# Optimal boundary control of a system of reaction diffusion equations

**Werner Barthel[1], Christian John[2,*], and Fredi Tröltzsch[2]**

[1] Technical University of Chemnitz, Faculty of Mathematics, D-09107 Chemnitz, Germany
[2] Technical University of Berlin, Institute of Mathematics, Straße des 17. Juni 136, 10623 Berlin, Germany

Received 11 September 2009, revised 9 March 2010, accepted 22 April 2010
Published online XXXX

**Key words** Optimal control, boundary control, semilinear parabolic system, reaction-diffusion equations, necessary optimality conditions, sufficient optimality conditions.

**MSC (2000)** Primary: 49J20; Secondary: 35K57

In this work, boundary control problems governed by a system of semilinear parabolic PDEs with pointwise control constraints are considered. This class of problems is related to applications in the chemical catalysis. After discussing existence and uniqueness of the state equation with both linear and nonlinear boundary conditions, the existence of an optimal solution is shown. Necessary and sufficient optimality conditions are derived to deal with numerical examples, which conclude the paper.



## 1 Introduction

We consider a class of optimal boundary control problems governed by a system of semilinear parabolic PDEs with application to chemical reaction. Very similar optimal control problems were discussed first in the PhD thesis [4] by Griesse and later extended by Griesse and Volkwein [5]. In contrast to them, we regard nonlinear boundary conditions in domains of arbitrary dimension. We consider the following class of systems with nonlinear boundary condition in a bounded domain $\Omega \subset \mathbb{R}^N$, $N \geq 1$, with Lipschitz-continuous boundary $\Gamma = \partial\Omega$:

$$\textbf{(P1)} \quad \min J(u,v,c) \;=\; \frac{\alpha_u}{2}\,\|u-u_Q\|^2_{L^2(Q)} + \frac{\alpha_v}{2}\,\|v-v_Q\|^2_{L^2(Q)} + \frac{\alpha_{TU}}{2}\,\|u(T)-u_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha_{TV}}{2}\,\|v(T)-v_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha_c}{2}\,\|c\|^2_{L^2(\Sigma)}$$

subject to

$$\textbf{(E1)}\left\{\begin{array}{rcll} u_t - D_1\Delta u + k_1 u &=& -\gamma_1 uv & \text{in } Q,\\ v_t - D_2\Delta v + k_2 v &=& -\gamma_2 uv & \text{in } Q,\\ D_1\partial_\nu u + b(x,t,u) &=& c & \text{on } \Sigma,\\ D_2\partial_\nu v + \varepsilon v &=& 0 & \text{on } \Sigma,\\ u(x,0) &=& u_0(x) & \text{in } \Omega,\\ v(x,0) &=& v_0(x) & \text{in } \Omega \end{array}\right.$$

and the box constraint

$$c \in C_{ad} = \{c \in L^\infty(\Sigma) : c_a \leq c \leq c_b \text{ a.e. on } \Sigma\}$$

for a fixed final time $T > 0$. In this setting, $Q = \Omega \times (0,T)$ denotes the space-time cylinder with its boundary $\Sigma = \Gamma \times (0,T)$. We denote by $D_1$, $D_2$ positive and by $k_1$, $k_2$, $\alpha_u, \alpha_v, \alpha_{TU}, \alpha_{TV}, \alpha_c, \varepsilon, \gamma_1$, and $\gamma_2$ nonnegative constants. The functions $c_a, c_b$ are given in $L^\infty(\Sigma)$, such that $c_a \leq c_b$ holds almost everywhere in $\Sigma$. The given desired terminal states $u_Q, v_Q, u_\Omega, v_\Omega$ are elements of $L^2(Q)$, $L^2(\Omega)$, respectively, and $u_0, v_0$ belong to $C(\bar{\Omega})$. We need further the following assumption on $b$:

* Corresponding author E-mail: john@math.tu-berlin.de, Phone: +00 49 30 31428035, Fax: +00 49 30 31478658

**Assumption 1** The nonlinear function $b : \Sigma \times \mathbb{R} \to \mathbb{R}$ is continuous in $(x, t, u)$ and monotone non-decreasing with respect to $u$ for all $(x, t) \in \Sigma$. Moreover, $b$ is twice continuously differentiable with respect to $u \in \mathbb{R}$ and $\partial^2 b(x, t, u)/\partial u^2$ is locally Lipschitz, i.e. for all $\rho > 0$ there exists $L(\rho) > 0$ such that

$$\left| \frac{\partial^2 b}{\partial u^2}(x, t, u_1) - \frac{\partial^2 b}{\partial u^2}(x, t, u_2) \right| \le L(\rho)\, |u_1 - u_2|$$

holds for all $u_1, u_2 \in \mathbb{R}$ with $|u_i| \le \rho,\ i = 1, 2$ and all $(x, t) \in \Sigma$.

Notice that, by the continuity of $b$ and its derivatives, there exists a constant $K$ with $|b(x, t, 0)| + |b_u(x, t, 0)| + |b_{uu}(x, t, 0)| \le K$ for almost all $(x, t) \in \Sigma$. Here and in the following, we denote by $b_u$ and $b_{uu}$ the first- and second-order partial derivative of $b$ w.r. to $u$, respectively. We denote by $\nu$ the outward normal vector at $\Gamma$.

This problem belongs to the class of optimal control problems for semilinear parabolic equations, where quite a number of publications were devoted to. We mention, for instance, [13], where a nonlinear boundary condition of Stefan-Boltzmann type was considered first, the contributions [3], [6], [7] to a nonlinear phase field model, and the papers [2], [12] on the Pontryagin principle for parabolic control problems. Further references on the control of nonlinear parabolic equations can be found in the monography [14].
Problem **(P1)** is close to a similar one investigated in [5]. Our starting point for discussing this matter was a problem of catalysis with a model, where the boundary conditions are linear. Motivated by the handling of the coupled parabolic system in [5], we became interested in dealing with more general nonlinearities. We considered a nonlinear boundary condition as a case study for more general nonlinearities in domains of arbitrary dimension, although we did not have an associated application in mind.

Compared with [5], this is our main novelty. Our main emphasis is laid on existence and uniqueness of a solution to the system of the state equations for **(P1)**, existence of a solution to the optimal control problem and necessary and sufficient optimality conditions. While this is mainly of theoretical interest, we also discuss some specific application to the optimal control of catalysis that was introduced in [8]. The corresponding numerical examples are also new, as the objective functional differs from that in [5].

## 2 The problem with nonlinear boundary conditions

### 2.1 Well-posedness of the state equation

In this section, we consider the problem **(P1)** with nonlinear boundary conditions. Systems of this type are interesting for the applications. For instance, the equations might model the diffusion of a substance with concentrations $v$ and temperature $u$, where a Stefan-Boltzmann type boundary condition for $u$ is given. However, we do not aim at discussing specific applications. We think that the system is interesting from a mathematical point of view. To show an existence and uniqueness theorem for the nonlinear system, we invoke the method of ordered upper and lower solutions. This method was also implied in [4] to the similar problem with linear boundary conditions mentioned in the introduction. Let us underline that, in contrast to our approach, in [5] the solutions of the PDEs are considered in the space $W(0, T)$. To deal with the nonlinear boundary condition or similar nonlinearities also in space dimensions $N > 1$, we need bounded state functions so that we have to show higher regularity.
For a Banach space $V$, the space $W(0, T)$ is defined by

$$W(0, T) := \{y \in L^2(0, T; V) : y' \in L^2(0, T; V^*)\}$$

and equipped with the norm

$$\|y\|_{W(0,T)} = \left( \int_0^T (\|y(t)\|_V^2 + \|y'(t)\|_{V^*}^2) dt \right)^{\frac{1}{2}},$$

where $V^*$ is the dual space of $V$, $y'$ denotes the distributional derivative of $y$ with respect to $t$ and $L^2(0, T; V)$ is the space of all (equivalence classes of) measurable abstract functions $y : [0, T] \to V$ with $\int_0^T \|y(t)\|_V^2\, dt < \infty$ and norm

$$\|y\|_{L^2(0,T;V)} = \left( \int_0^T \|y(t)\|_V^2\, dt \right)^{\frac{1}{2}}.$$

From now on, we consider the particular case $V := H^1(\Omega)$. Since then $W(0,T)$ is continuously embedded into $C([0,T]; L^2(\Omega))$, the space of all continuous functions from $[0,T]$ into $L^2(\Omega)$, there exists a constant $C > 0$ satisfying

$$\|y\|_{C([0,T];L^2(\Omega))} \leq C\|y\|_{W(0,T)} \text{ for all } y \in W(0,T).$$

We need higher regularity of $u$ and $v$ to make the nonlinearities well defined and to ensure the differentiability of the control-to-state mapping $c \mapsto (u,v)$.
Let us start by the definition of a weak solution and recall that $u_0 \in C(\bar{\Omega})$ and $v_0 \in C(\bar{\Omega})$ are given.

**Definition 2.1** A pair of functions $(u,v) \in (W(0,T) \cap L^\infty(Q))^2$ is called weak solution of the system **(E1)**, if the equations

$$u(\cdot,0) = u_0,\ v(\cdot,0) = v_0,$$
$$\int_0^T (u_t,\varphi)_{V^*,V}\, dt + \iint_\Sigma b(x,t,u)\varphi\, dsdt + \iint_Q (D_1\nabla u\cdot\nabla\varphi + k_1 u\varphi + \gamma_1 uv\varphi)\, dx\, dt = \iint_\Sigma c\,\varphi\, dsdt$$

and

$$\int_0^T (v_t,\varphi)_{V^*,V}\, dt + \iint_\Sigma \varepsilon\, v\, \varphi\, dsdt + \iint_Q (D_2\nabla v\cdot\nabla\varphi + k_2 v\varphi + \gamma_2 uv\varphi)\, dxdt = 0$$

are satisfied for all $\varphi \in L^2(0,T;H^1(\Omega))$, where $\nabla$ denotes the gradient with respect to the spatial variable and $ds$ denotes the Lebesgue surface measure. Here, the duality pairing between $V^* = H^1(\Omega)^*$ and $V = H^1(\Omega)$ is denoted by $(\cdot,\cdot)_{V^*,V}$ and $u_t$ is defined by $u_t := \frac{\partial u}{\partial t}$.
To prove the existence and uniqueness of a weak solution for **(E1)**, we apply the method of upper and lower solutions and follow the arguments of Pao [11], pp. 459-470. Since in this text classical solutions are considered, we transfer this method to our case of weak solutions rather than to directly apply the theorems of [11] for smooth data and to approximate the solutions to non-smooth data by passing to the limit.

Let us first introduce some preparatory constructions. We have assumed that $b$ is monotone non-decreasing in $u$. Additionally, we require

**Assumption 2** The function $b$ fulfills $b(x,t,0) \leq c_a(x,t)$ for all $(x,t) \in \Sigma$ and $\lim_{u\to\infty} \|b(\cdot,\cdot,u)\|_{C(\bar{Q})} = \infty$.

(Alternatively, we might assume that $\lim\limits_{u\to-\infty} b(x,t,u) = -\infty$ and $b(x,t,0) \geq c_b(x,t)$. We do not further discuss the associated changes.)

Under Assumption 2, there exists an $M \geq \max\{\|u_0\|_{C(\bar{\Omega})}, \|v_0\|_{C(\bar{\Omega})}\}$ such that

$$b(x,t,M) \geq \|c\|_{L^\infty(\Sigma)} \tag{1}$$

holds for all $c \in C_{ad}$. We now introduce functions $\tilde{u},\ \tilde{v},\ \hat{u},\ \hat{v} : \bar{Q} \to \mathbb{R}$ by

$$\tilde{u}(x,t) = \tilde{v}(x,t) = M, \quad \hat{u}(x,t) = \hat{v}(x,t) = 0 \quad \forall (x,t) \in \bar{Q}.$$

Then the pairs $(\tilde{u},\tilde{v})$ and $(\hat{u},\hat{v})$ obey the inequalities

$$\begin{array}{rcccll}
\tilde{u}_t - D_1\Delta\tilde{u} + k_1\tilde{u} + \gamma_1\tilde{u}\hat{v} & \geq & 0 & \geq & \hat{u}_t - D_1\Delta\hat{u} + k_1\hat{u} + \gamma_1\hat{u}\tilde{v} & \text{in } Q,\\
\tilde{v}_t - D_2\Delta\tilde{v} + k_2\tilde{v} + \gamma_2\hat{u}\tilde{v} & \geq & 0 & \geq & \hat{v}_t - D_2\Delta\hat{v} + k_2\hat{v} + \gamma_2\tilde{u}\hat{v} & \text{in } Q,\\
D_1\partial_\nu\tilde{u} - c(x,t) + b(x,t,\tilde{u}) & \geq & 0 & \geq & D_1\partial_\nu\hat{u} - c(x,t) + b(x,t,\hat{u}) & \text{on } \Sigma,\\
D_2\partial_\nu\tilde{v} + \varepsilon\tilde{v} & \geq & 0 & \geq & D_2\partial_\nu\hat{v} + \varepsilon\hat{v} & \text{on } \Sigma,\\
\tilde{u}(x,0) & \geq & u_0(x) & \geq & \hat{u}(x,0) & \text{in } \Omega,\\
\tilde{v}(x,0) & \geq & v_0(x) & \geq & \hat{v}(x,0) & \text{in } \Omega.
\end{array}$$

This means that $(\tilde{u},\tilde{v})$ and $(\hat{u},\hat{v})$ are pairs of ordered upper and lower classical solutions to **(E1)**, respectively, in the sense of [11]. We extend this notion in a natural way to weak solutions: For instance,$(\tilde{u},\tilde{v}) \in (W(0,T) \cap C(\bar{Q}))^2$ is said to be an upper solution, if there are $\psi \in L^\infty(Q)$, $\varphi \in L^\infty(\Sigma)$, $\chi \in C(\bar{\Omega})$ such that $\tilde{u}$ solves the system

$$\begin{array}{rcccl}
\tilde{u}_t - D_1\Delta\tilde{u} + k_1\tilde{u} + \gamma_1\tilde{u}\hat{v} & = & \psi & \geq & 0,\\
D_1\partial_\nu\tilde{u} - c + b(x,t,\tilde{u}) & = & \varphi & \geq & 0,\\
\tilde{u}(x,0) & = & \chi(x) & \geq & u_0(x)
\end{array}$$

in weak sense and analogous inequalities are satisfied by $\tilde{v}$. The notion of a lower solution is defined accordingly. Next, we rewrite the differential equations of **(E1)** in a different but equivalent form that is related to the settings in [11]:

$$\begin{array}{rcll} u_t - D_1\Delta u + (k_1+\gamma_1 M)u &=& \gamma_1 u(M-v) =: F_1(u,v) & \text{in } Q,\\ v_t - D_2\Delta v + (k_2+\gamma_2 M)v &=& \gamma_2 v(M-u) =: F_2(u,v) & \text{in } Q,\\ D_1\partial_\nu u + \alpha\, u &=& -b(x,t,u) + \alpha\, u + c(x,t) =: G(x,t,u) & \text{on } \Sigma,\\ D_2\partial_\nu v + \varepsilon v &=& 0 & \text{on } \Sigma, \end{array} \tag{2}$$

where $\alpha > 0$ is taken so large that

$$-b_u(x,t,u) + \alpha \geq 0$$

holds for all $(x,t) \in \Sigma$ and for all $u \in [0,M]$. Notice that $b$ is continuously differentiable with respect to $u$. Selecting $\alpha$ and $M$ in this way, we have

$$\begin{array}{rclrclrcl} \dfrac{\partial F_1(u,v)}{\partial u} &\geq& 0, & \dfrac{\partial F_1(u,v)}{\partial v} &\leq& 0, & \dfrac{\partial G(x,t,u)}{\partial u} &\geq& 0,\\ \dfrac{\partial F_2(u,v)}{\partial u} &\leq& 0, & \dfrac{\partial F_2(u,v)}{\partial v} &\geq& 0 & & & \end{array}$$

for all $u,\ v \in [0,M]$ and all $(x,t) \in \Sigma$.

Let us introduce for convenience the notation $Y := W(0,T) \cap C(\bar{Q})$.

**Theorem 2.2** *Assume that $u_0$ and $v_0$ are nonnegative functions and $b$ satisfies the additional Assumption 2. Then the system **(E1)** admits for each $c \in C_{ad}$ a unique solution $(u,v) \in Y \times Y$. For $s > N+1$, the mapping $c \mapsto (u,v)$ is continuous from $L^s(\Sigma)$ to $Y$.*

P r o o f. *(i) Construction of monotone sequences:* We adopt the iteration technique introduced in [11] and construct sequences $\{(\bar{u}^k, \bar{v}^k)\}_{k=0}^\infty$ and $\{(\underline{u}^k, \underline{v}^k)\}_{k=0}^\infty$ as follows: We define

$$\begin{array}{ll} \bar{u}^0 = \tilde{u} = M, & \bar{v}^0 = \tilde{v} = M,\\ \underline{u}^0 = \hat{u} = 0, & \underline{v}^0 = \hat{v} = 0. \end{array}$$

Initiating from $(\bar{u}^k, \bar{v}^k)$, $(\underline{u}^k, \underline{v}^k)$, the pair $(\bar{u}^{k+1}, \bar{v}^{k+1})$ is obtained by $(\bar{u}^{k+1}, \bar{v}^{k+1}) := (u^+, w^+)$, where $(u^+, w^+)$ is the solution of the system

$$\begin{array}{rcll} u_t^+ - D_1\Delta u^+ + (k_1+\gamma_1 M)u^+ &=& F_1(\bar{u}^k, \underline{v}^k) & \text{in } Q,\\ D_1\partial_\nu u^+ + \alpha\, u^+ &=& G(x,t,\bar{u}^k) & \text{on } \Sigma,\\ u^+(x,0) &=& u_0(x) & \text{in } \Omega,\\ v_t^+ - D_2\Delta v^+ + (k_2+\gamma_2 M)v^+ &=& F_2(\underline{u}^k, \bar{v}^k) & \text{in } Q,\\ D_2\partial_\nu v^+ + \varepsilon v^+ &=& 0 & \text{on } \Sigma,\\ v^+(x,0) &=& v_0(x) & \text{in } \Omega. \end{array} \tag{3}$$

Analogously, $(\underline{u}^{k+1}, \underline{v}^{k+1})$ is defined as solution of the system (3) above, but with right-hand sides $F_1(\underline{u}^k, \bar{v}^k)$, $G(x,t,\underline{u}^k)$, $u_0$, $F_2(\bar{u}^k, \underline{v}^k)$, $0$, $v_0$, respectively. All associated four systems are linear parabolic equations in a Lipschitz domain. Therefore, they have solutions in $Y = W(0,T) \cap C(\bar{Q})$. We refer, e.g. to Lemma 7.12 in [14], compare also Theorem 5.5 for the nonlinear equation. Now we obtain as in [11], Chap. 8 that the constructed sequences possess the following properties: For each $(x,t) \in \bar{Q}$, the sequences $\{\bar{u}^k(x,t)\}$ and $\{\bar{v}^k(x,t)\}$ are monotone non-increasing, while $\{\underline{u}^k(x,t)\}$ and $\{\underline{v}^k(x,t)\}$ are monotone non-decreasing. Moreover, $(\bar{u}^k, \bar{v}^k)$ is for all $k$ an upper solution while $(\underline{u}^k, \underline{v}^k)$ is a lower solution. Moreover, it holds

$$\underline{u}^k(x,t) \leq \bar{u}^k(x,t) \quad \text{and} \quad \underline{v}^k(x,t) \leq \bar{v}^k(x,t)$$

for all $k = 0, 1, \ldots$ and $(x,t) \in \bar{Q}$.

Let us show exemplarily the monotonicity of the sequence $\{\bar{u}^k(x,t)\}$ and that we have the inequality $\underline{u}^k \leq \bar{u}^k$ in $Q$. To verify the first, we use the definition of $\bar{u}^1$ and the property that $\bar{u}^0$ is an upper solution. This yields

$$\begin{array}{rcl} \bar{u}_t^1 - D_1\Delta\bar{u}^1 + (k_1+\gamma_1 M)\bar{u}^1 &=& F_1(\bar{u}^0, \underline{v}^0),\\ \bar{u}_t^0 - D_1\Delta\bar{u}^0 + (k_1+\gamma_1 M)\bar{u}^0 &\geq& F_1(\bar{u}^0, \underline{v}^0), \end{array}$$

hence

$$\bar{u}_t^0 - \bar{u}_t^1 - D_1\Delta(\bar{u}^0 - \bar{u}^1) + (k_1 + \gamma_1 M)(\bar{u}^0 - \bar{u}^1) \geq 0.$$

Analogously, we have in $\Sigma$

$$\begin{array}{rcl}\partial_\nu \bar{u}^1 &=& b(x,t,\bar{u}^0),\\ \partial_\nu \bar{u}^0 &\geq& b(x,t,\bar{u}^0),\end{array}$$

hence

$$\partial_\nu(\bar{u}^0 - \bar{u}^1) \geq 0.$$

In the same way we obtain in $\Omega$ the inequality $\bar{u}^0(\cdot,0) - \bar{u}^1(\cdot,0) \geq 0$. Therefore, $\bar{u}^0 - \bar{u}^1$ satisfies a linear initial-boundary value problem with nonnegative right-hand sides. From a known comparison principle, see [12], we obtain $\bar{u}^0 - \bar{u}^1 \geq 0$. The analogous inequality $\bar{u}^{k-1} - \bar{u}^k$ follows by induction. In the same way, we show the monotonicity of the other sequences.

The verification of $\underline{u}^k \leq \bar{u}^k$ differs only slightly from the reasoning above: We have

$$\begin{array}{rcl}\underline{u}_t^1 - D_1\Delta\underline{u}^1 + (k_1+\gamma_1 M)\underline{u}^1 &=& F_1(\underline{u}^0,\bar{v}^0) \leq F_1(\underline{u}^0,\underline{v}^0)\\ &\leq& F_1(\bar{u}^0,\underline{v}^0) = \underline{u}_t^1 - D_1\Delta\underline{u}^1 + (k_1+\gamma_1 M)\bar{u}^1.\end{array}$$

Proceeding in this way, we confirm that $\bar{u}^1 - \underline{u}^1$ satisfies also the associated initial-boundary value problem with nonnegative right-hand sides and obtain $\bar{u}^1 \geq \underline{u}^1$. The same holds for the other items of the sequence. We omit the verification of all other claimed inequalities, since their proofs follow by the same arguments.

*(ii) Convergence:* The sequence $\{(\bar{u}^k, \bar{v}^k)\}$ converges uniformly to a solution of **(E1)**. This is seen as follows: The sequence $\{\bar{u}^k(x,t)\}$ is monotone non-increasing and bounded from below by $\hat{u} = 0$. Therefore, it has a (pointwise) limit $u(x,t)$. The sequence is also bounded from above by $M$, hence the Lebesgue dominated convergence theorem shows

$$\lim_{k\to\infty} \iint_Q |\bar{u}^k(x,t) - u(x,t)|^p \, dxdt = 0 \quad \forall 1 \leq p < \infty.$$

The sequence $\{\bar{v}^k(x,t)\}$ has the same property and converges pointwise to a function $v(x,t)$. Since $p$ can be taken arbitrarily large, also the sequence $\{\bar{u}^k\bar{v}^k\}$ converges in any space $L^p(Q)$ with $p < \infty$, and the same holds for the sequence $\{b(\cdot,\cdot,\bar{u}^k)\}$ in any space $L^p(\Sigma)$. Notice that this sequence is uniformly bounded, too. Therefore, the right-hand sides of (3) converge in any $L^p$-space. The functions $\{\bar{v}^{k+1}\}$ and $\{\bar{u}^{k+1}\}$ are solutions to (3). In view of the $L^p$-convergence of the associated right-hand sides, they converge in $Y$ to their limit functions $u$ and $v$, respectively. Notice that the mapping from the right-hand sides to the solution is continuous from $L^r(Q) \times L^s(\Sigma) \times C(\bar{\Omega})$ to $Y$, if $r > N/2+1$, $s > N+1$, cf. Lemma 7.12 in [14]. Passing to the limit in (3), we see that $(u,v)$ solves the system **(E1)**.

*(iii) Uniqueness:* Let $(u_1,v_1),(u_2,v_2) \in Y \times Y$ be two pairs of weak solutions to **(E1)**. Then, $u := u_1 - u_2 \in Y$ and $v := v_1 - v_2 \in Y$ satisfy $u(0) = 0,\ v(0) = 0$,

$$\begin{aligned}(u_t(t),\varphi)_{V^*,V} &+ \int_\Gamma (b(x,t,u_1) - b(x,t,u_2))\varphi \, ds + \int_\Omega D_1\nabla u(t)\cdot\nabla\varphi \, dx\\ &+ \int_\Omega k_1 u(t)\varphi \, dx + \int_\Omega \gamma_1(u(t)v_1(t) + u_2(t)v(t))\varphi \, dx = 0, \end{aligned} \tag{4}$$

$$\begin{aligned}(v_t(t),\varphi)_{V^*,V} &+ \varepsilon\int_\Gamma v\varphi \, ds + \int_\Omega D_2\nabla v(t)\cdot\nabla\varphi \, dx\\ &+ \int_\Omega k_2 v(x,t)\varphi dx + \int_\Omega \gamma_2(u(t)v_1(t) + u_2(t)v(t))\varphi \, dx = 0 \end{aligned} \tag{5}$$

for all $\varphi \in H^1(\Omega)$ and almost all $t \in [0,T]$. Choosing $\varphi = u$ in (4), $\varphi = v$ in (5) and adding both equations, we obtain

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}(\|u(t)\|^2_{L^2(\Omega)} + \|v(t)\|^2_{L^2(\Omega)}) + D_1 \|u(t)\|^2_{H^1(\Omega)} + D_2 \|v(t)\|^2_{H^1(\Omega)} \\
&\quad \le D_1 \|u(t)\|^2_{L^2(\Omega)} + D_2 \|v(t)\|^2_{L^2(\Omega)} + \int_\Omega \gamma_1 \left(|u^2(t)v_1(t)| + |u_2(t)v(t)u(t)|\right)\, dx \\
&\quad\quad + \int_\Omega \gamma_2 \left(|u(t)v_1(t)v(t)| + |u_2(t)v^2(t)|\right)\, dx
\end{aligned}
$$

for almost all $t \in [0,T]$, because the monotonicity of $b$ w.r.t. to $u$ implies

$$
\int_\Gamma (b(x,t,u_1(x,t)) - b(x,t,u_2(x,t)))(u_1(x,t) - u_2(x,t))\, ds \ge 0.
$$

This inequality is equivalent to [5], (A.2). Now, we continue as in the proof [5], A.1.1, to show $u = 0$ and $v = 0$, using Hölder's, Gagliardo-Nirenberg's, Young's, and Gronwall's inequality.

*(iii) Continuity:* We shall prove in the next section by the implicit function theorem that the mapping $c \mapsto (u,v)$ is even twice continuously differentiable. This includes the claimed continuity. □

**Remark 2.3** A study of the proof reveals that $u \ge 0$ and $v \ge 0$, because we have chosen $(0,0)$ as lower solution.

**Theorem 2.4** *Problem **(P1)** admits at least one optimal control $\bar{c}$.*

Proof. Let us only briefly sketch the proof, which is along the lines of [14], Theorem 5.7. We know the uniform boundedness of the states by the proof of Theorem 2.2, because the functions $(\bar{u}^k, \bar{v}^k)$ are bounded by 0 and $M$, respectively. Therefore, also all solutions $(u,v)$ associated with controls $c \in C_{ad}$ obey these bounds. Consequently, the cost functional is bounded from below and we find a minimizing sequence $c_n \in C_{ad}$ converging weakly to a limit function $\bar{c} \in C_{ad}$ in $L^r(\Sigma)$, $N+1 < r < \infty$. One shows in a standard way that this limit is optimal and the associated pair of states fulfills the system **(E1)**. Here, the convexity and continuity of the reduced objective functional is needed that follows from the continuity of the mapping $c \mapsto (u,v)$ stated in Theorem 2.2. □

### 2.2 Necessary and sufficient optimality conditions

Let us define the control-to-state operator $S : L^r(\Sigma) \to Y$ by

$$
S:\ c \mapsto (u,v),
$$

where we fix $r > N+1$ throughout the following. Notice that, by the nonlinear coupling through $-\gamma_i uv$, $i = 1,2$, on the right-hand side of **(E1)**, the system of state equations is nonlinear. Since the cost functional is quadratic, we obtain the next Lemma by standard arguments.

**Lemma 2.1** *The cost functional $J$ is continuously Fréchet-differentiable from $Y \times Y \times L^r(\Sigma)$ to $\mathbb{R}$.*

We show instead:

**Theorem 2.5** *The control-to-state operator $S$ is twice continuously Fréchet-differentiable from $L^r(\Sigma)$ to $Y \times Y$.*

Proof. First, we derive an operator equation for $(u,v) = S(c)$. To this aim, shifting the nonlinearities to the right-hand sides, we transform the state system of **(E1)** to

$$
\begin{array}{rcll}
u_t - D_1\Delta u + k_1 u &=& -\gamma_1 uv & \text{in } Q,\\
v_t - D_2\Delta v + k_2 v &=& -\gamma_2 uv & \text{in } Q,\\
D_1\partial_\nu u &=& c(x,t) - b(x,t,u) & \text{on } \Sigma,\\
D_2\partial_\nu v + \varepsilon v &=& 0 & \text{on } \Sigma,\\
u(x,0) &=& u_0(x) & \text{in } \Omega,\\
v(x,0) &=& v_0(x) & \text{in } \Omega.
\end{array}
\tag{6}
$$

For the left linear part we establish linear and continuous solution operators $S_Q, G_Q : L^r(Q) \to Y$, $S_\Sigma : L^r(\Sigma) \to Y$, and $S_0, G_0 : C(\bar{\Omega}) \to Y$. $S_Q, S_\Sigma$ and $S_0$ are associated with the linear problem

$$\begin{aligned} u_t - D_1\Delta u + k_1 u &= d && \text{in } Q, \\ D_1\partial_\nu u &= c && \text{on } \Sigma, \\ u(x,0) &= e(x) && \text{in } \Omega \end{aligned}$$

in the following sense: $S_Q : d \mapsto u$ with $c = e = 0$; $S_\Sigma : c \mapsto u$ with $d = e = 0$, and $S_0 : e \mapsto u$ with $c = d = 0$. Analogously, we define $G_Q$ and $G_0$ by the system for $v$.
We consider these operators with image in $C(\bar{Q})$ and reformulate the nonlinear equation (6) as

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -S_Q\gamma_1 uv + S_\Sigma(c - b(\cdot,\cdot,u)) + S_0 u_0 \\ -G_Q\gamma_2 uv + G_0 v_0 \end{pmatrix}, \tag{7}$$

which is equivalent to

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} u + S_Q\gamma_1 uv - S_\Sigma(c - b(\cdot,\cdot,u)) - S_0 u_0 \\ v + G_Q\gamma_2 uv - G_0 v_0 \end{pmatrix} =: F(u,v,c).$$

Since $S_Q, G_Q, S_\Sigma, S_0$, and $G_0$ are linear and continuous and $-\gamma_1 uv, -\gamma_2 uv$, $b(\cdot,\cdot,u)$ are twice continuously Fréchet-differentiable from $C(\bar{Q})$ to $L^\infty(Q)$, respectively $L^\infty(\Sigma)$, $F$ is a twice continuously Fréchet-differentiable mapping from $C(\bar{Q}) \times C(\bar{Q}) \times L^r(\Sigma)$ to $C(\bar{Q}) \times C(\bar{Q})$.
To use the implicit function theorem, we have to show the bounded, i.e. continuous, invertibility of the partial Fréchet-derivative $F_{(u,v)}(u,v,c)$. To verify this property is true, we first mention that the equation $F_{(u,v)}(u,v,c)w = z$ is equivalent to the system

$$\begin{pmatrix} w_1 + S_Q\gamma_1(vw_1 + uw_2) + S_\Sigma b_u(x,t,u)w_1 \\ w_2 + G_Q\gamma_2(vw_1 + uw_2) \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

Because the mapping $z \mapsto w$ is not smoothing, we substitute $r_i = z_i - w_i$, $i = 1,2$, to obtain the equivalent system

$$\begin{aligned} (r_1)_t - D_1\Delta r_1 + k_1 r_1 + \gamma_1(vr_1 + ur_2) &= \gamma_1(vz_1 + uz_2) && \text{in } Q, \\ (r_2)_t - D_2\Delta r_2 + k_2 r_2 + \gamma_2(vr_1 + ur_2) &= \gamma_2(vz_1 + uz_2) && \text{in } Q, \\ D_1\partial_\nu r_1 + b_u(x,t,u)r_1 &= c(x,t) + b_u(x,t,u)z_1 && \text{on } \Sigma, \\ D_2\partial_\nu r_2 + \varepsilon r_2 &= 0 && \text{on } \Sigma, \\ u(x,0) &= 0 && \text{in } \Omega, \\ v(x,0) &= 0 && \text{in } \Omega. \end{aligned}$$

For every $(z_1, z_2) \in C(\bar{Q})^2$, this boundary value problem has a unique solution $(r_1, r_2) \in Y \times Y$, cf. e.g. Theorem 5.5 in [14]. The mapping $(z_1,z_2) \mapsto (r_1,r_2)$ is continuous, hence also the mapping $(z_1,z_2) \mapsto (w_1,w_2)$. Therefore, we can invoke the implicit function theorem and obtain that the control-to-state operator $S$ is twice continuously Fréchet-differentiable. □

In particular, this theorem covers the continuity of $S$. Having the existence of the first- and second-order derivatives, it is now easy to conclude their concrete form by implicit differentiation. We begin with computing the first derivative of $S$.

**Corollary 2.6** *The derivative of the control-to-state operator $S$ at $\bar{c}$ in direction $c$ is given by*

$$S'(\bar{c})c = (u,v),$$

*where $(u,v)$ is the weak solution of the linearized equation obtained by linearizing system **(E1)** at $(\bar{u},\bar{v}) := S(\bar{c})$,*

$$\begin{aligned} u_t - D_1\Delta u + k_1 u &= -\gamma_1(\bar{u}v + u\bar{v}) && \textit{in } Q, \\ v_t - D_2\Delta v + k_2 v &= -\gamma_2(\bar{u}v + u\bar{v}) && \textit{in } Q, \\ D_1\partial_\nu u + b_u(x,t,\bar{u})u &= c && \textit{on } \Sigma, \\ D_2\partial_\nu v + \alpha v &= 0 && \textit{on } \Sigma, \\ u(x,0) &= 0 && \textit{in } \Omega, \\ v(x,0) &= 0 && \textit{in } \Omega. \end{aligned} \tag{8}$$

P r o o f. Let us briefly sketch the proof in a slightly formal way. The system **(E1)** is of the form

$$A\, S(c) = B(S(c)) + C(S(c)) + D\, c + w_0,$$

where $S : c \mapsto (u(c), v(c))$, $B(u, v) = -(\gamma_1 uv, \gamma_2 uv)$, $C(u) = (-b(\cdot, \cdot, u), 0)$, $A$ stands for the linear differential operator on the left-hand side of **(E1)**, $w_0$ for the initial conditions and $D$ is a linear continuous operator. Therefore it holds

$$A\, S'(c)c_1 = B'(S(c))S'(c)c_1 + C'(S(c))S'(c)c_1 + D\, c_1. \tag{9}$$

The system (8) is obtained with $c_1 = c$, $(u, v) = S'(c)c_1$, $S(c) = (\bar{u}, \bar{v})$, since

$$B'(\bar{u}, \bar{v})(u, v) = -(\gamma_1(\bar{u}v + u\bar{v}), \gamma_2(\bar{u}v + u\bar{v}))$$

and

$$C'(\bar{u}, \bar{v})(u, v) = -(b_u(\cdot, \cdot, \bar{u})u, 0).$$

□

Similarly, we obtain the second derivative of $S$.

**Corollary 2.7** *The second derivative of $S$ at $\bar{c}$ in direction $(\hat{c}, \tilde{c})$ is given by*

$$S''(\bar{c})(\hat{c}, \tilde{c}) = (u, v),$$

*where $(u, v)$ is the weak solution of the system*

$$\begin{aligned}
u_t - D_1\Delta u + k_1 u + \gamma_1(\bar{u}v + u\bar{v}) &= -\gamma_1(\hat{u}\tilde{v} + \tilde{u}\hat{v}) && \text{in } Q,\\
v_t - D_2\Delta v + k_2 v + \gamma_2(\bar{u}v + u\bar{v}) &= -\gamma_2(\hat{u}\tilde{v} + \tilde{u}\hat{v}) && \text{in } Q,\\
D_1\partial_\nu u + b_u(x, t, \bar{u})u &= -b_{uu}(x, t, \bar{u})\hat{u}\tilde{u} && \text{on } \Sigma,\\
D_2\partial_\nu v + \varepsilon v &= 0 && \text{on } \Sigma,\\
u(x, 0) &= 0 && \text{in } \Omega,\\
v(x, 0) &= 0 && \text{in } \Omega
\end{aligned}$$

*where $(\bar{u}, \bar{v}) = S(\bar{c})$, $(\hat{u}, \hat{v}) = S'(\bar{c})(\hat{c})$ and $(\tilde{u}, \tilde{v}) = S'(\bar{c})(\tilde{c})$, see (8).*

P r o o f. Differentiating (9) with respect to $c$ in direction $c_2$ yields

$$\begin{aligned}
A\, S''(c)[c_1, c_2] =& B'(S(c))S''(c)[c_1, c_2] + B''(S(c))[S'(c)c_1, S'(c)c_2]\\
&+ C'(S(c))S''(c)[c_1, c_2] + C''(S(c))[S'(c)c_1, S'(c)c_2]
\end{aligned} \tag{10}$$

since $D'' = 0$ by linearity. Define $c_1 = \hat{c}$, $c_2 = \tilde{c}$, $(u, v) = S''(c)[c_1, c_2]$, $(\hat{u}, \hat{v}) = S'(c)c_1$, $(\tilde{u}, \tilde{v}) = S'(c)c_2$, then the claimed result on the second derivative of $S$ is obtained. Notice that $B''(u, v)[(\hat{u}, \hat{v}), (\tilde{u}, \tilde{v})] = -(\gamma_1(\hat{u}\tilde{v} + \tilde{u}\hat{v}), \gamma_2(\hat{u}\tilde{v} + \tilde{u}\hat{v}))$. For a less formal way of deriving the associated derivatives of $S'$ and $S''$ we refer, for instance, to [14], Sects. 5.4 and 5.7. □

To formulate necessary optimality conditions, let $\bar{c}$ be an optimal control of **(P1)** with states $(\bar{u}, \bar{v})$.
We have $(u, v) = S(c)$ with the control-to-state operator $S : L^r(\Sigma) \to Y \times Y$, hence we obtain the reduced functional $f$,

$$f(c) := J(u, v, c) = J(S(c), c).$$

Let us write for short $y = (u, v)$, $\bar{y} = (\bar{u}, \bar{v})$. The functional $f$ is Fréchet differentiable, because $S$ is differentiable by Theorem 2.5 and $J$ is differentiable by Lemma 2.1. Because the set of admissible controls $C_{ad}$ is convex, we obtain the following standard result, cf. for instance [14], Lemma 2.21.

**Lemma 2.2** *Every locally optimal control function $\bar{c}$ of **(P1)** satisfies the variational inequality*

$$f'(\bar{c})(c - \bar{c}) \geq 0 \qquad \forall c \in C_{ad}.$$

We determine $f'$ by the chain rule and obtain for the direction $c$

$$\begin{aligned} f'(\bar{c})c &= J_y(\bar{y},\bar{c})\, S'(\bar{c})c + J_c(\bar{y},\bar{c})c \\ &= \iint\limits_Q \Big(\alpha_u(\bar{u}-u_Q)u + \alpha_v(\bar{v}-v_Q)v\Big)\, dx\, dt \\ &\quad + \int\limits_\Omega \Big(\alpha_{TU}(\bar{u}(T)-u_{QT})u(T) + \alpha_{TV}(\bar{v}(T)-v_{QT})v(T)\Big)\, dx \\ &\quad + \iint\limits_\Sigma \alpha_c\, \bar{c}\, c\, dsdt, \end{aligned} \tag{11}$$

where, by Corollary 2.6, $(u,v) = S'(\bar{c})c$ is the solution of the linearized system (8). Now we apply this result to $c := c - \bar{c}$ with $c \in C_{ad}$. By Lemma 2.2, $f'(\bar{c})(c-\bar{c})$ is nonnegative. We can eliminate the states $u$ and $v$ in (11) by adjoint states $p$ and $q$, defined as the solutions of the adjoint system

$$\textbf{(A1)}\left\{\begin{array}{rcll} -p_t - D_1\Delta p + k_1 p + \gamma_1 v p + \gamma_2 v q &=& \alpha_u(u-u_Q) & \text{in } Q, \\ -q_t - D_2\Delta q + k_2 q + \gamma_1 u p + \gamma_2 u q &=& \alpha_v(v-v_Q) & \text{in } Q, \\ D_1\partial_\nu p + b_u(x,t,u)p &=& 0 & \text{on } \Sigma, \\ D_2\partial_\nu q + \varepsilon q &=& 0 & \text{on } \Sigma, \\ p(x,T) = \alpha_{TU}(u(x,T)-u_\Omega(x)) &&& \text{in } \Omega, \\ q(x,T) = \alpha_{TV}(v(x,T)-v_\Omega(x)) &&& \text{in } \Omega. \end{array}\right. \tag{12}$$

**Lemma 2.3** *If $(u,v)$ is the weak solution of the linearized system (8) and $(p,q)$ is the solution of the adjoint system **(A1)**, then it holds for all $c \in L^r(\Sigma)$*

$$\begin{aligned} &\iint\limits_Q \Big(\alpha_u(\bar{u}-u_Q)u + \alpha_v(\bar{v}-v_Q)v\Big)\, dx\, dt \\ &\quad + \int\limits_\Omega \Big(\alpha_{TU}(\bar{u}(T)-u_{QT})u(T) + \alpha_{TV}(\bar{v}(T)-v_{QT})v(T)\Big)\, dx \\ &= \iint\limits_\Sigma p(c-\bar{c})\, dsdt. \end{aligned}$$

The proof is analogous to the one of [14], Lemma 5.11, hence we skip it. In this way, (11) leads to

$$f'(\bar{c})\, c = \iint\limits_\Sigma (p + \alpha_c\bar{c})\, c\, dsdt. \tag{13}$$

Thanks to Lemma 2.2, it follows the

**Theorem 2.8** *Every locally optimal solution $\bar{c}$ of **(P1)** satisfies, together with the adjoint states $(p,q)$ of **(A1)**, the variational inequality*

$$\iint\limits_\Sigma (p+\alpha_c\bar{c})(c-\bar{c})\, dsdt \;\geq\; 0 \qquad \forall c \in C_{ad}.$$

An equivalent pointwise expression of the variational inequality is

$$\min_{c_a(x,t)\leq c\leq c_b(x,t)} (p(x,t)+\alpha_c\bar{c}(x,t))\, c = (p(x,t)+\alpha_c\bar{c}(x,t))\, \bar{c}(x,t),$$

i.e. the min on the left-hand side will be attained almost everywhere in $\Sigma$ by $c = \bar{c}(x,t)$. For $\alpha_c > 0$, this leads in a standard way to the projection formula

$$\bar{c}(x,t) = \mathbb{P}_{[c_a(x,t),c_b(x,t)]}\{-\frac{1}{\alpha_c}p(x,t)\}$$

for almost all $(x,t) \in \Sigma$, where $\mathbb{P}_{[c_a(x,t),c_b(x,t)]} : \mathbb{R} \to [c_a(x,t), c_b(x,t)]$ denotes the projection onto $[c_a(x,t), c_b(x,t)]$.

Next, we consider also sufficient second order optimality conditions for **(P1)**. Let $\bar{c} \in C_{ad}$ together with $(\bar{u}, \bar{v}) = S(\bar{c})$ and the adjoint states $(p, q)$ satisfy the first-order necessary optimality conditions, presented in Theorem 2.8. We want to set up sufficient conditions for $(\bar{c}, \bar{u}, \bar{v})$ to be a local optimum.

Because the cost functional $J$ and the control-to-state operator $S$ are twice continuously Fréchet-differentiable, the reduced functional $f$ is also twice continuously Fréchet-differentiable. By the chain rule, we derive

$$f'(c)\hat{c} = D_y J(S(c), c)S'(c)\hat{c} + D_c J(S(c), c)\hat{c}.$$

The derivative of $f'(c)\hat{c}$ with respect to $c$ in direction $\tilde{c}$ is

$$\begin{aligned} f''(c)[\hat{c}, \tilde{c}] &= D_y^2 J(S(c), c)[S'(c)\hat{c}, S'(c)\tilde{c}] + 2D_c D_y J(S(c), c)[S'(c)\hat{c}, \tilde{c}] \\ &\quad + D_y J(S(c), c)S''(c)[\hat{c}, \tilde{c}] + D_c^2 J(S(c), c)[\hat{c}, \tilde{c}] \\ &= J''(y, c)[(\hat{y}, \hat{c}), (\tilde{y}, \tilde{c})] + D_y J(y, c)z \end{aligned}$$

with states $z := (z_1, z_2) = S''(c)[\hat{c}, \tilde{c}]$, $\hat{y} := S'(c)\hat{c}$ and $\tilde{y} := S'(c)\tilde{c}$, see e.g. (9) and (10). Analogously to Lemma 2.3, the last expression,

$$\begin{aligned} D_y J(y, c)z = &\iint_Q \Big(\alpha_u (u - u_Q) z_1 + \alpha_v (v - v_Q) z_2\Big)\, dx\, dt \\ &+ \int_\Omega \Big(\alpha_{TU}(u(T) - u_{QT}) z_1(T) + \alpha_{TV}(v(T) - v_{QT}) z_2(T)\Big)\, dx \end{aligned}$$

containing yet the state $y = (u, v)$, can be transformed to

$$D_y J(y, c)z = \iint_Q \gamma_1(\hat{u}\tilde{v} + \tilde{u}\hat{v})p + \gamma_2(\hat{u}\tilde{v} + \tilde{u}\hat{v})q\; dxdt - \iint_\Sigma p b_{uu}(x, t, u)\hat{u}\tilde{u}\; dsdt$$

by using the adjoint states $(p, q)$ of **(A1)**. We derive

$$\begin{aligned} f''(c)[\hat{c}, \tilde{c}] = &J''(y, c)[(\hat{y}, \hat{c}), (\tilde{y}, \tilde{c})] + \iint_Q \gamma_1(\hat{u}\tilde{v} + \tilde{u}\hat{v})p + \gamma_2(\hat{u}\tilde{v} + \tilde{u}\hat{v})q\; dxdt. \\ &- \iint_\Sigma b_{uu}(x, t, u)[\hat{u}, \tilde{u}]p\; dsdt. \end{aligned} \tag{14}$$

To formulate our sufficient optimality conditions in a more convenient form, we introduce the Lagrange function $\mathcal{L} : Y \times Y \times L^\infty(\Sigma) \times Y \times Y \to \mathbb{R}$ by

$$\begin{aligned} \mathcal{L}(u, v, c, p, q) = \quad & J(u, v, c) + \iint_Q (u_t + k_1 u + \gamma_1 uv)p\; dxdt + \iint_Q D_1 \nabla u \cdot \nabla p\; dxdt \\ &+ \iint_\Sigma (b(x, t, u) - c)p\; dsdt + \int_\Omega (u(x, 0) - u_0)p(x, 0)\; dx \\ &+ \iint_Q (v_t + k_2 v + \gamma_2 uv)q\; dxdt + \iint_Q D_2 \nabla v \cdot \nabla q\; dxdt \\ &+ \iint_\Sigma \varepsilon v\, q\; dsdt + \int_\Omega (v(x, 0) - v_0)q(x, 0)\; dx. \end{aligned} \tag{15}$$

In view of (14), we obtain

$$\begin{aligned} &\mathcal{L}''(u, v, c, p, q)[(\hat{u}, \hat{v}, \hat{c}), (\tilde{u}, \tilde{v}, \tilde{c})] = J''(u, v, c)[(\hat{u}, \hat{v}, \hat{c}), (\tilde{u}, \tilde{v}, \tilde{c})] \\ &+ \iint_Q \gamma_1(\hat{u}\tilde{v} + \tilde{u}\hat{v})p + \gamma_2(\hat{u}\tilde{v} + \tilde{u}\hat{v})q\; dxdt - \iint_\Sigma b_{uu}(x, t, u)\hat{u}\tilde{u}p\; dsdt \\ &= f''(c)[(\hat{u}, \hat{v}, \hat{c}), (\tilde{u}, \tilde{v}, \tilde{c})] \end{aligned}$$

This result was to be expected, since the second derivative $f''$ of the reduced objective functional can be expressed in general by $\mathcal{L}''$ defined with the associated adjoint states, cf. [14]. By the variational inequality for the optimal solution $\bar{c}$ of **(P1)**, see Lemma 2.8, we obtain

$$\bar{c} = \begin{cases} c_a, \text{ if } p + \alpha_c \, \bar{c}(t) > 0 \\ c_b, \text{ if } p + \alpha_c \, \bar{c}(t) < 0. \end{cases} \tag{16}$$

Therefore, the first-order conditions fix the control function $\bar{c}$ in the set $\{(x,t) \in \Sigma : |(p + \alpha_c \bar{c})(x,t)| > 0\}$. Second-order sufficient conditions should be required on the remaining sets. For given $\tau > 0$, we define

$$A_\tau(\bar{c}) := \{(x,t) \in \Sigma : |p + \alpha_c \bar{c}| > \tau\}$$

as the *set of strongly active restrictions* for $\bar{c}$. The $\tau$*-critical cone* $C_\tau(\bar{c})$ is made up of all $c \in L^\infty(\Sigma)$ with

$$c(x,t) \begin{cases} = 0 & \text{for } (x,t) \in A_\tau(\bar{c}) \\ \geq 0 & \text{for } \bar{c}(x,t) = c_a \text{ and } (x,t) \notin A_\tau(\bar{c}) \\ \leq 0 & \text{for } \bar{c}(x,t) = c_b \text{ and } (x,t) \notin A_\tau(\bar{c}). \end{cases}$$

**Remark 2.9** This is the critical cone appearing in a natural way in second-order necessary conditions. In the case of sufficient conditions, where the $L^2$-norm occurs, one might consider also the same cone in $L^2(\Sigma)$. This however, will not give new conditions since $L^\infty(\Sigma)$ is dense in $L^2(\Sigma)$.

**Theorem 2.10** (Second-order sufficient conditions) *Suppose that the control function $\bar{c}$ satisfies the first-order necessary optimality conditions of Theorem 2.8. If there exist positive constants $\delta$ and $\tau$ such that*

$$\mathcal{L}''(\bar{u}, \bar{v}, \bar{c}, p, q)(u, v, c)^2 \geq \delta \|c\|^2_{L^2(0,T)}$$

*holds for all $c \in C_\tau(\bar{c})$ and all $(u,v) \in Y \times Y$ satisfying the linearized equation (8), then we find positive constants $\varepsilon$ and $\sigma$ such that the quadratic growth condition*

$$J(u, v, c) \geq J(\bar{u}, \bar{v}, \bar{c}) + \sigma \|c - \bar{c}\|^2_{L^2(Q)}$$

*holds for all $c \in C_{ad}$ with $\|c - \bar{c}\|_{L^\infty(Q)} \leq \varepsilon$. Therefore, the control function $\bar{c}$ is locally optimal in the sense of $L^\infty(Q)$.*

Notice that the so-called two-norm discrepancy occurs for parabolic control problems in space dimension larger than one. In the parabolic case, $L^2$ controls are only transformed continuously to bounded state functions, if they appear distributed and the space dimension is one. If the dimension is larger than one or the control appears in a boundary condition, then the two-norm discrepancy cannot be avoided.

For the technique of proving this theorem, we refer to [14], Theorem 5.17.

### 2.3 Numerical examples

The optimality system consists of the nonlinear state equations **(E1)**, the system of adjoint equations **(A1)** and the projection formula

$$\bar{c}(x,t) = \mathbb{P}_{[c_a(x,t), c_b(x,t)]}\{-\frac{1}{\alpha_c} p(x,t)\}$$

for $\alpha_c > 0$ and almost all $(x,t) \in \Sigma$. Inserting this in the system **(E1)**, we obtain a nonlinear and non-smooth coupled system of parabolic equations.

A direct numerical solution of this system by available commercial codes for PDEs that use multigrid techniques turned out to be very successful. We used the software package COMSOL MULTIPHYSICS[1] and interpreted the time as an additional space dimension, to solve the whole optimality system in the space-time cylinder by finite elements. Then the built-in damped Newton solver can be used, and fully coupled space-time adaptivity can be easily applied, see [9], [10]. We replaced the control function $c$ by the projection formula due to the control constraints. In this way, our optimal controls were determined in short running time and very precisely. We do not have a convergence theorem that the method will work for all problems posed. In other words, we cannot be sure a priori that the method converges. However, after applying it to our examples, the optimality systems were solved very precisely in a short time. Notice that, whenever the numerical method converges, the computed solution obeys the optimality system, since it is this system that is solved numerically.

[1] registered trademark of Comsol Ab

This evident fact is exemplarily confirmed in the next section in Fig. 6.

**Example 1:** We investigate the problem **(P1)** with the following data: $\Omega = [0,1] \times [0,1], T = 3, D_1 = D_2 = k_1 = k_2 = 1$, $\gamma_1 = \gamma_2 = 0.3$, $\alpha_{TU} = \alpha_{TV} = 10$, $\alpha_u = \alpha_v = 0$, $\alpha_c = 0.01$, $\varepsilon = 0.1$, $u_0(x) = 0$, $v_0(x) = 100$, $u_Q \equiv v_Q \equiv 0$, $v_\Omega \equiv 1$, $u_\Omega = \sin(2\pi x) + 1$. The control function is acting on the boundaries $y = 0$ and $y = 1$ and we take $c = 0$ on the boundaries $x = 0$ and $x = 1$. For the constraints of the control function $c$ we take $c_a \equiv 0$, $c_b \equiv 20$ and for the nonlinear boundary function $b$ we select $b = u^4$. We obtain the control functions and states presented in Figs. 1-3. The control function at $y = 1$ is exactly the same as on $y = 0$ but mirrored. The $z$-axis illustrates the variation in time. Since the distance of the computed $\bar{u}$ to $u_\Omega$ is very small, we can assume that $\bar{u}$ yields a minimum of $J$.

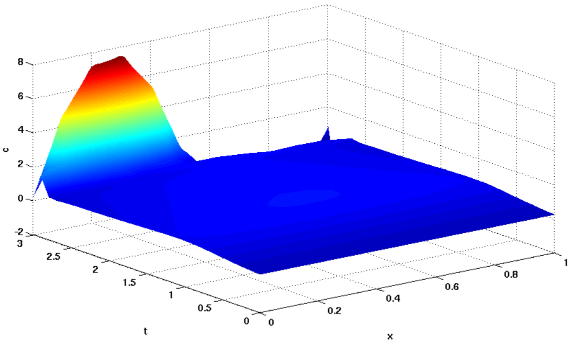


**Fig. 1** Optimal control $\bar{c}(x, 0, t)$ at $y = 0$ for Example 1

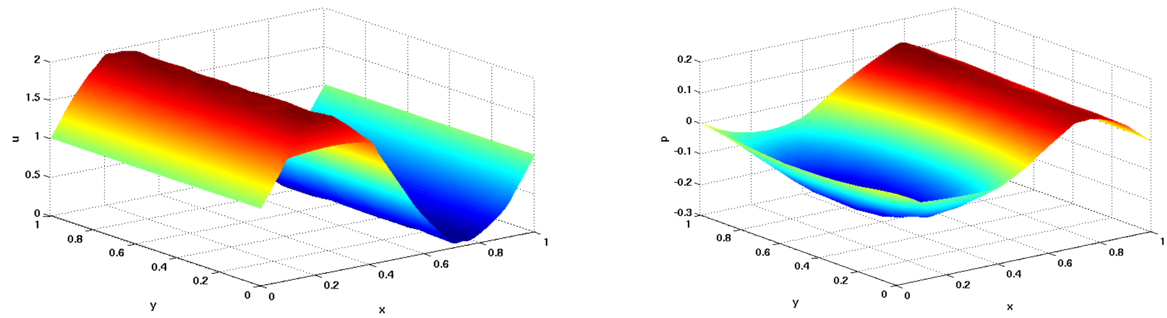


**Fig. 2** State $u$ at $t = T$ and adjoint $p$ at $t = T$ for Example 1

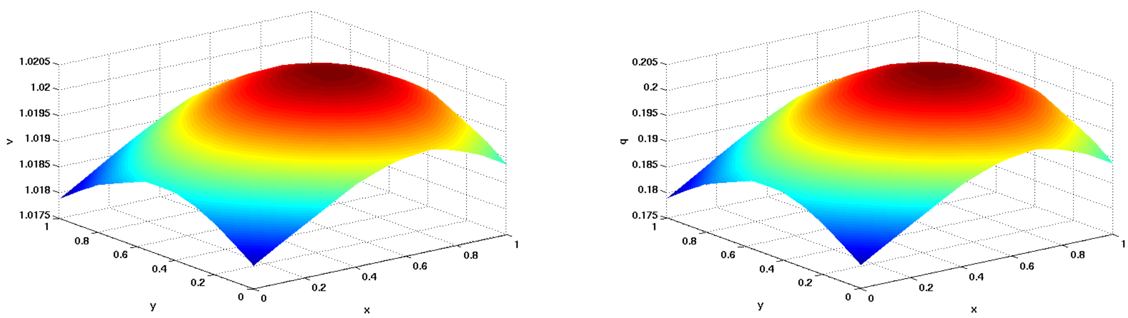


**Fig. 3** State $v$ at $t = T$ and adjoint $q$ at $t = T$ for Example 1

# 3 Special case: a problem of catalysis with linear boundary conditions

## 3.1 Existence of an optimal control

In this section, we consider a problem with linear boundary condition that is similar to one with almost identical state equations but different cost functional discussed by R. Griesse and S. Volkwein in [5].

Let us briefly explain this application to motivate our setting: In a catalyst, two substances are contained with concentrations $u$ and $v$. One of them is a harmful substance $v$ controlled by $d$, which we want to neutralize with the other substance $u$ controlled by $c$. In our examples, we consider $d$ as given and $c$ as the only control function. This function stands for a substrate with concentration $c$, which influences the catalyst at its boundary. The diffusion process is modelled by **(E2)** below. The term $uv$ in **(E2)** describes the chemical reaction in a low order approximation. For more accuracy, it is possible to replace this term by a stronger nonlinear coupling, see [1]. The catalyst will operate more efficiently, if the substances are inserted separately in alternating intervals of time. This alternative is considered in a variant of **(E2)** discussed in Section 3.2. In reality, the constants are depending on the temperature, the pressure, $u$ and $v$. We ignore these dependencies here. The process of neutralization is modeled by our cost functional. We assume that the ratio $u : v$ should be equal to $k$ in order to neutralize the substance $v$. These assumptions lead to the following optimal control problem **(P2)**, see [1], [8]:

$$\textbf{(P2)}\ \min J(u,v,c,d) := \frac{1}{2}\iint\limits_Q (u(x,t)-kv(x,t))^2\ dxdt + \frac{\lambda_1}{2}\int\limits_0^T c^2(t)\ dt + \frac{\lambda_2}{2}\int\limits_0^T d^2(t)\ dt$$

subject to the system of semilinear parabolic PDEs

$$\textbf{(E2)}\left\{\begin{array}{rcll} u_t - D_1 u_{xx} + k_1 u &=& -\gamma_1 uv & \text{in } Q, \\ v_t - D_2 v_{xx} + k_2 v &=& -\gamma_2 uv & \text{in } Q, \\ u(0,t) - D_1 u_x(0,t) &=& c(t) & \text{in } (0,T), \\ D_1 u_x(l,t) &=& 0 & \text{in } (0,T), \\ v(0,t) - D_2 v_x(0,t) &=& d(t) & \text{in } (0,T), \\ D_2 v_x(l,t) &=& 0 & \text{in } (0,T) \\ u(x,0) &=& u_0(x) & \text{in } \Omega, \\ v(x,0) &=& v_0(x) & \text{in } \Omega \end{array}\right.$$

and the box constraints

$$c \in C_{ad} = \{c \in L^2(0,T) | c_a(t) \le c(t) \le c_b(t) \text{ a.e. on } [0,T]\} \subset L^\infty(0,T),$$

$$d \in D_{ad} = \{d \in L^2(0,T) | d_a(t) \le d(t) \le d_b(t) \text{ a.e. on } [0,T]\} \subset L^\infty(0,T)$$

for a final time $T > 0$. In this setting, $\Omega$ denotes the open interval $(0,l)$ and $Q = \Omega \times (0,T)$ is the space-time cylinder. The functions $c_a$,$c_b$, $d_a$ and $d_b$ are given of $L^\infty(0,T)$, such that $c_a \le c_b$, $d_a \le d_b$ holds almost everywhere in $[0,T]$. We denote by $\lambda_1$, $\lambda_2$, $D_1$, $D_2$ positive and by $k_1$, $k_2$, $\gamma_1$ and $\gamma_2$ nonnegative constants. The control functions $c$ and $d$ are considered as elements of the space $L^\infty(0,T)$ and the fixed initial values $u_0$ and $v_0$ are elements of the space $L^2(\Omega)$.
In contrast to the last section with nonlinear boundary conditions, we consider a one-dimensional domain and two control functions $c(t)$ and $d(t)$. Let us define $b$ as

$$b(u,x,t) := \left\{\begin{array}{ll} u & \text{in } x=0 \\ 0 & \text{in } x=l \end{array}\right. \text{ and } \varepsilon := \left\{\begin{array}{ll} 1 & \text{in } x=0 \\ 0 & \text{in } x=l \end{array}\right. .$$

The control functions are acting in Robin boundary conditions with pointwise control constraints on both sides. A similar problem with almost identical state equations but a different cost functional was considered in [5] and [4]. In these papers, necessary and sufficient optimality conditions are derived and numerical techniques for this class of optimal control problem were suggested. We shortly mention the relevant theorems, which also can be deduced from the former sections.

**Theorem 3.1** *For each given pair of controls $(c,d) \in L^2(0,T) \times L^2(0,T)$, there exists a unique solution $(u,v) \in Y \times Y$ to **(E2)**.*

**Theorem 3.2** *Problem **(P2)** admits at least one optimal solution.*

The control-to-state operator $S$ of **(E2)** is now a mapping from $L^2(0,T) \times L^2(0,T)$ to $Y \times Y$.

**Theorem 3.3** *The control-to-state operator $S$ is twice continuously Fréchet-differentiable.*

This result follows by obvious modifications as in Theorem 2.5. Analogously to the last section, we obtain the following adjoint system

$$
\begin{array}{rcll}
-p_t - D_1 p_{xx} + k_1 p + \gamma_1 \bar{v} p + \gamma_2 \bar{v} q & = & \bar{u} - k\bar{v} & \text{in } Q, \\
p(0,t) - D_1 p_x(0,t) & = & 0 & \text{in } (0,T), \\
D_1 p_x(l,t) & = & 0 & \text{in } (0,T), \\
p(x,T) & = & 0 & \text{in } \Omega, \\
-q_t - D_2 q_{xx} + k_2 q + \gamma_1 \bar{u} p + \gamma_2 \bar{u} q & = & -k(\bar{u} - k\bar{v}) & \text{in } Q, \\
q(0,t) - D_2 q_x(0,t) & = & 0 & \text{in } (0,T), \\
D_2 q_x(l,t) & = & 0 & \text{in } (0,T), \\
q(x,T) & = & 0 & \text{in } \Omega.
\end{array}
\tag{17}
$$

**Theorem 3.4** *Every locally optimal pair of control functions $(\bar{c}, \bar{d})$ of **(P2)** satisfies, with a pair of adjoint states $(p, q)$ defined by (17) the variational inequalities*

$$
\int_0^T (p(0,t) + \lambda_1 \bar{c}(t))(c(t) - \bar{c}(t))\, dt \ \geq\ 0 \qquad \forall c \in C_{ad},
$$

$$
\int_0^T (q(0,t) + \lambda_2 \bar{d}(t))(d(t) - \bar{d}(t))\, dt \ \geq\ 0 \qquad \forall d \in D_{ad}.
$$

*If $\lambda_1$ and $\lambda_2$ are positive, then the inequalities are equivalent to the pointwise projection formulas*

$$
\begin{array}{rcl}
\bar{c}(t) & = & \mathbb{P}_{[c_a(t), c_b(t)]}\{-\frac{1}{\lambda_1} p(0,t)\}, \\
\bar{d}(t) & = & \mathbb{P}_{[d_a(t), d_b(t)]}\{-\frac{1}{\lambda_2} q(0,t)\}
\end{array}
$$

*for almost all $t \in [0, T]$.*

### 3.2 Numerical examples

Here, we consider examples related to the catalysis problem explained in the introduction. We consider $d$ as a periodic piecewise constant function that is given fixed. This means that the harmful substance is fed periodically into the catalyst by a certain quantity $d_0$ where $d$ has the form

$$
d(t) = \begin{cases} d_0 \text{ on } [T/4, T/2[ \cup [3T/4, T[ \\ 0 \text{ on } [0, T/4[ \cup [T/2, 3T/4[. \end{cases}
$$

We assume that we are able to inject the harmless substance only when the harmful substance is not injected. Hence, we define the bounds $c_a$ and $c_b$ as functions presented in Fig. 4 where

$$
c_i(t) = \begin{cases} 0 \text{ on } [T/4, T/2[ \cup [3T/4, T[ \\ \tilde{c}_i \text{ on } [0, T/4[ \cup [T/2, 3T/4[ \end{cases}
$$

with $i = a, b$ with $c_a \leq c_b$. They are periodic and piecewise constant functions with the only possible values 0 and $\tilde{c}_i$, $i = a, b$.

**Example 2**:
Setting $l = 1$, $T = 10$, $k = k_1 = k_2 = 1$, $\gamma_1 = \gamma_2 = 0.5$, $\alpha_1 = \alpha_2 = 0.3$, $\lambda_1 = \lambda_2 = 0.001$, $d_0 = 7$, $u_0 = v_0 \equiv 0$ and defining the control bounds by

$$
c_a(t) = \begin{cases} 0 \text{ on } [T/4, T/2[ \cup [3T/4, T[ \\ 1 \text{ on } [0, T/4[ \cup [T/2, 3T/4[ \end{cases}
$$

and

$$
c_b(t) = \begin{cases} 0 \text{ on } [T/4, T/2[ \cup [3T/4, T[ \\ 10 \text{ on } [0, T/4[ \cup [T/2, 3T/4[, \end{cases}
$$

we obtain the results, presented in the Figs. 5 and 6. Fig. 5 shows the compatibility of the control function $c$ and the projection $\mathbb{P}_{[c_a, c_b]}\{-\frac{1}{\lambda_1} p(0,t)\}$ with respect to the necessary optimality conditions. We see perfect coincidence of both

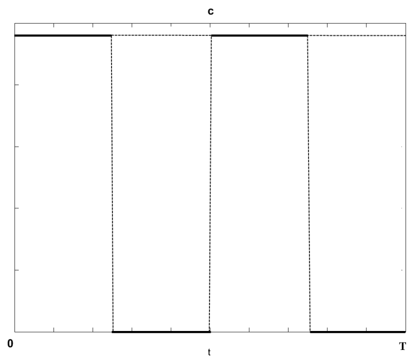


**Fig. 4** Periodic, piecewise constant function with a period of length $\frac{T}{2}$.

pictures, hence the necessary optimality conditions are satisfied. This was to be expected, since the numerical method converged. Therefore, the whole optimality system must have been solved.
All the solutions obtained in Sect. 3 were solved by a gradient method written in MATLAB so that we can be sure to have approximated (local) minima. We doublechecked them by solving again the optimality system with COMSOL MULTIPHYSICS in the same way as in Sect. 2. The result was the same.

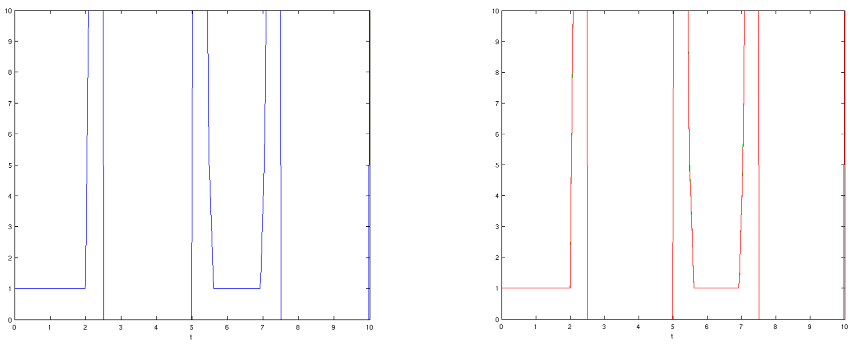

**Fig. 5** Optimal control $\bar{c}$ and projection $\mathbb{P}_{[c_a,c_b]}\{-\frac{1}{\lambda_1}p(0,t)\}$ for Example 2

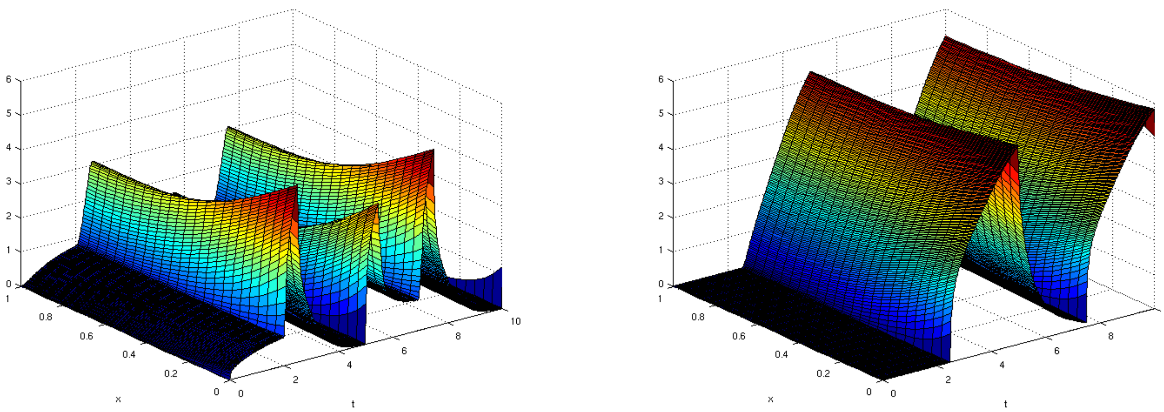

**Fig. 6** State $u$ (left) and state $v$ (right) for Example 2

Next, we consider the same data as in Example 1, but we choose $d_0 = 12$ in Fig. 7 left , while we set $\lambda_1 = \lambda_2 = 1$ in Fig. 7 right.

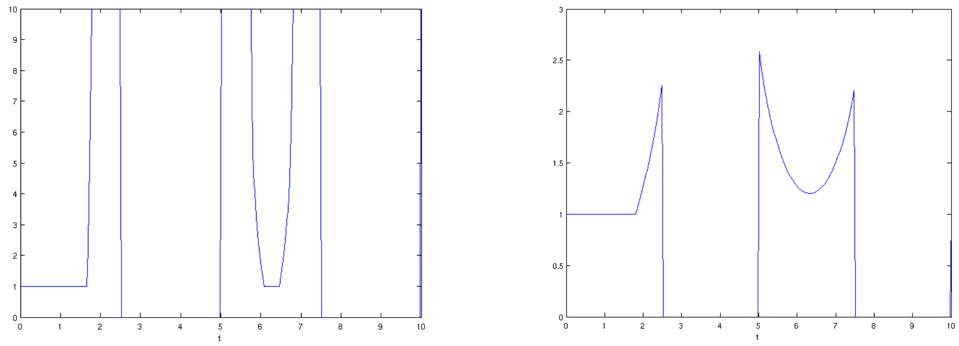

**Fig. 7** Optimal controls $\bar{c}$ for the second part of Example 2

**Example 3:** The computed optimal controls may exhibit a slightly irregular structure, cf. e.g. Fig. 7. From the application point of view, a control $c$ seems to be useful, which is piecewise constant and is concentrated on the intervals, where the harmful substrate is not injected. In view of this, we next reformulate our control problem. We consider the control function as a function described in Fig. 4 and optimize only the height $c_0$. This leads to:

$$\min\ J(c_0, u, v) = \frac{1}{2} \iint_Q (u - kv)^2\, dxdt + c_0^2 \frac{\lambda_1}{2} \int_0^T e\, dt,$$

where $e$ has the form

$$e(t) = \begin{cases} 0 \text{ on } [T/4, T/2[ \cup [3T/4, T[ \\ 1 \text{ on } [0, T/4[ \cup [T/2, 3T/4[. \end{cases}$$

We obtain for the associated reduced functional $f : \mathbb{R} \to \mathbb{R}$ the derivative

$$f'(c_0) = \int_0^T (p(0,t)e(t) + \lambda_1 c_0 e(t))\ dt$$

and the results, presented in Fig. 8 for the same data as in Example 2 with $d_0 = 2.4$.

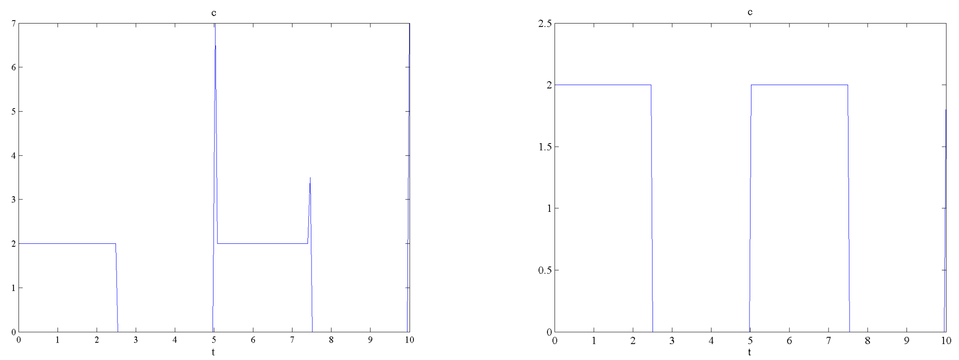

**Fig. 8** Optimal control function $\bar{c}$ for the original problem with the data of Example 2 and for Example 3

It turned out that the optimal values of the objective functionals in Example 3 and 4 are almost equal: The value 4.69 obtained for the restricted class of controls taken in Example 3 differs only by a relative error of 0.04 from the optimal value in Example 2. However, the computation needed only half the time (6 seconds instead of 12 seconds), since the degrees of freedom of the control is much smaller in this case. We selected in both examples 20 mesh points in space and 100 mesh points in time. This shows that, in our concrete application, it is justified to work with controls that are constant in each period of time.

**Acknowledgement.** We are grateful to R. Herzog (Chemnitz) for his helpful comments to a preparatory version of our work. Moreover we thank the unknown referees for their valuable hints which led to an essential improvement of the paper.

## References

[1] R. Aris, The mathematical theory of diffusion and reaction in permeable catalysts (Clarendon Press, Oxford, 1975).

[2] E. Casas, Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations, SIAM J. Control and Optimization **35**, 1297–1327 (1997).

[3] Z. Chen and K. H. Hoffmann, Numerical solutions of the optimal control problem governed by a phase field model, in: Estimation and Control of Distributed Parameter Systems, edited by W. Desch, F. Kappel, and K. Kunisch, ISNM Vol. 100 (Birkhäuser Verlag, Basel, 1991), pp. 79–97.

[4] R. Griesse, Parametric Sensitivity Analysis for Control-Constrained Optimal Control Problems Governed by Systems of Parabolic Partial Differential Equations, PhD thesis, University of Bayreuth, 2003.

[5] R. Griesse and S. Volkwein, A primal-dual active set strategy for optimal boundary control of a nonlinear reaction-diffusion system, SIAM Journal on Control and Optimization **44**(2), 467–494 (2005).

[6] M. Heinkenschloss, The Numerical Solution of a Control Problem Governed by a Phase Field Model, Optimization Methods and Software **7**, 211–263 (1997).

[7] M. Heinkenschloss and E. W. Sachs, Numerical solution of a constrained control problem for a phase field model, in: Control and Estimation of Distributed Parameter Systems, edited by W. Desch, F. Kappel, and K. Kunisch, ISNM Vol. 118 (Birkhäuser Verlag, Basel, 1994), pp. 171–188.

[8] M. A. Liauw, M. Baerns, R. Broucek, O. Buyevskaya, J. M. Commenge, J. P. Corriou, K. Gebauer, H. J. Hefter, O. U. Langer, M. Matlosz, A. Renken, A. Rouge, R. Schenk, N. Steinfeldt, and S. Walter, Periodic operation in microchannel reactors, in: Microreaction Technology: 3rd International Conference on Microreaction Technology, Proceedings of IMRET-3, edited by W. Ehrfeld (Springer, Berlin, 2000), pp. 224–234.

[9] I. Neitzel, U. Prüfert, and T. Slawig, Solving time-dependent optimal control problems in Comsol Multiphysics by space-time discretizations, Tech. Rep. 04-2009, Technische Universität Berlin, 2009.

[10] I. Neitzel, U. Prüfert, and T. Slawig, Strategies for time-dependent PDE control with inequality constraints using an integrated modeling and simulation environment, Numerical Algorithms **50**(3), 241–269 (2009).

[11] C. Pao, Nonlinear parabolic and elliptic equations (Plenum Press, New York, 1992).

[12] J. P. Raymond and H. Zidani, Hamiltonian Pontryagin's Principles for control Problems governed by Semilinear Parabolic Equations, Applied Mathematics and Optimization **39**, 143–177 (1999).

[13] E. Sachs, A parabolic control problem with a boundary condition of the Stefan–Boltzman type, ZAMM **58**, 443–449 (1978).

[14] F. Tröltzsch, Optimal Control of Partial Differential Equations. Theory, Methods and Applications. Graduate Studies in Mathematics, Volume 112 (American Mathematical Society, Providence, 2010).