Error estimates for the discretization of state constrained convex control problems

D. Tiba* F. Tröltzsch †

1 Introduction

We study the following standard convex control problem

$$(P) \quad Min \int_0^T \{\Theta(y) + \psi(u)\} dt$$

governed by linear evolution equations

$$y' + Ay = Bu + f \qquad \text{in }]0, T[, \tag{1.1}$$

$$y(0) = y_0 \tag{1.2}$$

and with restrictions on the control and on the state

$$u(t) \in K \qquad a.e. \text{ in } [0, T], \tag{1.3}$$

$$y(t) \in C \qquad \text{in } [0, T]. \tag{1.4}$$

We take U and $V \subset H \subset V^*$ compactly to be Hilbert spaces and the operators $B : U \to V^*$, $A : V \to V^*$ linear, bounded, such that

$$(Ay, y)_{V^* \times V} + \alpha |y|_H^2 \geq \beta |y|_V^2, \tag{1.5}$$

with certain $\beta > 0$. Here, $|\cdot|_X$ is the norm in the space X and $(\cdot,\cdot)_H$, $(\cdot,\cdot)_{V^*\times V}$ denote the inner product in H or the pairing between V and V^*. It is known that one can take without loss of generality $\alpha = 0$ by a simple changement of unknown function. If $y_0 \in H$, $u \in L^2(0, T; U)$, $f \in L^2(0, T; V^*)$, the state system (1.1), (1.2) has a unique solution $y \in L^2(0, T; V) \cap W^{1,2}(0, T; V^*)$. In particular, it yields $y \in C([0, T]; H)$, that is (1.3), (1.4) are meaningful.
We assume that $K \subset U$, $C \subset H$ are closed convex sets, $\Theta : H \to R_+$, $\psi : U \to R$ are convex continuous mappings.

*Institute of Mathematics, Bucharest
†Technical University of Chemnitz-Zwickau, Faculty of Mathematics

Let us notice that some of the above hypotheses are imposed for the sake of simplicity and are stronger than needed. For instance, Θ may be nonpositive as well and we can use in the subsequent proofs the fact that it is bounded from below by an affine mapping, but this would make the argument quite tedious. The continuity of Θ and ψ may be required since we keep the constraints explicit and do not apply the Moreau-Rockafellar technique to redefine Θ, ψ by $+\infty$ outside C, respectively K.

Under usual admissibility and coercivity conditions, (P) has an optimal pair denoted $[\overline{y}, \overline{u}]$ in $C([0,T]; H) \times L^2(0,T;U)$. As it is well known, an interiority (Slater) assumption is necessary in the setting of problem (P):

There is $[\tilde{y}, \tilde{u}]$ admissible for (P) such that:

$$\tilde{y}(t) \in \text{ int } C, \qquad t \in [0,T]. \tag{1.6}$$

By defining the new variables $z = y - \tilde{y}$ and $v = u - \tilde{u}$ and by shifting Θ, ψ, K, C, it is an elementary calculus to see that one can take $y_0 = 0$, $f = 0$ and to replace (1.6) by $[0,0]$ is admissible for (P) such that

$$0 \in int\ C. \tag*{(1.6)$'$}$$

The notations C, K, Θ, ψ will be preserved and we shall fix $f = 0$, $y_0 = 0$, $\alpha = 0$ (in (1.5)) $\tilde{y} = 0, \tilde{u} = 0$ throughout in the sequel.

Our specific assumption is that $\partial\psi : U \to U$ is bounded and strongly maximal monotone, that is $\partial\psi = aI + \gamma$, $a > 0$, $I : U \to U$, $Iu = u$, and $\gamma \subset U \times U$ is maximal monotone. This yields ψ coercive and uniformly convex. Then the optimal pair $[\overline{y}, \overline{u}]$ exists and is unique. The standard example for ψ satisfying the above conditions is $\psi(u) = \frac{1}{2}|u|_U^2$. As concerns the state equation (1.1), (1.2), we recall that this is an appropriate model for distributed control problems or for boundary control systems via Neumann or mixed type boundary conditions. The case of the control acting via Dirichlet boundary conditions is not included in (1.1), (1.2) and we quote the work of Lasiecka [6], where discretization error estimates are discussed in the setting of the time optimal control problem without state constraints.

General control problems involving constraints both in the state and control are studied in the paper by Alt and Mackenroth [2] by a different technique and under different assumptions. A nonlinear case (without state constraints) is discussed in the work by Pawlow [8], ch. 6. Our technique can be mainly compared with the method of Malanowski [7], but uses the optimality conditions in subdifferential form, not in the projection form. According to Barbu and Precupanu [4], the optimality of $[\overline{y}, \overline{u}]$ is characterized by the existence of $\overline{p} \in L^2(0,T;V) \cap L^\infty(0,T;H) \cap BV(0,T;V^*)$ such that

$$-\overline{p}' + A^*\overline{p} - \partial I_C(\overline{y}) \supset \partial\Theta(\overline{y}) \qquad \text{in }]0,T[, \tag{1.7}$$

$$\overline{p}(T) = 0, \tag{1.8}$$

$$B^*\overline{p}(t) \in \partial\psi(\overline{u}(t)) + \partial I_K(\overline{u}(t)) \quad a.e.\text{in} \quad [0,T]. \tag{1.9}$$

In the proof of (1.7) - (1.9) the Slater condition (1.6) plays an outstanding role. In the sequel, we shall construct the discrete analog of (P) and we shall study its convergence properties, the maximum principle. This will yield finally the desired error estimates.
The paper is organized as follows. In Section 2 we define the discretized problem and we prove a first stability result. Section 3 performs the error analysis for the finite dimensional approximation of problem (P). This is also the subject of the last section, when we renounce the convexity assumption. Then second order sufficient conditions have to be used and similar estimates are obtained.

2 Discretization

Let $V_h \subset V,\ \ U_h \subset U$ be finite dimensional subspaces associated to the discretization parameter $h > 0$ (with respect to the space variables in examples). The norms and the scalar products in V_h, U_h will be from H, U. Let $R_h^H : H \to V_h, R_h^U : U \to U_h$ be some restriction operators (interpolation or projection operators specific to the finite element discretization), such that $R_h^H|_{V_h}$ and $R_h^U|_{U_h}$ are the identity operators. Define $\Theta_h : V_h \to R_+,\ \ \psi_h : U_h \to R$ by
$\Theta_h = \Theta\ |_{V_h}, \psi_h = \psi \mid U_h$ and the discretized constraints sets by

$$K_h = R_h^U(K),\ \ C_h = R_h^H(C + S_H(0, s_h)) \tag{2.1}$$

where $s_h \to 0$ is some "small" constant to be precised later and $S_X(0,s)$ is the ball of center 0 and radius s in the space X. Let $\tau = \frac{T}{m}$ be the time discretization step, $m \in N$. The finite dimensional optimal control problem is to find $y = (y^k)_{k=1}^m, u = (u^k)_{k=1}^m$ in $V_h^m \times U_h^m$

$(P_{h,\tau})$ $\qquad Min \sum_{k=1}^m \tau[\Theta_h(y^k) + \psi_h(u^k)]$

$$\left(\frac{y^{k+1} - y^k}{\tau}, \varphi_h\right)_H + a(y^{k+1}, \varphi_h) = (Bu^{k+1}, \varphi_h)_{V^* \times V} \tag{2.2}$$

$\forall \varphi_h \in V_h,\ k = \overline{0, m-1}$

$$y^0 = 0 \in V_h \tag{2.3}$$

$$u^k \in K_h,\ y^k \in C_h \quad k = \overline{1,m}. \tag{2.4}$$

Remark It is possible to study this optimization problem as a finite dimensional mathematical programming problem with equality constraints (2.2) and other constraints (2.4). However, we discuss $(P_{h,\tau})$ as an optimal control problem, since this parallels the continuous case and is useful for the error analysis.

Remark The Slater assumption (1.6) or (1.6)' remains valid in this setting "uniformly" with respect to $h > 0$. Namely,

$$S_{V_h}(0,\rho) \subset C_h \quad \forall h > 0. \tag{2.5}$$

This follows since $C_h \supset [C + S_H(0,s_h)] \cap V_h \supset C \cap V_h \quad \supset S_H(0,\rho) \cap V_h = S_{V_h}(0,\rho)$ due to the use of the same norm on V_h as in H. Here, we have chosen $\rho > 0$ such that $S_H(0,\rho) \subset C$ according to (1.6)'. In particular, the pair $[0,0] \in V_h^m \times U_h^m$ is admissible for $(P_{h,\tau})$ and we obtain the existence of a unique finite dimensional optimal pair

$$[\overline{y}_{h,\tau}, \overline{u}_{h,\tau}] \in V_h^m \times U_h^m (\text{ or } C_h^m \times K_h^m).$$

Remark We underline that throughout the paper we identify freely the vectors from V_h^m or U_h^m with piecewise constant functions constructed on the given division of $[0,T]$ from these vectors, with values in V_h, respectively U_h.

Proposition 2.1 *The optimal pairs* $[\overline{y}_{h,\tau}, \overline{u}_{h,\tau}]$ *are bounded in* $L^2(0,T;V) \cap$ $\cap L^\infty(0,T;H) \times L^2(0,T;U)$ *with respect to* $h, \tau > 0$.

Proof: We obviously have

$$\sum_{k=1}^{m} \tau \left[\Theta_h(\overline{y}_{h,\tau}^k) + \psi_h(\overline{u}_{h,\tau}^k)\right] \leq T\left[\Theta(0) + \psi(0)\right], \tag{2.6}$$

since the pair $[0,0] \in C_h^m \times K_h^m$ is admissible for $(P_{h,\tau})$ any $h, \tau > 0$. As Θ_h is a positive mapping and $\psi_h = \psi|_{V_h}$ is coercive (with respect to h uniformly), we infer that

$$\sum_{k=1}^{m} \tau \mid \overline{u}_{h,\tau}^k \mid_U^2 \leq ct. \tag{2.7}$$

Fix $\varphi_h = \overline{y}_{h,\tau}^{k+1}$ in (2.2) and compute the first term

$$\begin{aligned} \sum_{k=0}^{l-1} \tau \left(\frac{\overline{y}_{h,\tau}^{k+1} - \overline{y}_{h,\tau}^k}{\tau}, \overline{y}_{h,\tau}^{k+1}\right)_H &\geq \sum_{k=0}^{l-1} \left[\frac{1}{2} \mid \overline{y}_{h,\tau}^{k+1} \mid_H^2 - \frac{1}{2} \mid \overline{y}_{h,\tau}^k \mid_H^2\right] \\ &= \frac{1}{2} \mid \overline{y}_{h,\tau}^l \mid_H^2, \quad 1 \leq l \leq m. \end{aligned} \tag{2.8}$$

Then, by (1.5)

$$\begin{aligned} \frac{1}{2} \mid \overline{y}_{h,\tau}^l \mid_H^2 + \beta \sum_{k=1}^{l} \tau \mid \overline{y}_{h,\tau}^k \mid_V^2 &\leq \sum_{k=1}^{l} \tau \left(B\overline{u}_{h,\tau}^k, \overline{y}_{h,\tau}^k\right)_{V^* \times V} \\ &\leq c \sum_{k=1}^{l} \left[\frac{2c}{\beta} \tau \mid \overline{u}_{h,\tau}^k \mid_U^2 + \frac{\beta}{2c} \tau \mid \overline{y}_{h,\tau}^k \mid_V^2\right] \end{aligned} \tag{2.9}$$

$1 \leq l \leq m$, and the proof is finished. □

Remark On a subsequence, we have $\overline{u}_{h,\tau} \to u^*$, $\overline{y}_{h,\tau} \to y^*$ weakly in $L^2(0,T;U)$, $L^2(0,T;V)$. It will follow later that $[y^*, u^*] = [\overline{y}, \overline{u}]$, the convergence is strong and on the whole sequence. This will be a consequence of the error estimates.

3 Error estimates - First approach

We establish first the optimality conditions for $(P_{h,\tau})$, h and τ fixed. We recall that the continuous case is described by (1.7) - (1.9). Due to (2.5) and the specific construction of K_h, C_h some Slater type condition is valid "uniformly" with respect to $h > 0$.
We start with a regularization of $(P_{h,\tau})$ defined by a penalization of the state constraint. For $\lambda > 0$, we introduce

$$(P_\lambda) \qquad Min \sum_{k=1}^{m} \tau \left[\Theta_h(y^k) + I_\lambda(y^k) + \psi_h(u^k)\right]$$

$$\left(\frac{y^{k+1} - y^k}{\tau}, \varphi_h\right)_H + a(y^{k+1}, \varphi_h) = (Bu^{k+1}, \varphi_h)_{V^* \times V} \quad \forall \varphi_h \in V_h, \tag{3.1}$$

$$k = \overline{0, m-1},$$

$$y^0 = 0, \tag{3.2}$$

$$u^k \in K_h, k = \overline{1, m}, \tag{3.3}$$

where

$$I_\lambda(y) = \inf \left\{ \frac{| y - v |_{V_h}^2}{2\lambda}, v \in C_h \right\}.$$

Above I_λ is the Yosida regularization of the indicator function I_{C_h} of C_h in V_h. We denote by $[\overline{y}_\lambda, \overline{u}_\lambda]$ the unique optimal pair for (P_λ), which obviously exists. Define admissible variations of $\overline{u}_\lambda$ by $u_\delta = \overline{u}_\lambda + \delta(v - \overline{u}_\lambda)$ for any $\delta \in [0,1], v \in K_h^m$ and let y_δ be the solution of (3.1), (3.2) associated to u_δ.
We have $y_\delta = \overline{y}_\lambda + \delta(w - \overline{y}_\lambda)$, $w \in V_h^m$ being the solution of (3.1), (3.2) corresponding to v. Then

$$\sum_{k=1}^{m} \tau \left[\Theta_h(\overline{y}_\lambda^k) + I_\lambda(\overline{y}_\lambda^k) + \psi_h(\overline{u}_\lambda^k)\right] \leq \tag{3.4}$$
$$\leq \sum_{k=1}^{m} \tau \left[\Theta_h(\overline{y}_\delta^k) + I_\lambda(y_\delta^k) + \psi_h(\overline{u}_\lambda^k + \delta(v^k - \overline{u}_\lambda^k))\right].$$

The definition of the subdifferential and the Gâteaux differentiability of I_λ give

$$\sum_{k=1}^{m} \tau \left[(\partial\Theta_h(y_\delta^k), y_\delta^k - \overline{y}_\lambda^k)_H + (\nabla I_\lambda(y_\delta^k), y_\delta^k - \overline{y}_\lambda^k)_H \right. \tag{3.5}$$
$$\left. + (\partial\psi_h(u_\delta^k), u_\delta^k - \overline{u}_\lambda^k)_U\right] \geq 0.$$

By the form of y_δ, u_δ, (3.5) is equivalent with

$$\sum_{k=1}^{m} \tau \left[(\partial\Theta_h(y_\delta^k), w^k - \overline{y}_\lambda^k)_H + (\nabla I_\lambda(y_\delta^k), w^k - \overline{y}_\lambda^k)_H + \right. \tag{3.6}$$

$$\left. + (\partial\psi_h(u_\delta^k), v^k - \overline{u}_\lambda^k)_U \right] \geq 0.$$

Clearly, $y_\delta \to \overline{y}_\lambda$, $u_\delta \to \overline{u}_\lambda$ strongly in V_h, U_h for $\delta \to 0$. As Θ_h, ψ_h, I_λ are continuous convex functions, then $\partial\Theta_h, \partial\psi_h, \nabla I_\lambda$ are locally bounded. On a subsequence we have $\partial\psi_h(\overline{u}_\lambda^k) \to \partial\Theta_h(\overline{y}_\lambda^k), \partial\psi_h(u_\delta^k) \to \partial\psi_h(\overline{u}_\lambda^k)$, $\nabla I_\lambda(y_\delta^k) \to \nabla I_\lambda(\overline{y}_\lambda^k)$ and we can take $\delta \to 0$ in (3.6)

$$\sum_{k=1}^{m} \tau \left[(\partial\Theta_h(\overline{y}_\lambda^k), w^k - \overline{y}_\lambda^k)_H + (\nabla I_\lambda(\overline{y}_\lambda^k), w^k - \overline{y}_\lambda^k)_H + \right. \tag{3.7}$$

$$\left. + \left(\partial\psi_h(\overline{u}_\lambda^k), v^k - \overline{u}_\lambda^k \right)_U \right] \geq 0, \ \forall v \in K_h.$$

Define the adjoint system

$$\left(\frac{\overline{p}_\lambda^{k+1} - \overline{p}_\lambda^k}{\tau}, \varphi_h \right)_H - a(\overline{p}_\lambda^k, \varphi_h) = (\partial\Theta_h(\overline{y}_\lambda^{k+1}) + \nabla I_\lambda(\overline{y}_\lambda^{k+1}), \varphi_h), \ \forall \varphi_h \in V_h, \tag{3.8}$$

$k = \overline{0, m-1}$,

$$\overline{p}_\lambda^m = 0. \tag{3.9}$$

The existence of the solution for (3.8), (3.9) is obvious since the stiffness matrix associated to $a(\cdot, \cdot)$ is nonsingular due to the ellipticity condition (1.5).
Take in (3.8) $\varphi_h = w^k - \overline{y}_\lambda^k$ and sum up over k

$$\begin{aligned} & \sum_{k=1}^{m} \left[(\overline{p}_\lambda^k - \overline{p}_\lambda^{k-1}, w^k - \overline{y}_\lambda^k)_H - \tau a(\overline{p}_\lambda^{k-1}, w^k - \overline{y}_\lambda^k) \right] = \\ = \ & -\sum_{k=1}^{m} \tau \left[a(\overline{p}_\lambda^{k-1}, w^k - \overline{y}_\lambda^k) + \left(\frac{w^k - w^{k-1}}{\tau} - \frac{\overline{y}_\lambda^k - \overline{y}_\lambda^{k-1}}{\tau}, \overline{p}_\lambda^{k-1} \right)_H \right] \\ = \ & -\sum_{k-1}^{m} \tau (\overline{p}_\lambda^{k-1}, Bv^k - B\overline{u}_\lambda^k)_{V \times V^*} \\ = \ & -\sum_{k=1}^{m} \tau (B^* \overline{p}_\lambda^{k-1}, v^k - \overline{u}_\lambda^k)_U. \end{aligned} \tag{3.10}$$

Combining (3.10), (3.8) and (3.7), we get

$$\sum_{k=1}^{m} \tau \left(\partial\psi_h(\overline{u}_\lambda^k) - B^* \overline{p}_\lambda^{k-1}, v^k - \overline{u}_\lambda^k \right)_U \geq 0 \tag{3.11}$$

for any $v \in K_h^m$. Relation (3.11) gives the so called "maximum principle" for the problem (P_λ) and is equivalent with

$$B^* \overline{p}_\lambda^{k-1} \in \partial\psi_h(\overline{u}_\lambda^k) + \partial I_{K_h}(\overline{u}_\lambda^k), \ \ k = \overline{1, m}, \tag{3.12}$$

where I_{K_h} is the indicator function of the convex set K_h in U_h. We recall that the relations (3.12), (3.8), (3.1) form the first order optimality conditions for the problem (P_λ). We shall establish some estimates with respect to $\lambda > 0$, which will enable us to pass to the limit $\lambda \to 0$ and to see the properties of the original discretized problem.
First, we notice that relation (2.5) yields that

$$I_\lambda(\rho v_h) = 0, \quad \forall v_h \in V_h, \quad \forall \lambda > 0, \tag{3.13}$$

if $\mid v_h \mid_{V_h} = 1$ and $h < h_0$, which is a direct consequence of the Slater condition.

Proposition 3.1 *There is $\overline{p}_{h,\tau} \in V_h^m$ such that it satisfies together with $\overline{y}_{h,\tau}$ and $\overline{u}_{h,\tau}$ the optimality conditions*

$$\left(\frac{\overline{p}_{h,\tau}^{k+1} - \overline{p}_{h,\tau}^{k}}{\tau}, \varphi_h\right)_H - a(\overline{p}_{h,\tau}^k, \varphi_h) = (\partial \Theta_h(\overline{y}_{h,\tau}^{k+1}) + \tag{3.14}$$
$$+ \partial I_{C_h}(\overline{y}_{h,\tau}^{k+1}), \varphi_h)_H, \quad k = \overline{0, m-1}, \forall \varphi_h \in V_h,$$

$$\overline{p}_{h,\tau}^m = 0, \tag{3.15}$$

$$B^* \overline{p}_{h,\tau}^{k-1} \in \partial \psi_h(\overline{u}_{h,\tau}^k) + \partial I_{k_h}(\overline{u}_{h,\tau}^k), \quad k = \overline{1, m}. \tag{3.16}$$

Proof We pass to the limit in (3.8), (3.9), (3.12) for $\lambda \to 0$ and h, τ fixed parameters. By (3.13), we have

$$\left(\nabla I_\lambda(\overline{y}_\lambda^{k+1}), \overline{y}_\lambda^{k+1} + \rho v_h\right)_H \geq I_\lambda(\overline{y}_\lambda^{k+1}) - I_\lambda(-\rho v_h) = \tag{3.17}$$
$$= I_\lambda(\overline{y}_\lambda^{k+1}) \geq 0, \quad \forall v_h \in V_h, \forall \lambda > 0, \mid v_h \mid_H = 1.$$

Take in (3.8) $\varphi_h = \overline{y}_\lambda^{k+1} + \rho v_h$ with the above choice

$$\begin{aligned}
\sum_{k=1}^m \tau \left[\left(\frac{\overline{p}_\lambda^k - \overline{p}_\lambda^{k-1}}{\tau}, \overline{y}_\lambda^k\right)_H - a(\overline{p}_\lambda^{k-1}, \overline{y}_\lambda^k) \right] &= \\
= \; - \sum_{k=1}^m \tau \Big[\Big(\frac{\overline{y}_\lambda^k - \overline{y}_\lambda^{k-1}}{\tau}, \overline{p}_\lambda^{k-1}\Big)_H + a(\overline{y}_\lambda^k, \overline{p}_\lambda^{k-1}) \Big] & \\
= \; - \sum_{k-1}^m \tau \left(B^* \overline{p}_\lambda^{k-1}, \overline{u}_\lambda^k\right)_U & \\
= \; - \sum_{k=1}^m \tau \left(\partial \psi_h(\overline{u}_\lambda^k) + \partial I_{K_h}(\overline{u}_\lambda^k), \overline{u}_\lambda^k\right)_U & \\
\leq \; \sum_{k=1}^m \tau [\psi_h(0) - \psi_h(\overline{u}_\lambda^k)]. &
\end{aligned} \tag{3.18}$$

Here, we use that $0 \in K_h$ by the original Slater assumption and the construction of K_h. By (3.17), (3.18) and again (3.8), we obtain

$$\rho \sum_{k=1}^{m} \left[\left(\frac{\overline{p}_\lambda^k - \overline{p}_\lambda^{k-1}}{\tau}, v_h \right)_H - a(\overline{p}_\lambda^{k-1}, v_h) \right] \geq \tag{3.19}$$

$$\geq \sum_{k=1}^{m} \tau \left[\Theta_h(\overline{y}_\lambda^k) - \Theta_h(\rho v_h) + I_\lambda(\overline{y}_\lambda^k) + \Psi_h(\overline{u}_\lambda^k) - \Psi_h(0) \right] \geq$$

$$\geq \sum_{k=1}^{m} \tau \left[\Psi_h(\overline{u}_\lambda^k) - \Psi_h(0) - \Theta_h(\rho v_h) \right] \geq c.$$

Obviously, as in the previous section, $\{\overline{u}_\lambda\}$ is bounded with respect to all parameters λ, h, τ. Since Ψ can be majorized from below by an affine function and Θ is locally bounded, the constant c in (3.19) is independent of λ, τ, h. Applying (3.19) to (3.8), it yields

$$\sum_{k=1}^{m} \tau (\partial \Theta_h(\overline{y}_\lambda^k) + \nabla I_\lambda(\overline{y}_\lambda^k), v_h)_H \geq \frac{c}{\rho}. \tag{3.20}$$

Since h, τ are fixed, the boundedness established in the previous section (which many be extended straightforwardly here) shows that $\overline{y}_\lambda \to \overline{y}_{h,\tau}$ strongly in $L^\infty(0,T;H)$ as $\lambda \to 0$. This is in fact some convergence of a vector sequence in a finite dimensional space. As $\partial\Theta$ is locally bounded, we infer that $\{\partial\Theta_h(\overline{y}_\lambda^k)\}$ is bounded in H for $k = \overline{1,m}$ and $\lambda > 0$. By (3.20), this is also true for $\{\nabla I_\lambda(\overline{y}_\lambda^k)\}$, $k = \overline{1,m}, \lambda > 0$. Let us now choose $\varphi_h = \overline{p}_\lambda^k$ in (3.8)

$$\sum_{k=l}^{m-1} \tau \left(\frac{\overline{p}_\lambda^{k+1} - \overline{p}_\lambda^k}{\tau}, \overline{p}_\lambda^k \right)_H = \sum_{k=l}^{m-1} \left[(\overline{p}_\lambda^{k+1}, \overline{p}_\lambda^k)_H - | \overline{p}_\lambda^k |_H^2 \right] \leq$$

$$\leq \sum_{k=l}^{m-1} \left[\frac{1}{2} | \overline{p}_\lambda^{k+1} |_H^2 - \frac{1}{2} | \overline{p}_\lambda^k |_H^2 \right] = -\frac{1}{2} | \overline{p}_\lambda^l |_H^2 .$$

Since $a(\overline{p}_\lambda^k, \overline{p}_\lambda^k) \geq \beta \mid \overline{p}_\lambda^k \mid_v^2 \geq 0$, the above inequality gives

$$-\frac{1}{2} | \overline{p}_\lambda^l |_H^2 \geq \sum_{k=l}^{m-1} \tau (\partial\Theta_h(\overline{y}_\lambda^{k+1}) + \nabla I_\lambda(\overline{y}^{k+1}), \overline{p}_\lambda^{k+1})_H \geq$$

$$\geq \left(- \max_{l \leq k \leq m} | \overline{p}_\lambda^k |_H \right) \sum_{k=l}^{m-1} \tau \mid \nabla I_\lambda(\overline{y}_\lambda^k) + \partial\Theta_h(\overline{y}^{k+1}) \mid_H .$$

Due to (3.20), where v_h may be choosen

$$v_h = (\partial\Theta_h(\overline{y}_\lambda^k) + \nabla I_\lambda(\overline{y}_\lambda^k)) / \mid \partial\Theta_h(\overline{y}_\lambda^k) + \nabla I_\lambda(\overline{y}_\lambda^k) \mid_H,$$

we obtain

$$\frac{1}{2} | \overline{p}_\lambda^l |_H^2 \leq c \max_{l \leq k \leq m} | \overline{p}_\lambda^k |_H, \forall l = \overline{1,m}, \tag{3.21}$$

where c is independent of λ, h, τ. Let $k_\lambda \in \{1, 2, ..., m\}$ be the index where the maximum occurs in (3.21). Fixing $l = k_\lambda$, we can write

$$\frac{1}{2}\left(\max_{1\le k\le m} \mid \overline{p}^k \mid_H\right)^2 \le c \max_{1\le k\le m} \mid \overline{p}^k \mid_H,$$

and we see that $\{\overline{p}_\lambda\}$ is bounded with respect to λ, h, τ in $L^\infty(0, T; H)$. Taking into account the coercivity of $a(\cdot,\cdot)$, we also see that $\{\overline{p}_\lambda\}$ is bounded in $L^2(0, T; V)$. Since h, τ are fixed it is very easy to pass to the limit in (3.8), (3.9), (3.12) and to get (3.14) - (3.16). We use the demiclosedness of the subdifferential as well. $\square$

Remark As it has been pointed out in the above proof, some of the estimates are independent of $h, \tau > 0$ too.
Denote:

$$\overline{z} = B^*\overline{p} - \partial\Psi(\overline{u}) \in \partial I_k(\overline{u}), \tag{3.22}$$

$$\overline{z}_{h,\tau} = B^*\overline{p}_{h,\tau} - \partial\Psi_h(\overline{u}_{h,\tau}) \in \partial I_{k_h}(\overline{u}_{h,\tau}), \tag{3.23}$$

where the sections of the multivalued mappings occuring in (1.9), respectively (3.16), are taken into account. We have

$$\overline{z} + \partial\Psi(\overline{u}) - B^*\overline{p} = \overline{z}_{h,\tau} + \partial\Psi_h(\overline{u}_{h,\tau}) - B^*\overline{p}_{h,\tau} = 0 \tag{3.24}$$

for a.e. $t \in [0, T]$ and with the same remark as above. Denote $\overline{u}^{k+1} = \overline{u}|_{]k\tau,(k+1)\tau]}$, $k = \overline{0, m-1}$. Define $q_{h,\tau}, r_{h,\tau}$ by

$$\begin{aligned} \left(\frac{q_{h,\tau}^{k+1} - q_{h,\tau}^k}{\tau}, \varphi_h\right)_H - a(q_{h,\tau}^k, \varphi_h) &= \left(\partial\Theta_h(r_{h,\tau}^{k+1}) + \partial I_{C_h}(r_{h,\tau}^{k+1}), \varphi_h\right)_H; \quad (3.25)\\ k &= \overline{0, m-1}, \forall\varphi_h \in V_h, \end{aligned}$$

$$\begin{aligned} \left(\frac{r_{h,\tau}^{k+1} - r_{h,\tau}^k}{\tau}, \varphi_h\right)_H + a(r_{h,\tau}^{k+1}, \varphi_h) &= \left(\frac{1}{\tau}B\int_{k\tau}^{(k+1)\tau} \overline{u}^{k+1}\,dt, \varphi_h\right)_{V^*\times V}, \quad (3.26)\\ k &= \overline{0, m-1}, \forall\varphi_h \in V_h,\\ q_{h,\tau}^m = 0, r_{h,\tau}^0 = 0 & \quad . \qquad\qquad\qquad\qquad\qquad\qquad (3.27) \end{aligned}$$

Here we should notice that the distance $\mid r_{h,\tau} - \overline{y} \mid_{C([0,T];H)}$ is supposed to be small. This defines the constant denoted (without mention of τ) by s_h in (2.1). Then $r_{h,\tau}^k \in C_h$ and (3.25) is meaningful. The choice of the elements in $\partial\Theta_h(r_{h,\tau}^{k+1}), \partial I_{C_h}(r_{h,\tau}^{k+1})$ will be specified in the last Remark. We establish some important inequalities

$$\begin{aligned} &-\int_0^T \left(B^*\overline{p}_{h,\tau} - B^*q_{h,\tau}, \overline{u}_{h,\tau} - \overline{u}\right)_U dt = \qquad\qquad (3.28)\\ &= \; -\sum_{k=1}^m \tau\left(\overline{p}_{h,\tau}^{k-1} - q_{h,\tau}^{k-1}, B\overline{u}_{h,\tau}^k - B\overline{u}^k\right)_{V\times V^*} = \end{aligned}$$

$$
\begin{aligned}
&= -\sum_{k=1}^{m}\tau\left(\frac{\overline{y}^k_{h,\tau}-\overline{y}^{k-1}_{h,\tau}}{\tau}-\frac{r^k_{h,\tau}-r^{k-1}_{h,\tau}}{\tau},\overline{p}^{k-1}_{h,\tau}-q^{k-1}_{h,\tau}\right)_H - \\
&\quad -\sum_{k=1}^{m}\tau a\left(\overline{y}^k_{h,\tau}-r^k_{h,\tau},\overline{p}^{k-1}_{h,\tau}-q^{k-1}_{h,\tau}\right) = \\
&= -\sum_{k=1}^{m}\tau\left(\overline{y}^k_{h,\tau}-r^k_{h,\tau},\frac{\overline{p}^{k-1}_{h,\tau}-\overline{p}^{k}_{h,\tau}}{\tau}-\frac{q^{k-1}_{h,\tau}-q^{k}_{h,\tau}}{\tau}\right)_H - \\
&\quad -\sum_{k=1}^{m}\tau a\left(\overline{y}^k_{h,\tau}-r^k_{h,\tau},\overline{p}^{k-1}_{h,\tau}-q^{k-1}_{h,\tau}\right) = \\
&= \sum_{k=1}^{m}\tau(\overline{y}^k_{h,\tau}-r^k_{h,\tau},\partial\Theta_h(\overline{y}^k_{h,\tau})+\partial I_{C_h}(\overline{y}^k_{h,\tau}) - \\
&\quad -\partial\Theta_h(r^k_{h,\tau})-\partial I^k_{C_h}(r^k_{h,\tau}))_H \geq 0,
\end{aligned}
$$

due to the monotonicity of $\partial\Theta_h+\partial I_{C_h}$.

$$
\begin{aligned}
&\int_0^T(\partial\Psi_h(\overline{u}_{h,\tau})-\partial\Psi(\overline{u}),\overline{u}_{h,\tau}-\overline{u})_U\,dt = \\
&= \int_0^T(\partial\Psi(\overline{u}_{h,\tau})-\partial\Psi(\overline{u}),\overline{u}_{h,\tau}-\overline{u})_U\,dt + \\
&\quad +\int_0^T(\partial\Psi_h(\overline{u}_{h,\tau})-\partial\Psi(\overline{u}_{h,\tau}),\overline{u}_{h,\tau}-\overline{u})_U\,dt \geq \\
&\geq a\int_0^T|\,\overline{u}_{h,\tau}-\overline{u}\,|_U^2\,dt-e_{h,\tau}\,|\,\overline{u}_{h,\tau}-\overline{u}\,|_{L^2(0,T;V)}\,.
\end{aligned}
\tag{3.29}
$$

We have used the strong monotonicity of $\partial\Psi$, and $e_{h,\tau}$ is some error term $e_{h,\tau} := |\,\partial\Psi_h(\overline{u}_{h,\tau})$ -$\partial\Psi(\overline{u}_{h,\tau})\,|_{L^2(0,T;U)}$. If Ψ is some elementary convex function, we may have $e_{h,\tau}=0$ in (3.29).

The last inequality we use is the following

$$
\begin{aligned}
&\int_0^T(\overline{z}_{h,\tau}-\overline{z},\overline{u}_{h,\tau}-\overline{u})_U\,dt \geq \int_0^T(\overline{z}_{h,\tau},\overline{u}_{h,\tau}-\overline{u})_U\,dt = \\
&= \int_0^T(\overline{z}_{h,\tau},\overline{u}_{h,\tau}-R_h^U\overline{u})_U\,dt+\int_0^T(\overline{z}_{h,\tau},R_h^U\overline{u}-\overline{u})_U\,dt \geq \\
&\geq \int_0^T(\overline{z}_{h,\tau},R_h^U\overline{u}-\overline{u})_U\,dt \geq -\tilde{e}_{h,\tau}=-c\,|\,R_h^U\overline{u}-\overline{u}\,|_{L^2(0,T;U)}\,.
\end{aligned}
\tag{3.30}
$$

Above, we have applied (3.22), (3.23) and the definition of the subdifferential. The last inequality is given by the interpolation error $|\,R_h^U\overline{u}-\overline{u}\,|_{L^2(0,T;U)}$ and the boundedness of (3.23) since the term $\{\partial\Psi_h(\overline{u}_{h,\tau})\}$ is bounded by the boundedness of $\{\overline{u}_{h,\tau}\}$ established in section 2 and the assumption that $\partial\Psi$ is a bounded operator. The term $\{B^*\overline{p}_{h,\tau}\}$ is bounded in $L^2(0,T;U)$ for $h,\tau>0$ by the estimates from the proof of Proposition 3.1.

Theorem 3.2 *We have the estimate*

$$
|\,\overline{u}_{h,\tau}-\overline{u}\,|_{L^2(0,T;U)} \leq \frac{1}{a}[|\,\partial\Psi_h(\overline{u}_{h,\tau})-\partial\Psi(\overline{u}_{h,\tau})\,|_{L^2(0,T;U)} + \tag{3.31}
$$

$$+ \mid B^*\overline{p} - B^*q_{h,\tau} \mid_{L^2(0,T;U)} + c \mid R_h^U \overline{u} - \overline{u} \mid^{\frac{1}{2}}_{L^2(0,T;U)}].$$

Proof The inequalities (3.28) - (3.30) give

$$\int_0^T (\overline{z}_{h,\tau} - \overline{z}, \overline{u}_{h,\tau} - \overline{u})_U \, dt - \int_0^T (B^*\overline{p}_{h,\tau} - B^*q_{h,\tau}, \overline{u}_{h,\tau} - \overline{u})_U \, dt +$$
$$+ \int_0^T (\partial\Psi_h(\overline{u}_{h,\tau}) - \partial\Psi(\overline{u}), \overline{u}_{h,\tau} - \overline{u})_U \, dt \geq a \int_0^T \mid \overline{u}_{h,\tau} - \overline{u} \mid_U^2 \, dt -$$
$$- e_{h,\tau} \mid \overline{u}_{h,\tau} - \overline{u} \mid_{L^2(0,T;U)} - \tilde{e}_{h,\tau}.$$

That is, we obtain, due to (3.24), the following

$$\begin{aligned} a \textstyle\int_0^T \mid \overline{u}_{h,\tau} - \overline{u} \mid_U^2 \, dt &\leq e_{h,\tau} \mid \overline{u}_{h,\tau} - \overline{u} \mid_{L^2(0,T;U)} + \tilde{e}_{h,\tau} + \mid \overline{u}_{h,\tau} - \overline{u} \mid_{L^2(0,T;U)} \cdot \\ &\quad \cdot \mid \overline{z}_{h,\tau} - \overline{z} - B^*\overline{p}_{h,\tau} + B^*q_{h,\tau} + \partial\Psi_h(\overline{u}) - \partial\Psi(\overline{u}) \mid_{L^2(0,T;U)} \\ &= e_{h,\tau} \mid \overline{u}_{h,\tau} - \overline{u} \mid_{L^2(0,T;U)} + \tilde{e}_{h,\tau} + \mid \overline{u}_{h,\tau} - \overline{u} \mid_{L^2(0,T;U)} \cdot \\ &\quad \cdot \mid B^*\overline{p} - B^*q_{h,\tau} \mid_{L^2(0,T;U)} . \end{aligned}$$

We have taken $a \leq 1$ without loss of generality. □

Remark All the terms in the right-hand side of (3.31), except one, depend of various interpolation error of $\overline{u}, \partial\Psi$, that is on the regularity properties.
If $\Psi(u) = \frac{1}{2} \mid u \mid_U^2$ then

$$\partial\Psi_h(\overline{u}_{h,\tau}) - \partial\Psi(\overline{u}_{h,\tau}) = 0,$$

and sufficient regularity is available for $\overline{u}$ in some simple cases (Malanowski [7], Tiba and Neittaanmäki [9], ch. V.1.).
The term $\mid B^*\overline{p} - B^*q_{h,\tau} \mid_{L^2(0,T;U)}$ includes the discretization error for the linear parabolic equation. Moreover, one has to fix the section of the multivalued right-hand side in (3.25) such that $\mid \partial\Theta(\overline{y}) + \partial I_C(\overline{y}) - \partial\Theta_h(r_{h,\tau}) - \partial I_{C_h}(r_{h,\tau}) \mid_{L^2(0,T;H)}$ is small. Similar considerations for $\partial\Psi$ can be made in this setting as well.

4 Error estimates and sufficient second order optimality conditions

The theory of the preceding sections relies heavily on the linearity of the state-equation and the convexity of the objective. If one of these basic assumptions is not true, then (P) behaves totally different. To derive error estimates, second order optimality conditions may be helpful in this case. We mention the results by Alt [1] with application to control problems governed by nonlinear ODE and the convergence analysis of Tröltzsch [10], [11] for nonlinear parabolic boundary control problems.

In this section, we shall discuss the error analysis for problems, where Θ is nonconvex while the equation of state is still linear. We are able to derive similar results for semilinear equations, too. However, the presentation would be much more technical. For the same reason, we omit the state constraint (1.4) setting $C = H$. Further, K is required to be a bounded convex set of U. Then feasible controls u are automatically uniformly bounded.

We assume Θ and Ψ to be twice continuously Fréchet differentiable on V and U, respectively. Moreover, we assume that

$$F(y,u) = \int_0^T \{\Theta(y(t)) + \Psi(u(t))\}\, dt$$

is twice continuously Fréchet differentiable on $L^\infty(0,T;V) \times L^p(0,T;U)$ for some $p \in [2,\infty)$. It should be underlined that these requirements of differentiability are often too strong, in particular for parabolic control problems in domains of higher dimension. Then we must work in a dense subspace $L^\infty(0,T;\hat{V}) \times L^p(0,T;\hat{U})$, where F is sufficiently smooth. However, the discussion of the corresponding two-norm technique would go beyond the scope of this paper. We restrict ourselves to explain the main ideas in a simplified setting.

Let r^2_Θ denote the second order remainder term of Θ,

$$r^2_\Theta(y,h) = \frac{1}{2}\int_0^1 (\Theta''(y+sh) - \Theta''(y))[h,h]ds.$$

Analogously, r^2_Ψ is defined. In order to cope with the well known two-norm discrepancy we assume that

$$\lim_{\|y\|_{L^\infty(0,T;V)}\to 0} \int_0^T \mid r^2_\Theta(\overline{y}(t), y(t)) \mid\, dt \cdot \| y \|^{-2}_{L_2(0,T;H)} = 0 \tag{4.1}$$

and

$$\lim_{\|u\|_{L^p(0,T;U)}\to 0} \int_0^T \mid r^2_\Psi(\overline{u}(t), u(t)) \mid\, dt \cdot \| u \|^{-2}_{L_2(0,T;U)} = 0. \tag{4.2}$$

The equation of state must be well posed in $C([0,T];V) \times L^p(0,T;U)$. This means higher regularity of solutions to parabolic equations. We refer, for instance, to Amann [3]. In our setting, we require the *continuity property* that the mapping $u \mapsto y$ assigning the state to the control is continuous from $L^p(0,T;U)$ to $C([0,T];V)$. The parameter p must be chosen sufficiently large to enhance this property. Let $[\overline{y},\overline{u}]$ be optimal for (P). The existence of an optimal pair can be shown under additional assumptions, among them the convexity of Ψ is most essential.

We just assume that $[\overline{y}, \overline{u}]$ is feasible for (P) and satisfies the *optimality conditions* (1.7) - (1.9). Owing to the differentiability of Ψ and Θ, they simplify as follows: The *adjoint state* $\overline{p} \in L_2(0,T;V) \cap W^{1,2}(0,T;V^*)$ solves the *adjoint equation*

$$\overline{p}'(t) - A^*\overline{p}(t) = \Theta'(\overline{y}(t)), \quad \overline{p}(T) = 0 \tag{4.3}$$

(note that $\Theta'(\overline{y}(\cdot)) \in C[0,T;V^*]$). The control $\overline{u}$ satisfies the *variational inequality*

$$\int_0^T (\Psi'(\overline{u}(t)) - B^*\overline{p}(t), u(t) - \overline{u}(t))_{U^* \times U}\, dt \geq 0 \tag{4.4}$$

for all $u \in L_2(0,T;U)$ such that $u(t) \in K$ a.e. on $[0,T]$. (4.3) - (4.4) form the standard, expected system of first order necessary optimality conditions. To derive them from optimality is not an easy task in the nonconvex case. Moreover, $[\overline{y}, \overline{u}]$ is supposed to fulfil the following *second order condition:*
There is a $\delta > 0$ such that

$$\begin{aligned} \int_0^T \{\Theta''(\overline{y}(t))[y(t), y(t)] \;&+\; \Psi''(\overline{u}(t))[u(t), u(t)]\}\, dt \\ &\geq\; \delta\{\| y \|^2_{L_2(0,T;H)} + \| u \|^2_{L_2(0,T;U)}\} \end{aligned} \tag{4.5}$$

for all y, u solving the parabolic equation

$$\begin{aligned} y'(t) + Ay(t) &= Bu(t) \\ y(0) &= 0. \end{aligned} \tag{4.6}$$

Proposition 4.1 *If $[\overline{y}, \overline{u}]$ satisfies the first order necessary optimality conditions (4.3) - (4.4) together with the second order condition (4.5) - (4.6), then $[\overline{y}, \overline{u}]$ affords to F a strong local minimum in the following sense:*
There is a constant $r > 0$ such that

$$F(y,u) > F(\overline{y}, \overline{u}) \tag{4.7}$$

for all $[y,u]$ solving (4.6), $u(t) \in K$ a.e. $[0,T]$, and $\| u - \overline{u} \|_{L^p(0,T;U)} \leq r$.

Proof We have

$$F(y,u) = \int_0^T \{\Theta(y(t)) + \Psi(u(t))\}\, dt + \int_0^T (\overline{p}(t), y'(t) + Ay(t) - Bu(t))_{V \times V^*}\, dt$$

for $[y,u]$ satisfying (4.6). Owing to our continuity assumption, $[y,u]$ and $[\overline{y},\overline{u}]$ belong to $C([0,T];V)\times L^p(0,T;U)$, where F is supposed to be twice continuously differentiable. Hence, by a Taylor expansion

$$\begin{aligned}
F(y,u)=&\int_0^T\{\Theta(\overline{y}(t))+\Psi(\overline{u}(t))\}\,dt+\int_0^T(\overline{p}(t),\overline{y}'(t)+A\overline{y}(t)-B\overline{u}(t))_{V\times V^*}\,dt\\
&+\int_0^T\{(\Theta'(\overline{y}(t)),y(t)-\overline{y}(t))_{V^*\times V}+(\Psi'(\overline{u}(t)),u(t)-\overline{u}(t))_{U^*\times U}\}\,dt\\
&+\int_0^T(\overline{p}(t),(y-\overline{y})'(t)+A(y-\overline{y})(t)-B(u-\overline{u})(t))_{V\times V^*}\,dt\\
&+\frac{1}{2}\int_0^T\{\Theta''(\overline{y}(t))[(y-\overline{y})(t),(y-\overline{y})(t)]\\
&\qquad+\Psi''(\overline{u}(t))[(u-\overline{u})(t),(u-\overline{u})(t))]\}\,dt\\
&+\int_0^T\{r_\Theta^2(\overline{y}(t),(y-\overline{y})(t))+r_\Psi^2(\overline{u}(t),(u-\overline{u})(t))\}\,dt\\
=&F(\overline{y},\overline{u})+\int_0^T(\Psi'(\overline{u}(t))-B^*\overline{p}(t),u(t)-\overline{u}(t))_{U^*\times U}\,dt\\
&+\frac{1}{2}\int_0^T\{\Theta''(\overline{y}(t))[...]+\Psi''(\overline{u}(t))[...]\}\,dt\\
&+\int_0^T\{r_\Theta^2+r_\Psi^2\}\,dt
\end{aligned}$$

by (4.3) and the state-equation for $[\overline{y},\overline{u}]$. Taking advantage of (4.4) and the second order condition (4.5) - (4.6), we continue

$$\begin{aligned}
&F(y,u)\geq F(\overline{y},\overline{u})+(\|\, y-\overline{y}\,\|^2_{L_2(0,T;H)}+\|\, u-\overline{u}\,\|^2_{L_2(0,T;U)})\cdot\\
&\cdot[\frac{\delta}{2}+\int_0^T\{r_\Theta^2+r_\Psi^2\}\,dt\cdot(\|\, y-\overline{y}\,\|^2_{L_2(0,T;H)}+\|\, u-\overline{u}\,\|^2_{L_2(0,T;U)})^{-1}].
\end{aligned}$$

By (4.1) - (4.2), the part [...] is greater than $\delta/4$ provided that r is sufficiently small. This proves the assertion. □

Remark Proposition 4.1 shows, why $p<\infty$ is supposed. The choice $p=\infty$ would restrict the neighbourhood of optimality of $\overline{u}$ to functions u having exactly the same jumps as $\overline{u}$ (provided there are some).
Next, we shall analyze the convergence of discretized solutions $[\overline{y}_{h,\tau},\overline{u}_{h,\tau}]$ to $[\overline{y},\overline{u}]$. Since $[\overline{y},\overline{u}]$ is only locally optimal, we can expect this convergence only in a sufficiently small $L^\infty(0,T;V)\times L^p(0,T;U)$ - neighbourhood N of $[\overline{y},\overline{u}]$. In the following, this neighbourhood N is fixed, independent from h and τ. A pair $[\overline{y}_{h,\tau},\overline{u}_{h,\tau}]$ is said to be *optimal for* $(P_{h,\tau})$ *in* N, if $[\overline{y}_{h,\tau},\overline{u}_{h,\tau}]$ affords to F the minimum among all pairs being feasible for $(P_{h,\tau})$ and contained in N.
We need the following functions on $[0,T]$:
As before, $\overline{y}_{h,\tau}(t)=\overline{y}^k_{h,\tau}$ and $\overline{u}_{h,\tau}(t)=\overline{u}^k_{h,\tau}$ on $]t_{k-1},t_k]$, $k=1,...,m$. Moreover we

introduce

$$\begin{aligned}
\tilde{p}_{h,\tau}(t) &= \frac{1}{\tau}\int\limits_{t_{k-1}}^{t_k} R_h^H \overline{p}(t)\,dt = \tilde{p}_{h,\tau}^{k-1} \text{ on } [t_{k-1}, t_k[\\
\tilde{u}_{h,\tau}(t) &= \frac{1}{\tau}\int\limits_{t_{k-1}}^{t_k} R_h^U \overline{u}(t)\,dt = \tilde{u}_{h,\tau}^{k} \text{ on }]t_{k-1}, t_k] \\
\hat{y}_{h,\tau}(t) &= \overline{y}_{h,\tau}^{k-1} + \frac{1}{\tau}(t - t_{k-1})(\overline{y}_{h,\tau}^{k} - \overline{y}_{h,\tau}^{k-1}) \text{ on } [t_{k-1}, t_k] \\
\hat{p}_{h,\tau}(t) &= \tilde{p}_{h,\tau}^{k-1} + \frac{1}{\tau}(t - t_{k-1})(\tilde{p}_{h,\tau}^{k} - \tilde{p}_{h,\tau}^{k-1}) \text{ on } [t_{k-1}, t_k],
\end{aligned}$$

$k = 1, ..., m$. Note that $\hat{y}_{h,\tau}$ and $\hat{p}_{h,\tau}$ are piecewise linear and continuous functions, while $\tilde{p}_{h,\tau}, \overline{y}_{h,\tau}$, and $\overline{u}_{h,\tau}$ are step functions. We further mention that

$$\begin{aligned}
F_{h,\tau}(y_{h,\tau}, u_{h,\tau}) &= \sum_{k=1}^{m} \tau(\Theta(y_{h,\tau}^k) + \Psi(u_{h,\tau}^k)) \\
&= \int\limits_0^T \{\Theta(y_{h,\tau}(t)) + \Psi(u_{h,\tau}(t))\}\,dt = F(y_{h,\tau}, u_{h,\tau})
\end{aligned}$$

for step functions $y_{h,\tau}$ and $u_{h,\tau}$.

Proposition 4.2 *(Upper estimate): Let $[\overline{y}_{h,\tau}, \overline{u}_{h,\tau}]$ be optimal for $(P_{h,\tau})$ in N. Assume that $[\tilde{y}_{h,\tau}, \tilde{u}_{h,\tau}]$ belongs to N for all sufficiently small h and τ, where $\tilde{y}_{h,\tau}$ is the solution of the discrete system (2.2) - (2.3) associated to $\tilde{u}_{h,\tau}$.*
Then a constant $c > 0$ exists, which does not depend on h and τ, such that

$$F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) \le F(\overline{y}, \overline{u}) + c(\| \overline{y} - \tilde{y}_{h,\tau} \|_{L_1(0,T;V)} + \| \overline{u} - \tilde{u}_{h,\tau} \|_{L_1(0,T;U)}) \quad (4.8)$$

for all sufficiently small h and τ.

Proof: By definition of K_h, we have $\tilde{u}_{h,\tau}(t) \in K_h$ for all t, hence $\tilde{u}_{h,\tau}$ is feasible for $(P_{h,\tau})$. As $[\tilde{y}_{h,\tau}, \tilde{u}_{h,\tau}]$ belongs to N, we have $F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) \le F_{h,\tau}(\tilde{y}_{h,\tau}, \tilde{u}_{h,\tau})$. Hence

$$\begin{aligned}
F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) &\le \sum_{k=1}^{m} \tau\{\Theta(\tilde{y}_{h,\tau}^k) + \Psi(\tilde{u}_{h,\tau}^k)\} \\
&= \int\limits_0^T \{\Theta(\tilde{y}_{h,\tau}(t)) + \Psi(\tilde{u}_{h,\tau}(t))\}\,dt \\
&= \int\limits_0^T \{\Theta(\overline{y}(t)) + \Psi(\overline{u}(t))\}\,dt + \int\limits_0^T \{\Theta(\tilde{y}_{h,\tau}(t)) - \Theta(\overline{y}(t))\}\,dt \\
&+ \int\limits_0^T \{\Psi(\tilde{u}_{h,\tau}(t)) - \Psi(\overline{u}(t))\}\,dt \\
&\le F(\overline{y}, \overline{u}) + c(\| \overline{y} - \tilde{y}_{h,\tau} \|_{L_1(0,T;V)} + \| \overline{u} - \tilde{u}_{h,\tau} \|_{L_1(0,T;U)})
\end{aligned}$$

owing to the Lipschitz property of Θ and Ψ. $\square$

To apply this estimate we need $[\tilde{y}_{h,\tau}, \tilde{u}_{h,\tau}] \to [\overline{y}, \overline{u}]$ as $h \to 0, \tau \to 0$. This property holds provided that $\overline{u}$ is sufficiently smooth. Moreover, we should mention that (4.8) implies the same estimate in any L_q-norm, $q \in]1, \infty]$.
Let $e(h)$ denote the one-sided distance of K_h to K,

$$e(h) = \sup_{u_h \in K_h} (\inf_{u \in K} \| u - u_h \|_U).$$

Proposition 4.3 *(Lower estimate)*
a) Suppose that $[\overline{y}, \overline{u}]$ satisfies the first order conditions (4.3) - (4.4) together with the second order condition (4.5), which is assumed to hold for all $[y, u] \in L_2(0,T;H) \times L_2(0,T;U)$ (strengthened second order condition). Let $[\overline{y}_{h,\tau} \overline{u}_{h,\tau}]$ be a solution of $(P_{h,\tau})$ in $N \subset L^\infty(0,T;V) \times L^p(0,T;U)$ having a sufficiently small diameter $r > 0$.
Then constants $c > 0$ and $\sigma > 0$ exist, which are independent from h and τ, such that

$$\begin{aligned} F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) &\geq F(\overline{y}, \overline{u}) + \sigma(\| \overline{y}_{h,\tau} - \overline{y} \|^2_{L_2(0,T;H)} + \| \overline{u}_{h,\tau} - \overline{u} \|^2_{L_2(0,T;U)}) \\ &\quad -c(e(h) + \| \frac{d}{dt}(\hat{p}_{h,\tau} - \overline{p}) \|_{L_1(0,T;V^*)} + \| \tilde{p}_{h,\tau} - \overline{p} \|_{L_1(0,T;V)}). \end{aligned} \tag{4.9}$$

b) Let the same conditions hold as above with the exception that (4.5) is only required for all $[y, u]$ satisfying the linearized equation (4.6). Then analogously

$$\begin{aligned} & F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) \geq \\ &\geq F(\overline{y}, \overline{u}) + \sigma(\| \overline{y}_{h,\tau} - \overline{y} \|^2_{L_2(0,T;H)} + \| \overline{u}_{h,\tau} - \overline{u} \|^2_{L_2(0,T;U)}) \\ &\quad -c(e(h) + \| \frac{d}{dt}(\hat{p}_{h,\tau} - \overline{p}) \|_{L_1(0,T;V^*)} + \| \tilde{p}_{h,\tau} - \overline{p} \|_{L_1(0,T;V)} \\ &\quad + \| \tilde{y}_{h,\tau} - \overline{y} \|^2_{L_2(0,T;H)} + \| \tilde{y}_{h,\tau} - \overline{y} \|_{L_2(0,T;H)}) \end{aligned} \tag{4.10}$$

holds.

Proof:
a) Adding a zero to $F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau})$ we get

$$\begin{aligned} F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) &= F(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) \\ = F(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) &+ \sum_{k=1}^{m} \tau \Big\{ (\tilde{p}^{k-1}_{h,\tau}, \frac{1}{\tau}(\overline{y}^k_{h,\tau} - \overline{y}^{k-1}_{h,\tau}))_H \\ &+ a(\overline{y}^k_{h,\tau}, \tilde{p}^{k-1}_{h,\tau}) - (\tilde{p}^k_{h,\tau}, B\overline{u}^{k-1}_{h,\tau})_{V^* \times V} \Big\} \\ = F(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) &+ \int_0^T \Big\{ (-\frac{d}{dt}\hat{p}_{h,\tau}, \overline{y}_{h,\tau}(t))_H + a(\overline{y}_{h,\tau}(t), \tilde{p}_{h,\tau}(t)) \\ &- (B^* \tilde{p}_{h,\tau}(t), \overline{u}_{h,\tau}(t))_{U^* \times U} \Big\} \, dt. \end{aligned} \tag{4.11}$$

Here we took advantage of the simple calculation

$$\sum_{k=1}^{m} \tau(\tilde{p}_{h,\tau}^{k-1}, \frac{1}{\tau}(\overline{y}_{h,\tau}^{k} - \overline{y}_{h,\tau}^{k-1}))_H = \sum_{k=1}^{m} \tau(\frac{1}{\tau}(\tilde{p}_{h,\tau}^{k-1} - \tilde{p}_{h,\tau}^{k}), \overline{y}_{h,\tau}^{k})_H.$$

Rewriting the integral in (4.11) we find

$$\int_0^T \{...\}\, dt = \int_0^T \{(-\frac{d}{dt}\overline{p}(t) + A^*\overline{p}(t), \overline{y}_{h,\tau}(t))_{V^*\times V}$$
$$-(B^*\overline{p}(t), \overline{u}_{h,\tau}(t))_{U^*\times U}\}\, dt + R_1,$$

where

$$R_1 = \int_0^T \Big\{(-\frac{d}{dt}(\hat{p}_{h,\tau} - \overline{p}) + A^*(\tilde{p}_{h,\tau} - \overline{p}), \overline{y}_{h,\tau})_{V^*\times V} - (B^*(\tilde{p}_{h,\tau} - \overline{p}), \overline{u}_{h,\tau})_{U^*\times U}\Big\} dt.$$

Hence

$$\begin{aligned}
\int_0^T \{...\}\, dt &= \int_0^T \Big\{(-\frac{d}{dt}\overline{p} + A^*\overline{p}, \overline{y}_{h,\tau} - \overline{y})_{V^*\times V} - (B^*\overline{p}, \overline{u}_{h,\tau} - \overline{u})_{U^*\times U}\Big\}\, dt \\
&+ \int_0^T \Big\{(-\frac{d}{dt}\overline{p} + A^*\overline{p}, \overline{y})_{V^*\times V} - (B^*\overline{p}, \overline{u})_{U^*\times U}\Big\}\, dt + R_1 \\
&= \int_0^T \Big\{(-\Theta'(\overline{y}), \overline{y}_{h,\tau} - \overline{y})_{V^*\times V} - (B^*\overline{p}, \overline{u}_{h,\tau} - \overline{u})_{U^*\times U}\Big\}\, dt + R_1
\end{aligned}$$

by means of the adjoint equation (4.3) for $\overline{p}$ and the equations (1.1) - (1.2) for $\overline{y}$ (note that $y_0 = 0, f = 0$). To insert (1.1) - (1.2), an integration by parts must be performed under the second integral on the right hand side. Now we return to (4.11). After a Taylor expansion around $[\overline{y}, \overline{u}]$ we continue inserting the final expression of $\int_0^T \{...\}\, dt$ in (4.11)

$$\begin{aligned}
&F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) \\
&\quad = F(\overline{y}, \overline{u}) + \tfrac{1}{2}\int_0^T \Theta''(\overline{y}(t))[\overline{y}_{h,\tau}(t) - \overline{y}(t), \overline{y}_{h,\tau}(t) - \overline{y}(t)]\, dt \\
&\qquad + \int_0^T r_\Theta^2(\overline{y}(t), \overline{y}_{h,\tau}(t) - \overline{y}(t))\, dt \\
&\qquad + \int_0^T \Psi'(\overline{u}(t))(\overline{u}_{h,\tau}(t) - \overline{u}(t))\, dt \\
&\qquad + \tfrac{1}{2}\int_0^T \Psi''(\overline{u}(t))[\overline{u}_{h,\tau}(t) - \overline{u}(t), \overline{u}_{h,\tau}(t) - \overline{u}(t)]\, dt \\
&\qquad + \int_0^T r_\Theta^2(\overline{u}(t), \overline{u}_{h,\tau}(t) - \overline{u}(t))\, dt \\
&\qquad - \int_0^T (B^*\overline{p}(t), \overline{u}_{h,\tau}(t) - \overline{u}(t))\, dt \quad + R_1
\end{aligned} \tag{4.12}$$

(the parts containing Θ' disappear by opposite signs). Let $u_{h,\tau}(t)$ be the function with values in K realizing the distance between $\overline{u}_{h,\tau}(t) \in K_h$ and K.
Therefore,

$$\| u_{h,\tau}(t) - \overline{u}_{h,\tau}(t) \|_U \le e(h) \quad a.e.\text{in } [0,T]$$

and

$$\int_0^T \Big(\Psi'(\overline{u}(t)) - B^*\overline{p}(t), \overline{u}_{h,\tau}(t) - \overline{u}(t)\Big)_{U^*\times U} dt = \int_0^T \Big(\Psi'(\overline{u}(t)) - B^*\overline{p}(t), u_{h,\tau}(t) - $$
$$ - \overline{u}(t)\Big)_{U^*\times U} dt + \int_0^T \Big(\Psi'(\overline{u}(t)) - B^*\overline{p}(t), \overline{u}_{h,\tau}(t) - u_{h,\tau}(t)\Big)_{U^*\times U} dt$$
$$\ge -ce(h)$$

by the variational inequality (4.4). Invoking the strenghened second order condition we conclude

$$F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) \ge F(\overline{y}, \overline{u}) + \tfrac{\delta}{2}\{\| \overline{y}_{h,\tau} - \overline{y} \|^2_{L_2(0,T;H)} + \| \overline{u}_{h,\tau} - \overline{u} \|^2_{L_2(0,T;U)}\}$$
$$+ \int_0^T \{r_\Theta^2 + r_\Psi^2\}\, dt + R_1 - ce(h)$$
$$= F(\overline{y}, \overline{u}) + \{\| \overline{y}_{h,\tau} - \overline{y} \|^2_{L_2(0,T;U)} + \| \overline{u}_{h,\tau} - \overline{u} \|^2_{L_2(0,T;U)}\} \cdot \{\tfrac{\delta}{2} + \int_0^T (r_\Theta^2 + r_\Psi^2)\, dt \cdot$$
$$\cdot (\| \overline{y}_{h,\tau} - \overline{y} \|^2_{L_2(0,T;H)} + \| \overline{u}_{h,\tau} - \overline{u} \|^2_{L_2(0,T;U)}))^{-1}\} + R_1 - ce(h).$$

The functions $\overline{y}_{h,\tau}$ and $\overline{u}_{h,\tau}$ belong to a bounded set of $L^\infty(0,T;V)$ and $L^\infty(0,T;U)$, respectively, as N and K_h are bounded. Therefore,

$$R_1 \ge -c(\| \frac{d}{dt}(\hat{p}_{h,\tau} - \overline{p}) \|_{L_1(0,T;V^*)} + \| \tilde{p}_{h,\tau} - \overline{p} \|_{L_1(0,T;V)}). \tag{4.13}$$

If r is sufficiently small, then the term in braces is not less than $\delta/4$ by (4.1) - (4.2). Owing to (4.12), this proves the part a) of our assertion, where $\sigma = \delta/4$.
b) In case (4.5) holds only in the weaker form, we are not able to estimate Θ'' and Ψ'' directly, as $[\overline{y}_{h,\tau}, \overline{u}_{h,\tau}]$ does not satisfy the linearized equation (4.6).
Let $\tilde{y}_{h,\tau}(t)$ denote the (continuous) solution of (4.6) associated to the discrete optimal control function $\overline{u}_{h,\tau}$.
Initiating from (4.12), we substitute $\overline{y}_{h,\tau} = \tilde{y}_{h,\tau} + (\overline{y}_{h,\tau} - \tilde{y}_{h,\tau})$ in the part connected with Θ''. Then, $\tilde{y}_{h,\tau} - \overline{y}$ satisfies the linearized equation, while $\overline{y}_{h,\tau} - \tilde{y}_{h,\tau}$ is expected to be small. Arguing as before,

$$F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) \ge F(\overline{y}, \overline{u}) + \frac{\delta}{2}\{\| \tilde{y}_{h,\tau} - \overline{y} \|^2_{L_2(0,T;H)} + \| \overline{u}_{h,\tau} - \overline{u} \|^2_{L_2(0,T;U)}\}$$
$$+ \int_0^T \{r_\Theta^2 + r_\Psi^2\}\, dt + R_1 - e(h) + R_2,$$

where

$$R_2 \geq -c \parallel \tilde{y}_{h,\tau} - \overline{y}_{h,\tau} \parallel^2_{L_2(0,T;H)} -c \parallel \tilde{y}_{h,\tau} - \overline{y}_{h,\tau} \parallel_{L_2(0,T;H)} . \tag{4.14}$$

Next, we re-substitute $\tilde{y}_{h,\tau} = \overline{y}_{h,\tau} + (\tilde{y}_{h,\tau} - \overline{y}_{h,\tau})$ in the norms standing with $\delta/2$. Finally, we arrive at

$$\begin{aligned} F_{h,\tau}(\overline{y}_{h,\tau}, \overline{u}_{h,\tau}) &\geq F(\overline{y}, \overline{u}) + \tfrac{\delta}{2}\{\parallel \overline{y}_{h,\tau} - \overline{y} \parallel^2_{L_2(0,T;H)} + \parallel \overline{u}_{h,\tau} - \overline{u} \parallel^2_{L_2(0,T;U)}\} \\ &+ \int_0^T \{r_\Theta^2 + r_\Psi^2\}\, dt + R_1 - e(h) + R_2. \end{aligned}$$

Now we are able to finish the proof completely analogous to the further arguments of a). □

As an immediate conclusion we obtain the

Theorem 4.4 *Let the assumptions of Proposition 4.2 be fulfilled together with those of Proposition 4.3* a) *or* b)*, respectively. Then*

$$\parallel \overline{y}_{h,\tau} - \overline{y} \parallel^2_{L_2(0,T;H)} + \parallel \overline{u}_{h,\tau} - \overline{u} \parallel^2_{L_2(0,T;U)} \leq R(h,\tau),$$

where

$$\begin{aligned} R(h,\tau) &= R_a(h,\tau) \\ &:= c(e(h) + \parallel \tfrac{d}{dt}(\hat{p}_{h,\tau} - \overline{p}) \parallel_{L_1(0,T;V^*)} + \parallel \tilde{p}_{h,\tau} - \overline{p} \parallel_{L_1(0,T;V)} \\ &+ \parallel \overline{y} - \tilde{y}_{h,\tau} \parallel_{L_1(0,T;V)} + \parallel \overline{u} - \tilde{u}_{h,\tau} \parallel_{L_1(0,T;U)}) \end{aligned}$$

in case a) *and*

$$R(h,\tau) = R_a(h,\tau) + c(\parallel \tilde{y}_{h,\tau} - \overline{y}_{h,\tau} \parallel^2_{L_2(0,T;H)} + \parallel \tilde{y}_{h,\tau} - \overline{y}_{h,\tau} \parallel_{L_2(0,T;H)}),$$

in case b)*, where $c > 0$ does not depend on h and τ.*

Proof The result follows from combining (4.8) with (4.9) and (4.10), respectively. □

Remark We should underline, that in case b) an error estimate for the distance of $\tilde{y}_{h,\tau}$ to $\overline{y}_{h,\tau}$ is necessary, although $\overline{u}_{h,\tau}$ is given as a (nonsmooth) step function. To make this result useful, the discretization in time of (1.1) with respect to $y(t)$ should be finer than that for $u(t)$. Then the theory can be developed similarly and ends up with the estimate of Theorem 4.4.

5 Example

Let us now regard a slightly changed situation. We delete the state constraints (5.5) and consider the minimization of the functional

$$\int_Q \vartheta(x, y(t,x))\, dxdt + \frac{1}{2}\int_Q u^2(t,x)\, dxdt. \tag{5.1}$$

Here, $\vartheta = \vartheta(x, y)$ is real-valued and twice continuously differentiable on $\bar{\Omega} \times I\!R$ with respect to y. We assume $\Omega \subset I\!R$ to have the continuous embedding $V = H_o^1(\Omega) \subset C(\bar{\Omega})$. The functional θ in (P) is

$$\theta(y) = \int_\Omega \vartheta(x, y(x))dx.$$

θ is twice continuously Frechet-differentiable on V. The same holds true for the functional (5.1) on $C([0,T]; V) \times L_2(0,T; H)$, where $H = L_2(\Omega)$. We assume that a feasible pair $[\bar{y}, \bar{u}]$ is given such that the condition

$$\vartheta_{yy}(x, \bar{y}(t,x)) \geq \delta' > 0 \tag{5.2}$$

holds for almost all $(t, x) \in Q$. Then

$$\begin{array}{rcl} \int_0^T \theta''(\bar{y}(t))[h(t), h(t)]\,dt & = & \int_0^T \int_\Omega \vartheta_{yy}(x, \bar{y}(t,x))h^2(t,x)\,dxdt \\ & \geq & \delta' \|h\|^2_{L_2(0,T;H)}, \end{array} \tag{5.3}$$

hence the strengthened second order condition is satisfied with $\delta = min(\delta', 1/2)$. From $\bar{u} \in L_\infty(Q)$ we obtain $\bar{y} \in C(\bar{Q})$ by parabolic regularity, thus $\vartheta_y(x, \bar{y}(t,x))$ belongs to $C(\bar{Q})$, too. Therefore, the solution $\bar{p}$ of

$$\begin{array}{rcll} -p' - \Delta p & = & \vartheta_y(\cdot, \bar{y}) & \text{in } Q \\ p(T, \cdot) & = & 0 & \text{in } \Omega \\ p & = & 0 & \text{on } \Sigma, \end{array} \tag{5.4}$$

belongs to $H^1(Q)$. Repeating the same arguments as before, the regularity $\bar{u} \in H^1(Q)$ is obtained.

References

[1] Alt, W.: On the approximation of infinite optimization problems with an application to optimal control problems. Applied Math. Optimization 22(1991), 83-98.

[2] Alt, W., Mackenroth, U.: Convergence of finite element approximations to state constrained convex parabolic boundary control problems, SIAM J. Control and Optimiz., vol. 27(4), p. 718-736 (1989).

[3] Amann, H.: Parabolic evolution equation with nonlinear boundary conditions. In: F.E. Browder, Editor, Nonlinear Functional Analysis Proc. Sympos. Pure Math. Vol. 45, Part. I.

[4] Barbu, V., Precupanu, Th.: Convexity and optimization in Banach spaces, Sijthoff & Nordhoff, Leyden (1978).

[5] Knowles, G.: Finite element approximation of parabolic time optimal control problems, SIAM J. Control Opt. 20(1982), 414-427.

[6] Lasiecka, I.: Ritz-Galerkin approximation of the time optimal boundary control problem for parabolic systems with Dirichlet boundary conditions, SIAM J. Control Optim. 22(1984), 477-500.

[7] Malanowski, K.: Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal control problems, Appl. Math. Optim. 8(1981), 69-95.

[8] Pawlow, I.: Analysis and control of evolution multi-phase problems with free boundaries, Ossolineum, Warszawa (1987).

[9] Tiba, D., Neittaanmäki, P.: Optimal control of nonlinear parabolic equations, Theory, algorithms and applications, Marcel Dekker, New York (1994).

[10] Tröltzsch, F.: Approximation of nonlinear parabolic boundary control problems by the Fourier method - convergence of optimal controls, Optimization 22(1991), 83-98.

[11] Tröltzsch, F.: Semidiscrete Ritz-Galerkin approximation of nonlinear parabolic boundary control problems - strong convergence of optimal controls, To appear 1994 in Applied Math. Optimization.