ANALYSIS FOR OPTIMAL BOUNDARY CONTROL FOR A THREE-DIMENSIONAL REACTION-DIFFUSION SYSTEM

R. GRIESSE AND S. VOLKWEIN

Abstract. This paper is concerned with optimal boundary control of an instationary reaction-diffusion system in three spatial dimensions. This problem involves a coupled nonlinear system of parabolic differential equations with bilateral as well as integral control constraints. The integral constraint is included in the cost by a penalty term whereas the bilateral control constraints are handled explicitly. First- and second-order conditions for the optimization problem, which involves bilateral and integral control constraints, are analyzed. A primal-dual active set strategy is utilized to compute optimal solutions numerically. The algorithm is compared to a semi-smooth Newton method.

1. Introduction

The subject matter of the present paper is an optimal control problem for a coupled system of semi-linear parabolic reaction-diffusion equations. The equations model a chemical or biological process where the species involved are subject to diffusion and reaction among each other. As an example, we consider the reaction $A + B \to C$ which obeys the law of mass action. To simplify the discussion, we assume that the backward reaction $C \to A + B$ is negligible and that the forward reaction proceeds with a constant (e.g., not temperature-dependent) rate. This leads to a coupled semilinear parabolic system for the respective concentrations; see (2.3) later on. We consider the state equation in three spatial dimensions. Simplifications to the two- or even one-dimensional situation are of course possible in a straightforward way.

The control function acts as the Neumann boundary values for one of the reaction components on some subset of the two-dimensional boundary manifold. It is natural to impose bilateral pointwise bounds on the control function: On the one hand, the substance can never be extracted through the boundary, i.e., the lower control bound should be nonnegative. On the other hand, only a limited amount may be added *at any given time*. In addition, we impose a constraint on the *total amount* of control action. This scalar integral constraint (see (2.11)) is very much in contrast with the usual pointwise bounds.

The integral constraint is included into the cost functional by a penalty term, whereas the bilateral control constraints are treated explicitly by a primal-dual active set strategy for nonlinear problems. The primal-dual active set method has

Date: December 2, 2003.

1991 *Mathematics Subject Classification.* 35Kxx, 49Lxx, 65Kxx.

Key words and phrases. Reaction-diffusion equations, optimal boundary control, bilateral and integral control constraints, nonlinear primal-dual active set strategy, semi-smooth Newton methods.

proved to be an efficient numerical tool in the context of diverse applications; see, for instance, [BIK99, BHHK00, Hin01]. So far it was mainly investigated for linear-quadratic problems, which arise, for example, as a subproblem within SQP- or Newton methods (compare, e.g., [Hin03, HH02, KR99, TV01]). Ito and Kunisch studied the primal-dual active set algorithm for nonlinear problems and bilateral control constraints in [IK02]. Utilizing the close relationship between the primal-dual active set strategy and semi-smooth Newton methods, local superlinear convergence was shown as well. Let us mention that the primal-dual active set strategy for nonlinear problems was already applied numerically combined with Newton-, SQP- and nonlinear conjugate gradient in [Gri03], [Rey02], [Vol03], respectively. Semi-smooth Newton methods for general purpose nonlinear finite-dimensional optimal control problems are well studied, see, for instance, [LPR96] and [GK02, Section 7.5]. Much less is known about such methods in infinite dimensions, and specifically in the context of optimal control problems. We refer here , e.g., to [HIK03, HS03, Ulb03]. We will compare the nonlinear primal-dual active set strategy with a semi-smooth Newton method in the numerical test examples.

The article is organized in the following manner: In Section 2, the state equations are analyzed and the optimal control problem is investigated. The integral control constraint is treated using a penalization approach. Section 3 is devoted to the optimality conditions for the penalized optimization problem. The primal-dual active set algorithm and its relationship to a semi-smooth Newton method is discussed in Section 4. Numerical examples are presented in the fifth section and we draw some conclusions in the last section.

2. The problem formulation

The goal of this section is to introduce the infinite dimensional optimal control problem. The cost functional is of tracking type, the equality constraints are given by a coupled nonlinear parabolic system and the inequality constraints are bilateral control constraints as well as an integral constraint for the control. We study the state equations, propose the optimal control problem, and prove existence of optimal controls.

2.1. Preliminaries. Let Ω denote an open and bounded subset of $\mathbb{R}^3$ with Lipschitz-continuous boundary $\Gamma = \partial\Omega$ such that Γ is decomposed into two parts $\Gamma = \Gamma_n \cup \Gamma_c$ with $\Gamma_n \cap \Gamma_c = \emptyset$. For terminal time $T > 0$ let $Q = (0,T) \times \Omega$, $\Sigma = (0,T) \times \Gamma$ and $\Sigma_c = (0,T) \times \Gamma_c$.

By $L^2(0,T;H^1(\Omega))$ we denote the space of all measurable functions $\varphi : [0,T] \to H^1(\Omega)$, which are square integrable, i.e.,

$$\int_0^T \|\varphi(t)\|^2_{H^1(\Omega)}\,\mathrm{d}t < \infty,$$

where $\varphi(t)$ stands for the function $\varphi(t,\cdot)$ considered as a function in Ω only. The space $W(0,T)$ is defined by

$$W(0,T) = \big\{\varphi \in L^2(0,T;H^1(\Omega)) : \varphi_t \in L^2(0,T;H^1(\Omega)')\big\}. \tag{2.1}$$

Here $H^1(\Omega)'$ denotes the dual space of $H^1(\Omega)$. Recall that $W(0,T)$ is a Hilbert space endowed with the common inner product and the induced norm; see, e.g., [DL92, pp. 472-479]. Since $W(0,T)$ 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 $C_W > 0$ satisfying

$$\|\varphi\|_{C([0,T];L^2(\Omega))} \le C_W \|\varphi\|_{W(0,T)} \quad \text{for all } \varphi \in W(0,T); \tag{2.2}$$

see [DL92, p. 473].

Since we will often use the Gagliardo-Nierenberg, Gronwall and Young inequalities, we give complete formulation of them here.

Gagliardo-Nierenberg's or interpolation inequality (see, e.g., [Tan96, p. 81]): For $\Omega \subset \mathbb{R}^3$ there exists a constant $C_{GN} > 0$ such that

$$\|\varphi\|_{L^p(\Omega)} \le C_{GN} \|\varphi\|^{\theta}_{H^1(\Omega)} \|\varphi\|^{1-\theta}_{L^2(\Omega)} \quad \text{for all } \varphi \in H^1(\Omega),$$

where $p = 6/(3-2\theta) \in [2,6]$ and $\theta \in [0,1]$.

Gronwall's inequality (see, e.g., [Wal86, p. 219]): Let c be a positive constant. Suppose that $\varphi \in L^1(0,T)$ is non-negative in $[0,T]$ a.e. If $\psi \in C([0,T])$ satisfies

$$\psi(t) \le c + \int_0^t \varphi(s)\psi(s)\,\mathrm{d}s \quad \text{for all } t \in [0,T],$$

then we have

$$\psi(t) \le c \exp\Big(\int_0^t \varphi(s)\,\mathrm{d}s\Big) \quad \text{for all } t \in [0,T].$$

Young's inequality (see, e.g., [Alt92, p. 28]): For all $a, b, \varepsilon > 0$ and for all $p \in (1,\infty)$ we have

$$ab \le \frac{\varepsilon a^p}{p} + \frac{b^q}{q\varepsilon^{q/p}} \quad \text{for } q = \frac{p}{p-1}.$$

2.2. The state equations. Suppose that d_1, d_2, d_3 and k_1, k_2, k_3 are positive constants. Moreover, let $\alpha \in L^\infty(0,T;L^2(\Gamma_c))$ denote a shape function with $\alpha \ge 0$ on Σ_c almost everywhere (a.e.). We consider the following system of semi-linear parabolic equations, where c_i denotes the concentration of the ith substance:

$$(c_1)_t(t,x) = d_1 \Delta c_1(t,x) - k_1 c_1(t,x) c_2(t,x) \quad \text{for all } (t,x) \in Q, \tag{2.3a}$$

$$(c_2)_t(t,x) = d_2 \Delta c_2(t,x) - k_2 c_1(t,x) c_2(t,x) \quad \text{for all } (t,x) \in Q, \tag{2.3b}$$

$$(c_3)_t(t,x) = d_3 \Delta c_3(t,x) + k_3 c_1(t,x) c_2(t,x) \quad \text{for all } (t,x) \in Q \tag{2.3c}$$

together with the Neumann boundary conditions

$$d_1 \frac{\partial c_1}{\partial n}(t,x) = 0 \quad \text{for all } (t,x) \in \Sigma, \tag{2.3d}$$

$$d_2 \frac{\partial c_2}{\partial n}(t,x) = u(t)\alpha(t,x) \quad \text{for all } (t,x) \in \Sigma_c, \tag{2.3e}$$

$$d_2 \frac{\partial c_2}{\partial n}(t,x) = 0 \quad \text{for all } (t,x) \in \Sigma_n = \Sigma \setminus \Sigma_c, \tag{2.3f}$$

$$d_3 \frac{\partial c_3}{\partial n}(t,x) = 0 \quad \text{for all } (t,x) \in \Sigma \tag{2.3g}$$

and the initial conditions

$$c_1(0,x) = c_{10}(x) \quad \text{for all } x \in \Omega, \tag{2.3h}$$

$$c_2(0,x) = c_{20}(x) \quad \text{for all } x \in \Omega, \tag{2.3i}$$

$$c_3(0,x) = c_{30}(x) \quad \text{for all } x \in \Omega, \tag{2.3j}$$

where $c_{i0} \in L^2(\Omega)$ for $i = 1,2,3$.

The control $u \in L^2(0,T)$ enters the right-hand side of (2.3e) in the inhomogeneous Neumann condition. For instance, the function α models a spray nozzle moving over the control part Γ_c, and $u(t)$ denotes the intensity of the spray.

Remark 2.1. The parabolic problem for c_3, i.e., (2.3c) together with the Neumann boundary condition (2.3g) and initial condition (2.3j) can be solved independently of the problem for (c_1, c_2). Therefore, we will focus on the computation of c_1 and c_2 and, in particular, we are interested in weak solutions for c_1 and c_2. ♢

Definition 2.2. *The two functions c_1 and c_2 in $W(0,T)$ are called* weak solutions *to the system* (2.3a)*,* (2.3b)*,* (2.3d)*–*(2.3f)*,* (2.3h) *and* (2.3i) *provided the initial conditions*

$$c_1(0) = c_{10} \quad \text{and} \quad c_2(0) = c_{20} \quad \text{in } L^2(\Omega) \tag{2.4a}$$

hold and

$$\langle (c_1)_t(t), \varphi \rangle_{H^1(\Omega)', H^1(\Omega)} + \int_\Omega d_1 \nabla c_1(t) \cdot \nabla \varphi + k_1 c_1(t) c_2(t) \varphi \,\mathrm{d}x = 0, \tag{2.4b}$$

$$\begin{aligned} \langle (c_2)_t(t), \varphi \rangle_{H^1(\Omega)', H^1(\Omega)} + \int_\Omega d_2 \nabla c_2(t) \cdot \nabla \varphi + k_2 c_1(t) c_2(t) \varphi \,\mathrm{d}x \\ = u(t) \int_{\Gamma_c} \alpha(t) \varphi \,\mathrm{d}x \end{aligned} \tag{2.4c}$$

for all $\varphi \in H^1(\Omega)$ and almost all $t \in [0,T]$. In (2.4b) *and* (2.4c)*, $\langle \cdot\,, \cdot \rangle_{H^1(\Omega)', H^1(\Omega)}$ denotes the duality pairing between $H^1(\Omega)$ and its dual $H^1(\Omega)'$.*

The following theorem ensures that (2.4) possesses a unique solution. The proof is given in the Appendix A.1.

Theorem 2.3. *For every control $u \in L^2(0,T)$, there exists a unique pair $(c_1, c_2) \in W(0,T) \times W(0,T)$ satisfying* (2.4)*. Moreover, the estimate*

$$\|c_1\|_{W(0,T)} + \|c_2\|_{W(0,T)} \leq C \left(1 + \|c_{10}\|_{L^2(\Omega)} + \|c_{20}\|_{L^2(\Omega)} + \|u\|_{L^2(0,T)}\right) \tag{2.5}$$

holds for a constant $C > 0$ depending on d_1, d_2, k_1, k_2, C_{GN}, $\|\alpha\|_{L^\infty(0,T;L^2(\Gamma_c))}$, $\|c_i\|_{C([0,T];L^2(\Omega)))}$, and $\|c_i\|_{L^2(0,T;H^1(\Omega))}$, $i = 1, 2$.

Theorem 2.3 also implies the unique solvability of the partial differential equation for the reaction product (2.3c), (2.3g), and (2.3j). This is formulated in the following corollary, which is proved in Appendix A.2.

Corollary 2.4. *Let $c_{10}, c_{20} \in L^2(\Omega)$ and $u \in L^2(0,T)$ be given and $(c_1, c_2) \in W(0,T) \times W(0,T)$ denote the solution pair to* (2.4)*. Then there exists a unique $c_3 \in W(0,T)$ satisfying*

$$c_3(0) = c_{30} \quad \text{in } L^2(\Omega) \tag{2.6a}$$

and

$$\langle (c_3)_t(t), \varphi \rangle_{H^1(\Omega)', H^1(\Omega)} + \int_\Omega d_3 \nabla c_3(t) \cdot \nabla \varphi \,\mathrm{d}x = \int_\Omega k_3 c_1(t) c_2(t) \varphi \,\mathrm{d}x \tag{2.6b}$$

for all $\varphi \in H^1(\Omega)$ and almost all $t \in [0,T]$.

In the next result, which is proved in Appendix A.3, we present sufficient conditions so that the L^2-norm of the sum of concentrations $c_1 + c_2 + c_3$ does not increase with time.

Proposition 2.5. *Suppose that* $c_i \in W(0,T)$, $i = 1,2,3$, *are the solutions to* (2.4) *and* (2.6)*. If* $k_3 = k_1 + k_2$ *and* $d_1 = d_2 = d_3$ *hold, we obtain*

$$\|(c_1 + c_2 + c_3)(t)\|^2_{L^2(\Omega)} \le \|c_{10} + c_{20} + c_{30}\|^2_{L^2(\Omega)} \quad \text{for almost all } t \in [0,T].$$

To write the state equations as a non-linear operator equation, we introduce the two Hilbert product spaces

$$\begin{aligned} X &= W(0,T) \times W(0,T) \times L^2(0,T), \\ Y &= L^2(0,T;H^1(\Omega)) \times L^2(0,T;H^1(\Omega)) \times L^2(\Omega) \times L^2(\Omega) \end{aligned}$$

endowed with their product topology and identify

$$Y' \equiv L^2(0,T;H^1(\Omega)') \times L^2(0,T;H^1(\Omega)') \times L^2(\Omega) \times L^2(\Omega).$$

Then we introduce the mapping $e : X \to Y'$ by

$$e(x) = \begin{pmatrix} e_1(x) \\ e_2(x) \\ c_1(0) - c_{10} \\ c_2(0) - c_{20} \end{pmatrix} \quad \text{for } x = (c_1, c_2, u) \in X,$$

where

$$\begin{aligned} &\langle e_1(x), \varphi \rangle_{L^2(0,T;H^1(\Omega)'),L^2(0,T;H^1(\Omega))} \\ &\quad = \int_0^T \Big(\langle (c_1)_t(t), \varphi(t) \rangle_{H^1(\Omega)',H^1(\Omega)} + \int_\Omega d_1 \nabla c_1 \cdot \nabla \varphi + k_1 c_1 c_2 \varphi \,\mathrm{d}x \Big) \mathrm{d}t \end{aligned}$$

and

$$\begin{aligned} \langle e_2(x), \varphi \rangle_{L^2(0,T;H^1(\Omega)'),L^2(0,T;H^1(\Omega))} &= \int_0^T \Big(\langle (c_2)_t(t), \varphi(t) \rangle_{H^1(\Omega)',H^1(\Omega)} \,\mathrm{d}t \\ &\quad + \int_\Omega d_2 \nabla c_2 \cdot \nabla \varphi + k_2 c_1 c_2 \varphi \,\mathrm{d}x - u \int_{\Gamma_c} \alpha \varphi \,\mathrm{d}x \Big) \mathrm{d}t \end{aligned}$$

for $\varphi \in L^2(0,T;H^1(\Omega))$. Now, (2.4) is equivalent with the operator equation $e(x) = 0$ in Y' for $x = (c_1, c_2, u) \in X$.

2.3. The optimal control problem. Our goal is to drive the reaction-diffusion system from the given initial state near a desired terminal state. Hence, we introduce the cost functional

$$J(c_1, c_2, u) = \frac{1}{2} \int_\Omega \beta_1 \,|c_1(T) - c_{1T}|^2 + \beta_2 \,|c_2(T) - c_{2T}|^2 \,\mathrm{d}x + \frac{\gamma}{2} \int_0^T |u - u_d|^2 \,\mathrm{d}t,$$

where $\beta_1, \beta_2 \ge 0$, $\beta_1 + \beta_2, \gamma > 0$, $c_{1T}, c_{2T} \in L^2(\Omega)$ are given desired terminal states and $u_d \in L^2(0,T)$ denotes some nominal (or expected) control.

The closed and bounded convex set of admissible control parameters involves an integral constraint as well as bilateral control constraints:

$$\bar{U}_{\mathsf{ad}} = \Big\{ u \in L^2(0,T) : \int_0^T u(t) \,\mathrm{d}t \le u_c \text{ and } u_a \le u \le u_b \text{ in } [0,T] \Big\} \subset L^\infty(0,T),$$

where u_a and u_b are given functions in $L^\infty(0,T)$ satisfying $u_a \le u_b$ in $[0,T]$ almost everywhere (a.e.), and u_c is a positive constant.

Furthermore, let us define the closed convex set

$$\bar{K}_{\mathsf{ad}} = W(0,T) \times W(0,T) \times \bar{U}_{\mathsf{ad}}.$$

The infinite dimensional optimal control problem can be expressed as

$$(\mathbf{P}) \qquad \min J(x) \quad \text{s.t.} \quad x \in \bar{K}_{\mathsf{ad}} \text{ and } e(x) = 0.$$

The following theorem guarantees that $(\mathbf{P})$ has a solution.

Theorem 2.6. *Problem* $(\mathbf{P})$ *possesses at least one optimal control.*

Proof. The claim follows by standard arguments: Let $\{x^n\}_{n=1}^\infty$, $x^n = (c_1^n, c_2^n, u^n)$, be a minimizing sequence in $\bar{K}_{\mathsf{ad}}$ for the non-negative cost J. Since J is radially unbounded, it follows from Theorem 2.3 that this sequence is bounded in X. Therefore, there exists an element $x^* = (c_1^*, c_2^*, u^*) \in X$ such that

$$c_1^n \rightharpoonup c_1^* \quad \text{in } W(0,T) \text{ as } n \to \infty, \tag{2.7}$$

$$c_2^n \rightharpoonup c_2^* \quad \text{in } W(0,T) \text{ as } n \to \infty, \tag{2.8}$$

$$u^n \rightharpoonup u^* \quad \text{in } L^2(0,T) \text{ as } n \to \infty. \tag{2.9}$$

By assumption, $\alpha \in L^\infty(0,T;L^2(\Gamma_c))$ holds. Recall that there exists a constant $K_1 > 0$ such that

$$\|\psi\|_{L^2(\Gamma_c)} \le K_1 \|\psi\|_{H^1(\Omega)} \quad \text{for all } \psi \in H^1(\Omega);$$

see, e.g., [Eva98, p. 258]. Thus, for $\varphi \in L^2(0,T;H^1(\Omega))$, the mapping $t \mapsto \int_{\Gamma_c} \alpha(t)\varphi(t)\,\mathrm{d}x$ belongs to $L^2(0,T)$ and

$$\lim_{n\to\infty} \int_0^T \big(u^n(t) - u^*(t)\big)\Big(\int_{\Gamma_c} \alpha(t)\varphi(t)\,\mathrm{d}x\Big)\mathrm{d}t = 0 \quad \text{for all } \varphi \in L^2(0,T;H^1(\Omega)).$$

From (2.7) and (2.8) we find for $i = 1,2$

$$\lim_{n\to\infty} \int_0^T \big\langle (c_i^n - c_i^*)_t(t), \varphi(t) \big\rangle_{H^1(\Omega)', H^1(\Omega)}\,\mathrm{d}t = 0 \quad \text{for all } \varphi \in L^2(0,T;H^1(\Omega))$$

and

$$\lim_{n\to\infty} \int_0^T \int_\Omega d_i \nabla\big(c_i^n - c_i^*\big)(t) \cdot \nabla\varphi(t)\,\mathrm{d}x\mathrm{d}t = 0 \quad \text{for all } \varphi \in L^2(0,T;H^1(\Omega)).$$

Next we consider the non-linear terms. Using Hölder's inequality we infer that for $\varphi \in L^2(0,T;H^1(\Omega))$,

$$\begin{aligned}
\int_0^T \int_\Omega \big(c_1^n c_2^n - c_1^* c_2^*\big)\varphi\,\mathrm{d}x\mathrm{d}t &= \int_0^T \int_\Omega (c_1^n - c_1^*)c_2^n + c_1^*\big(c_2^n - c_2^*\big)\varphi\,\mathrm{d}x\mathrm{d}t \\
&\le \int_0^T \|c_1^n(t) - c_1^*(t)\|_{L^3(\Omega)} \|c_2^n(t)\|_{L^2(\Omega)} \|\varphi(t)\|_{L^6(\Omega)}\,\mathrm{d}t \\
&\quad + \int_0^T \|c_1^*(t)\|_{L^2(\Omega)} \|c_2^n(t) - c_2^*(t)\|_{L^3(\Omega)} \|\varphi(t)\|_{L^6(\Omega)}\,\mathrm{d}t \\
&\le \|c_1^n - c_1^*\|_{L^2(0,T;L^3(\Omega))} \|c_2^n\|_{C([0,T];L^2(\Omega))} \|\varphi\|_{L^2(0,T;L^6(\Omega))} \\
&\quad + \|c_1^*\|_{C([0,T];L^2(\Omega))} \|c_2^n - c_2^*\|_{L^2(0,T;L^3(\Omega))} \|\varphi\|_{L^2(0,T;L^6(\Omega))}.
\end{aligned} \tag{2.10}$$

Since $W(0,T)$ is continuously embedded into $C([0,T];L^2(\Omega))$ and compactly into $L^2(0,T;L^3(\Omega))$ (see, for instance, [Tem79, p. 271]), the sequence $\|c_i^n\|_{C([0,T];L^2(\Omega))}$ is bounded and $\lim_{n\to\infty} \|c_i^n - c_i^*\|_{L^2(0,T;L^3(\Omega))} = 0$ for $i = 1,2$. Thus, (2.10) yields

$$\lim_{n\to\infty} \int_0^T \int_\Omega \big(c_1^n c_2^n - c_1^* c_2^*\big)\varphi\,\mathrm{d}x\mathrm{d}t = 0 \quad \text{for all } \varphi \in L^2(0,T;H^1(\Omega)).$$

Using

$$\int_\Omega \big(c_i^n(0) - c_i^*(0)\big)\psi \,\mathrm{d}x = 0 \quad \text{for all } \psi \in L^2(\Omega) \text{ and } i = 1, 2$$

we have $e(x^*) = 0$ in Y'. Since $\bar{U}_{\mathsf{ad}}$ is bounded, closed and convex, $\bar{U}_{\mathsf{ad}}$ is weakly closed. This implies that $\bar{K}_{\mathsf{ad}}$ is also weakly closed. As J is weakly lower semi-continuous, the claim follows. □

2.4. The penalized optimization problem

In (**P**), we have two different types of inequality constraints for the control variable: a scalar integral constraint and an infinite-dimensional box-constraint. To handle the integral constraint

$$\int_0^T u \,\mathrm{d}t \le u_c \tag{2.11}$$

numerically, we introduce the penalized cost functional

$$J_\varepsilon(x) = J(x) + \frac{1}{\varepsilon}\, I(u) \quad \text{for all } x = (c_1, c_2, u) \in X \text{ and } \varepsilon > 0,$$

where the mapping $I : L^2(0,T) \to \mathbb{R}$ is defined as

$$I(u) = g\Big(\int_0^T u \,\mathrm{d}t - u_c\Big),$$

where we choose $g = [\,\cdot\,]_+^3$ in $\mathbb{R}$ and $[s]_+ = \max\{0, s\}$, $s \in \mathbb{R}$, denotes the positive part function.

The goal of this section is to analyze the optimal control problem with the penalized cost. The bilateral control constraints are treated explicitly and realized numerically by nonlinear primal-dual active set strategies, see Sections 4 and 5.

Lemma 2.7. *The function g is twice differentiable and its second derivative is Lipschitz-continuous.*

Lemma 2.7 is proved in Appendix A.4. For a proof of the following lemma we refer the reader to Appendix A.5.

Lemma 2.8. *The mapping $I : L^2(0,T) \to \mathbb{R}$ is weakly continuous. Moreover, I is twice continuously Fréchet-differentiable and its second Fréchet-derivative is Lipschitz-continuous in $L^2(0,T)$.*

As the integral constraint is already included in the cost J_ε by a penalty term, we replace $\bar{U}_{\mathsf{ad}}$ by

$$U_{\mathsf{ad}} = \Big\{u \in L^2(0,T) : u_a \le u \le u_b \text{ in } [0,T]\Big\} \subset L^\infty(0,T)$$

and set $K_{\mathsf{ad}} = W(0,T) \times W(0,T) \times U_{\mathsf{ad}}$. Now the penalized optimal control problem has the form

$$(\mathbf{P}_\varepsilon) \qquad \min J_\varepsilon(x) \quad \text{s.t.} \quad x \in K_{\mathsf{ad}} \text{ and } e(x) = 0.$$

Utilizing Lemma 2.8, the next result can be proved analogously to Theorem 2.6.

Theorem 2.9. *There exists at least one optimal solution to* $(\mathbf{P}_\varepsilon)$.

Proof. Due to Lemma 2.8 the mapping I is weakly continuous, so that the penalized cost functional J_ε is weakly lower semi-continuous on X. Thus, the proof is analogous to that of Theorem 2.6. □

In the next proposition we turn to the question whether solutions of $(\mathbf{P}_\varepsilon)$ converges to a solution to $(\mathbf{P})$ if ε tends to zero.

Proposition 2.10. *Assume that* $\{\varepsilon_n\}_{n=0}^\infty \subset \mathbb{R}$ *is a sequence converging to zero from above. Let* $\{x_n\}_{n=0}^\infty$ *denote a sequence of optimal solutions to* $(\mathbf{P}_{\varepsilon_{n'}})$. *Then there exists at least one weak accumulation point* $x^* \in X$ *for* $\{x_n\}_{n=0}^\infty$. *That is,* $x_{n'} \rightharpoonup x^*$ *in* X *as* $n' \to \infty$ *for some subsequence* $\{x_{n'}\}_{n'=0}^\infty$. *In addition, every weak accumulation point* x^* *solves* $(\mathbf{P})$.

Proof. Since $x_n = (c_{1n}, c_{2n}, u_n) \in X$ solves $(\mathbf{P}_{\varepsilon_n})$, the sequence $\{u_n\}_{n=0}^\infty$ belongs to U_{ad}. Thus, $\|u_n\|_{L^2(0,T)}$ is bounded by a constant which does not depend on n. Due to the a priori bound (2.5), the family of pairs $\{(c_{1n}, c_{2n})\}_{n=0}^\infty$ is also bounded in $W(0,T) \times W(0,T)$. Since X is reflexive, there exists a subsequence in K_{ad}, denoted by $\{x_{n'}\}_{n'=0}^\infty$, and an element $x^* = (c_1^*, c_2^*, u^*) \in X$, such that

$$x_{n'} \rightharpoonup x^* \text{ in } X \text{ as } n' \to \infty. \tag{2.12}$$

Reasoning as in the proof of Theorem 2.6, we find that $x^* \in K_{\mathsf{ad}}$ and $e(x^*) = 0$ in Y'. Since $x_{n'}$ solves $(\mathbf{P}_{\varepsilon_{n'}})$, we have

$$J_{\varepsilon_{n'}}(x_{n'}) = J(x_{n'}) + \frac{1}{\varepsilon_{n'}}\, I(u_{n'}) \le J(x) + \frac{1}{\varepsilon_{n'}}\, I(u) = J(x) \tag{2.13}$$

for all $x = (c_1, c_2, u) \in \bar{K}_{\mathsf{ad}}$. Hence

$$0 \le I(u_{n'}) \le \varepsilon_{n'}\big(J(x) - J(x_{n'})\big) \quad \text{for all } x \in \bar{K}_{\mathsf{ad}}. \tag{2.14}$$

Choosing $x \in \bar{K}_{\mathsf{ad}}$ arbitrarily and using the boundedness of $x_{n'}$ and thus of $J(x_{n'})$, we obtain $I(u^*) = 0$ from passing to the limit in (2.14), applying Lemma 2.8. Thus, the integral constraint (2.11) is satisfied for u^*, hence $x^* \in K_{\mathsf{ad}}$ holds. Finally, it follows from weak lower semicontinuity of J and from (2.13) that

$$J(x^*) \le \liminf_{n' \to \infty} J(x_{n'}) \le \liminf_{n' \to \infty} J(x_{n'}) + \frac{1}{\varepsilon_{n'}} I(u_{n'}) \le J(x) \quad \text{for all } x \in \bar{K}_{\mathsf{ad}}$$

so that x^* solves $(\mathbf{P})$. □

3. Optimality conditions

In Section 2 we have introduced the optimal control problem $(\mathbf{P}_\varepsilon)$ and proved existence of optimal controls. This section is devoted to analyze necessary and sufficient optimality conditions for $(\mathbf{P}_\varepsilon)$.

3.1. Smoothness properties for J and e. We start by investigating differentiability properties of the cost functional as well as of the mapping describing the equality constraints. A proof of the following result is given in Appendix B.1.

Proposition 3.1. *The penalized cost functional* J_ε *and the mapping* e *are twice continuously Fréchet-differentiable and their second Fréchet-derivatives are Lipschitz-continuous on* X.

The linear operator $\nabla_{(c_1,c_2)} e(x) : W(0,T) \times W(0,T) \to Y'$ has the following property:

Proposition 3.2. *For all* $x \in X$, *the linearization* $\nabla_{(c_1,c_2)} e(x)$ *is bijective. Moreover, for all* $\delta x = (\delta c_1, \delta c_2, \delta u) \in N(\nabla e(x))$ *we have*

$$\|\delta c_1\|_{W(0,T)}^2 + \|\delta c_2\|_{W(0,T)}^2 \le C_N \,\|\delta u\|_{L^2(0,T)}^2 \tag{3.1}$$

for all $C_N > 0$, where $N(\nabla e(x))$ denotes the null space of the operator $\nabla e(x)$.

For a proof we refer the reader to Appendix B.2.

Remark 3.3. Proposition 3.2 implies the standard constraint qualification condition for x^* (see [Rob76], for example), which in our case has the form

$$\begin{aligned} \begin{pmatrix} 0 \\ 0 \end{pmatrix} &\in \operatorname{int}\left\{ \begin{pmatrix} X \\ \nabla e(x^*)X \end{pmatrix} - \begin{pmatrix} K_{\mathsf{ad}} - x^* \\ Y - e(x^*)) \end{pmatrix} \right\} \\ &= \operatorname{int}\{X - (K_{\mathsf{ad}} - x^*)\} \times \operatorname{int}\{\nabla e(x^*)X\} \end{aligned} \tag{3.2}$$

where int S denotes the interior of a set S and $e'(x^*)$ is the Fréchet-derivative of the operator e at x^*. It follows from (3.2) that the set of Lagrange multipliers is non-empty and bounded [MZ79]. ♢

For every $\varepsilon > 0$, the Lagrange functional $L_\varepsilon : X \times Y \to \mathbb{R}$ associated with $(\mathbf{P}_\varepsilon)$ is given by

$$\begin{aligned} L_\varepsilon(x,p) &= J_\varepsilon(x) + \langle e(x), p\rangle_{Y',Y} \\ &= J_\varepsilon(x) + \int_0^T \langle (c_1)_t(t), \lambda_1(t)\rangle_{H^1(\Omega)',H^1(\Omega)} \mathrm{d}t \\ &\quad + \int_0^T \int_\Omega d_1 \nabla c_1 \cdot \nabla \lambda_1 + k_1 c_1 c_2 \lambda_1 \,\mathrm{d}x\mathrm{d}t \\ &\quad + \int_0^T \langle (c_2)_t(t), \lambda_2(t)\rangle_{H^1(\Omega)',H^1(\Omega)} \mathrm{d}t \\ &\quad + \int_0^T \int_\Omega d_2 \nabla c_2 \cdot \nabla \lambda_2 + k_2 c_1 c_2 \lambda_2 \,\mathrm{d}x - u \int_{\Gamma_c} \alpha \lambda_2 \,\mathrm{d}x\mathrm{d}t \\ &\quad + \int_\Omega \big(c_1(0) - c_{10}\big)\mu_1 + \big(c_2(0) - c_{20}\big)\mu_2 \,\mathrm{d}x \end{aligned}$$

for $x = (c_1, c_2, u) \in X$ and $p = (\lambda_1, \lambda_2, \mu_1, \mu_2) \in Y$. From Proposition 3.1 we conclude that L_ε twice continuously Fréchet-differentiable and its second Fréchet-derivative is Lipschitz-continuous.

3.2. First-order necessary optimality conditions. This subsection is devoted to present the first-order necessary optimality conditions for $(\mathbf{P}_\varepsilon)$. Problem $(\mathbf{P}_\varepsilon)$ is a non-convex programming problem so that different local minima might occur. Numerical methods will produce a local minimum close to their starting point. Therefore, we do not restrict our investigations to global solutions of $(\mathbf{P}_\varepsilon)$. We will assume that a fixed reference solution $x^* = (c_1^*, c_2^*, u^*) \in K_{\mathsf{ad}}$ is given satisfying first- and second-order optimality conditions. Let us define the active sets at x^* by $A^* = A^*_- \cup A^*_+$, where

$$A^*_- = \{t \in [0,T] : u^*(t) = u_a(t) \text{ a.e.}\} \quad \text{and} \quad A^*_+ = \{t \in [0,T] : u^*(t) = u_b(t) \text{ a.e.}\}.$$

The corresponding inactive set at x^* is given by $I^* = [0,T] \setminus A^*$. First-order necessary optimality conditions are presented in the next theorem.

Theorem 3.4. *Let $x^* = (c_1^*, c_2^*, u^*) \in K_{\mathsf{ad}}$ be a local solution to $(\mathbf{P}_\varepsilon)$. Then there exists a unique Lagrange multiplier $p^* = (\lambda_1^*, \lambda_2^*, \mu_1^*, \mu_2^*) \in W(0,T) \times W(0,T) \times$*

$L^2(\Omega) \times L^2(\Omega) \subsetneq Y$ such that the pair $(\lambda_1^*, \lambda_2^*)$ are weak solutions to the adjoint (or dual) equations

$$-(\lambda_1^*)_t - d_1\Delta\lambda_1^* = -k_1 c_2^*\lambda_1^* - k_2 c_2^*\lambda_2^* \quad \textit{in } Q, \tag{3.3a}$$

$$-(\lambda_2^*)_t - d_2\Delta\lambda_2^* = -k_1 c_1^*\lambda_1^* - k_2 c_1^*\lambda_2^* \quad \textit{in } Q, \tag{3.3b}$$

$$d_1\frac{\partial\lambda_1^*}{\partial n} = 0 \quad \textit{on } \Sigma, \tag{3.3c}$$

$$d_2\frac{\partial\lambda_2^*}{\partial n} = 0 \quad \textit{on } \Sigma, \tag{3.3d}$$

$$\lambda_1^*(T) = -\beta_1(c_1^*(T) - c_{1T}) \quad \textit{in } \Omega, \tag{3.3e}$$

$$\lambda_2^*(T) = -\beta_2(c_2^*(T) - c_{2T}) \quad \textit{in } \Omega. \tag{3.3f}$$

and for $i = 1, 2$ we have

$$\mu_i^* = \lambda_i^*(0) \quad \textit{in } \Omega. \tag{3.4}$$

Moreover, there is a Lagrange multiplier $\xi^* \in L^2(0,T)$ associated with the bilateral inequality constraint with

$$\xi^*|_{A_-^*} \le 0, \quad \xi^*|_{A_+^*} \ge 0 \tag{3.5}$$

and the optimality condition

$$\gamma(u^*(t) - u_d(t)) + \frac{1}{\varepsilon}\, g'\Big(\int_0^T u^*\,\mathrm{d}t - u_c\Big) - \int_{\Gamma_c} \alpha(t)\lambda_2^*(t)\,\mathrm{d}x + \xi^*(t) = 0 \tag{3.6}$$

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

Proof. Due to Remark 3.3, there exists $p^* \in Y$ such that

$$\begin{aligned} \nabla_{(c_1,c_2)} L_\varepsilon(x^*, p^*) &= \nabla_{(c_1,c_2)} J(x^*) + \nabla_{(c_1,c_2)} e(x^*)^* p^* \\ &= 0 \quad \text{in } W(0,T) \times W(0,T). \end{aligned} \tag{3.7}$$

By Proposition 3.2, $\nabla_{(c_1,c_2)} e(x^*)^*$ is injective, so p^* is unique. Next we prove that λ_1^* and λ_2^* are more regular and belong to $W(0,T)$. Condition (3.7) is equivalent to

$$\begin{aligned} 0 &= \nabla_{(c_1,c_2)} L_\varepsilon(x^*, p^*)(\delta c_1, \delta c_2) \\ &= \int_\Omega \beta_1(c_1^*(T) - c_{1T})\delta c_1(T) + \beta_2(c_2^*(T) - c_{2T})\delta c_2(T)\,\mathrm{d}x \\ &\quad + \int_0^T \langle (\delta c_1)_t(t), \lambda_1^*(t)\rangle_{H^1(\Omega)', H^1(\Omega)}\mathrm{d}t \\ &\quad + \int_0^T \int_\Omega d_1 \nabla\delta c_1 \cdot \nabla\lambda_1^* + k_1(\delta c_1 c_2^* + c_1^*\delta c_2)\lambda_1^*\,\mathrm{d}x\mathrm{d}t \\ &\quad + \int_0^T \langle (\delta c_2)_t(t), \lambda_2^*(t)\rangle_{H^1(\Omega)', H^1(\Omega)}\mathrm{d}t \\ &\quad + \int_0^T \int_\Omega d_2 \nabla\delta c_2 \cdot \nabla\lambda_2^* + k_2(\delta c_1 c_2^* + c_1^*\delta c_2)\lambda_2^*\,\mathrm{d}x\mathrm{d}t \\ &\quad + \int_\Omega \delta c_1(0)\mu_1^* + \delta c_2(0)\mu_2^*\,\mathrm{d}x \end{aligned} \tag{3.8}$$

for all $(\delta c_1, \delta c_2) \in W(0,T) \times W(0,T)$, specifically for $\delta c_i(t) = \chi(t)\varphi$, where $\chi \in C_c^\infty(0,T)$ and $\psi \in H_0^1(\Omega)$. Here, $C_c^\infty(0,T)$ denotes the space of infinitely differentiable functions on $(0,T)$ with compact support. We find

$$\begin{aligned}\int_0^T \langle (\delta c_i)_t(t), \lambda_i^*(t)\rangle_{H^1(\Omega)',H^1(\Omega)} \mathrm{d}t &= \Big\langle \int_0^T \lambda_i^*(t)\dot{\chi}(t)\,\mathrm{d}t, \varphi\Big\rangle_{L^2(\Omega)} \\ &= -\Big\langle \int_0^T (\lambda_i^*)_t(t)\chi(t)\,\mathrm{d}t, \varphi\Big\rangle_{H^1(\Omega)',H^1(\Omega)}\end{aligned}$$

for $i = 1, 2$, where $(\lambda_i^*)_t$ denotes the distributional derivative of λ_i^* with respect to t. The remaining terms in (3.8) lead to

$$\begin{aligned}&\int_0^T \int_\Omega d_1 \nabla\delta c_1 \cdot \nabla\lambda_1 + k_1(\delta c_1 c_2^* + c_1^* \delta c_2)\lambda_1^* \,\mathrm{d}x\mathrm{d}t \\ &\quad + \int_0^T \int_\Omega d_2 \nabla\delta c_2 \cdot \nabla\lambda_2^* + k_2(\delta c_1 c_2^* + c_1^* \delta c_2)\lambda_2^* \,\mathrm{d}x\mathrm{d}t \\ &= \Big\langle \int_0^T -d_1\Delta\lambda_1^* - d_2\Delta\lambda_2^* + (c_1^* + c_2^*)\big(k_1\lambda_1^* + k_2\lambda_2^*\big)\chi\,\mathrm{d}t, \varphi\Big\rangle_{H^1(\Omega)',H^1(\Omega)}.\end{aligned}$$

Inserting these expressions into (3.8) yields

$$\begin{aligned}&\Big\langle \int_0^T (\lambda_1^* + \lambda_2^*)_t\chi\,\mathrm{d}t, \varphi\Big\rangle_{H^1(\Omega)',H^1(\Omega)} \\ &\quad = \Big\langle \int_0^T d_1\Delta\lambda_1^* + d_2\Delta\lambda_2^* - (c_1^* + c_2^*)\big(k_1\lambda_1^* + k_2\lambda_2^*\big)\chi\,\mathrm{d}t, \varphi\Big\rangle_{H^1(\Omega)',H^1(\Omega)}\end{aligned}$$

for all $\chi \in C_c^\infty(0,T)$ and $\varphi \in H_0^1(\Omega)$. Since

$$d_1\Delta\lambda_1^* + d_2\Delta\lambda_2^* - (c_1^* + c_2^*)\big(k_1\lambda_1^* + k_2\lambda_2^*\big) \in L^2(0,T; H^1(\Omega))$$

holds and the set

$$\{\delta c : \delta c(t) = \chi(t)\varphi \text{ for } \chi \in C_c^\infty(0,T) \text{ and } \varphi \in H_0^1(\Omega)\}$$

is dense in $L^2(0,T;H^1(\Omega))$, we conclude that $(\lambda_i^*)_t \in L^2(0,T;H^1(\Omega)')$ and consequently $\lambda_i^* \in W(0,T)$ for $i = 1, 2$. Hence, (3.3a) and (3.3b) are proved. We notice that for all $\delta c_i \in W(0,T)$, $i = 1, 2$, we have

$$\begin{aligned}&\int_0^T \langle (\lambda_i^*)_t(t), \delta c_i(t)\rangle_{H^1(\Omega)',H^1(\Omega)}\,\mathrm{d}t + \int_0^T \langle (\delta c_i)_t(t), \lambda_i^*(t)\rangle_{H^1(\Omega)',H^1(\Omega)}\,\mathrm{d}t \\ &= \int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\langle \lambda_i^*(t), \delta c_i(t)\rangle_{L^2(\Omega)}\,\mathrm{d}t \\ &= \langle \lambda_i^*(T), \delta c_i(T)\rangle_{L^2(\Omega)} + \langle \lambda_i^*(0), \delta c_i(0)\rangle_{L^2(\Omega)}.\end{aligned} \tag{3.9}$$

Choosing appropriate test functions in $W(0,T)$, we infer from (3.3a), (3.3b), (3.8) and (3.9) that (3.3c)-(3.3f) as well as (3.4) are satisfied. Due to optimality the following optimality inequality holds

$$\nabla_u L_\varepsilon(x^*, p^*)(u - u^*) \geq 0 \quad \text{for all } u \in U_{\mathsf{ad}}.$$

Setting

$$
\begin{aligned}
-\langle \xi^*, u-u^* \rangle_{L^2(0,T)} &:= \nabla_u L_\varepsilon(x^*,p^*)(u-u^*) \\
&= \Big\langle \gamma(u^*-u_d) - \int_{\Gamma_c} \alpha\lambda_2^* \,\mathrm{d}x, u-u^* \Big\rangle_{L^2(0,T)} \\
&\quad + \frac{1}{\varepsilon}\, g'\Big(\int_0^T u^* \,\mathrm{d}t - u_c \Big) \int_0^T u - u^* \,\mathrm{d}t \\
&= \Big\langle \gamma(u^*-u_d) + \frac{1}{\varepsilon}\, g'\Big(\int_0^T u^* \,\mathrm{d}t - u_c \Big) - \int_{\Gamma_c} \alpha\lambda_2^* \,\mathrm{d}x, u-u^* \Big\rangle_{L^2(0,T)}
\end{aligned} \tag{3.10}
$$

for all $u \in U$ we obtain (3.6). For the proof of (3.5) we refer the reader to [Hin03]. □

Remark 3.5. From Remark 3.13 uniqueness of the Lagrange multiplier ξ^* will follow later on. ◊

The following corollary provides an a-priori estimate for the Lagrange multipliers λ_1^* and λ_2^* that will be used for the second-order optimality conditions; see Section 3.3, Theorem 3.15.

Corollary 3.6. *There exists a constant $C_\lambda > 0$ such that*

$$
\begin{aligned}
&\|\lambda_1^*\|_{L^2(0,T;L^4(\Omega))} + \|\lambda_2^*\|_{L^2(0,T;L^4(\Omega))} \\
&\le C_\lambda \big(\|c_1^*(T) - c_{1T}\|_{L^2(\Omega)} + \|c_2^*(T) - c_{2T}\|_{L^2(\Omega)} \big).
\end{aligned} \tag{3.11}
$$

For the proof we refer to Appendix B.3.

3.3. Second-order optimality conditions. In Section 3.2 we have investigated the first-order necessary optimality conditions for $(\mathbf{P}_\varepsilon)$. To ensure that a solution (x^*, p^*) satisfying (2.4), $x^* \in K_{\mathsf{ad}}$, (3.3) and (3.6) solves $(\mathbf{P}_\varepsilon)$, we have to guarantee second-order sufficient optimality. This is the focus of this section. To introduce the critical cone later on, we recall the tangent and normal cones.

Definition 3.7. *Let K be a convex subset of a Hilbert space Z and $z \in K$. The set*

$$T_K(z) = \{\tilde z \in Z : \text{ there exists } z(\sigma) = z + \sigma \tilde z + o(\sigma) \in K \text{ as } \sigma \searrow 0\}$$

is called the tangent cone *at the point z. Moreover, the* normal cone *N_K at the* point *z is given by*

$$N_K(z) = \{\tilde z \in Z : \langle \tilde z, \hat z - z \rangle_Z \le 0 \text{ for all } \hat z \in K\}.$$

In case of $z \notin K$ these two cones are set equal to the empty set.

For $K = K_{\mathsf{ad}}$ we have the following characterizations.

Lemma 3.8. *Let $x = (c_1, c_2, u) \in K_{\mathsf{ad}}$.*

a) $T_{K_{\mathsf{ad}}}(x) = W(0,T) \times W(0,T) \times T_{U_{\mathsf{ad}}}(u)$*, where*

$$T_{U_{\mathsf{ad}}}(u) = \{\tilde u \in L^2(0,T) : \tilde u(t) \in T_{[u_a(t),u_b(t)]}(u(t)) \text{ for } t \in [0,T] \text{ a.e.}\},$$

where for $a, b, s \in \mathbb{R}$ with $a \le b$

$$T_{[a,b]}(s) = \begin{cases} \mathbb{R}^+ = \{t \in \mathbb{R} : t \ge 0\} & \text{if } s = a, \\ \mathbb{R}^- = \{t \in \mathbb{R} : t \le 0\} & \text{if } s = b, \\ \mathbb{R} & \text{otherwise.} \end{cases}$$

b) $N_{K_{\mathsf{ad}}}(x) = \{0\} \times \{0\} \times N_{U_{\mathsf{ad}}}(u)$, *where*

$N_{U_{\mathsf{ad}}}(u) = \{\tilde{u} \in L^2(0,T) : \tilde{u}(t) \in N_{[u_a(t),u_b(t)]}(u(t))$ *for* $t \in [0,T]$ *a.e.*$\}$.

That is,

$$\tilde{u}(t) \in \begin{cases} \mathbb{R}^+ = \{t \in \mathbb{R} : t \geq 0\} & \text{if } u(t) = u_a(t), \\ \mathbb{R}^- = \{t \in \mathbb{R} : t \leq 0\} & \text{if } u(t) = u_b(t), \\ \mathbb{R} & \text{otherwise.} \end{cases}$$

c) *Moreover,*

$$\begin{aligned} &T_{U_{\mathsf{ad}}}(u^*) \cap \{\xi^*\}^\perp \\ &\quad = \{u \in L^2(0,T) : u \geq 0 \text{ on } A^*_-,\ u \leq 0 \text{ on } A^*_+ \text{ and } u = 0 \text{ on } A^*_\pm\}, \end{aligned} \tag{3.12}$$

where $\xi^* \in N_{U_{\mathsf{ad}}}$ *is the Lagrange multipliers introduced in Theorem* 3.4, $S^\perp$ *denotes the orthogonal complement of a set* S, *and* $A^*_\pm = \{t \in [0,T] : \xi^* > 0$ *or* $\xi^* < 0$ *a.e.*$\} \subset A^*$.

Proof. The characterization of the tangent and normal cones is a classical result. For a proof we refer to [Roc74]. What remains to show is (3.12). Due to (3.10) we have $\xi^* \in N_{U_{\mathsf{ad}}}(u^*)$ satisfying $\xi^* = 0$ on the set $[0,T] \setminus A^*_\pm$. Suppose that $t \in A^*_-$. We conclude $T_{[u_a(t),u_b(t)]}(u^*(t)) = \mathbb{R}^+$. Thus, $u \in T_{U_{\mathsf{ad}}}(u^*)$ implies $u \geq 0$ on A^*_- by part a). Analogously, $u \leq 0$ on A^*_+ holds. Hence,

$$\begin{aligned} T_{U_{\mathsf{ad}}}(u^*) \cap \{\xi^*\}^\perp &= \{\tilde{u} \in L^2(0,T) : \tilde{u}(t) \in T_{[u_a(t),u_b(t)]}(u^*(t)) \text{ for } t \in [0,T] \text{ a.e.}\} \\ &\cap \Big\{u \in L^2(0,T) : \int_{A^*_\pm} \xi^* u \,\mathrm{d}t = 0\Big\}. \end{aligned}$$

Since $\xi^* > 0$ and $u \geq 0$ on $A^*_- \cap A^*_\pm$ and $\xi^* < 0$ and $u \leq 0$ on $A^*_+ \cap A^*_\pm$, (3.12) holds. □

Suppose that the point $\bar{x} = (\bar{c}_1, \bar{c}_2, \bar{u}) \in X$ satisfies the first-order necessary optimality conditions. By Proposition 3.2 there exists unique Lagrange multipliers $\bar{p} = (\bar{\lambda}_1, \bar{\lambda}_2, \bar{\mu}_1, \bar{\mu}_2) \in Y$ and $\bar{\xi} \in N_{U_{\mathsf{ad}}}$ satisfying the first-order necessary optimality conditions

$$\nabla_x L_\varepsilon(\bar{x}, \bar{p}) + (0, \bar{\xi})^\mathsf{T} = 0, \quad \bar{x} \in K_{\mathsf{ad}} \quad \text{and} \quad e(\bar{x}) = 0. \tag{3.13}$$

Now we introduce the critical cone at $\bar{x}$.

Definition 3.9. *The* critical cone at $\bar{x}$ *is defined by*

$$C(\bar{x}) = \{\delta x \in T_{K_{\mathsf{ad}}}(\bar{x}) \cap \{(0,0,\bar{\xi})\}^\perp : \delta x \in N(\nabla e(\bar{x}))\}.$$

The critical cone at $\bar{x}$ is the set of directions of non-increase of the cost that are tangent to the feasible set $\{x \in K_{\mathsf{ad}} :\ e(x) = 0\}$. This is formulated in the next lemma.

Lemma 3.10. *It follows that* $\nabla J_\varepsilon(\bar{x})\delta x = 0$ *for all* $\delta x \in C(\bar{x})$.

Proof. Let $\delta x = (\delta c_1, \delta c_2, \delta u) \in C(\bar{x})$. From (3.13), $\delta x \in N(\nabla e(\bar{x}))$ and $\delta x \in \{(0,0,\bar{\xi})\}^\perp$ we infer that

$$0 = \big(\nabla_x L_\varepsilon(\bar{x}, \bar{p}) + (0,0,\bar{\xi})^\mathsf{T}\big)\delta x = \nabla J_\varepsilon(\bar{x})\delta x,$$

which gives the proof. □

Now we turn to the second-order necessary optimality conditions. Let $\delta x = (\delta c_1, \delta c_2, \delta u) \in X$. We find

$$\begin{aligned}\nabla_x^2 L_\varepsilon(\bar{x}, \bar{p})(\delta x, \delta x) &= \int_\Omega \beta_1 |\delta c_1(T)|^2 + \beta_2 |\delta c_2(T)|^2 \,\mathrm{d}x \\ &+ \frac{1}{\varepsilon} \nabla^2 I(\bar{u})(\delta u, \delta u) + \int_0^T \gamma |\delta u|^2 \,\mathrm{d}t \\ &+ \int_0^T \int_\Omega (2k_1 \bar{\lambda}_1 + 2k_2 \bar{\lambda}_2) \delta c_1 \delta c_2 \,\mathrm{d}x\mathrm{d}t.\end{aligned} \tag{3.14}$$

In Theorem 2.9 we have denoted by x^* the local solution to $(\mathbf{P}_\varepsilon)$. The associated unique Lagrange multipliers are p^* and ξ^*, see Theorem 3.4.

Definition 3.11. *The* second-order necessary optimality conditions *are defined as*

$$\nabla_x^2 L_\varepsilon(x^*, p^*)(\delta x, \delta x) \geq 0 \quad \textit{for all } \delta x \in C(x^*). \tag{3.15}$$

Now let $\bar{x} = x^*$ be a local solution to $(\mathbf{P}_\varepsilon)$.

Theorem 3.12. *The point (x^*, p^*) satisfies the second-order necessary optimality condition* (3.15).

Proof. We apply analogous arguments as in the proof of Theorem 2.7 in [BZ99]. The equality constraints can be written as

$$e(x) \in K_Y = \{0\} \subset Y,$$

where, of course, K_Y is a closed convex set. Thus, the result follows provided the following strict semi-linearized qualification condition

$$0 \in \operatorname{int} \{\nabla e(x^*)((K_{\mathsf{ad}} - x^*) \cap \{(0, 0, \xi^*)\}^\perp)\} \subset Y \tag{CQA}$$

holds. In our case we have

$$(K_{\mathsf{ad}} - x^*) \cap \{(0, 0, \xi^*)\}^\perp = W(0, T) \times W(0, T) \times \left((U_{\mathsf{ad}} - u^*) \cap \{\xi^*\}^\perp \right).$$

Let $z \in Y$ be arbitrary, close enough to zero. Then (CQA) follows if there exists an element $\delta x = (\delta c_1, \delta c_2, \delta u) \in W(0, T) \times W(0, T) \times (U_{\mathsf{ad}} - u^*) \cap \{\xi^*\}^\perp)$ satisfying

$$\nabla e(x^*) \delta x = z. \tag{3.16}$$

Due to Proposition 3.2 the operator $\nabla_{(c_1, c_2)} e(x^*)$ is bijective. Thus, there exists even a unique pair $(\delta c_1, \delta c_2) \in W(0, T) \times W(0, T)$ such that

$$\nabla_{(c_1, c_2)} e(x^*)(\delta c_1, \delta c_2) = z - \nabla_u e(x^*) \delta u$$

for arbitrary $\delta u \in L^2(0, T)$. This gives (3.16) so that the claim follows. □

Remark 3.13. As is proved in [BZ99], condition (CQA) implies uniqueness of the Lagrange multipliers p^* and ξ^*. ♢

Definition 3.14. *Suppose that $\bar{x}$ satisfies the first-order necessary optimality conditions with associated unique Lagrange multipliers $\bar{p} \in Y$ and $\bar{\xi} \in N_{U_{\mathsf{ad}}}(\bar{u})$. At $(\bar{x}, \bar{p})$, the* second-order sufficient optimality condition *holds if there exists $\kappa > 0$ such that*

$$\nabla_x^2 L_\varepsilon(\bar{x}, \bar{p})(\delta x, \delta x) \geq \kappa \|\delta x\|_X^2 \quad \textit{for all } \delta x \in C(\bar{x}).$$

Theorem 3.15. *If $\|c_1^*(T) - c_{1T}\|_{L^2(\Omega)} + \|c_2^*(T) - c_{2T}\|_{L^2(\Omega)}$ is sufficiently small, the second-order sufficient optimality condition is satisfied.*

Proof. Let $\delta x = (\delta c_1, \delta c_2, \delta u) \in C(x^*) \setminus \{0\}$. Then we have

$$\nabla^2 I(\bar u)(\delta u, \delta u) = 6 \Big[\int_0^T \bar u \,\mathrm{d}t - u_c \Big]_+ \Big(\int_0^T \delta u \,\mathrm{d}t \Big)^2 \geq 0.$$

Using Hölder's and Gagliardo-Nierenberg's inequality, (2.2), (3.1) and (3.14), we estimate

$$\begin{aligned}
&\nabla_x^2 L_\varepsilon(x^*, p^*)(\delta x, \delta x)\\
&\geq \frac{\gamma}{2} \|\delta u\|^2_{L^2(0,T)} + \frac{\gamma}{2C_N^2} \big(\|\delta c_1\|^2_{W(0,T)} + \|\delta c_2\|^2_{W(0,T)} \big)\\
&\quad - 2C_{GN}^2 C_W^2 \big(k_1 \|\lambda_1^*\|_{L^2(0,T;L^4(\Omega))} + k_2 \|\lambda_2^*\|_{L^2(0,T;L^4(\Omega))} \big) \|\delta c_1\|_{W(0,T)} \|\delta c_2\|_{W(0,T)}\\
&\geq \frac{\gamma}{2} \|\delta u\|^2_{L^2(0,T)} + \big(\|\delta c_1\|^2_{W(0,T)} + \|\delta c_2\|^2_{W(0,T)} \big)\\
&\qquad \cdot \Big(\frac{\gamma}{2C_N^2} - K_1 \big(\|\lambda_1^*\|_{L^2(0,T;L^4(\Omega))} + \|\lambda_2^*\|_{L^2(0,T;L^4(\Omega))} \big) \Big),
\end{aligned}$$

where we set $K_1 = \max(k_1, k_2) C_{GN}^2 C_W^2 > 0$. Due to (3.11) we find

$$\begin{aligned}
&\nabla_x^2 L_\varepsilon(x^*, p^*)(\delta x, \delta x)\\
&\geq \frac{\gamma}{2} \|\delta u\|^2_{L^2(0,T)} + \big(\|\delta c_1\|^2_{W(0,T)} + \|\delta c_2\|^2_{W(0,T)} \big)\\
&\qquad \cdot \Big(\frac{\gamma}{2C_N} - K_2 \big(\|c_1^*(T) - c_{1T}\|_{L^2(\Omega)} + \|c_2^*(T) - c_{2T}\|_{L^2(\Omega)} \big) \Big)
\end{aligned}$$

with the constant $K_2 = K_1 C_\lambda > 0$. Now, for instance, if

$$\|c_1^*(T) - c_{1T}\|_{L^2(\Omega)} + \|c_2^*(T) - c_{2T}\|_{L^2(\Omega)} \leq \frac{\gamma}{4K_2 C_N}$$

holds, the second-order sufficient optimality condition is satisfied for the coercivity constant $\kappa = \gamma \min(1, 1/C_N)/2$. □

Remark 3.16. 1) Notice that smallness of the two terms $\|c_i^*(T) - c_{iT}\|_{L^2(\Omega)}$, $i = 1, 2$, ensures that

$$\nabla_x^2 L_\varepsilon(x^*, p^*)(\delta x, \delta x) \geq \kappa \|\delta x\|_X^2 \quad \text{for all } \delta x \in N(\nabla e(x^*)).$$

Since $C(x^*) \subset N(\nabla e(x^*))$ holds, the second-order sufficient optimality condition is satisfied.

2) Utilizing polyhedricity of the feasible set and the theory of Legendre forms the second-order optimality condition

$$\nabla_x^2 L_\varepsilon(\bar x, \bar p)(\delta x, \delta x) > 0 \quad \text{for all } \delta x \in C(\bar x) \setminus \{0\} \tag{3.17}$$

is already sufficient [BS03, Chapter 3]. Note that (3.17) is very close the second-order sufficient condition introduced in Definition 3.14.

2) Arguing as in the proof of Theorem 4.12 in [Vol01] it follows that the second-order sufficient optimality condition is equivalent to the *quadratic growth condition*, i.e., equivalent to the existence of a constant $\varrho > 0$ such that

$$J_\varepsilon(x) \geq J_\varepsilon(\bar x) + \varrho \|x - \bar x\|_X^2 + o(\|x - \bar x\|_X) \quad \text{for all } x \in \mathcal{K}$$

with $\mathcal{K} = \{x \in X : x \in K_{\mathsf{ad}} \text{ and } e(x) = 0\}$. ◊

4. The primal-dual active set method

In this section we describe the primal-dual active set strategy for nonlinear problems and review convergence results. For more details and for the proofs we refer the reader to [IK02].

Due to Theorem 2.3 we can define the solution operator

$$\mathcal{S} : L^2(0,T) \to W(0,T) \times W(0,T)$$

by $(c_1, c_2) = \mathcal{S}(u)$ for $u \in L^2(0,T)$, where the pair $(c_1, c_2) \in W(0,T) \times W(0,T)$ is the solution to (2.4). Introducing the reduced cost functional

$$\hat{J}_\varepsilon(u) = J_\varepsilon(\mathcal{S}(u), u),$$

problem $(\mathbf{P}_\varepsilon)$ can be expressed as

$$(\hat{\mathbf{P}}_\varepsilon) \qquad \min \hat{J}_\varepsilon(u) \quad \text{s.t.} \quad u \in U_{\mathsf{ad}}.$$

Notice that $(\hat{\mathbf{P}}_\varepsilon)$ is a minimization problem with bilateral control constraints but with no equality constraints. The gradient of $\hat{J}_\varepsilon$ at a point $u \in L^2(0,T)$ is given by

$$\hat{J}_\varepsilon(u) = \gamma(u^* - u_d) + \frac{1}{\varepsilon}\, g'\Big(\int_0^T u^*\,\mathrm{d}t - u_c\Big) - \int_{\Gamma_c} \alpha\lambda_2^*\,\mathrm{d}x \in L^2(0,T), \tag{4.1}$$

where $(\lambda_1, \lambda_2) \in W(0,T) \times W(0,T)$ solves (3.3) for the state pair (c_1, c_2), which is the solution to (2.4) for the control input u.

From Theorem 3.4 we derive that the first-order necessary optimality conditions

$$\langle \hat{J}_\varepsilon(u^*), u - u^* \rangle_{L^2(0,T)} \geq 0 \quad \text{for all } u \in U_{\mathsf{ad}}$$

are equivalent to

$$\gamma(u^* - u_d) + \frac{1}{\varepsilon}\, g'\Big(\int_0^T u^*\,\mathrm{d}t - u_c\Big) - \int_{\Gamma_c} \alpha\lambda_2^*\,\mathrm{d}x + \xi^* = 0 \quad \text{in } L^2(0,T) \tag{4.2a}$$

where the Lagrange multiplier $\xi^* \in N_{U_{\mathsf{ad}}}(u^*)$ associated to the bilateral control constraints satisfies

$$\xi^* = \max\big\{0, \xi^* + (u^* - u_b)\big\} + \min\big\{0, \xi^* + (u^* - u_a)\big\} \quad \text{in } L^2(0,T). \tag{4.2b}$$

In (4.2b) the functions max and min are interpreted as pointwise a.e. operations. We next specify the primal-dual active set method.

Algorithm 4.1 (Primal-dual active set strategy)**.**

1) *Choose* $(u^0, \xi^0) \in L^2(0,T) \times L^2(0,T)$*,* $\sigma > 0$ *and set* $k = 0$*.*
2) *Determine the active sets*

$$\mathcal{A}_-^k = \Big\{t \in [0,T] : u^k(t) + \frac{\xi^k(t)}{\sigma} < u_a(t)\Big\},$$
$$\mathcal{A}_+^k = \Big\{t \in [0,T] : u^k(t) + \frac{\xi^k(t)}{\sigma} > u_b(t)\Big\}$$

and set $\mathcal{I}^k = [0,T] \setminus (\mathcal{A}_-^k \cup \mathcal{A}_+^k)$*.*

3) *If* $k \geq 1$ *and* $\mathcal{A}_+^k = \mathcal{A}_+^{k-1}$*,* $\mathcal{A}_-^k = \mathcal{A}_-^{k-1}$*, then STOP.*

4) *Solve for* $(c_1, c_2, u, \lambda_1, \lambda_2) \in W(0,T)\times W(0,T)\times L^2(I^k)\times W(0,T)\times W(0,T)$ *the coupled linear system*

$$
\left\{
\begin{array}{lll}
(c_1)_t &= d_1\Delta c_1 - k_1 c_1 c_2 & \text{in } Q,\\
(c_2)_t &= d_2\Delta c_2 - k_2 c_1 c_2 & \text{in } Q,\\
d_1\frac{\partial c_1}{\partial n} &= 0 & \text{on } \Sigma,\\
d_2\frac{\partial c_2}{\partial n} &= u_a\alpha & \text{on } A_-^k\times\Gamma_c,\\
d_2\frac{\partial c_2}{\partial n} &= u_b\alpha & \text{on } A_-^k\times\Gamma_c,\\
d_2\frac{\partial c_2}{\partial n} &= u\alpha & \text{on } I^k\times\Gamma_c,\\
c_1(0) &= c_{10} & \text{in } \Omega,\\
c_2(0) &= c_{20} & \text{in } \Omega,\\
-(\lambda_1^*)_t &= d_1\Delta\lambda_1^* - k_1 c_2^*\lambda_1^* - k_2 c_2^*\lambda_2^* & \text{in } Q,\\
-(\lambda_2^*)_t &= d_2\Delta\lambda_2^* - k_1 c_1^*\lambda_1^* - k_2 c_1^*\lambda_2^* & \text{in } Q,\\
d_1\frac{\partial\lambda_1^*}{\partial n} &= 0 & \text{on } \Sigma,\\
d_2\frac{\partial\lambda_2^*}{\partial n} &= 0 & \text{on } \Sigma,\\
\lambda_1^*(T) &= -\beta_1(c_1^*(T) - c_{1T}) & \text{in } \Omega,\\
\lambda_2^*(T) &= -\beta_2(c_2^*(T) - c_{2T}) & \text{in } \Omega,\\
\gamma(u-u_d) &= \int_{\Gamma_c}\alpha\lambda_2\,\mathrm{d}x - \frac{1}{\varepsilon}\,g'\big(\int_0^T \bar u^k\,\mathrm{d}t - u_c\big) & \text{in } I^k
\end{array}
\right.
\tag{4.3}
$$

with $\bar u^k = u_a$ *in* A_-^k, $\bar u^k = u_b$ *in* A_+^k *and* $\bar u^k = u$ *in* I^k.

5) *Set* $(c_1^{k+1}, c_2^{k+1}, \lambda_1^{k+1}, \lambda_2^{k+1}) = (c_1, c_2, \lambda_1, \lambda_2)$, $u^{k+1} = u_a$ *in* A_-^k, $u^{k+1} = u_b$ *in* A_+^k *and* $u^{k+1} = u$ *in* I^k *and*

$$
\xi^{k+1} = \int_{\Gamma_c}\alpha\lambda_2^{k+1}\,\mathrm{d}x - \frac{1}{\varepsilon}\,g'\big(\int_0^T u^{k+1}\,\mathrm{d}t - u_c\big) - \gamma(u^{k+1} - u_d) \quad \text{in } [0,T]. \tag{4.4}
$$

Set $k = k+1$ *and go back to step* 2).

Remark 4.2.

1) Notice that (4.3) are the first-order necessary optimality conditions for

$$
\min \hat J_\varepsilon(u) \quad \text{s.t.} \quad u = u_a \text{ in } A_-^k \text{ and } u = u_b \text{ in } A_+^k, \tag{$\hat{\mathbf{P}}_\varepsilon^k$}
$$

that is a minimization problem without any inequality constraints.

2) In Section 5 we solve (4.3) by applying an inexact Newton method to ($\hat{\mathbf{P}}_\varepsilon^k$). The inexact Newton step is computed by utilizing the CG method with negative curvature test (see, e.g., [NW99, Section 6.2]) to the Newton system (restricted to the inactive set I^k)

$$
\nabla^2\hat J_\varepsilon(u_i)_{|I^k}\,\delta u_{i|I^k} = -\nabla\hat J_\varepsilon(u_i)_{|I^k} \quad \text{for } i \ge 0 \tag{4.5}
$$

with zero initial guess $\delta u_i \equiv 0$ in $A_-^k \cup A_+^k$ for all i. More precisely, if $R : L^2(I^k) \to L^2(0,T)$ is the linear extension-by-zero operator from the inactive set I^k to $(0,T)$, then in fact (4.5) reads

$$
\nabla^2\hat J_\varepsilon(u_i)(R\,\delta u_{i|I^k}, R\,\cdot) = -\nabla\hat J_\varepsilon(u_i)(R\,\cdot) \quad \text{in } L^2(I^k) \text{ for } i \ge 0 \tag{4.6}
$$

To compute the right-hand side in (4.5) for given current iterate u_i we have to solve the state and adjoint equations. For each CG step, the solutions to a linearized state and an adjoint problem have to be computed.

3) The strategy for choosing the current active sets A_-^k and A_+^k is based on convex analysis techniques (see [BIK99, IK00]). Due to the simple nature of the box constraints in $(\hat{\mathbf{P}}_\varepsilon)$ this strategy is related to strategies already used in [Ber92, HI84]. $\Diamond$

In the case of $g \equiv 0$, sufficient conditions for global convergence of Algorithm 4.1 were given in [IK02]. The proof is based theoretically on descent properties of the merit function $\Phi : L^2(0,T) \times L^2(0,T) \to \mathbb{R}$ defined as

$$\Phi(u,\xi) = \gamma^2 \int_0^T \big([u-u_a]_+^2 + [u_b - u]_+^2\big)\,\mathrm{d}t + \int_{A_-(u)} [\xi]_-^2\,\mathrm{d}t + \int_{A_+(u)} [\xi]_+^2\,\mathrm{d}t,$$

where $[x]_+ = \max\{0,x\}$ and $[x]_- = -\min\{0,x\}$ denote the positive and negative part functions, respectively, and $A_-(u) = \{t \in [0,T] : u \le u_a\}$, $A_+(u) = \{t \in [0,T] : u \ge u_b\}$.

By expressing the primal-dual active set method as a partial semi-smooth Newton algorithm for (4.2), sufficient conditions for super-linear convergence were also derived in [IK02]. Since only the nonlinearity due to the max and min operations are linearized, whereas $u \mapsto \mathcal{S}(u)$ is not, the method is called a partial semi-smooth Newton algorithm. Next we specify the nonlinear equation, which is the starting point for proving superlinear convergence of Algorithm 4.1. First notice that (4.2b) is equivalent to

$$\xi = \max\big\{0, \xi + \sigma(u-u_b)\big\} + \min\big\{0, \xi + \sigma(u-u_a)\big\} \quad \text{for every } \sigma > 0. \tag{4.7}$$

Choosing $\sigma = \gamma$ (an essential prerequisite in proving superlinear convergence) in (4.7) we find

$$-\xi^* + \max\big\{0, \xi^* + \gamma(u-u_b)\big\} + \min\big\{0, \xi^* + \gamma(u-u_a)\big\} = 0. \tag{4.8}$$

Inserting the optimality condition (4.2a) into (4.8) we derive

$$\begin{aligned} 0 &= \gamma(u-u_d) + \frac{1}{\varepsilon}\, g'\Big(\int_0^T u\,\mathrm{d}t - u_c\Big) - \int_{\Gamma_c} \alpha\lambda_2\,\mathrm{d}x \\ &\quad + \max\Big\{0, \int_{\Gamma_c} \alpha\lambda_2\,\mathrm{d}x - \frac{1}{\varepsilon}\, g'\Big(\int_0^T u\,\mathrm{d}t - u_c\Big) + \gamma(u_d - u_b)\Big\} \\ &\quad + \min\Big\{0, \int_{\Gamma_c} \alpha\lambda_2\,\mathrm{d}x - \frac{1}{\varepsilon}\, g'\Big(\int_0^T u\,\mathrm{d}t - u_c\Big) - \gamma(u_a - u_d)\Big\}. \end{aligned} \tag{4.9}$$

Notice that

$$a - b + \max\{0, b-c\} = a - c + \max\{0, c-b\} \quad \text{for all } a,b,c \in \mathbb{R}.$$

Taking

$$a = \gamma(u-u_d),\ b = \int_{\Gamma_c} \alpha\lambda_2\,\mathrm{d}x - \frac{1}{\varepsilon}\, g'\Big(\int_0^T u\,\mathrm{d}t - u_c\Big),\ c = \gamma(u_b - u_d)$$

we infer from (4.9) that (4.2) is equivalent to

$$\mathcal{F}(u) = 0 \quad \text{in } L^2(0,T) \tag{4.10}$$

with the nonlinear mapping $\mathcal{F} : L^2(0,T) \to L^2(0,T)$ given by

$$\begin{aligned}\mathcal{F}(u) &= \gamma(u-u_b)\\ &\quad + \max\Big\{0, \gamma(u_b-u_d) - \int_{\Gamma_c} \alpha\lambda_2\,\mathrm{d}x + \frac{1}{\varepsilon}\, g'\Big(\int_0^T u\,\mathrm{d}t - u_c\Big)\Big\}\\ &\quad + \min\Big\{0, \int_{\Gamma_c} \alpha\lambda_2\,\mathrm{d}x - \frac{1}{\varepsilon}\, g'\Big(\int_0^T u\,\mathrm{d}t - u_c\Big) - \gamma(u_a-u_d)\Big\}.\end{aligned}$$

where $\lambda_2 = \lambda_2(u)$ depends on u via the nonlinear state and the linear adjoint equations.

Remark 4.3. Note that $\mathcal{F}$ is generalized differentiable in the sense of [HIK03]; see also [Ulb03] for a related approach. This infinite-dimensional generalized differentiability concept of the max and min functions requires a norm gap which is guaranteed by the smoothing properties of the mapping $u \mapsto \lambda_2(u)$. For each $u \in L^2(0,T)$, the argument under the max and min functions which depends on u is an element of $L^4(0,T)$, which can be seen as follows. First of all, $g'(\int_0^T u\,\mathrm{d}t - u_c)$ is a scalar. Secondly, with $\alpha \in L^\infty(0,T;L^2(\Gamma_c))$ we have

$$\int_0^T \left|\int_{\Gamma_c} \alpha\lambda_2\,\mathrm{d}x\right|^p \mathrm{d}t \le \|\alpha\|^p_{L^\infty(0,T;L^2(\Gamma_c))} \int_0^T \|\lambda_2(t)\|^p_{L^2(\Gamma_c)}\,\mathrm{d}t.$$

As the trace operator is continuous from $H^{1/2}(\Omega)$ to $L^2(\Gamma_c)$, say with operator norm C_T, and by the interpolation inequality, see, e.g., [LM72], $\|\varphi\|_{H^{1/2}(\Omega)} \le C_I\|\varphi\|^{1/2}_{H^1(\Omega)}\|\varphi\|^{1/2}_{L^2(\Omega)}$ holds, we have

$$\int_0^T \left|\int_{\Gamma_c} \alpha\lambda_2\,\mathrm{d}x\right|^p \mathrm{d}t \le (C_T C_I)^p\,\|\alpha\|^p_{L^\infty(0,T;L^2(\Gamma_c))} \int_0^T \|\lambda_2(t)\|^{p/2}_{H^1(\Omega)}\|\lambda_2(t)\|^{p/2}_{L^2(\Omega)}\,\mathrm{d}t.$$

Since λ_2 is an element of $W(0,T)$, we can estimate the second term in the integral by $\|\lambda_2\|^{p/2}_{L^\infty(0,T;L^2(\Omega))}$ and find that the right hand side remains finite for $1 \le p \le 4$. $\diamondsuit$

Next we turn to the semi-smooth Newton method, which — in contrast to Algorithm 4.1 — also linearizes the mapping $u \mapsto \mathcal{S}(u)$. Suppose that δu^k is the computed inexact Newton step solving (4.5). Then, instead of using (4.4), the Lagrange multiplier is updated by

$$\xi^{k+1} = \begin{cases} \displaystyle\int_{\Gamma_c} \alpha\lambda_2^{k+1}\,\mathrm{d}x - \frac{1}{\varepsilon}\, g'(\int_0^T u^{k+1}\,\mathrm{d}t - u_c) - \gamma(u^{k+1}-u_d) \\ \qquad \displaystyle -\frac{1}{\varepsilon}\, g''(\int_0^T u^{k+1}\,\mathrm{d}t - u_c)\cdot \int_0^T \delta u^k(t)\,\mathrm{d}t \qquad \text{in } A_-^k \cup A_+^k, \\ 0 \qquad \text{in } I^k. \end{cases} \tag{4.11}$$

Note that (4.11) is one Newton step for the optimality condition (4.2a). Commonly, condition (4.2a) is linear in u and λ (and y), so δu^k does not appear in its linearization, see, for example, [IK02]. We also mention that the choice of $\sigma = \gamma$ is only of theoretical interest. In the numerical implementation, usually σ is set to a larger value in order to prevent optimization variables to jump from the upper to the lower bound (or vice versa) in consecutive iterations, see [Sta03].

5. Numerical Examples

In this section, we describe the behavior of the primal-dual and the semi-smooth Newton methods by means of some examples of our penalized problem ($\mathbf{P}_\varepsilon$). All coding is done in MATLAB using routines from the FEMLAB 2.2 package concerning the finite element implementation. The given CPU times were obtained on a standard 1700 MHz desktop PC. They include only the run time for the core algorithm, excluding the generation of the mesh, pre-computing integrals and incomplete Cholesky decompositions. The three-dimensional geometry of the problem is given by the annular cylinder between the planes $z = 0$ and $z = 0.5$ with inner radius 0.4 and outer radius 1.0 whose rotational axis is the z-axis (Figure 5.2). The control boundary Γ_c is the upper annulus, and we use the control shape function

$$\alpha(t, x) = \exp\left(-5\left[(x - 0.7\cos(2\pi t))^2 + (y - 0.7\sin(2\pi t))^2\right]\right), \tag{5.1}$$

see Figure 5.1. Note that α corresponds to a nozzle circling for $t \in [0, 1]$ once around in counter-clockwise direction at a radius of 0.7. For fixed t, α is a function which decays exponentially with the square of the distance from the current location of the nozzle.

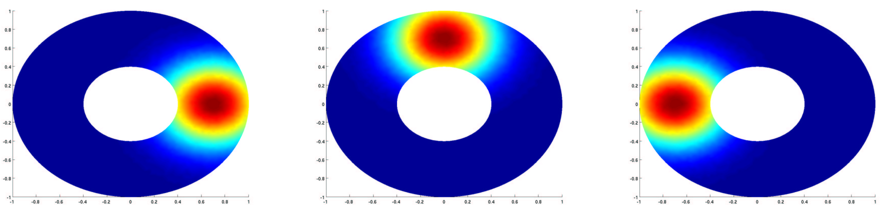

FIGURE 5.1. Control shape function $\alpha(t, x)$ at $t = 0.0$, $t = 0.25$, and $t = 0.5$.

The 'triangulation' of the domain Ω by tetrahedra is also shown in Figure 5.2. It was created using an initial mesh obtained from `meshinit(fem,'Hmax',0.4)`. As the geometry suggests that much of the reaction will take place near the top surface Γ_c of the annular cylinder, we refine this initial mesh near the top using the command `meshinit(fem,'Hexpr','0.4*(0.5-z)+0.10',···)`. The final mesh consists of 1797 points and 7519 tetrahedra. In the time direction, we use $T = 1$ and partition the interval into 100 subintervals of equal lengths. We use the semi-implicit Euler time integration scheme for the state equation (2.3a)-(2.3b), where the nonlinearities are treated as explicit terms, i.e., they are always taken from the previous time step. In the adjoint equation, we use the same semi-implicit scheme, i.e., the right hand side terms in (3.3a)-(3.3b) are taken from the previously computed time step. The elliptic problem which arises on each time level in the state and adjoint equations is solved using the conjugate gradient method with incomplete Cholesky preconditioning. Note that the preconditioner needs to be computed only once since the coefficient matrices are the same in each time step, provided the time step lengths are all identical.

The gradient of the reduced cost functional $\nabla\hat{J}_\varepsilon$ given in (4.1) is then assembled using the pre-computed expressions $\int_{\Gamma_c} \alpha(t^i, x)\,dx$. This strategy clearly follows the paradigm of optimize-then-discretize. Consequently, on the discrete level, the

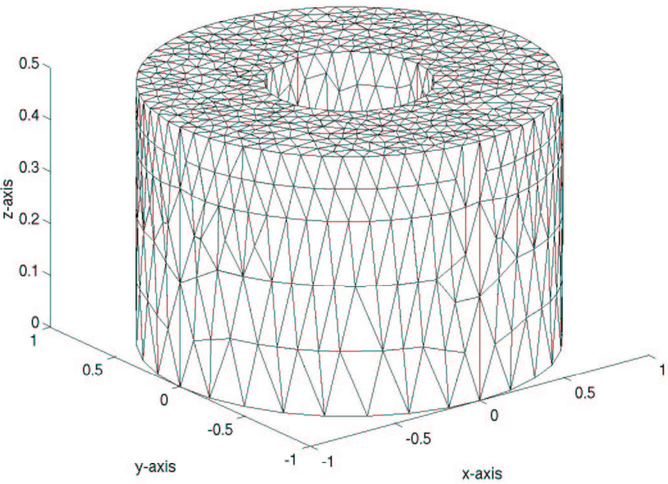

FIGURE 5.2. Domain $\Omega \subset \mathbb{R}^3$ and its triangulation with tetrahedra

gradient and also the Hessian of the reduced cost functional are evaluated only to a certain accuracy, which depends on the level of discretization. Hence, even if the reduced Hessian system (4.5) was solved to full machine precision, the discrete solution of (3.6) would in general only be found up to a residual whose size depends on the level of discretization. For the discretization parameters given above, we used $\|\nabla \hat{J}_\varepsilon(u^n)_{|I^n}\|_{L^2(0,T)} \leq 3 \cdot 10^{-3}$ as a termination criterion, which is evaluated by setting the components of $\nabla \hat{J}_\varepsilon(u^n)$ to zero which correspond to either of the active sets A_+^n or A_-^n. Of course, coincidence of the active sets on two consecutive iterations is also required for the algorithm to terminate.

5.1 Example 1. In the first example, we use the uniform control bounds

$$u_a \equiv 1 \qquad u_b \equiv 5. \tag{5.2}$$

Controlling the second substance, we wish to steer the concentration of the first substance to zero at terminal time $T = 1$, i.e., we choose

$$\beta_1 = 1 \qquad \beta_2 = 0 \qquad c_{1T} \equiv 0 \tag{5.3}$$

The control cost parameter is $\gamma = 10^{-2}$.

The chemical reaction is governed by equations (2.3a)–(2.3b) with parameters

$$d_1 = 0.15, \qquad d_2 = 0.20, \qquad k_1 = 1.0, \qquad k_2 = 1.0. \tag{5.4}$$

As initial concentrations, we use

$$c_{10} \equiv 1.0 \qquad c_{20} \equiv 0.0. \tag{5.5}$$

The discrete optimal solution without integral constraint (2.11) yield

$$\int_0^T u^*(t)\,\mathrm{d}t = 4.2401, \qquad \hat{J}_\varepsilon(u^*) = 0.2413. \tag{5.6}$$

In order for this constraint to become relevant, we choose $u_c = 3.5$ and enforce it using the penalization parameter $\varepsilon = 1$.

The numerical solution is obtained once using the primal-dual (Table 5.1) and once using the semi-smooth Newton method (Table 5.2), from an initial guess $u^0 \equiv 3$. In both cases, the reduced Hessian system (4.5) was solved using the conjugate gradient method, although due to the use of an optimize–then discretize strategy, the matrix we obtain as an approximation to the reduced Hessian $\nabla^2 \hat{J}_\varepsilon(u^n)$ is only approximately symmetric. The discrete analogue of (4.6) is

$$R^\top \nabla^2 \hat{J}_\varepsilon(u_i) R \, \delta u_{i|I^n} = -R^\top \nabla \hat{J}_\varepsilon(u_i) \quad \text{for } i \geq 0. \tag{5.7}$$

This linear system of equations is of size 100 minus the cardinality of the active sets, $|A_+^n|$ and $|A_-^n|$. Here, R is derived from the identity matrix by canceling the rows whose indices belong to either of the active sets. Our CG algorithm solves (5.7) by evaluating $\nabla^2 \hat{J}_\varepsilon(u_i)\, \delta u_i$ and then disregarding the active components. It also respects the discrete $L^2(0,T)$-inner product. As the solution of (5.7) required only a few CG steps (Tables 5.1 and 5.2), no preconditioning was used here. Our CG method also features a coercivity check: Should a direction of negative curvature be encountered, the CG iteration is terminated prematurely. Otherwise, if (5.7) was solved to sufficient accuracy, we still check the coercivity constant in the direction of δu: If $\nabla^2 \hat{J}_\varepsilon(u_i)(\delta u, \delta u)/\|\delta u\|^2$ is too small, then the direction δu is computed once again with a modification of the reduced Hessian obtained by omitting the terms which may spoil positive definiteness. These are the terms originating in the PDE's nonlinearity, corresponding to $\int_0^T \int_\Omega (2k_1 \bar{\lambda}_1 + 2k_2 \bar{\lambda}_2) \delta c_1 \delta c_2 \, \mathrm{d}x \mathrm{d}t$ in the full Hessian (3.14). In the numerical examples presented, however, negative curvature or insufficient coercivity did not occur.

The parameter σ which enters step 2) of Algorithm 4.1 (determining the active sets) was chosen as $\sigma = 10^2$. The solution is shown in Figure 5.3. Recall that

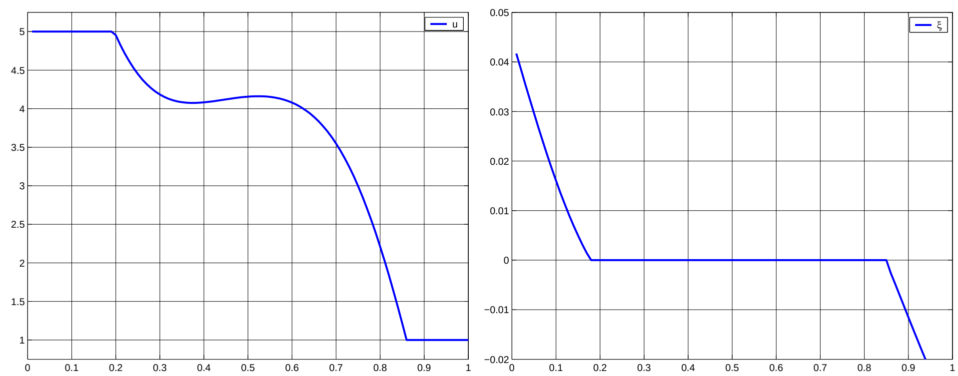

FIGURE 5.3. Example 1: Optimal control u (left) and control constraint multiplier ξ

without penalization, the optimal solution yielded an integral value of $\int_0^T u^*(t)\, \mathrm{d}t \approx 4.24$, while with penalization parameter $\varepsilon = 1$, we are down to $\int_0^T u^*(t)\, \mathrm{d}t \approx 3.58$.

The primal-dual algorithm. In order to achieve local quadratic convergence, the conjugate gradient method inside the primal-dual algorithm for (4.5) was terminated when $\|r\|_{L^2(0,T)} \leq \eta \cdot \|\nabla \hat{J}_\varepsilon(u^n)\|_{L^2(0,T)}$, where $\|r\|$ denotes the norm of the residual and $\eta = \min\{0.5, \|\nabla \hat{J}_\varepsilon(u^n)\|\}$. Finally, an Armijo backtracking

n	$\lvert A_+^n\rvert$	$\lvert A_-^n\rvert$	$\lVert\nabla\hat{J}_\varepsilon(u^n)_{\lvert I^n}\rVert$	CG	$\lVert r\rVert$	α
1	0	0	5.8320E–02	2	2.3835E–03	5.0000E–01
			3.2408E–02	3	7.2951E–04	1.0000E+00
			4.6224E–03	4	1.7267E–05	1.0000E+00
			2.3110E–04			
2	15	15	7.3888E–03	3	4.0316E–05	1.0000E+00
			1.0929E–03			
3	19	15	3.3424E–03	3	1.4936E–06	1.0000E+00
			1.9500E–04			
Run time: 517 seconds						Objective: 0.2702
						$\int_0^T u(t)\,\mathrm{d}t = 3.5803$

TABLE 5.1. Example 1 solved with the primal-dual active set method. $\lVert\nabla\hat{J}_\varepsilon(u^n)_{\lvert I^n}\rVert$ is the $L^2(0,T)$-norm of the right hand side in the reduced Hessian system.

line search was employed in the direction δu^n obtained from (4.5), with trial step lengths $\alpha_k = 2^{-k}$, $k = 0, 1, 2$, etc. A step of length α_k was accepted whenever $\hat{J}_\varepsilon(u^n + \alpha_k \delta u^n) \le \hat{J}_\varepsilon(u^n) + 10^{-4} \cdot \frac{d}{d\alpha}\hat{J}_\varepsilon(u^n + \alpha\delta u^n)|_{\alpha=0}$.

n	$\lvert A_+^n\rvert$	$\lvert A_-^n\rvert$	$\lVert\nabla\hat{J}_\varepsilon(u^n)_{\lvert I^n}\rVert$	CG	$\lVert r\rVert$	α
1	0	0	5.8320E–02	4	6.8630E–05	5.0000E–01
2	5	0	1.9816E–02	5	7.2135E–05	1.0000E–00
3	17	14	2.5667E–02	2	2.0780E–04	5.0000E–01
4	14	14	1.3587E–02	3	4.1075E–05	1.0000E–00
5	19	15	7.0329E–03	1	2.3267E–05	1.0000E–00
6	18	15	1.1982E–03			
Run time: 506 seconds						Objective: 0.2702
						$\int_0^T u(t)\,\mathrm{d}t = 3.5826$

TABLE 5.2. Example 1 solved with the semi-smooth Newton method. $\lVert\nabla\hat{J}_\varepsilon(u^n)_{\lvert I^n}\rVert$ is the $L^2(0,T)$-norm of the right hand side in the reduced Hessian system.

The semi-smooth Newton method. In case of the semi-smooth Newton method (Table 5.2), the reduced Hessian system (4.5) was solved also inexactly. Here, the conjugate gradient iteration was terminated when $\lVert r\rVert \le 3 \cdot 10^{-4}$. That is, we use a constant threshold here which does not depend on the progress in the outer iteration. The same Armijo line search procedure was applied in the search direction δu found in the CG iteration.

Comparing the solutions. Both methods obtained the solution depicted in Figure 5.3 within 15 CG iterations and roughly the same run time. Note that the final residual $\lVert\nabla\hat{J}_\varepsilon(u^n)_{\lvert I^n}\rVert$ achieved in the primal-dual run is smaller since in the next to last Newton step, the desired tolerance of $3 \cdot 10^{-3}$ is only scarcely missed. The primal-dual and the semi-smooth Newton methods appear equally well-suited for this problem. They proved to be robust with respect to the choice of the initial guess.

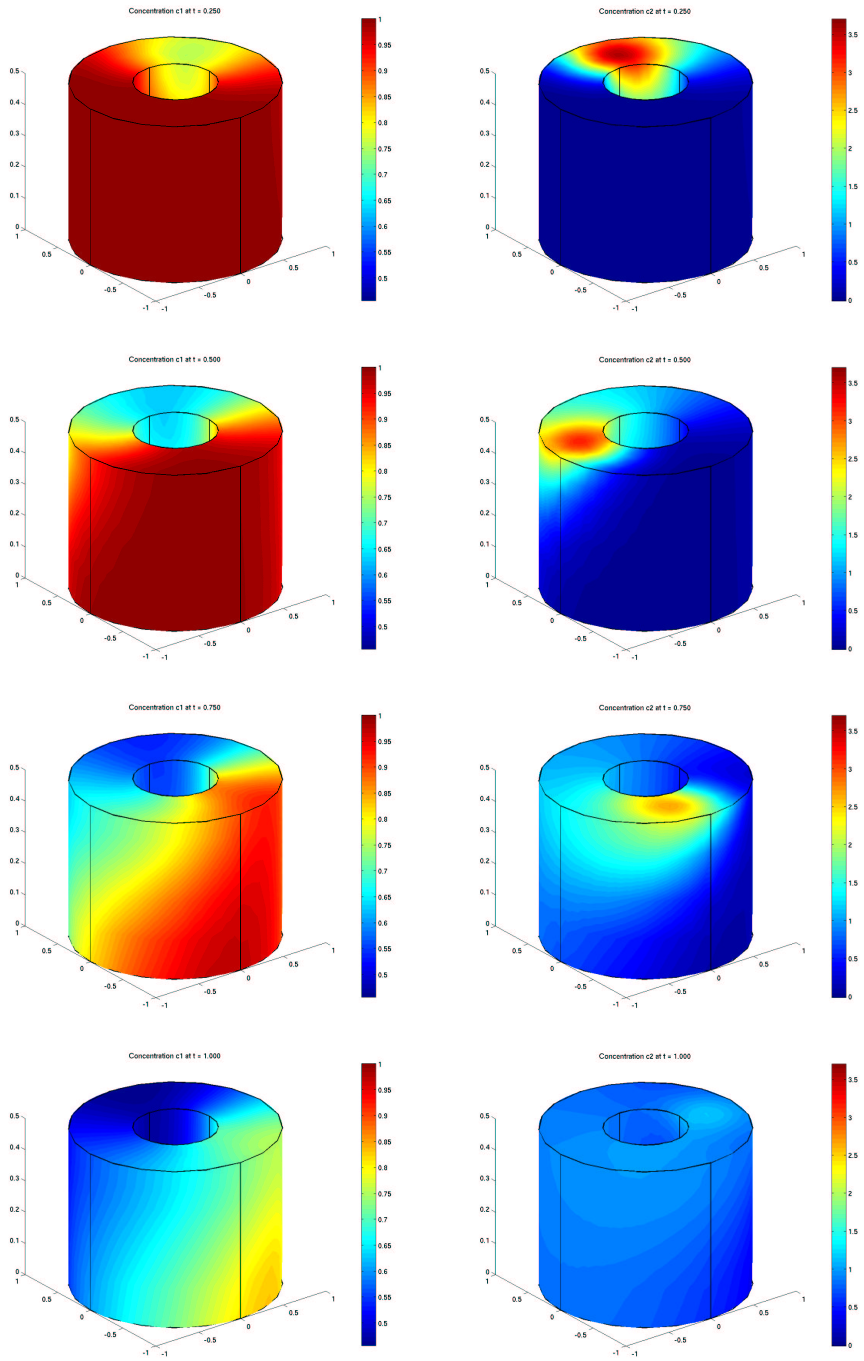

FIGURE 5.4. Example 1: Concentrations of substances 1 (left) and 2 (right) at times $t = 0.25$, $t = 0.50$, $t = 0.75$, and $t = 1.00$.

5.2 Example 2. The second example was designed to be numerically more challenging. Recall that in the first example, the integral constraint was chosen such that it required a rather moderate modification of the optimal solution without integral constraint. To obtain a contrasting situation, we now use $\gamma = 10^{-3}$ as a weight for the control cost and $u_c = 1.8$ as a bound in the integral constraint. All other data are taken over from Example 1. Note that the smaller γ leads to an increased control action, while the lowered integral constraint bound has the opposite effect. The optimal solution is depicted in Figure 5.5. It has been ob-

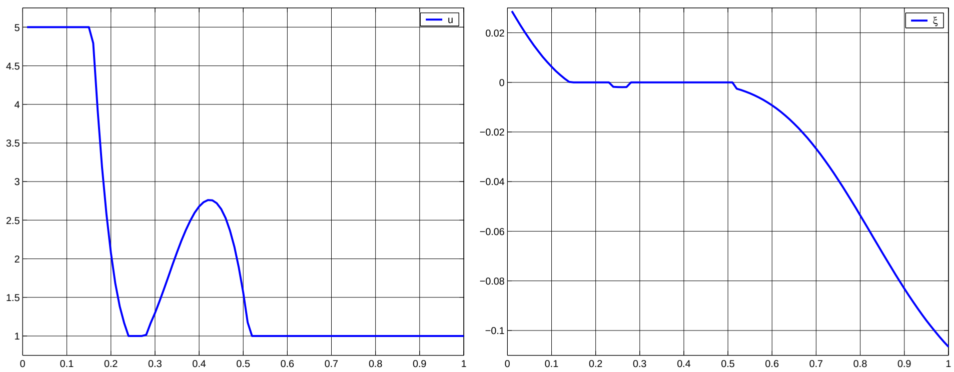

FIGURE 5.5. Example 2: Optimal control u (left) and control constraint multiplier ξ.

tained with the primal-dual active set method as described for Example 1. One immediately observes that the active sets have an interesting structure in this example. In particular, when the control enters the lower bound for the first time, the constraint is only very weakly active. That is, the corresponding multiplier is close to zero, which makes proper identification of the active sets a challenge for both the nonlinear primal-dual and semi-smooth Newton methods.

Table 5.3 shows the convergence history of the primal-dual active set algorithm again from an initial guess of $u_0 \equiv 3$. We observe that significantly more outer iterations are necessary in order to determine the active sets. Also, in every first Newton step, the primal-dual method starts over with a fairly large contribution of the penalty part to the objective. The first Newton steps taken all aim to reduce mainly this part of the objective. This can be seen from the zero angle between the search direction and the negative objective gradient. In other words, the reduced Hessian matrix is a multiple of the identity matrix, being dominated by $(1/\varepsilon)\nabla^2 I(u)$ (see (A.28)).

6. Conclusion

We have presented an optimal control problem for a three-dimensional time-dependent system of semilinear reaction-diffusion equations. The problem is subject to pointwise control constraints as well as a scalar integral control constraint. The latter has not been treated as an explicit constraint by rather added to the objective as a penalty term. We have investigated the state equation and the problem theoretically, presenting first- and second-order necessary and sufficient optimality conditions as well as a convergence result as the penalty parameter tends to zero.

n	$\|A_+^n\|$	$\|A_-^n\|$	$\\|\nabla \hat{J}_\varepsilon(u^n)_{\|I^n}\\|$	CG	$\\|r\\|$	angle	objective	penalty
1	0	0	4.2473E+00	1	4.1407E–02	0.00°	6.0656E-01	2.2882E-01
			1.0379E+00	1	4.4998E–02	0.00°	4.3893E-01	3.6003E-02
			2.3973E–01	1	4.7176E–02	0.00°	4.2233E-01	9.0226E-03
			5.8519E–02	6	2.3392E–03	52.52°	2.0473E-01	2.0718E-03
			1.2424E–02	7	7.7037E–05	36.58°	1.6440E-01	1.2367E-03
			2.2656E–03					
2	52	21	5.3326E+00	1	1.5598E–02	0.00°	1.1584E+00	8.2217E-01
			1.3115E+00	1	2.3360E–02	0.00°	5.0886E-01	1.1913E-01
			3.0751E–01	1	2.7113E–02	0.00°	4.4234E-01	2.5718E-02
			6.4022E–02	2	2.4967E–03	62.19°	3.6867E-01	7.5459E-03
			8.0378E–03	5	5.4365E–06	61.35°	3.6154E-01	5.9262E-03
			1.6748E–04					
3	17	38	4.9186E+00	1	8.4897E–03	0.00°	8.2320E-01	4.9814E-01
			1.2121E+00	1	9.4278E–03	0.00°	4.2938E-01	7.1600E-02
			2.8619E–01	1	9.9664E–03	0.00°	3.8880E-01	1.4939E-02
			5.7817E–02	2	2.7543E–03	76.73°	3.7773E-01	6.5030E-03
			6.4547E–03	7	1.5756E–05	66.44°	3.7424E-01	5.5492E-03
			5.4716E–05					
4	18	35	7.0924E–01	1	5.0034E–03	0.00°	3.9872E-01	3.6886E-02
			1.5946E–01	1	5.6519E–03	0.00°	3.8415E-01	1.0484E-02
			2.7040E–02	3	5.5633E–04	76.80°	3.8037E-01	6.3845E-03
			1.6097E–03					
5	14	50	2.1543E–01	1	1.2975E–03	0.00°	3.8378E-01	1.3592E-02
			4.0717E–02	1	1.3730E–03	0.00°	3.8222E-01	7.1675E-03
			3.9585E–03	7	3.8834E–06	72.65°	3.8183E-01	6.5111E-03
			3.7167E–05					
6	15	52	3.5305E–02	1	6.8442E–04	0.00°	3.8212E-01	7.0777E-03
			3.1016E–03	5	8.4937E–07	75.66°	3.8205E-01	6.5734E-03
			2.8178E–05					
7	15	53	6.0990E–03	2	2.2282E–05	49.47°	3.8207E-01	6.6022E-03
			1.2872E–04					
Run time: 2113 seconds								Objective: 0.38207
								$\int_0^T u(t)\,\mathrm{d}t = 1.9876$

TABLE 5.3. Example 2 solved with the primal-dual active set method. $\|\nabla \hat{J}_\varepsilon(u^n)_{|I^n}\|$ is the $L^2(0,T)$-norm of the right hand side in the reduced Hessian system. The angle between the search direction and the negative reduced objective's gradient is also given. The step length in the line search algorithm was always 1.

Primal-dual active set and semi-smooth Newton methods are known to handle the resulting pointwise control-constrained problem efficiently. We have described in detail how to apply them in the present situation. Finally, we have presented two numerical examples of unequal difficulty.

Acknowledgements

The authors would like to thank Georg Stadler for helpful comments.

Appendix A. Proofs for Section 2

A.1. Proof of Theorem 2.3. In this section we prove that (2.4) possesses a unique solution pair $(c_1, c_2) \in W(0,T) \times W(0,T)$. Uniqueness is shown in Section A.1.1 and Section A.1.2 is devoted to ensuring the existence of a solution by applying the Leray-Schauder fixed-point theorem.

A.1.1. Proof of uniqueness. Let $(c_1, c_2), (\tilde{c}_1, \tilde{c}_2) \in W(0,T) \times W(0,T)$ be two pairs of weak solutions to (2.4). Then, $\delta c_1 = c_1 - \tilde{c}_1 \in W(0,T)$ and $\delta c_2 = c_2 - c_1 \in W(0,T)$ satisfy

$$\delta c_1(0) = 0 \quad \text{and} \quad \delta c_2(0) = 0 \quad \text{in } L^2(\Omega) \tag{A.1a}$$

and

$$\begin{aligned} &\langle (\delta c_1)_t(t), \varphi \rangle_{H^1(\Omega)', H^1(\Omega)} + \int_\Omega d_1 \nabla \delta c_1(t) \cdot \nabla \varphi \, \mathrm{d}x \\ &\qquad + \int_\Omega k_1 \big(\delta c_1(t) c_2(t) + \tilde{c}_1(t) \delta c_2(t) \big) \varphi \, \mathrm{d}x = 0, \end{aligned} \tag{A.1b}$$

$$\begin{aligned} &\langle (\delta c_2)_t(t), \varphi \rangle_{H^1(\Omega)', H^1(\Omega)} + \int_\Omega d_2 \nabla \delta c_2(t) \cdot \nabla \varphi \, \mathrm{d}x \\ &\qquad + \int_\Omega k_2 \big(\delta c_1(t) c_2(t) + \tilde{c}_1(t) \delta c_2(t) \big) \varphi \, \mathrm{d}x = 0 \end{aligned} \tag{A.1c}$$

for all $\varphi \in H^1(\Omega)$ and almost all $t \in [0,T]$. Upon choosing $\varphi = \delta c_1(t)$ in (A.1b), $\varphi = \delta c_2(t)$ in (A.1c) and adding both equations we obtain the inequality

$$\begin{aligned} &\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \big(\|\delta c_1(t)\|^2_{L^2(\Omega)} + \|\delta c_2(t)\|^2_{L^2(\Omega)} \big) + d_1 \|\delta c_1(t)\|^2_{H^1(\Omega)} + d_2 \|\delta c_2(t)\|^2_{H^1(\Omega)} \\ &\quad \le d_1 \|\delta c_1(t)\|^2_{L^2(\Omega)} + d_2 \|\delta c_2(t)\|^2_{L^2(\Omega)} \\ &\qquad + \int_\Omega k_1 \big(|\delta c_1(t)^2 c_2(t)| + |\tilde{c}_1(t) \delta c_2(t) \delta c_1(t)| \big) \, \mathrm{d}x \\ &\qquad + \int_\Omega k_2 \big(|\delta c_1(t) c_2(t) \delta c_2(t)| + |\tilde{c}_1(t) \delta c_2(t)^2| \big) \, \mathrm{d}x \end{aligned} \tag{A.2}$$

for almost all $t \in [0,T]$. Next we estimate the two integrals on the right-hand side of (A.2). Using Hölder's, Gagliardo-Nierenberg's and Young's inequality we find

$$\begin{aligned} \int_\Omega |\delta c_1(t)^2 c_2(t)| \, \mathrm{d}x &\le \|\delta c_1(t)\|_{L^2(\Omega)} \|\delta c_1(t)\|_{L^4(\Omega)} \|c_2(t)\|_{L^4(\Omega)} \\ &\le C_{GN} \|\delta c_1(t)\|^{5/4}_{L^2(\Omega)} \|\delta c_1(t)\|^{3/4}_{H^1(\Omega)} \|c_2(t)\|_{L^4(\Omega)} \\ &\le \frac{d_1}{2k_1} \|\delta c_1(t)\|^2_{H^1(\Omega)} + K_1 \|c_2(t)\|^{8/5}_{L^4(\Omega)} \|\delta c_1(t)\|^2_{L^2(\Omega)} \end{aligned} \tag{A.3}$$

for almost all $t \in [0,T]$ and for a constant $K_1 > 0$ depending on d_1, k_1, and C_{GN}. Analogously, it follows that

$$\begin{aligned} &\int_\Omega |\tilde{c}_1(t) \delta c_2(t) \delta c_1(t)| \, \mathrm{d}x \le \|\tilde{c}_1(t)\|_{L^4(\Omega)} \|\delta c_2(t)\|_{L^2(\Omega)} \|\delta c_1(t)\|_{L^4(\Omega)} \\ &\le C_{GN} \|\tilde{c}_1(t)\|_{L^4(\Omega)} \|\delta c_2(t)\|_{L^2(\Omega)} \|\delta c_1(t)\|^{3/4}_{H^1(\Omega)} \|\delta c_1(t)\|^{1/4}_{L^2(\Omega)} \\ &\le \frac{d_1}{2k_1} \|\delta c_1(t)\|^2_{H^1(\Omega)} + K_2 \|\tilde{c}_1(t)\|^{8/5}_{L^4(\Omega)} \big(\|\delta c_1(t)\|^2_{L^2(\Omega)} + \|\delta c_2(t)\|^2_{L^2(\Omega)} \big) \end{aligned} \tag{A.4}$$

for almost all $t \in [0,T]$ and for a constant $K_2 > 0$, which depends on d_1, k_1, and C_{GN},

$$\begin{aligned}\int_\Omega |\delta c_1(t) c_2(t) \delta c_2(t)| \,\mathrm{d}x &\le \|\delta c_1(t)\|_{L^2(\Omega)} \|c_2(t)\|_{L^4(\Omega)} \|\delta c_2(t)\|_{L^4(\Omega)} \\ &\le \frac{d_2}{2k_2} \|\delta c_2(t)\|^2_{H^1(\Omega)} + K_3 \|c_2(t)\|^{8/5}_{L^4(\Omega)} \big(\|\delta c_1(t)\|^2_{L^2(\Omega)} + \|\delta c_2(t)\|^2_{L^2(\Omega)}\big)\end{aligned} \tag{A.5}$$

for almost all $t \in [0,T]$ and for a constant $K_3 > 0$ depending on d_2, k_2, and C_{GN} and

$$\int_\Omega |\tilde{c}_1(t) \delta c_2(t)^2| \,\mathrm{d}x \le \frac{d_2}{2k_2} \|\delta c_2(t)\|^2_{H^1(\Omega)} + K_4 \|\tilde{c}_1(t)\|^{8/5}_{L^4(\Omega)} \|\delta c_2(t)\|^2_{L^2(\Omega)} \tag{A.6}$$

for almost all $t \in [0,T]$ and for a constant $K_4 > 0$, which depends on d_2, k_2, and C_{GN}. Inserting (A.3)–(A.6) into (A.2), we find that

$$\begin{aligned}&\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \big(\|\delta c_1(t)\|^2_{L^2(\Omega)} + \|\delta c_2(t)\|^2_{L^2(\Omega)}\big) \\ &\quad\le K_5 \big(1 + \|\tilde{c}_1(t)\|^{8/5}_{L^4(\Omega)} + \|c_2(t)\|^{8/5}_{L^4(\Omega)}\big) \big(\|\delta c_1(t)\|^2_{L^2(\Omega)} + \|\delta c_2(t)\|^2_{L^2(\Omega)}\big)\end{aligned} \tag{A.7}$$

for almost all $t \in [0,T]$ and for a constant K_5 depending on K_1 through K_4 and d_1 and d_2. Recall that the space $L^2(0,T;H^1(\Omega))$ is continuously embedded into the space $L^{8/5}(0,T;L^4(\Omega))$; see, e.g., [DL92, p. 470 and p. 471]. Therefore, there exists a constant $K_6 > 0$ such that

$$\int_0^T \|\tilde{c}_1(t)\|^{8/5}_{L^4(\Omega)} + \|c_2(t)\|^{8/5}_{L^4(\Omega)} \,\mathrm{d}t \le K_6.$$

Hence, by Gronwall's inequality we derive from (A.7) that

$$\big(\|\delta c_1(t)\|^2_{L^2(\Omega)} + \|\delta c_2(t)\|^2_{L^2(\Omega)}\big) \le K_7 \big(\|\delta c_1(0)\|^2_{L^2(\Omega)} + \|\delta c_2(0)\|^2_{L^2(\Omega)}\big),$$

where $K_7 = \exp(K_5(T + K_6))$. Due to (A.1a) we have $\delta c_1 = \delta c_2 = 0$. Thus, $c_1 = \tilde{c}_1$ and $c_2 = \tilde{c}_2$ so that the claim follows.

A.1.2. Proof of existence. To prove existence of a weak solution, we apply the Leray-Schauder fixed-point theorem. For a proof we refer to [GT77, p. 222].

Theorem A.1. *Let $\mathcal{T}$ be a compact mapping of a Banach space B into itself and suppose that there exists a constant $M > 0$ such that*

$$\|\varphi\|_B < M \text{ for all } \varphi \in B \text{ and } s \in [0,1] \text{ satisfying } \varphi = s\mathcal{T}\varphi.$$

Then $\mathcal{T}$ has a fixed-point.

Here we choose the Banach space $B = W(0,T) \times W(0,T)$ and introduce the linear operator $\mathcal{T} : B \to B$ as follows: $(c_1, c_2) = \mathcal{T}(y_1, y_2)$ solves

$$c_1(0) = c_{10} \quad \text{and} \quad c_2(0) = c_{20} \quad \text{in } L^2(\Omega) \tag{A.8a}$$

and

$$\langle (c_1)_t(t), \varphi \rangle_{H^1(\Omega)', H^1(\Omega)} + \int_\Omega d_1 \nabla c_1(t) \cdot \nabla \varphi + k_1 c_1(t) y_2(t) \varphi \,\mathrm{d}x = 0, \tag{A.8b}$$

$$\begin{aligned}\langle (c_2)_t(t), \varphi \rangle_{H^1(\Omega)', H^1(\Omega)} &+ \int_\Omega d_2 \nabla c_2(t) \cdot \nabla \varphi + k_2 y_1(t) c_2(t) \varphi \,\mathrm{d}x \\ &= u(t) \int_{\Gamma_c} \alpha(t) \varphi \,\mathrm{d}x\end{aligned} \tag{A.8c}$$

for all $\varphi \in H^1(\Omega)$ and almost all $t \in [0,T]$.

Remark A.2. Notice that for given $(y_1, y_2) \in B$ the parabolic problems for c_1 and c_2 are decoupled. $\Diamond$

Notice that the solvability of (2.4) is equivalent to the existence of a solution pair $(c_1, c_2) \in B$ to the problem $(c_1, c_2) = \mathcal{T}(c_1, c_2)$. Moreover, the equation $(c_1, c_2) = s\mathcal{T}(c_1, c_2)$, $s \in [0,1]$, is equivalent to

$$c_1(0) = sc_{10} \quad \text{and} \quad c_2(0) = sc_{20} \quad \text{in } L^2(\Omega) \tag{A.9a}$$

and

$$\langle (c_1)_t(t), \varphi\rangle_{H^1(\Omega)', H^1(\Omega)} + \int_\Omega d_1\, \nabla c_1(t) \cdot \nabla\varphi + k_1 c_1(t) y_2(t)\varphi \,\mathrm{d}x = 0, \tag{A.9b}$$

$$\begin{aligned} \langle (c_2)_t(t), \varphi\rangle_{H^1(\Omega)', H^1(\Omega)} + \int_\Omega d_2\, \nabla c_2(t) \cdot \nabla\varphi + k_2 y_1(t) c_2(t)\varphi \,\mathrm{d}x \\ = su(t)\int_{\Gamma_c} \alpha(t)\varphi\,\mathrm{d}x \end{aligned} \tag{A.9c}$$

for all $\varphi \in H^1(\Omega)$ and almost all $t \in [0,T]$. The following proposition ensures that $\mathcal{T}$ is well-defined and maps B into itself.

Proposition A.3. *Suppose that $c_{10}, c_{20} \in L^2(\Omega)$ and $u \in L^2(0,T)$. Then there exists a unique solution pair $(c_1, c_2) \in B$ to* (A.8) *for every pair $(y_1, y_2) \in B$.*

Proof. Due to Remark A.2 we show that the function c_1 is uniquely determined for given pair $(y_1, y_2) \in B$. The proof for c_2 is analogous. Let us introduce the bilinear form

$$a(t;\varphi,\psi) = \int_\Omega d_1 \nabla\varphi\cdot\nabla\psi + k_1 y_2(t)\varphi\psi\,\mathrm{d}x \quad \text{for } \varphi,\psi \in H^1(\Omega) \text{ and } t \in [0,T].$$

Since

$$|a(t;\varphi,\psi)| \le \big(d_1 + k_1 C_{L^4}^2 \|y_2\|_{C([0,T];L^2(\Omega))}\big)\|\varphi\|_{H^1(\Omega)}\|\psi\|_{H^1(\Omega)}$$

holds, it follows that $a(t;\cdot,\cdot)$ is continuous for all $t \in [0,T]$. Here, $C_{L^4} > 0$ denotes the embedding constant satisfying

$$\|\varphi\|_{L^4(\Omega)} \le C_{L^4}\|\varphi\|_{H^1(\Omega)} \quad \text{for all } \varphi \in H^1(\Omega).$$

Furthermore, using Hölder's, Gagliardo-Nierenberg's, and Young's inequality, we estimate

$$\begin{aligned} -\int_\Omega y_2(t)\varphi^2\,\mathrm{d}x &\le \|y_2(t)\|_{L^2(\Omega)}\|\varphi\|_{L^4(\Omega)}^2 \\ &\le C_{GN}^2\|y_2\|_{C([0,T];L^2(\Omega)}\|\varphi\|_{H^1(\Omega)}^{3/2}\|\varphi\|_{L^2(\Omega)}^{1/2} \\ &\le \frac{d_1}{2k_1}\|\varphi\|_{H^1(\Omega)}^2 + K_8\|y_2\|_{C([0,T];L^2(\Omega)}^4\|\varphi\|_{L^2(\Omega)}^2 \end{aligned}$$

for a constant $K_8 > 0$ depending on d_1, k_1, and C_{GN}. Thus,

$$a(t;\varphi,\varphi) \ge \frac{d_1}{2k_1}\|\varphi\|_{H^1(\Omega)}^2 - K_9\|\varphi\|_{L^2(\Omega)}^2 \quad \text{for all } \varphi \in H^1(\Omega) \text{ and } t \in [0,T],$$

where $K_9 = d_1 + K_8\|y_2\|_{C([0,T];L^2(\Omega)}^4 > 0$. By Theorems 1 and 2 in [DL92, pp. 512-513] we infer that there exists a unique $c_1 \in W(0,T)$ solving $c_1(0) = c_{10}$ in $L^2(\Omega)$ and (A.8b). Analogously, we derive the existence of a unique $c_2 \in W(0,T)$ satisfying

$c_2(0) = c_{20}$ in $L^2(\Omega)$ and (A.8c). Consequently, the claim of the proposition is shown. □

Proposition A.4. *The operator $\mathcal{T}$ is compact.*

Proof. Let $(y_1, y_2) \in B$ and $\{(y_1^n, y_2^n)\}_{n=1}^\infty$ be a sequence in B satisfying $(y_1^n, y_2^n) \rightharpoonup (y_1, y_2)$ in B as n tends to infinity. Then we prove that the sequence of pairs $(c_1^n, c_2^n) = \mathcal{T}(y_1^n, y_2^n)$, $n \in \mathbb{N}$, converges strongly in B to $(c_1, c_2) = \mathcal{T}(y_1, y_2)$. The functions $z_1^n = c_1^n - c_1 \in W(0,T)$ and $z_1^n = c_2^n - c_2 \in W(0,T)$ solve

$$z_1^n(0) = 0 \quad \text{and} \quad z_2^n(0) = 0 \quad \text{in } L^2(\Omega) \tag{A.10a}$$

and

$$\begin{aligned}\langle (z_1^n)_t(t), \varphi\rangle_{H^1(\Omega)', H^1(\Omega)} &+ \int_\Omega d_1 \nabla z_1^n(t) \cdot \nabla\varphi \,\mathrm{d}x \\ + \int_\Omega k_1 \big((y_2^n(t) - y_2(t))c_1^n(t) &+ z_1^n(t) y_2(t)\big)\varphi \,\mathrm{d}x = 0,\end{aligned} \tag{A.10b}$$

$$\begin{aligned}\langle (z_2^n)_t(t), \varphi\rangle_{H^1(\Omega)', H^1(\Omega)} &+ \int_\Omega d_2 \nabla z_2^n(t) \cdot \nabla\varphi \,\mathrm{d}x \\ + \int_\Omega k_2 \big((y_1^n(t) - y_1(t))c_2^n(t) &+ z_2^n(t) y_1(t)\big)\varphi \,\mathrm{d}x = 0\end{aligned} \tag{A.10c}$$

for all $\varphi \in H^1(\Omega)$ and almost all $t \in [0,T]$. Taking $\varphi = z_1^n(t)$ in (A.10b) and utilizing Hölder's inequality we estimate

$$\begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|z_1^n(t)\|_{L^2(\Omega)}^2 + d_1\big(\|z_1^n(t)\|_{H^1(\Omega)}^2 - \|z_1^n(t)\|_{L^2(\Omega)}^2\big) \\ &\quad\le k_1 \|y_1^n(t) - y_1(t)\|_{L^4(\Omega)} \|c_2^n(t)\|_{L^4(\Omega)} \|z_1^n(t)\|_{L^2(\Omega)} \\ &\qquad + k_1 \|z_1^n(t)\|_{L^4(\Omega)} \|z_1^n(t)\|_{L^2(\Omega)} \|y_2(t)\|_{L^4(\Omega)}.\end{aligned} \tag{A.11}$$

Due to the Gagliardo-Nierenberg inequality and the Young inequality we find

$$\begin{aligned}&\|y_1^n(t) - y_1(t)\|_{L^4(\Omega)} \|c_2^n(t)\|_{L^4(\Omega)} \|z_1^n(t)\|_{L^2(\Omega)} \\ &\quad\le \frac{1}{2}\|y_1^n(t) - y_1(t)\|_{L^4(\Omega)}^2 + \frac{1}{2}\|c_2^n(t)\|_{L^4(\Omega)}^2 \|z_1^n(t)\|_{L^2(\Omega)}^2\end{aligned}$$

and

$$\begin{aligned}&\|z_1^n(t)\|_{L^4(\Omega)} \|z_1^n(t)\|_{L^2(\Omega)} \|y_2(t)\|_{L^4(\Omega)} \\ &\quad\le C_{GN} \|z_1^n(t)\|_{H^1(\Omega)}^{3/4} \|z_1^n(t)\|_{L^2(\Omega)}^{5/4} \|y_2(t)\|_{L^4(\Omega)} \\ &\quad\le \frac{d_1}{2k_1}\|z_1^n(t)\|_{H^1(\Omega)}^2 + K_{10}\|y_2(t)\|_{L^4(\Omega)}^{8/5} \|z_1^n(t)\|_{L^2(\Omega)}^2\end{aligned}$$

for a constant $K_{10} > 0$ depending on d_1, k_1, and C_{GN}. Inserting these two inequalities into (A.11) we derive

$$\begin{aligned}&\|z_1^n(t)\|_{L^2(\Omega)}^2 + d_1 \int_0^t \|z_1^n(\tau)\|_{H^1(\Omega)}^2 \,\mathrm{d}\tau \\ &\quad\le k_1 \int_0^t \|y_1^n(t)(\tau) - y_2^n(\tau)\|_{L^4(\Omega)}^2 \,\mathrm{d}\tau.\end{aligned} \tag{A.12}$$

Since $W(0,T)$ is compactly embedded into $L^{8/5}(0,T;L^4(\Omega))$ and $L^2(0,T;L^4(\Omega))$ there exist constants $K_{11}, K_{12} > 0$ such that

$$\int_0^T k_1 \|c_2^n(t)\|^2_{L^4(\Omega)} \,\mathrm{d}t \le K_{11} \quad \text{and} \quad \int_0^T K_{10} \|y_2(t)\|^{8/5}_{L^4(\Omega)} \,\mathrm{d}t \le K_{12}.$$

Therefore, we infer from the Gronwall inequality

$$\|z_1^n\|^2_{L^2(\Omega)} \le k_1 \|y_1^n y_1\|^2_{L^2(0,T;L^4(\Omega))} \exp(d_1 + K_{11} + K_{12}).$$

By Aubin's lemma (see, e.g., [Tem79, p. 271]) the space $W(0,T)$ is compactly embedded into $L^2(0,T;L^4(\Omega))$. Thus, $z_1^n \to 0$ in $L^\infty(0,T;L^2(\Omega)$ as n tends to infinity and (A.12) yields that $\to$ in L^2 as $n \to \infty$. From (A.10b) we find also that $(z_1^n)_t \to 0$ in $L^2(0,T;H^1(\Omega)')$ so that we have $\lim_{n\to\infty} \|z_1^n\|_{W(0,T)} = 0$. Analogously, we obtain $z_2^n \to 0$ in W as $n \to \infty$ so that the claim of the proposition is shown. □

Proposition A.5. *Let $(c_1, c_2) \in B$ satisfy the fixed-point equation $(c_1, c_2) = s\mathcal{T}(c_1, c_2)$ for $s \in [0,1]$. Then there exists a $T_0 \in (0,T]$ and a constant $M > 0$ such that $\|(c_1, c_2\|_{W(0,T_0)} \le M$.*

Proof. We choose $\varphi = c_1(t)$ in (A.9b) and $\varphi = c_2(t)$ in (A.9c). Adding both equations we obtain

$$\begin{aligned} &\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\|c_1(t)\|^2_{L^2(\Omega)} + \|c_2(t)\|^2_{L^2(\Omega)}\right) \\ &+ d_1\left(\|c_1\|^2_{H^1(\Omega)} - \|c_1\|^2_{L^2(\Omega)}\right) + d_2\left(\|c_2\|^2_{H^1(\Omega)} - \|c_2\|^2_{L^2(\Omega)}\right) \\ &+ \int_\Omega k_1 c_1(t)^2 c_2(t) + k_2 c_1(t) c_2(t)^2 \,\mathrm{d}x = u(t)\int_{\Gamma_c} \alpha(t) c_2(t)\,\mathrm{d}s \end{aligned} \tag{A.13}$$

for almost all $t \in [0,T]$. Applying Hölder's inequality we conclude that

$$\int_{\Gamma_c} \alpha(t) c_2(t)\,\mathrm{d}s \le \|\alpha(t)\|_{L^2(\Gamma_c)} \|c_2(t)\|_{L^2(\Gamma_c)} \tag{A.14}$$

for almost all $t \in [0,T]$. Since there exists a constant $C_\Gamma > 0$ satisfying

$$\|\varphi\|_{L^2(\Gamma)} \le C_\Gamma \|\varphi\|_{H^1(\Omega)} \quad \text{for all } \varphi \in H^1(\Omega),$$

we derive from (A.14) and Young's inequality

$$\begin{aligned} u(t)\int_{\Gamma_c} \alpha(t) c_2(t)\,\mathrm{d}s &\le C_\Gamma\, |u(t)|\, \|\alpha(t)\|_{L^2(\Gamma_c)} \|c_2(t)\|_{H^1(\Omega)} \\ &\le \frac{d_{\mathsf{min}}}{4} \|c_2(t)\|^2_{H^1(\Omega)} + K_{13} \|\alpha(t)\|^2_{L^2(\Gamma_c)} |u(t)|^2 \end{aligned} \tag{A.15}$$

with $d_{\mathsf{min}} = \min(d_1, d_2)$ and $K_{13} = C_\Gamma^2/(d_{\mathsf{min}})$. Next we estimate the integral term on the left-hand side of equation (A.13). From Hölder's, Gagliardo-Nierenberg's and Young's inequality

$$\begin{aligned}\int_\Omega c_1(t)^2 c_2(t)\,\mathrm{d}x &\le \|c_1(t)\|_{L^2(\Omega)}\|c_1(t)\|_{L^4(\Omega)}\|c_2(t)\|_{L^4(\Omega)}\\ &\le C_{GN}^2\,\|c_1(t)\|_{H^1(\Omega)}^{3/4}\|c_1(t)\|_{L^2(\Omega)}^{5/4}\|c_2(t)\|_{H^1(\Omega)}^{3/4}\|c_2(t)\|_{L^2(\Omega)}^{1/4}\\ &\le \frac{3\varepsilon^{4/3}}{4}\big(\|c_1(t)\|_{H^1(\Omega)}\|c_2(t)\|_{H^1(\Omega)}\big)\\ &\quad+\frac{C_{GN}^8}{4\varepsilon^3}\big(\|c_1(t)\|_{L^2(\Omega)}^5\|c_2(t)\|_{L^2(\Omega)}\big)\\ &\le \frac{3\varepsilon^{4/3}}{8}\big(\|c_1(t)\|_{H^1(\Omega)}^2+\|c_2(t)\|_{H^1(\Omega)}^2\big)\\ &\quad+\frac{C_{GN}^8}{4\varepsilon^3}\Big(\frac{5\,\|c_1(t)\|_{L^2(\Omega)}^6}{6}+\frac{\|c_2(t)\|_{L^2(\Omega)}^6}{6}\Big)\end{aligned}$$

and

$$\begin{aligned}\int_\Omega c_1(t)c_2(t)^2\,\mathrm{d}x &\le \frac{3\varepsilon^{4/3}}{8}\big(\|c_1(t)\|_{H^1(\Omega)}^2+\|c_2(t)\|_{H^1(\Omega)}^2\big)\\ &\quad+\frac{C_{GN}^8}{4\varepsilon^3}\Big(\frac{\|c_1(t)\|_{L^2(\Omega)}^6}{6}+\frac{5\,\|c_2(t)\|_{L^2(\Omega)}^6}{6}\Big).\end{aligned}$$

for arbitrary $\varepsilon > 0$. Hence, we find

$$\begin{aligned}&\int_\Omega k_1c_1(t)^2c_2(t)+k_2c_1(t)c_2(t)^2\,\mathrm{d}x\\ &\quad\le\frac{3\varepsilon^{4/3}k_{\mathsf{max}}}{4}\big(\|c_1(t)\|_{H^1(\Omega)}^2+\|c_2(t)\|_{H^1(\Omega)}^2\big)\\ &\quad+\frac{K_{14}}{\varepsilon^3}\big(\|c_1(t)\|_{L^2(\Omega)}^6+\|c_2(t)\|_{L^2(\Omega)}^6\big)\end{aligned}\tag{A.16}$$

with $k_{\mathsf{max}} = \max(k_1, k_2) > 0$ and $K_{14} = 5k_{\mathsf{max}}C_{GN}^8/24$. Next we choose $\varepsilon = \sqrt[4]{d_{\mathsf{min}}^3/(27k_{\mathsf{max}}^3)}$ and set $K_{15} = 2K_{14}/\varepsilon^3$, $K_{16} = 2K_{13}$. Combining (A.13), (A.15), and (A.16) it follows that

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\big(\|c_1(t)\|_{L^2(\Omega)}^2+\|c_2(t)\|_{L^2(\Omega)}^2\big)+d_{\mathsf{min}}\big(\|c_1\|_{H^1(\Omega)}^2+\|c_2\|_{H^1(\Omega)}^2\big)\\ &\quad\le K_{15}\big(\|c_1(t)\|_{L^2(\Omega)}^6+\|c_2(t)\|_{L^2(\Omega)}^6\big)+2\big(\|c_1(t)\|_{L^2(\Omega)}^2+\|c_2(t)\|_{L^2(\Omega)}^2\big)\\ &\qquad+K_{16}\|\alpha(t)\|_{L^2(\Gamma_c)}^2|u(t)|^2\\ &\quad\le K_{17}\big(\|c_1(t)\|_{L^2(\Omega)}^6+\|c_2(t)\|_{L^2(\Omega)}^6+\|c_1(t)\|_{L^2(\Omega)}^2+\|c_2(t)\|_{L^2(\Omega)}^2\big)\\ &\qquad+K_{17}|u(t)|^2\end{aligned}\tag{A.17}$$

for almost all $t \in [0, T]$ and $K_{17} = \max(K_{15}, 2, K_{16}\|\alpha\|_{L^\infty(0,T;L^2(\Gamma_c))}^2)$. Notice that

$$\|c_1(t)\|_{L^2(\Omega)}^6+\|c_1(t)\|_{L^2(\Omega)}^2\le\|c_1(t)\|_{L^2(\Omega)}^2\big(1+\|c_1(t)\|_{L^2(\Omega)}^2\big)^2$$

and

$$\|c_2(t)\|_{L^2(\Omega)}^6+\|c_2(t)\|_{L^2(\Omega)}^2\le\|c_2(t)\|_{L^2(\Omega)}^2\big(1+\|c_2(t)\|_{L^2(\Omega)}^2\big)^2.$$

Consequently,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\|c_1(t)\|^2_{L^2(\Omega)} + \|c_2(t)\|^2_{L^2(\Omega)}\right) + d_{\mathsf{min}}\left(\|c_1\|^2_{H^1(\Omega)} + \|c_2\|^2_{H^1(\Omega)}\right)\\
&\qquad \le K_{17}\|c_1(t)\|^2_{L^2(\Omega)}\big(1+\|c_1(t)\|^2_{L^2(\Omega)}\big)^2 \\
&\qquad\quad + K_{17}\left(\|c_1(t)\|^2_{L^2(\Omega)}\big(1+\|c_1(t)\|^2_{L^2(\Omega)}\big)^2 + |u(t)|^2\right).
\end{aligned} \tag{A.18}$$

Let us define $w_i(t) = 1 + \|c_i(t)\|^2_{L^2(\Omega)} \ge 1$ for $i = 1, 2$. Then it follows from (A.18) that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\big(w_1(t)+w_2(t)\big)^2 &\le K_{17}\left(\big(w_1(t)+w_2(t)\big)^3 + |u(t)|^2\right)\\
&\le K_{17}\big(1+|u(t)|^2\big)\big(w_1(t)+w_2(t)\big)^3
\end{aligned}$$

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

$$\int_{w_1(t)+w_2(t)}^{w_1(0)+w_2(0)} \frac{\mathrm{d}z}{z^3} \le K_{17}\int_0^t \big(1+|u(\tau)|^2\big)\,\mathrm{d}\tau,$$

which yields

$$\frac{1}{2(w_1(0)+w_2(0))^2} - \frac{1}{2(w_1(t)+w_2(t))^2} \le K_{17}\big(t + \|u\|^2_{L^2(0,t)}\big).$$

Consequently,

$$\frac{1}{2(w_1(t)+w_2(t))^2} \ge \frac{1}{2(w_1(0)+w_2(0))^2} - K_{17}\big(t + \|u\|^2_{L^2(0,t)}\big).$$

By the dominated convergence theorem (see, e.g., [RS80, p. 17]) there exists $\bar{T}_1 \in (0, T]$ such that

$$\|u\|^2_{L^2(0,t)} \le \frac{1}{4K_{17}(w_1(0)+w_2(0))^2} \quad \text{for almost all } t \in (0, \bar{T}_1].$$

From this we derive

$$\begin{aligned}
&\frac{1}{2(w_1(t)+w_2(t))^2}\\
&\quad\ge \frac{1 - 2K_{17}(t + \|u\|^2_{L^2(0,t)})(w_1(0)+w_2(0))^2}{2(w_1(0)+w_2(0))^2}\\
&\quad\ge \frac{1 - 2K_{17}t - 1/2}{2(w_1(0)+w_2(0))^2} = \frac{1-2K_{17}t}{4(w_1(0)+w_2(0))^2} \quad \text{for almost all } t \in (0, \bar{T}_1].
\end{aligned}$$

Finally, we obtain

$$\big(w_1(t)+w_2(t)\big)^2 \le \frac{2(w_1(0)+w_2(0))^2}{1 - 4K_{17}(w_1(0)+w_2(0))^2 t} \quad \text{for almost all } t \in (0, \bar{T}_1].$$

Setting $T_1 = \min(\bar{T}_1, 1/(8K_{17}(w_1(0)+w_2(t))^2)$ we infer

$$\begin{aligned}
\|c_1(t)\|^2_{L^2(\Omega)} + \|c_2(t)\|^2_{L^2(\Omega)} &\le \big(w_1(t)+w_2(t)\big)^2 \le 4\big(w_1(0)+w_2(0)\big)^2\\
&= 4\big(2 + \|c_{10}\|^2_{L^2(\Omega)} + \|c_{20}\|^2_{L^2(\Omega)}\big)^2 =: K_{18}
\end{aligned}$$

for almost all $t \in [0, T_1]$. Thus,

$$c_1, c_2 \in L^\infty(0, T_1; L^2(\Omega)). \tag{A.19}$$

By integrating (A.18) over the integral $[0, T_1]$ and using (A.19) we get

$$\begin{aligned}\|c_1(T_1)\|^2_{L^2(\Omega)} + \|c_2(T_1)\|^2_{L^2(\Omega)} + d_{\mathsf{min}} \int_0^{T_1} \|c_1(t)\|^2_{H^1(\Omega)} + \|c_2(t)\|^2_{H^1(\Omega)}\,\mathrm{d}t \\ \le \|c_{10}\|^2_{L^2(\Omega)} + \|c_{20}\|^2_{L^2(\Omega)} + K_{17}\left(2TK_{18}\big(1+K_{18}\big)^2 + \|u\|^2_{L^2(0,T_1)}\right).\end{aligned}$$

Hence, c_1 and c_2 are uniformly bounded in the $L^2(0, T_1; H^1(\Omega))$-norm. This fact together with (A.9b), (A.9c), (A.19) imply that $\|(c_i)_t\|_{L^2(0,T_1;H^1(\Omega)')}$ is uniformly bounded for $i = 1, 2$. Therefore, c_1 and c_2 are uniformly bounded in $W(0, T)$ by a constant M, which gives the claim. □

Now we prove the existence of a solution to (2.4). By applying Theorem A.1 we infer the existence of a solution pair $(c_1^1, c_2^1) \in W(0, T_1) \times W(0, T_1)$ from Propositions A.3-A.5. Let us define the operator $\tilde{\mathcal{T}} : B \to B$ by

$$(c_1, c_2) = \tilde{\mathcal{T}}(y_1, y_2) = \begin{cases} (c_1^1, c_2^1) & \text{on } [0, T_1] \times \Omega, \\ (c_1^2, c_2^2) & \text{on } (T_1, T] \times \Omega, \end{cases}$$

where c_1^2, c_2^2 solve

(A.20a) $$c_1^2(T_1) = c_1^1(T_1) \quad \text{and} \quad c_2^2(T_1) = c_1^1(T_1) \quad \text{in } L^2(\Omega)$$

and

(A.20b) $$\langle (c_1^2)_t(t), \varphi\rangle_{H^1(\Omega)',H^1(\Omega)} + \int_\Omega d_1\,\nabla c_1^2(t)\cdot\nabla\varphi + k_1 c_1^2(t) y_2(t)\varphi\,\mathrm{d}x = 0,$$

(A.20c) $$\begin{aligned}\langle (c_2^2)_t(t), \varphi\rangle_{H^1(\Omega)',H^1(\Omega)} + \int_\Omega d_2\,\nabla c_2^2(t)\cdot\nabla\varphi + k_2 y_1(t) c_2^2(t)\varphi\,\mathrm{d}x \\ = u(t)\int_{\Gamma_c} \alpha(t)\varphi\,\mathrm{d}x\end{aligned}$$

for all $\varphi \in H^1(\Omega)$ and almost all $t \in [T_1, T]$. Note that Propositions A.3 and A.4 also holds for the operator $\tilde{\mathcal{T}}$. Let

$$z_1 = \begin{cases} c_1^1 & \text{on } [0, T_1] \times \Omega, \\ 0 & \text{on } (T_1, T] \times \Omega \end{cases} \quad \text{and} \quad z_2 = \begin{cases} c_2^1 & \text{on } [0, T_1] \times \Omega, \\ 0 & \text{on } (T_1, T] \times \Omega. \end{cases}$$

Then we have $c_i - z_i = 0$ in $[0, T_1] \times \Omega$ for $i = 1, 2$. In particular, $(c_i - z_i)(T_1) = 0$ in $L^2(\Omega)$ for $i = 1, 2$. We take $w_i^2(t) = 1 + \|(c_i - z_i)(t)\|^2_{L^2(\Omega)}$ for $i = 1, 2$ and $t \in [0, T]$. Then,

(A.21) $$w_i^2(T_1) = 1 + \|(c_i - z_i)(T_1)\|^2_{L^2(\Omega)} = 1.$$

Let $\bar{T}_2 \in (T_1, T]$ be chosen in such a way that

(A.22) $$\|u\|_{L^2(T_1,t)} \le \frac{1}{16K_{17}} \quad \text{for almost all } t \in (T_1, \bar{T}_2].$$

Moreover, we define $T_2 = \min(\bar{T}_2, T_1 + 1/(32K_{17}))$. Then we obtain

$$\begin{aligned}\big(w_1^2(t) + w_2^2(t)\big)^2 &\le \frac{2(w_1^2(T_1) + w_2^2(T_1))^2}{1 - 2K_{17}(t + \|u\|_{L^2(T_1,T_2)})(w_1^2(T_1) + w_2^2(T_1))^2} \\ &= \frac{8}{1 - 8K_{17} - 1/2} = \frac{16}{1 - 16K_{17}t} \le 32\end{aligned}$$

for almost all $t \in [0, T_2]$, where we used (A.21), (A.22). Arguing as in the proof of Proposition A.5 there exists a constant $\bar{M} > 0$ such that

$$\|c_1 - z_1\|_{W(0,T_2)} + \|c_2 - z_2\|_{W(0,T_2)} \le \bar{M}.$$

Hence,

$$\begin{aligned} \|c_1\|_{W(0,T_2)} + \|c_2\|_{W(0,T_2)} &\le \|c_1 - z_1\|_{W(0,T_2)} + \|c_2 - z_2\|_{W(0,T_2)} \\ &\quad + \|c_1^1\|_{W(0,T_1)} + \|c_2^1\|_{W(0,T_1)} \le \bar{M} + M. \end{aligned}$$

This implies existence of a solution pair $(c_1, c_2) \in W(0, T_2) \times W(0, T_2)$. Now we can use an introduction argument to get existence of a pair $(c_1, c_2) \in B$. Notice that this induction is based on the existence of a finite decomposition

$$0 < T_1 < \ldots < T_m = T$$

of the interval $[0, T_m]$ such that $T_i = \min(\bar{T}_i, T_{i-1} + 1/(32K_{17})$ for $i = 2, \ldots, m$, where $\bar{T}_i$ ensures

$$\|u\|_{L^2(T_{i-1},t)} \le \frac{1}{16K_{17}} \quad \text{for almost all } t \in (T_{i-1}, \bar{T}_i].$$

Thus,

$$\|c_1\|_{W(0,T_i)} + \|c_2\|_{W(0,T_i)} \le M + (i-1)\bar{M} \quad \text{for } i = 1, \ldots, m.$$

A.2. Proof of Corollary 2.4. Let $(c_1, c_2) \in W(0, T) \times W(0, T)$ be the unique solution to (2.4). Then the claim follows from Theorems 1 and 2 in [DL92, pp. 512-513] provided the linear functional

$$F(\varphi) = \int_0^T \int_\Omega c_1 c_2 \varphi \,\mathrm{d}x\mathrm{d}t \quad \text{for } \varphi \in L^2(0, T; H^1(\Omega))$$

is continuous. Applying Hölder's and Gagliardo-Nierenberg's inequality we have

$$\begin{aligned} \int_0^T \int_\Omega c_1 c_2 \varphi \,\mathrm{d}x\mathrm{d}t &\le \int_0^T \|c_1(t)\|_{L^2(\Omega)} \|c_2(t)\|_{L^3(\Omega)} \|\varphi(t)\|_{L^6(\Omega)} \,\mathrm{d}t \\ &\le C_{GN} \|c_1\|_{C([0,T];L^2(\Omega))} \|c_2\|^{1/2}_{C([0,T];L^2(\Omega))} \int_0^T \|c_2(t)\|^{1/2}_{H^1(\Omega)} \|\varphi(t)\|_{L^6(\Omega)} \,\mathrm{d}t \\ &\le C_{GN} \|c_1\|_{C([0,T];L^2(\Omega))} \|c_2\|^{1/2}_{C([0,T];L^2(\Omega))} \|c_2\|_{L^1(0,T;H^1(\Omega))} \|\varphi\|_{L^2(0,T;L^6(\Omega))} . \end{aligned}$$

As $W(0, T)$ is continuously embedded into $L^1(0, T; H^1(\Omega))$ and as $L^2(0, T; H^1(\Omega))$ is continuously embedded into $L^2(0, T; L^6(\Omega))$, it follows that $F : L^2(0, T; L^2(\Omega)) \to \mathbb{R}$ is continuous.

A.3. Proof of Proposition 2.5. We take $\varphi = c_1(t)+c_2(t)+c_3(t)$ in (2.4b), (2.4b) and (2.6b) for $t \in [0, T]$ and these three equations. Setting $d = d_i$ for $i = 1, 2, 3$ we find

$$\begin{aligned} &\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} \|(c_1 + c_2 + c_3)(t)\|^2_{L^2(\Omega)} + d \int_\Omega |\nabla(c_1 + c_2 + c_3)(t)|_2^2 \,\mathrm{d}x \\ &\qquad = \int_\Omega (-k_1 - k_2 + k_3) c_1(t) c_2(t) (c_1 + c_2 + c_3)(t) \,\mathrm{d}x, \end{aligned} \tag{A.23}$$

where $|\cdot|_2$ is Euclidean norm in $\mathbb{R}^3$. By assumption we have $k_3 = k_1 + k_2$ so that the right-hand side in (A.23) is zero. Hence, by integrating over $[0, t] \subset [0, T]$, we conclude

$$\begin{aligned}\|c_1+c_2+c_3)(t)\|^2_{L^2(\Omega)} + 2d\int_\Omega |\nabla(c_1+c_2+c_3)|_2^2\,\mathrm{d}x \\ = \|c_{10}+c_{20}+c_{30}\|^2_{L^2(\Omega)},\end{aligned} \tag{A.24}$$

for almost all $t \in [0,T]$. From (A.24) the claim follows.

A.4. Proof of Lemma 2.7. For $s \in \mathbb{R}$ it follows that

$$g'(s) = \left\{\begin{array}{ll} 3s^2 & \text{for } s \geq 0, \\ 0 & \text{otherwise} \end{array}\right\} = 3\,[s]_+^2$$

and

$$g''(s) = \left\{\begin{array}{ll} 6s & \text{for } s \geq 0, \\ 0 & \text{otherwise} \end{array}\right\} = 6\,[s]_+.$$

Hence, g'' exists and is continuous. To prove the Lipschitz-continuity we take $s_1, s_2 \in \mathbb{R}$ and estimate

$$|g''(s_1) - g''(s_2)| = 6\,|[s_1]_+ - [s_2]_+|0 \left\{\begin{array}{ll} |0-0| & \text{if } s_1, s_2 \leq 0, \\ 6\,|s_1 - s_2| & \text{if } s_1, S_2 > 0, \\ 6\,|-s_2| & \text{if } s_1 \leq, s_2 > 0, \\ 6\,|s_1| & \text{if } s_1 >, s_2 \leq 0. \end{array}\right. \tag{A.25}$$

Since $|-s_2| \leq |s_1 - s_2|$ for $s_1 \leq 0$, $s_2 > 0$ and $||$ for s, s, we infer from (A.25) that

$$|g''(s_1) - g''(s_2)| \leq 6\,|s_1 - s_2| \quad \text{for all } s_1, s_2 \in \mathbb{R}$$

so that g'' is Lipschitz-continuous in $\mathbb{R}$ with Lipschitz-constant 6.

A.5. Proof of Lemma 2.8. Suppose that $u \in L^2(0,T)$ and the sequence $\{u^n\}_{n=1}^\infty$ are given satisfying $u^n \rightharpoonup u$ in $L^2(0,T)$ as $n \to \infty$. Then we have

$$\lim_{n\to\infty} \int_0^T u^n \varphi\,\mathrm{d}t = \int_0^T u\varphi\,\mathrm{d}t \quad \text{for all } \varphi \in L^2(0,T)$$

implies that

$$\lim_{n\to\infty} \int_0^T u^n\,\mathrm{d}t = \lim_{n\to\infty} \int_0^T u^n \cdot 1\,\mathrm{d}t = \int_0^T u \cdot 1\,\mathrm{d}t = \int_0^T u\,\mathrm{d}t. \tag{A.26}$$

Since g is continuous, we infer from (A.26) that

$$\lim_{n\to\infty} I(u^n) = \lim_{n\to\infty} g\Big(\int_0^T u^n\,\mathrm{d}t - u_c\Big) = g\Big(\lim_{n\to\infty}\int_0^T u^n\,\mathrm{d}t - u_c\Big) = I(u).$$

Thus, I is weakly continuous. Next we prove that I is twice Fréchet-differentiable. For that purpose let $\delta u, \widetilde{\delta u} \in L^2(0,T)$ denote two arbitrary directions. The the directional derivatives of I at $u \in L^2(0,T)$ are given by

$$\nabla I(u)\delta u = g'\Big(\int_0^T u\,\mathrm{d}t - u_c\Big)\int_0^T \delta u\,\mathrm{d}t \tag{A.27}$$

and

$$\nabla^2 I(u)(\delta u, \widetilde{\delta u}) = g'\Big(\int_0^T u\,\mathrm{d}t - u_c\Big)\int_0^T \delta u\,\mathrm{d}t \int_0^T \widetilde{\delta u}\,\mathrm{d}t. \tag{A.28}$$

We prove that $\nabla I(u)$ is the first Fréchet-derivative of I at u. Using Taylor expansion for g at $s+\delta s$ with $s=\int_0^T u\,\mathrm{d}t - u_c$ and $\delta s = \int_0^T \delta u\,\mathrm{d}t$ we find

$$|I(u+\delta u) - I(u) - \nabla I(u)\delta u| = |g(s+\delta s) - g(s) - g'(s)\delta s + O(|\delta s|^2)|. \tag{A.29}$$

As

$$|\delta s|^2 = \Big|\int_0^T \delta u\,\mathrm{d}t\Big|^2 \le T\,\|\delta u\|_{L^2(0,T)} \quad \text{for } \delta u \in L^2(0,T) \tag{A.30}$$

holds, we derive from (A.29) that

$$\lim_{\|\delta u\|_{L^2(0,T)}\to 0} \frac{1}{\|\delta u\|_{L^2(0,T)}}\,|I(u+\delta u) - I(u) - \nabla I(u)\delta u| = 0.$$

Utilizing analogous arguments it also follows that $\nabla^2 I(u)$ is the second Fréchet-derivative of I at u. Lemma 2.7 implies that $\nabla^2 I$ is Lipschitz-continuous in $L^2(0,T)$; in fact, for arbitrary $u_1, u_2 \in L^2(0,T)$ we deduce from (A.28) and (A.30) that

$$\begin{aligned}
&|\nabla^2 I(u_1)(\delta u, \widetilde{\delta u}) - \nabla^2 I(u_2)(\delta u, \widetilde{\delta u})| \\
&\quad \le 6\Big|\int_0^T \delta u\,\mathrm{d}t\Big|\,\Big|\int_0^T \widetilde{\delta u}\,\mathrm{d}t\Big|\int_0^T |u_1 - u_2|\,\mathrm{d}t \\
&\quad \le 6T^{3/2}\,\|\delta u\|_{L^2(0,T)}\|\widetilde{\delta u}\|_{L^2(0,T)}\|u_1 - u_2\|_{L^2(0,T)},
\end{aligned}$$

which gives the claim.

Appendix B. Proofs of Section 3

B.1. Proof of Proposition 3.1. Since the cost functional J is a quadratic cost functional and I is twice continuously Fréchet-differentiable with Lipschitz-continuous second Fréchet-derivative (see Lemma 2.8), the assertion for J_ε follows by standard arguments. Let us present the first and second derivatives of J_ε at the point $x = (c_1, c_2, u) \in X$ for later reference. We choose arbitrary directions $\delta x = (\delta c_1, \delta c_2, \delta u)$ and $\widetilde{\delta x} = (\widetilde{\delta c_1}, \widetilde{\delta c_2}, \widetilde{\delta u})$ in X and obtain

$$\begin{aligned}
\nabla J_\varepsilon(x)\delta x &= \int_\Omega \beta_1\big(c_1(T) - c_{1T}\big)\delta c_1(T) + \beta_2\big(c_2(T) - c_{2T}\big)\delta c_2(T)\,\mathrm{d}x \\
&\quad + \int_0^T \gamma u \delta u\,\mathrm{d}t + \frac{1}{\varepsilon}\,\nabla I(u)\delta u,
\end{aligned} \tag{B.1}$$

$$\begin{aligned}
\nabla^2 J_\varepsilon(u)\big(\delta x, \widetilde{\delta x}\big) &= \int_\Omega \beta_1 \delta c_1(T)\widetilde{\delta c_1}(T) + \beta_2 \delta c_2(T)\widetilde{\delta c_2}\,\mathrm{d}x \\
&\quad + \int_0^T \gamma \delta u \widetilde{\delta u}\,\mathrm{d}t + \frac{1}{\varepsilon}\,\nabla^2 I(u)\big(\delta u, \widetilde{\delta u}\big),
\end{aligned} \tag{B.2}$$

where $\nabla I(u)$ and $\nabla I(u)$ are given in (A.27) and (A.28), respectively. Next we turn to the operator e. For arbitrary $\delta x = (\delta c_1, \delta c_2, \delta u), \widetilde{\delta x} = (\widetilde{\delta c_1}, \widetilde{\delta c_2}, \widetilde{\delta u}) \in X$ we compute the directional derivatives at a point $x = (c_1, c_2, u) \in X$. The first derivative is given by

$$\begin{aligned}&\langle \nabla e_1(x)\delta x, \varphi\rangle_{L^2(0,T;H^1(\Omega)'),L^2(0,T;H^1(\Omega))}\\ &\quad= \int_0^T \langle (\delta c_1)_t(t), \varphi(t)\rangle_{H^1(\Omega)',H^1(\Omega)}\,\mathrm{d}t\\ &\qquad+\int_0^T\int_\Omega d_1\nabla\delta c_1\cdot\nabla\varphi + k_1(\delta c_1 c_2 + c_1\delta c_2)\varphi\,\mathrm{d}x\mathrm{d}t,\end{aligned}\tag{B.3a}$$

$$\begin{aligned}&\langle \nabla e_2(x)\delta x, \varphi\rangle_{L^2(0,T;H^1(\Omega)'),L^2(0,T;H^1(\Omega))}\\ &\quad= \int_0^T \Big(\langle (\delta c_2)_t(t), \varphi(t)\rangle_{H^1(\Omega)',H^1(\Omega)} - \delta u\int_{\Gamma_c}\alpha\varphi\,\mathrm{d}x\Big)\mathrm{d}t\\ &\qquad+\int_0^T\int_\Omega d_2\nabla\delta c_2\cdot\nabla\varphi + k_2(\delta c_1 c_2 + c_1\delta c_2)\varphi\,\mathrm{d}x\mathrm{d}x\mathrm{d}t\end{aligned}\tag{B.3b}$$

for $\varphi \in L^2(0,T;H^1(\Omega))$ and

$$\nabla e_3(x)\delta x = \delta c_1(0) \quad\text{and}\quad \nabla e_4(x)\delta x = \delta c_2(0).\tag{B.3c}$$

For the second derivative we compute

$$\begin{aligned}&\langle \nabla^2 e_1(x)(\delta x,\widetilde{\delta x}), \varphi\rangle_{L^2(0,T;H^1(\Omega)'),L^2(0,T;H^1(\Omega))}\\ &\quad= \int_0^T\int_\Omega k_1(\delta c_1\widetilde{\delta c_2} + \widetilde{\delta c_1}\delta c_2)\varphi\,\mathrm{d}x\mathrm{d}t,\end{aligned}\tag{B.4a}$$

$$\begin{aligned}&\langle \nabla^2 e_2(x)(\delta x,\widetilde{\delta x}), \varphi\rangle_{L^2(0,T;H^1(\Omega)'),L^2(0,T;H^1(\Omega))}\\ &\quad= \int_0^T\int_\Omega k_2(\delta c_1\widetilde{\delta c_2} + \widetilde{\delta c_1}\delta c_2)\varphi\,\mathrm{d}x\mathrm{d}t\end{aligned}\tag{B.4b}$$

for $\varphi \in L^2(0,T;H^1(\Omega))$ and

$$\nabla^2 e_3(x)(\delta x,\widetilde{\delta x}) = \nabla^2 e_4(x)(\delta x,\widetilde{\delta x}) = 0.\tag{B.4c}$$

Utilizing Hölder's and Gagliardo-Nierenberg's inequality, (2.1) and $\|\varphi\|_{L^2(\Omega)} \le \|\varphi\|_{H^1(\Omega)}$ for all $\varphi \in H^1(\Omega)$ we have

$$\begin{aligned}&\|e_1(x+\delta x) - e_1(x) - \nabla e_1(x)\delta x\|_{L^2(0,T;H^1(\Omega)')}\\ &= \sup\Big\{k_1\int_0^T\int_\Omega \delta c_1\delta c_2\varphi\,\mathrm{d}x\mathrm{d}t : \|\varphi\|_{L^2(0,T;H^1(\Omega))=1}\Big\}\\ &\le k_1\,\|\delta c_1\|_{C([0,T];L^2(\Omega))}\|\delta c_2\|_{L^2(0,T;L^4(\Omega))}\\ &\qquad\cdot\sup\big\{\|\varphi\|_{L^2(0,T;L^4(\Omega))} : \|\varphi\|_{L^2(0,T;H^1(\Omega))} = 1\big\}\\ &\le k_1 C_W\|\delta c_1\|_{W(0,T)}\|\delta c_2\|^{1/2}_{L^2(0,T;L^2(\Omega))}C_{GN}\|\delta c_2\|^{1/2}_{L^2(0,T;H^1(\Omega))}\\ &\qquad\cdot C_{GN}\sup\big\{\|\varphi\|^{1/2}_{L^2(0,T;H^1(\Omega))} : \|\varphi\|_{L^2(0,T;H^1(\Omega))} = 1\big\}\\ &\le k_1 C_W\|\delta c_1\|_{W(0,T)}\|\delta c_2\|_{W(0,T)} \le \frac{k_1 C_W C_{GN}^2}{2}\,\Big(\|\delta c_1\|^2_{W(0,T)} + \|\delta c_2\|^2_{W(0,T)}\Big).\end{aligned}$$

This yields

$$\lim_{\|\delta x\|_X \to 0} \frac{1}{\|\delta x\|_X} \|e_1(x+\delta x) - e_1(x) - \nabla e_1(x)\delta x\|_{L^2(0,T;H^1(\Omega)')} = 0$$

so that (B.3a) is the first Fréchet-derivative of e_1 in direction δx at x. Analogously, we estimate

$$\begin{aligned}
&\|e_2(x+\delta x) - e_2(x) - \nabla e_2(x)\delta x\|_{L^2(0,T;H^1(\Omega)')} \\
&\qquad \le \frac{k_2 C_W C_{GN}^2}{2} \left(\|\delta c_1\|^2_{W(0,T)} + \|\delta c_2\|^2_{W(0,T)} \right),
\end{aligned}$$

which implies that (B.3b) is also the first Fréchet-derivative of e_2 in direction δx at x. Since e_3 and e_4 are linear and bounded operators, their directional derivatives (B.3c) are their first Fréchet-derivatives. From

$$\begin{aligned}
\|\nabla e_1(x+\widetilde{\delta x}) - \nabla e_1(x) - \nabla^2 e_1(x)\widetilde{\delta x}\|_{\mathcal{L}(L^2(0,T;H^1(\Omega)),L^2(0,T;H^1(\Omega)'))} &= 0, \\
\|\nabla e_2(x+\widetilde{\delta x}) - \nabla e_2(x) - \nabla^2 e_2(x)\widetilde{\delta x}\|_{\mathcal{L}(L^2(0,T;H^1(\Omega)),L^2(0,T;H^1(\Omega)'))} &= 0
\end{aligned}$$

we infer that (B.4a) and (B.4b) are the second Fréchet-derivatives of e_1 and e_2, respectively. Here, $\mathcal{L}(L^2(0,T;H^1(\Omega)),L^2(0,T;H^1(\Omega)'))$ denotes the Banach space of all bounded and linear operators from $L^2(0,T;H^1(\Omega))$ into $L^2(0,T;H^1(\Omega)')$ supplied with the common norm. Clearly, (B.4c) are the second Fréchet-derivatives of e_3 and e_4, respectively. Since $\nabla^2 e$ does not depend on the point x, it follows immediately that $\nabla^2 e$ is Lipschitz-continuous on X.

B.2. Proof of Proposition 3.2. The operator $\nabla_{(c_1,c_2)} e(x) : W(0,T) \times W(0,T) \to Y'$ is bijective if for any $(f_1, f_2, \phi_1, \phi_2) \in Y'$ there exists a unique pair $(\delta c_1, \delta c_2) \in W(0,T) \times W(0,T)$ solving

$$\delta c_1(0) = \phi_1 \quad \text{and} \quad \delta c_2(0) = \phi_2 \quad \text{in } L^2(\Omega) \tag{B.5a}$$

hold and

$$\begin{aligned}
&\langle (\delta c_1)_t(t), \varphi \rangle_{H^1(\Omega)',H^1(\Omega)} + \int_\Omega d_1 \nabla \delta c_1(t) \cdot \nabla \varphi \, \mathrm{d}x \\
&\quad = -k_1 \int_\Omega \big(\delta c_1(t) c_2(t) + c_1(t) \delta c_2(t)\big) \varphi \, \mathrm{d}x + \langle f_1(t), \varphi \rangle_{H^1(\Omega)',H^1(\Omega)},
\end{aligned} \tag{B.5b}$$

$$\begin{aligned}
&\langle (\delta c_2)_t(t), \varphi \rangle_{H^1(\Omega)',H^1(\Omega)} + \int_\Omega d_2 \nabla \delta c_2(t) \cdot \nabla \varphi \, \mathrm{d}x \\
&\quad = -k_2 \int_\Omega \big(\delta c_1(t) c_2(t) + c_1(t) \delta c_2(t)\big) \varphi \, \mathrm{d}x + \langle f_2(t), \varphi \rangle_{H^1(\Omega)',H^1(\Omega)}
\end{aligned} \tag{B.5c}$$

for all $\varphi \in H^1(\Omega)$ and almost all $t \in [0,T]$. To prove uniqueness of a solution to (B.5) we can proceed as in Appendix A.1.1. Moreover, the existence follows from an a-priori estimate for δc_1 and δc_2 in $W(0,T) \times W(0,T)$. Again, we argue similarly as in Appendix A.1.1, but instead of (A.3)–(A.6) we estimate

$$\text{(B.6)}\quad \begin{aligned}\int_\Omega |\delta c_1(t)^2 c_2(t)|\,\mathrm{d}x &\le \|\delta c_1(t)\|_{L^2(\Omega)}\|\delta c_1(t)\|_{L^4(\Omega)}\|c_2(t)\|_{L^4(\Omega)}\\ &\le \frac{d_1}{6k_1}\|\delta c_1(t)\|^2_{H^1(\Omega)} + \tilde{K}_1\|c_2(t)\|^{8/5}_{L^4(\Omega)}\|\delta c_1(t)\|^2_{L^2(\Omega)},\end{aligned}$$

$$\text{(B.7)}\quad \begin{aligned}\int_\Omega |\tilde{c}_1(t)\delta c_2(t)\delta c_1(t)|\,\mathrm{d}x &\le \|\tilde{c}_1(t)\|_{L^4(\Omega)}\|\delta c_2(t)\|_{L^2(\Omega)}\|\delta c_1(t)\|_{L^4(\Omega)}\\ &\le \frac{d_1}{6k_1}\|\delta c_1(t)\|^2_{H^1(\Omega)} + \tilde{K}_2\|\tilde{c}_1(t)\|^{8/5}_{L^4(\Omega)}\big(\|\delta c_1(t)\|^2_{L^2(\Omega)} + \|\delta c_2(t)\|^2_{L^2(\Omega)}\big),\end{aligned}$$

$$\text{(B.8)}\quad \begin{aligned}\int_\Omega |\delta c_1(t)c_2(t)\delta c_2(t)|\,\mathrm{d}x &\le \|\delta c_1(t)\|_{L^2(\Omega)}\|c_2(t)\|_{L^4(\Omega)}\|\delta c_2(t)\|_{L^4(\Omega)}\\ &\le \frac{d_2}{6k_2}\|\delta c_2(t)\|^2_{H^1(\Omega)} + \tilde{K}_3\|c_2(t)\|^{8/5}_{L^4(\Omega)}\big(\|\delta c_1(t)\|^2_{L^2(\Omega)} + \|\delta c_2(t)\|^2_{L^2(\Omega)}\big),\end{aligned}$$

$$\text{(B.9)}\quad \int_\Omega |\tilde{c}_1(t)\delta c_2(t)^2|\,\mathrm{d}x \le \frac{d_2}{6k_2}\|\delta c_2(t)\|^2_{H^1(\Omega)} + \tilde{K}_4\|\tilde{c}_1(t)\|^{8/5}_{L^4(\Omega)}\|\delta c_2(t)\|^2_{L^2(\Omega)}$$

for almost all $t \in [0,T]$ and for positive constants $\tilde{K}_i$, $i = 1,\dots,4$, depending on, which depends on d_1, d_2, k_1, k_2, and C_{GN}. Furthermore, we have

$$\text{(B.10)}\quad \langle f_i(t), \delta c_i(t)\rangle_{H^1(\Omega)',H^1(\Omega)} \le \frac{d_1}{6}\|\delta c_i(t)\|^2_{H^1(\Omega)} + \frac{1}{6d_1}\|f_i(t)\|^2_{H^1(\Omega)'}, \quad i = 1,2.$$

Inserting (B.6)-(B.10) we find

$$\text{(B.11)}\quad \begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\big(\|\delta c_1(t)\|^2_{L^2(\Omega)} + \|\delta c_2(t)\|^2_{L^2(\Omega)}\big)d_1\|\delta c_1(t)\|^2_{L^2(\Omega)} + d_2\|\delta c_2(t)\|^2_{L^2(\Omega)}\\ &\le \tilde{K}_5\big(1 + \|c_1(t)\|^{8/5}_{L^2(\Omega)} + \|c_2(t)\|^{8/5}_{L^2(\Omega)}\big)\big(\|\delta c_1(t)\|^2_{L^2(\Omega)} + \|\delta c_2(t)\|^2_{L^2(\Omega)}\big)\\ &\quad + \tilde{K}_6\big(\|f_1(t)\|^2_{H^1(\Omega)'} + \|f_2(t)\|^2_{H^1(\Omega)'}\big)\end{aligned}$$

for almost all $t \in [0,T]$ and for positive constants $\tilde{K}_5, \tilde{K}_6$ depending on $\tilde{K}_i$, $i = 1,\dots,4$, and d_1, d_2. From (B.11) we derive that δc_1 and δc_2 are bounded in $L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;H^1(\Omega))$. By (B.5b) and (B.5b) it follows that $(\delta c_i)_t$ is bounded in $L^2(0,T;H^1(\Omega)')$ for $i = 1,2$. Consequently, there exists a constant $\tilde{K}_7 > 0$ such that

$$\text{(B.12)}\quad \begin{aligned}&\|\delta c_1\|_{W(0,T)} + \|\delta c_2\|_{W(0,T)}\\ &\le \tilde{K}_7\big(\|\phi_1\|_{L^2(\Omega)} + \|\phi_2\|_{L^2(\Omega)} + \|f_1\|_{L^2(0,T;H^1(\Omega)')} + \|f_2\|_{L^2(0,T;H^1(\Omega)')}\big).\end{aligned}$$

Now existence of a solution to (B.5) can be shown by utilizing (B.12) and standard Galerkin techniques. We obtain (3.1) by taking $f_i = 0$, $\phi_i = 0$ for $i = 1,2$ and estimating similarly as in (A.15).

B.3. Proof of Corollary 3.6. Multiplying (3.3a) by $\lambda_1^*(t)$ and (3.3b) by $\lambda_2^*(t)$, integrating over Ω and adding both equations we obtain the inequality

$$\begin{aligned}
&-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\|\lambda_1^*(t)\|^2_{L^2(\Omega)} + \|\lambda_2^*(t)\|^2_{L^2(\Omega)}\right) + d_1\,\|\lambda_1^*(t)\|^2_{H^1(\Omega)} \\
&\quad + d_2\,\|\lambda_2^*(t)\|^2_{H^1(\Omega)} \le d_1\,\|\lambda_1^*(t)\|^2_{L^2(\Omega)} + d_2\,\|\lambda_2^*(t)\|^2_{L^2(\Omega)} \\
&\qquad + \int_\Omega k_1\left(|c_2^*(t)\lambda_1^*(t)^2| + |c_1^*(t)\lambda_1^*(t)\lambda_2^*(t)|\right)\mathrm{d}x \\
&\qquad + \int_\Omega k_2\left(|c_2^*(t)\lambda_1^*(t)\lambda_2^*(t)| + |c_1^*(t)\lambda_2^*(t)^2|\right)\mathrm{d}x
\end{aligned} \tag{B.13}$$

for almost all $t \in [0,T]$. Next we estimate the two integrals on the right-hand side of (B.13). Using Hölder's, Gagliardo-Nierenberg's and Young's inequalities we find

$$\begin{aligned}
&\int_\Omega \left(|c_2^*(t)\lambda_1^*(t)^2| + |c_1^*(t)\lambda_1^*(t)\lambda_2^*(t)|\right)\mathrm{d}x \le \frac{d_1}{2k_1}\,\|\lambda_1^*(t)\|^2_{H^1(\Omega)} \\
&\quad + \tilde{K}_1\left(\|c_1^*(t)\|^{8/5}_{L^4(\Omega)} + \|c_2^*(t)\|^{8/5}_{L^4(\Omega)}\right)\left(\|\lambda_1^*(t)\|^2_{L^2(\Omega)} + \|\lambda_2^*(t)\|^2_{L^2(\Omega)}\right)
\end{aligned} \tag{B.14}$$

for almost all $t \in [0,T]$ and for a constant $\tilde{K}_1 > 0$, which depends on d_1, k_1 and C_{GN} (compare estimates (A.3) and (A.4)) and

$$\begin{aligned}
&\int_\Omega \left(|c_2^*(t)\lambda_1^*(t)\lambda_2^*(t)| + |c_1^*(t)\lambda_2^*(t)^2|\right)\mathrm{d}x \le \frac{d_2}{2k_2}\,\|\lambda_2^*(t)\|^2_{H^1(\Omega)} \\
&\quad + \tilde{K}_2\left(\|c_1^*(t)\|^{8/5}_{L^4(\Omega)} + \|c_2^*(t)\|^{8/5}_{L^4(\Omega)}\right)\left(\|\lambda_1^*(t)\|^2_{L^2(\Omega)} + \|\lambda_2^*(t)\|^2_{L^2(\Omega)}\right)
\end{aligned} \tag{B.15}$$

for almost all $t \in [0,T]$ and for a constant $\tilde{K}_2 > 0$ depending on d_2, k_2 and C_{GN} (compare estimates (A.5) and (A.6)). Inserting (B.14) and (B.15) into (B.13) we conclude

$$\begin{aligned}
&-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\|\lambda_1^*(t)\|^2_{L^2(\Omega)} + \|\lambda_2^*(t)\|^2_{L^2(\Omega)}\right) + d_1\,\|\lambda_1^*(t)\|^2_{H^1(\Omega)} \\
&\quad + d_2\,\|\lambda_2^*(t)\|^2_{H^1(\Omega)} \le d_1\,\|\lambda_1^*(t)\|^2_{L^2(\Omega)} + d_2\,\|\lambda_2^*(t)\|^2_{L^2(\Omega)} \\
&\le \tilde{K}_3\left(1 + \|c_1^*(t)\|^{8/5}_{L^4(\Omega)} + \|c_2^*(t)\|^{8/5}_{L^4(\Omega)}\right)\left(\|\lambda_1^*(t)\|^2_{L^2(\Omega)} + \|\lambda_2^*(t)\|^2_{L^2(\Omega)}\right)
\end{aligned} \tag{B.16}$$

for almost all $t \in [0,T]$ and for a constant $\tilde{K}_3 > 0$ depending on $\tilde{K}_1$, $\tilde{K}_2$, d_i and k_i for $i = 1,2$. To achieve an estimate for the λ_i^*'s in the $L^\infty(0,T;L^2(\Omega))$-norm we integrate (B.16) over $[t,T] \subset [0,T]$ and deduce

$$\begin{aligned}
&\|\lambda_1^*(t)\|^2_{L^2(\Omega)} + \|\lambda_2^*(t)\|^2_{L^2(\Omega)} \le \|\lambda_1^*(T)\|^2_{L^2(\Omega)} + \|\lambda_2^*(T)\|^2_{L^2(\Omega)} \\
&\quad + \int_t^T \tilde{K}_3\left(1 + \|c_1^*(t)\|^{8/5}_{L^4(\Omega)} + \|c_2^*(t)\|^{8/5}_{L^4(\Omega)}\right) \\
&\qquad \cdot\left(\|\lambda_1^*(t)\|^2_{L^2(\Omega)} + \|\lambda_2^*(t)\|^2_{L^2(\Omega)}\right)\mathrm{d}s
\end{aligned} \tag{B.17}$$

for almost all $t \in [0,T]$. Using Gronwall's inequality, (3.3e) and (3.3f) it follows that

$$\begin{aligned}
&\|\lambda_1^*(t)\|^2_{L^2(\Omega)} + \|\lambda_2^*(t)\|^2_{L^2(\Omega)} \\
&\quad \le \tilde{K}_4\left(\|c_1^*(T) - c_{1T}\|^2_{L^2(\Omega)} + \|c_2^*(T) - c_{2T}\|^2_{L^2(\Omega)}\right)
\end{aligned} \tag{B.18}$$

for almost all $t \in [0,T]$, where $\tilde{K}_4 > 0$ depends on $\tilde{K}_3$, T, β_i and $\|c_i^*\|_{L^{8/5}(0,T;L^4(\Omega))}$ for $i = 1, 2$. Thus, using

$$\frac{1}{2}(a+b)^2 \le a^2 + b^2 \quad \text{for all } a, b \in \mathbb{R}$$

the functions λ_i^*, $i = 1, 2$, are bounded in $L^\infty(0,T;L^2(\Omega))$ by

$$\begin{aligned} &\|\lambda_1^*\|^2_{L^\infty(0,T;L^2(\Omega))} + \|\lambda_2^*\|^2_{L^\infty(0,T;L^2(\Omega))} \\ &\qquad \le \tilde{K}_5\big(\|c_1^*(T) - c_{1T}\|^2_{L^2(\Omega)} + \|c_2^*(T) - c_{2T}\|^2_{L^2(\Omega)}\big) \end{aligned} \tag{B.19}$$

with $\tilde{K}_5 = \sqrt{2}\sqrt{\tilde{K}_4}$. From (B.17) and (B.18) we infer that

$$\begin{aligned} &\|\lambda_1^*\|^2_{L^2(0,T;H^1(\Omega))} + \|\lambda_2^*\|^2_{L^2(0,T;H^1(\Omega))} \\ &\qquad \le \tilde{K}_6\big(\|c_1^*(T) - c_{1T}\|^2_{L^2(\Omega)} + \|c_2^*(T) - c_{2T}\|^2_{L^2(\Omega)}\big) \end{aligned}$$

for a constant $\tilde{K}_6 > 0$. Finally, (3.11) is satisfied, since $H^1(\Omega)$ is continuously embedded into $L^4(\Omega)$.

References

[Alt92] H. W. Alt. *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung.* Springer-Verlag, Berlin, 1992.

[BHHK00] M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch. A comparison of interior point methods and a Moreau-Yosida based active set strategy for constrained optimal control problems. *SIAM J. Optimization*, 11:495-521, 2000.

[BIK99] M. Bergounioux, K. Ito, and K. Kunisch. Primal-dual strategy for constrained optimal control problems. *SIAM J. Control and Optimization*, 37:1176-1194, 1999.

[Ber92] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods.* Academic Press, New York, 1982.

[BZ99] J. F. Bonnans and H. Zidani. Optimal control problems with partially polyhedric constraints. *SIAM J. Control Optim.*, 37:1726-1741, 1999.

[BS03] J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems.* Springer-Verlag, New York, 2000.

[DL92] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution Problems I.* Springer-Verlag, Berlin, 1992.

[DiB93] E. DiBenedetto. *Degenerate Parabolic Equations.* Springer-Verlag, New York, 1993.

[Eva98] L. C. Evans. *Partial Differential Equations.* Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, Rhode Island, 1998.

[GK02] C. Geiger and C. Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben.* Springer-Verlag, Berlin, 2002.

[GT77] D. Gilbarg and N. S. Trudinger. *Elliptic Differential Equations of Second-Order.* Springer-Verlag, Berlin, 1977.

[Gri03] 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.

[HI84] W. W. Hager and G. Ianculescu. Dual approximations in optimal control. *SIAM J. Control Optim.*, 22:423-465, 1984.

[Hin01] M. Hintermüller. On a globalized augmented Lagrangian-SQP algorithm for nonlinear optimal control problems with box constraints. In K.-H. Hoffmann, R. H. W. Hoppe, V. Schulz, editors, *Fast solution of discretized optimization problems*, International Series of Numerical Mathematics, 138:139-153, 2001.

[Hin03] M. Hintermüller. A primal-dual active set algorithm for bilaterally control constrained optimal control problems. *Quarterly of Applied Mathematics*, 61:131-160, 2003.

[HH02] M. Hintermüller and M Hinze. Globalization of SQP methods in control of the instationary Navier-Stokes equations *Math. Model. Numer. Anal.* 36:725-746, 2002.

[HIK03] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semi-smooth Newton method. *SIAM J. Optimization,* 13:865-888, 2003.

[HS03] M. Hintermüller and G. Stadler. A semi-smooth Newton method for constrained linear-quadratic control problems. *Z. Angew. Math. Mech.*, 83:219-237, 2003.

[IK00] K. Ito and K. Kunisch. Augmented Lagrangian methods for nonsmooth convex optimization in Hilbert spaces. *Nonlinear Analysis – Theory, Methods and Applicatitions*, 41:573-589, 2000.

[IK02] K. Ito and K. Kunisch. The primal-dual active set method for nonlinear problems with bilaterally constraints. Technical Report No. 214, Special Research Center F 003 *Optimization and Control*, Project area *Continuous Optimization and Control*, University of Graz & Technical University of Graz, 2001, submitted.

[KR99] K. Kunisch and A. Rösch. Primal-dual strategy for constrained parabolic optimal control problems. *SIAM J. Optimization*, 13:321-224, 2002.

[LM72] J. L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications. Volume 1.* Springer, 1972.

[LPR96] Z. Q. Luo, J. S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints.* Cambridge University Press, Cambridge, 1996.

[MZ79] H. Maurer and J. Zowe. First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Math. Programming*, 16:98-110, 1979.

[NW99] J. Nocedal and S. J. Wright. *Numerical Optimization.* Springer Series in Operation Research. Springer-Verlag, New York, 1999.

[RS80] M. Reed and B. Simon. *Methods of Modern Mathematical PhysicsI: Functional Analysis.* Academic Press, New York, 1980.

[Rey02] J. C. de los Reyes. *Constrained optimal control of stationary viscous incompressible fluids by primal-dual active set methods.* PhD thesis, Institut für Mathematik, Karl-Franzens-Universität Graz, 2003.

[Rob76] S. M. Robinson. Stability theorems for systems of inequalities, Part II: differentiable nonlinear systems. *SIAM J. Numer. Anal.*, 13:497-513, 1976.

[Roc74] R. T. Rockafellar. Conjugate duality and optimization, 1974. Regional conference Series in Applied Mathematics.

[Sta03] G. Stadler. Semi-smooth Newton Methods and Augmented Lagrangian Methods for a Simplified Friction Problem. Technical Report No. 277, Special Research Center F 003 *Optimization and Control*, Project area *Continuous Optimization and Control*, University of Graz & Technical University of Graz, June 2003, submitted.

[Tan96] H. Tanabe. *Functional Analytic Methods for Partial Differential Equations.* Dekkar, New York, 1996.

[Tem79] R. Temam. *Navier-Stokes Equations.* North-Holland, Amsterdam, 1979.

[TV01] F. Tröltzsch and S. Volkwein. The SQP method for bilaterally control constrained optimal control of the Burgers equation. *ESAIM, Control Optim. Calc. Var.*, 6:649-674, 2001.

[Wal86] W. Walter. *Gewöhnliche Differentialgleichungen. Eine Einführung.* Springer-Verlag, Berlin, 1986.

[Ulb03] M. Ulbrich. Semismooth Newton methods for operator equations in function spaces. *SIAM J. Optimization*, 13:805-842, 2003.

[Vol01] S. Volkwein. Second-order conditions for boundary control problems of the Burgers equation. *Control and Cybernetics*, 30:249-278, 2001.

[Vol03] S. Volkwein. Optimal control of laser surface hardening by utilizing a nonlinear primal-dual active set strategy. Technical Report No. 277, Special Research Center F 003 *Optimization and Control*, Project area *Continuous Optimization and Control*, University of Graz & Technical University of Graz, June 2003, submitted.

R. Griesse and S. Volkwein, Karl-Franzens-Universität Graz, Institut für Mathematik, Heinrichstrasse 36, A-8010 Graz, Austria

E-mail address: {roland.griesse,stefan.volkwein}@uni-graz.at