{true}}\|_{L^{4/3}}$, $i \in \{\sigma, \mathrm{vd}\}$ grow.

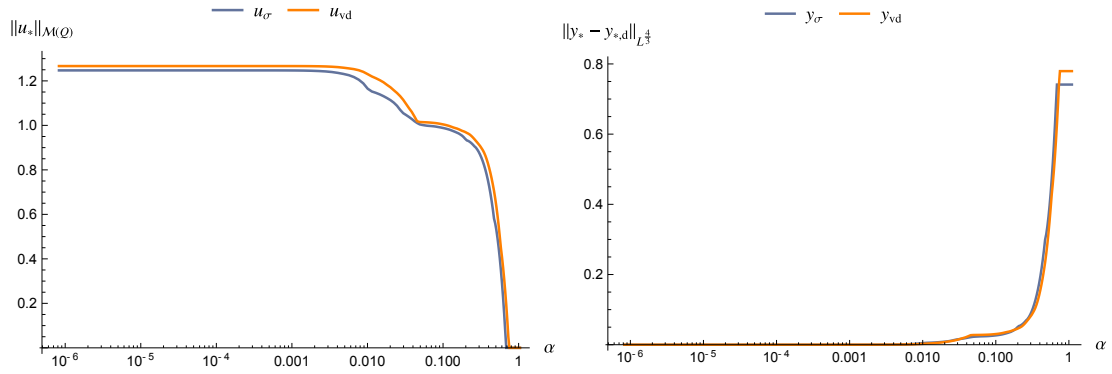

Figure 2: The dependence on the penalty parameter α of the measure norm of u_σ and u_{vd} (left) and the errors $y_\sigma - y_{\mathrm{true}}$ and $y_{\mathrm{vd}} - y_{\mathrm{true}}$ in the $L^{4/3}$ norm (right).

The graphs for both strategies look very similar, which makes sense, as we discretized the same problem and both discretization strategies converge towards the true solution.

To compare the two discretization strategies, we choose a value of α that leads to a norm of the controls, which is not zero nor maximal. The reconstructed controls and states are displayed in Figure 3. If the control u_{true} is not located on our space-time grid, it will be not possible to reproduce its support exactly. In the variational discretization approach a remedy might be choosing a test space $\mathcal{V}_\sigma$ consisting of piecewise quadratic – or even higher order – functions in time. Thereby the maximal values of the test functions $\pm\alpha$ could be attained not only at grid points, but also inside the time intervals. Determining the location of these maximal values would mean to determine the exact position in time of the potential support of the control. This will be part of further research.

While deriving the algorithms to solve the discrete problems, we observe many similarities. The implementation and the level of difficulty in programming is comparable for both approaches and using a homotopy we also observe similar iteration counts.

The main advantage of the variational discretization compared to the discontinuous Galerkin discretization is the *maximal* discrete sparsity of the control achieved by choosing a suitable Petrov-Galerkin-Ansatz and -Test space.

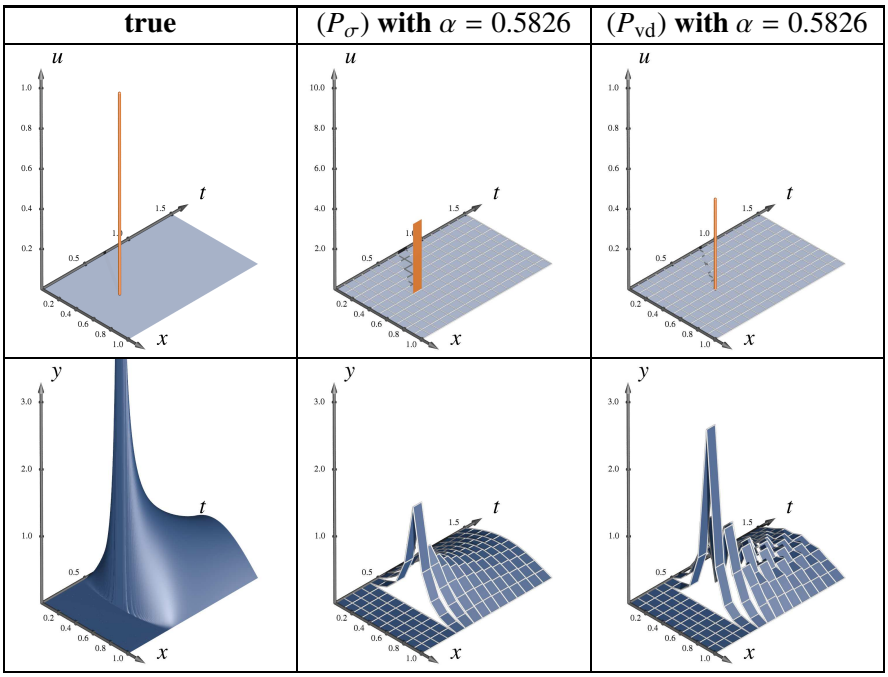

Figure 3: **Top row:** The true control and the optimal controls u_σ and u_{vd} for $\alpha = 0.5826$. **Bottom row:** The true state y_{true} (sampled from the analytic solution with spacial Fourier modes) and the corresponding calculated states y_σ and y_{vd} for $\alpha = 0.5826$.

References

[1] Casas, E. (1997). Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations. SIAM Journal on Control and Optimization, 35(4), 1297-1327.

[2] Casas, E., Clason, C., & Kunisch, K. (2012). Approximation of elliptic control problems in measure spaces with sparse solutions. SIAM Journal on Control and Optimization, 50(4), 1735-1752.

[3] Casas, E., Clason, C., & Kunisch, K. (2013). Parabolic control problems in measure spaces with sparse solutions. SIAM Journal on Control and Optimization, 51(1), 28-63.

[4] Casas, E., & Kunisch, K. (2016). Parabolic control problems in space-time measure spaces. ESAIM: Control, Optimisation and Calculus of Variations, 22(2), 355-370.

[5] Casas, E., Vexler, B., & Zuazua, E. (2015). Sparse initial data identification for parabolic PDE and its finite element approximations.

[6] Clason, C., & Kunisch, K. (2011). A duality-based approach to elliptic control problems in non-reflexive Banach spaces. ESAIM: Control, Optimisation and Calculus of Variations, 17(1), 243-266.

[7] Daniels, N. von, Hinze, M., & Vierling, M. (2015). Crank–Nicolson Time Stepping and Variational Discretization of Control-Constrained Parabolic Optimal Control Problems. SIAM Journal on Control and Optimization, 53(3), 1182-1198.

[8] Geiger, C., & Kanzow, C. (2013). Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben. Springer-Verlag.

[9] Hinze, M. (2005). A variational discretization concept in control constrained optimization: the linear-quadratic case. Computational Optimization and Applications, 30(1), 45-61.

[10] Hinze, M., Pinnau, R., Ulbrich, M., & Ulbrich, S. (2008). Optimization with PDE constraints (Vol. 23). Springer Science & Business Media.

[11] Kunisch, K., Pieper, K., & Vexler, B. (2014). Measure valued directional sparsity for parabolic optimal control problems. SIAM Journal on Control and Optimization, 52(5), 3078-3108.

[12] Pieper, K., & Vexler, B. (2013). A priori error analysis for discretization of sparse elliptic optimal control problems in measure space. SIAM Journal on Control and Optimization, 51(4), 2788-2808.

[13] Schirotzek, W. (2007). Nonsmooth analysis. Springer Science & Business Media.

[14] Tröltzsch, F. (2005). Optimale Steuerung partieller Differentialgleichungen. Vieweg, Wiesbaden.

[15] Wloka, J. (1987). Partial differential equations. Cambridge University.