Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

Optimal control of elastic vector-valued Allen–Cahn variational inequalities

Mohammad Hassan Farshbaf-Shaker[1], Claudia Hecht[2]

submitted: October 23, 2013

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: Hassan.Farshbaf-Shaker@wias-berlin.de

[2] Fakultät für Mathematik
Universität Regensburg
93040 Regensburg
Germany
E-Mail: Claudia.Hecht@mathematik.uni-regensburg.de

No. 1858
Berlin 2013

2010 *Mathematics Subject Classification.* 35K86, 49K20, 49K21, 49J20, 35R35.

Key words and phrases. Vector-valued Allen–Cahn system, parabolic obstacle problems, linear elasticity, MPCCs, mathematical programs with complementarity constraints, optimality conditions.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

In this paper we consider a elastic vector-valued Allen–Cahn MPCC (Mathematical Programs with Complementarity Constraints) problem. We use a regularization approach to get the optimality system for the subproblems. By passing to the limit in the optimality conditions for the regularized subproblems, we derive certain generalized first-order necessary optimality conditions for the original problem.

1 Introduction

Optimization problems with interfaces and free boundaries, see [7], frequently appear in materials science, fluid dynamics and biology, for example phase separation in alloys, epitaxial growth, dynamics of multiphase fluids, evolution of cell membranes and in industrial processes such as crystal growth. The mathematical modelling of these phenomena often yields variational problems and highly nonlinear partial differential equations or inclusions. The governing equations for the dynamics of the interfaces in many of these applications involve surface tension expressed in terms of the mean curvature and a driving force. Often in applications of these mathematical models, suitable performance indices and appropriate control actions have to be specified. Mathematically this leads to optimization problems with partial differential equation constraints including free boundaries. The analysis of these problems including optimization of variational inequalities and geometric PDEs is a notoriously difficult task. Surveys and articles concerning the mathematical and numerical approaches to optimal control of free boundary problems may be found in [6, 10]. In this paper we use a phase field approximation for the dynamics of an interface optimization problem. More precisely we consider a multi-component Allen–Cahn model which additionally takes elastic effects into account. Phase field methods provide a natural method for dealing with the complex topological changes that occur. The interface between the phases is replaced by a thin transitional layer of width $\mathcal{O}(\varepsilon)$ where ε is a small parameter, and the different phases are described by the phase field variable. The underlying non-convex elastic interfacial energy is based on the well-known elastic Ginzburg-Landau energy, see [12, 13],

$$E(\mathbf{c},\mathbf{u}) := \int_\Omega \left\{ \frac{\varepsilon}{2}|\nabla \mathbf{c}|^2 + \frac{1}{\varepsilon}\Psi(\mathbf{c}) + W(\mathbf{c},\mathcal{E}(\boldsymbol{u})) \right\} \mathrm{d}x, \qquad \varepsilon > 0 \tag{1.1}$$

where $\Omega \subset \mathbb{R}^d$, $1 \le d \le 3$ is a bounded domain, $\boldsymbol{c} : (0,T) \times \Omega \to \mathbb{R}^N$ is the phase field vector (in our setting the state variable), $\boldsymbol{u} : (0,T)\times\Omega \to \mathbb{R}^d$ is the displacement field and Ψ is the bulk potential. Hence, d denotes the dimension of our working domain Ω and N stands for the number of materials. Since each component of $\boldsymbol{c} := (c_1,\dots,c_N)^T$ stands for the fraction of one phase, the phase space for the order parameter $\boldsymbol{c}$ is the Gibbs simplex

$$\boldsymbol{G} := \{\boldsymbol{v} \in \mathbb{R}^N : \boldsymbol{v} \ge \boldsymbol{0}, \boldsymbol{v}\cdot\boldsymbol{1} = 1\}. \tag{1.2}$$

Note that we use the notation $\boldsymbol{v} \ge \boldsymbol{0}$ for $v_i \ge 0$ for all $i \in \{1,\dots,N\}$, $\boldsymbol{1} = (1,\dots,1)^T$. For the bulk potential $\Psi : \mathbb{R}^N \to \mathbb{R}\cup\{\infty\}$ we consider the multi obstacle potential

$$\Psi(\boldsymbol{v}) := \Psi_0(\boldsymbol{v}) + I_{\boldsymbol{G}}(\boldsymbol{v}) = \begin{cases} \Psi_0(\boldsymbol{v}) := -\frac{1}{2}\|\boldsymbol{v}\|^2 & \text{for } \boldsymbol{v} \in \boldsymbol{G}, \\ \infty & \text{otherwise,} \end{cases} \tag{1.3}$$

where $I_{\boldsymbol{G}}$ is the indicator function of the Gibbs simplex. The last term in (1.1) is the elastic free energy density $W(\mathbf{c}, \mathcal{E})$. Since in phase separation processes of alloys the deformations are typically small we choose a theory based on the linearized strain tensor which is given by $\mathcal{E} := \mathcal{E}(\boldsymbol{u})$, where $\mathcal{E}(\boldsymbol{u}) = \frac{1}{2}(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^T)$ is the symmetric part of $\nabla\boldsymbol{u}$. Moreover, the linear theory leads to a quadratic form of the elastic free energy, namely

$$W(\boldsymbol{c}, \mathcal{E}) = \frac{1}{2}(\mathcal{E} - \mathcal{E}^*(\boldsymbol{c})) : \mathcal{C}(\mathcal{E} - \mathcal{E}^*(\boldsymbol{c})). \tag{1.4}$$

Here $\mathcal{C}$ is the symmetric and positive definite elasticity tensor mapping from symmetric tensors in $\mathbb{R}^{d\times d}$ into itself. Let us note explicitly that we do not assume that $\mathcal{C}$ is isotropic. This takes into account that in applications $\mathcal{C}$ in general will be an anisotropic tensor. The quantity $\mathcal{E}^*(\boldsymbol{c})$ is the eigenstrain at concentration $\boldsymbol{c}$ and we assume (Vegard's law)

$$\mathcal{E}^*(\boldsymbol{c}) = \sum_{i=1}^{N} c_i \mathcal{E}^*(\boldsymbol{e}_i), \tag{1.5}$$

where $\mathcal{E}^*(\boldsymbol{e}_i)$ is the value of the strain tensor if the material consists only of component i and is unstressed. Here $(\boldsymbol{e}_i)_{i=1}^N$ denote the standard coordinate vectors in $\mathbb{R}^N$.

Since we are interested in phase kinetics, the interface motion can be modelled by the steepest descent of (1.1) under the constraint (1.2) with respect to the L^2-norm; for details we refer the reader to [4, 11, 12]. The mechanical equilibrium is obtained on a much faster time scale and therefore we assume quasi-static equilibrium for the mechanical variable $\boldsymbol{u}$. This results, after suitable rescaling of time, in the following elastic Allen–Cahn equation

$$\begin{pmatrix} \varepsilon\partial_t \boldsymbol{c} \\ \mathbf{0} \end{pmatrix} = -\mathrm{grad}_{L^2} E(\boldsymbol{c}, \boldsymbol{u}) = \begin{pmatrix} \varepsilon\Delta\boldsymbol{c} + \frac{1}{\varepsilon}(\boldsymbol{c} - \boldsymbol{\xi}) - D_{\boldsymbol{c}} W(\boldsymbol{c}, \mathcal{E}(\boldsymbol{u})) \\ -\nabla\cdot D_{\mathcal{E}} W(\boldsymbol{c}, \mathcal{E}(\boldsymbol{u})) \end{pmatrix}, \tag{1.6}$$

where $\boldsymbol{\xi} \in \partial I_{\boldsymbol{G}}$ and $\partial I_{\boldsymbol{G}}$ denotes the subdifferential of $I_{\boldsymbol{G}}$. Moreover, D_c and $D_{\mathcal{E}}$ denote the differential with respect to $\boldsymbol{c}$ and $\mathcal{E}$, respectively. We have

$$D_c W(\boldsymbol{c}, \mathcal{E}) = -\mathcal{E}^* : \mathcal{C}(\mathcal{E} - \mathcal{E}^*(\boldsymbol{c})) \text{ and } D_{\mathcal{E}} W(\boldsymbol{c}, \mathcal{E}) = \mathcal{C}(\mathcal{E} - \mathcal{E}^*(\boldsymbol{c})). \tag{1.7}$$

Note, that the first component in (1.6) is in fact an inclusion and hence we later will rewrite this in its complementarity formulation.

1.1 Notations and general assumptions

For simplicity we set $\varepsilon = 1$ in the remainder of this paper. In the sequel "generic" positive constants are denoted by $C_i,\ i \in \mathbb{N}$.
Moreover we define $\mathbb{R}^N_+ := \{\boldsymbol{v} \in \mathbb{R}^N \mid \boldsymbol{v} \geq \mathbf{0}\}$ and introduce the affine hyperplane

$$\boldsymbol{\Sigma} := \{\boldsymbol{v} \in \mathbb{R}^N \mid \boldsymbol{v}\cdot\mathbf{1} = 1\},$$

which is indeed a convex subset of $\mathbb{R}^N$. Its tangential space

$$\boldsymbol{T\Sigma} := \{\boldsymbol{v} \in \mathbb{R}^N \mid \boldsymbol{v}\cdot\mathbf{1} = 0\}$$

is a linear subspace of $\mathbb{R}^N$. With these definitions we obtain for the Gibbs simplex $\boldsymbol{G} = \mathbb{R}^N_+ \cap \boldsymbol{\Sigma}$.

We denote by $L^p(\Omega), W^{k,p}(\Omega)$ for $1 \le p \le \infty$ and $k \in \mathbb{N}$ the Lebesgue- and Sobolev spaces of functions on Ω with the usual norms $\|\cdot\|_{L^p(\Omega)}, \|\cdot\|_{W^{k,p}(\Omega)}$, and we write $H^k(\Omega) = W^{k,2}(\Omega)$.
For a Banach space X we denote its dual by X^*, the dual pairing between $f \in X^*$, $g \in X$ will be denoted by $\langle f, g\rangle_{X^*,X}$. If X is a Banach space with norm $\|\cdot\|_X$, we denote for $T > 0$ by $L^p(0,T;X)$ $(1 \le p \le \infty)$ the Banach space of all (equivalence classes of) Bochner measurable functions $u : (0,T) \longrightarrow X$ such that $\|u(\cdot)\|_X \in L^p(0,T)$. Similarly, we define the space $H^1(0,T;X)$ as the space of functions $u \in L^2(0,T;X)$ whose distributional time derivative is an element in $L^2(0,T;X)$.
We set $\Omega_T := (0,T) \times \Omega, \Gamma_T := (0,T) \times \Gamma$.
Furthermore we denote vector-valued function spaces by boldface letters, $\boldsymbol{L}^2(\Omega) := L^2(\Omega, \mathbb{R}^N) \simeq L^2(\Omega,\mathbb{R})^N$. Moreover we define $\boldsymbol{L}^2_+(\Omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \mid \boldsymbol{v} \in \mathbb{R}^N_+$ a.e. in $\Omega\}$ which is a convex cone in $\boldsymbol{L}^2(\Omega)$; $\boldsymbol{L}^2_{\boldsymbol{\Sigma}}(\Omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \mid \boldsymbol{v} \in \boldsymbol{\Sigma}$ a.e. in $\Omega\}$ which is a convex subset of $\boldsymbol{L}^2(\Omega)$ and $\boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \mid \boldsymbol{v} \in \boldsymbol{T\Sigma}$ a.e. in $\Omega\}$ which is a closed subspace of $\boldsymbol{L}^2(\Omega)$ and hence also a Hilbert space. Furthermore we have $\boldsymbol{L}^2_{\boldsymbol{G}}(\Omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \mid \boldsymbol{v} \in \boldsymbol{G}$ a.e. in $\Omega\}$ and $\boldsymbol{H}^1_i(\Omega) = \boldsymbol{H}^1(\Omega) \cap \boldsymbol{L}^2_i(\Omega)$ where $i \in \{+, \boldsymbol{\Sigma}, \boldsymbol{T\Sigma}, \boldsymbol{G}\}$. Later we also use the following special time dependent spaces: $\boldsymbol{L}^2(\Omega_T) := L^2(0,T;\boldsymbol{L}^2(\Omega))$,

$$\boldsymbol{\mathcal{V}} := L^\infty(0,T;\boldsymbol{H}^1(\Omega)) \cap H^1(0,T;\boldsymbol{L}^2(\Omega)) \cap L^2(0,T;\boldsymbol{H}^2(\Omega))$$

and $\boldsymbol{\mathcal{W}}(0,T) := L^2(0,T;\boldsymbol{H}^1(\Omega)) \cap H^1(0,T;\boldsymbol{H}^1(\Omega)^*)$. Moreover, we use $\boldsymbol{L}^2_i(\Omega_T) := L^2(0,T;$ $\boldsymbol{L}^2_i(\Omega))$, where $i \in \{+, \boldsymbol{\Sigma}, \boldsymbol{T\Sigma}\}$ and $\boldsymbol{\mathcal{V}}_i := \boldsymbol{\mathcal{V}} \cap \boldsymbol{L}^2_i(\Omega_T)$ where $i \in \{\boldsymbol{\Sigma}, \boldsymbol{T\Sigma}\}$. For vector-valued functions $\boldsymbol{\xi} := (\xi_1, \ldots, \xi_N)^T$ and $\boldsymbol{c} := (c_1, \ldots, c_N)^T$, we define the $\boldsymbol{L}^2$-inner product by

$$(\boldsymbol{\xi}, \boldsymbol{c})_{\boldsymbol{L}^2(\Omega_T)} := \sum_{i=1}^N \int_0^T \int_\Omega \xi_i c_i \,\mathrm{d}x\,\mathrm{d}t \tag{1.8}$$

and for two matrices $A, B \in \mathbb{R}^{d\times d}$ we denote by $A : B := \mathrm{tr}(A^T B)$ the standard scalar product for matrices.
We make the following general assumptions, which are assumed to hold throughout the paper:

(A1) $\Omega \subset \mathbb{R}^d$, $1 \le d \le 3$, is a bounded domain with either convex or $C^{1,1}$-boundary and let $T > 0$. The boundary Γ of Ω is divided into a Dirichlet part Γ_D with positive $(d-1)-$dimensional Hausdorff measure, i.e. $\mathcal{H}^{d-1}(\Gamma_D) > 0$, and a non-homogeneous Neumann part Γ_g.

(A2) elasticity tensor:

- **(A2.1)** $\mathcal{C} = (\mathcal{C}_{ijkl})_{i,j,k,l=1}^d, \mathcal{C}_{ijkl} \in \mathbb{R}$
  $\mathcal{C}_{ijkl} = \mathcal{C}_{jikl} = \mathcal{C}_{klij}$,
- **(A2.2)** $\exists \theta, \vartheta > 0$ such that for all symmetric $A, B \in \mathbb{R}^{d\times d}$
  $|\mathcal{C}A : B| \le \vartheta\,|A|\,|B|, \mathcal{C}A : A \ge \theta\,|A|^2$.

For the physical justification of these assumptions we refer the reader to [11]. Let us introduce the boundary conditions, which will be involved in our state equations:

(BC) boundary conditions:

$$\begin{aligned} \nabla \boldsymbol{c} \cdot \boldsymbol{n} &= \boldsymbol{0} \quad \text{on } \Gamma, \\ D_{\mathcal{E}} W(\boldsymbol{c}, \mathcal{E}(\boldsymbol{u})) \cdot \boldsymbol{n} &= \boldsymbol{g} \quad \text{on } \Gamma_g, \\ \boldsymbol{u} &= \boldsymbol{0} \quad \text{on } \Gamma_D. \end{aligned}$$

The function $\boldsymbol{g} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ will in the sequel act as a control.

To write the elastic terms more conveniently, we introduce for a given tensor $\mathcal{C}$ the following scalar product of two matrix-valued functions $\mathcal{A}$ and $\mathcal{B}$: $\langle \mathcal{A},\mathcal{B}\rangle_{\mathcal{C}} := \int_\Omega \mathcal{A} : \mathcal{C}\mathcal{B}$. Furthermore we introduce the projection operator $\boldsymbol{P}_\Sigma : \mathbb{R}^N \to \boldsymbol{T\Sigma}$ defined by $\boldsymbol{P}_\Sigma \boldsymbol{v} := \boldsymbol{v} - \boldsymbol{1} \sum \boldsymbol{v} := \boldsymbol{v} - \boldsymbol{1}\frac{1}{N}\sum_{i=1}^N v_i$. Besides we use the function space

$$H^1_D(\Omega,\mathbb{R}^d) := \{\boldsymbol{u} \in H^1(\Omega,\mathbb{R}^d) \mid \boldsymbol{u}|_{\Gamma_D} = \boldsymbol{0}\}.$$

1.2 Allen–Cahn MPEC problem

Now we introduce our overall optimization problem. Our aim is to transform an initial phase distribution $\boldsymbol{c}_0 : \Omega \to \mathbb{R}^N$ with minimal cost of control, which is given by the applied surface load $\boldsymbol{g}$, to some desired phase pattern $\boldsymbol{c}_T : \Omega \to \mathbb{R}^N$ at a given final time $T > 0$ with $\boldsymbol{c}_T \in \boldsymbol{L}^2(\Omega)$. Besides we track a desired evolution $\boldsymbol{c}_d \in \boldsymbol{L}^2(\Omega_T)$ by choosing $\nu_d > 0$, where $\nu_d, \nu_T \geq 0$ and $\nu_g > 0$ are given constants. Then, defining the tracking type functional

$$\begin{aligned} J(\boldsymbol{c},\boldsymbol{g}) :=& \frac{\nu_T}{2}\|\boldsymbol{c}(T,\cdot) - \boldsymbol{c}_T\|^2_{\boldsymbol{L}^2(\Omega)} + \frac{\nu_d}{2}\|\boldsymbol{c} - \boldsymbol{c}_d\|^2_{\boldsymbol{L}^2(\Omega_T)} + \\ &+ \frac{\nu_g}{2}\|\boldsymbol{g}\|^2_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))} \end{aligned} \tag{1.9}$$

as well as the vector-valued elastic Allen–Cahn variational inequality in its complementarity formulation (**CC**):

For given $(\boldsymbol{c}_0,\boldsymbol{g}) \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega) \times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ find $(\boldsymbol{c},\boldsymbol{u},\boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}_\Sigma} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$ such that $\boldsymbol{c}(0,\cdot) = \boldsymbol{c}_0(\cdot)$ a.e. in Ω and

$$\begin{aligned} &\int_0^T \int_\Omega \partial_t \boldsymbol{c} \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_\Omega \nabla \boldsymbol{c} \cdot \nabla \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t + \\ &\quad - \int_0^T \int_\Omega (\boldsymbol{c} + \boldsymbol{\xi} - D_{\boldsymbol{c}} W(\boldsymbol{c}, \mathcal{E}(\boldsymbol{u}))) \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t = 0, \end{aligned} \tag{1.10}$$

$$\int_0^T \langle \mathcal{E}(\boldsymbol{u}) - \mathcal{E}^*(\boldsymbol{c}), \mathcal{E}(\boldsymbol{\eta})\rangle_{\mathcal{C}} \,\mathrm{d}t = \int_0^T \int_{\Gamma_g} \boldsymbol{g} \cdot \boldsymbol{\eta} \,\mathrm{d}s\,\mathrm{d}t, \tag{1.11}$$

which has to hold for all $\boldsymbol{\chi} \in L^2(0,T;\boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega))$ and $\boldsymbol{\eta} \in L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ and we have the complementarity conditions

$$\boldsymbol{c} \geq \boldsymbol{0} \text{ a.e. in } \Omega_T, \tag{1.12}$$

$$\boldsymbol{\xi} \geq \boldsymbol{0} \text{ a.e. in } \Omega_T, \tag{1.13}$$

$$(\boldsymbol{\xi}, \boldsymbol{c})_{\boldsymbol{L}^2(\Omega_T)} = 0, \tag{1.14}$$

our overall optimization problem reads as follows:

$$(\mathcal{P}_0) \quad \begin{cases} \min & J(\boldsymbol{c},\boldsymbol{g}) \\ \text{over} & (\boldsymbol{c},\boldsymbol{g}) \in \boldsymbol{\mathcal{V}_\Sigma} \times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d)) \\ \text{s.t.} & (\textbf{CC}) \text{ holds.} \end{cases} \tag{1.15}$$

The system (1.10)-(1.14) is an elastic vector-valued Allen–Cahn variational inequality problem in its complementarity formulation. As we will see in Section 2 this problem admits for fixed initial distribution $\boldsymbol{c}_0 \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)$ and given surface load $\boldsymbol{g} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ a unique solution

$$(\boldsymbol{c},\boldsymbol{u},\boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}_\Sigma} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T).$$

Hence, the solution operator

$$\boldsymbol{S}_0 : L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d)) \to \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$$

with its components $\boldsymbol{S}_0(\boldsymbol{g}) := (\boldsymbol{S}_{0|1}(\boldsymbol{g}), \boldsymbol{S}_{0|2}(\boldsymbol{g}), \boldsymbol{S}_{0|3}(\boldsymbol{g}))$ is well-defined, and the control problem $(\mathcal{P}_0)$ is equivalent to minimizing the reduced cost functional $j_0(\boldsymbol{g}) := J(\boldsymbol{S}_{0|1}(\boldsymbol{g}), \boldsymbol{g})$ over $L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$. Given a desired target material distribution $\boldsymbol{c}_T$ at final time, the optimization problem $(\mathcal{P}_0)$ should find the optimal material distribution with minimal cost such that the final time error compared to the target distribution and the mean time error compared to a given track distribution is minimal.

The optimization problem $(\mathcal{P}_0)$ belongs to the problem class of so-called MPECs (Mathematical Programs with Equilibrium Constraints) and in particular to the MPCCs (Mathematical Programs with Complementarity Constraints). It is a well-known fact that the variational inequality condition or in the MPCC case the complementarity conditions occurring as constraints in the minimization problem violates all the known classical NLP (nonlinear programming) constraint qualifications. Hence, the existence of Lagrange multipliers cannot be inferred from standard theory, and the derivation of first-order necessary conditions becomes very difficult, as the treatments in [9, 15, 16, 17, 18] show (note that [18] deals with the more difficult case of the Cahn–Hillard equation). The difference of this present paper with [9] is: In [9] the scalar Allen–Cahn variational inequality with distributed control was considered. Here, we not only have a boundary control but also treat the multi-component, e.g. vectorial, case, which additionally couples with an elastic system. This clearly makes the analysis more difficult.

Now following [5], we replace the indicator function in (1.3) by a convex function $\psi^{\ominus}_{\sigma} \in C^2(\mathbb{R})$, $\sigma \in (0, \frac{1}{4})$, given by

$$\psi^{\ominus}_{\sigma}(r) := \begin{cases} 0 & \text{for} \quad r \geq 0, \\ -\frac{1}{6\sigma^2} r^3 & \text{for} \quad -\sigma < r < 0, \\ \frac{1}{2\sigma}\left(r + \frac{\sigma}{2}\right)^2 + \frac{\sigma}{24} & \text{for} \quad r \leq -\sigma, \end{cases}$$

and define the regularized potential function by

$$\Psi_\sigma(\boldsymbol{c}) = \Psi_0(\boldsymbol{c}) + \hat{\Psi}^{\ominus}(\boldsymbol{c}), \quad \hat{\Psi}^{\ominus}(\boldsymbol{c}) = \sum_{i=1}^{N} \psi^{\ominus}_{\sigma}(c_i). \tag{1.16}$$

For the resulting optimal control problem (later to be denoted by $(\mathcal{P}_\sigma)$) we then derive for any $\sigma \in (0, \frac{1}{4})$ first-order necessary optimality conditions using techniques presented in [21]. Proving a priori estimates (uniform in $\sigma \in (0, \frac{1}{4})$), and employing compactness and monotonicity arguments, we will be able to show the following existence and approximation result: whenever $\{\boldsymbol{g}_\sigma\} \subset L^2(0,T;$ $L^2(\Gamma_g,\mathbb{R}^d))$ is a family of optimal controls for $(\mathcal{P}_\sigma)$ then there exists a subsequence $\{\sigma_n\}$, where $\sigma_n \searrow 0$ as $n \to \infty$, and an optimal control $\bar{\boldsymbol{g}} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ of $(\mathcal{P}_0)$ such that

$$\boldsymbol{g}_{\sigma_n} \to \bar{\boldsymbol{g}} \quad \text{weakly} \quad \text{in} \quad L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d)).$$

In other words, optimal controls for $(\mathcal{P}_\sigma)$ are for small $\sigma > 0$ likely to be "close" to optimal controls for $(\mathcal{P}_0)$. It is natural to ask if the reverse holds, i. e., whether every optimal control for $(\mathcal{P}_0)$ can be approximated by a sequence $\{\boldsymbol{g}_{\sigma_n}\}$ of optimal controls for $(\mathcal{P}_{\sigma_n})$ for some sequence $\sigma_n \searrow 0$.

Unfortunately, we will not be able to prove such a "global" result that applies to all optimal controls for $(\mathcal{P}_0)$. However, a "local" result can be established. To this end, let $\bar{\boldsymbol{g}}$ be any optimal control for $(\mathcal{P}_0)$. We introduce the *"adapted cost functional"*

$$\widetilde{J}(\boldsymbol{c},\boldsymbol{g}) = J(\boldsymbol{c},\boldsymbol{g}) + \frac{1}{2}\|\boldsymbol{g} - \bar{\boldsymbol{g}}\|^2_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))} \tag{1.17}$$

and consider for every $\sigma \in (0, \frac{1}{4})$ the *"adapted control problem"* of minimizing $\widetilde{J}$ over $L^2(0,T; L^2(\Gamma_g, \mathbb{R}^d))$ subject to the constraint that $(\boldsymbol{c}, \boldsymbol{u}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T; H^1_D(\Omega, \mathbb{R}^d))$ solves the system approximating (1.10)–(1.14). It will then turn out that the following is true:

(i) There are some sequence $\sigma_n \searrow 0$ and minimizers $\bar{\boldsymbol{g}}_{\sigma_n} \in L^2(0,T; L^2(\Gamma_g, \mathbb{R}^d))$ of the adapted control problem associated with σ_n, $n \in \mathbb{N}$, such that as $n \to \infty$

$$\bar{\boldsymbol{g}}_{\sigma_n} \quad \to \bar{\boldsymbol{g}} \quad \text{strongly} \quad \text{in} \quad L^2(0,T; L^2(\Gamma_g, \mathbb{R}^d)).$$

(ii) It is possible to pass to the limit as $\sigma \searrow 0$ in the first-order necessary optimality conditions corresponding to the adapted control problems associated with $\sigma \in (0, \frac{1}{4})$ in order to derive first-order necessary optimality conditions for problem $(\mathcal{P}_0)$.

The paper is organized as follows: in Section 2, we derive some results concerning the state system (1.10)–(1.14) and its σ-approximation which is obtained if in (1.3) the indicator function is approximated as in (1.16). In Section 3, we then prove the existence of optimal controls and the approximation result formulated above in (i). The final Section 4 is devoted to the derivation of the first-order necessary optimality conditions, where the strategy outlined in (ii) is employed.

2 Analysis of the vector-valued elastic Allen–Cahn variational inequality

In this section we prove the existence and uniqueness of the solution to the state system (1.10)–(1.14) using its σ-approximation which is obtained if in (1.3) the indicator function is replaced by terms penalizing deviations of $\boldsymbol{c}$ from $\boldsymbol{c} \geq \boldsymbol{0}$, see (1.16).

Theorem 1 *There exists a unique solution to (**CC**).*

The proof of Theorem 1 is established using the following two lemmata. To make notations simpler, we define the function $\frac{1}{\sigma}\hat{\Phi}(r) = \frac{\partial}{\partial r}\psi^{\ominus}_{\sigma}(r)$ for all $r \in \mathbb{R}$ and note that $D\hat{\Psi}^{\ominus}(\boldsymbol{c}) = \frac{1}{\sigma}\hat{\boldsymbol{\Phi}}(\boldsymbol{c}) = \frac{1}{\sigma}\{\hat{\Phi}(c_i)\}_{i=1}^N$. Moreover, we use $D\Psi_\sigma(\boldsymbol{c}_\sigma) = \frac{1}{\sigma}\hat{\boldsymbol{\Phi}}(\boldsymbol{c}_\sigma) - \boldsymbol{c}_\sigma$ and define $\boldsymbol{\xi}_\sigma := -\frac{1}{\sigma}\hat{\boldsymbol{\Phi}}(\boldsymbol{c}_\sigma)$.

The following lemma introduces the regularized elastic vector-valued Allen–Cahn equation $(\mathbf{CC}_\sigma)$. It can be proven using similar techniques used in the papers [11, 20]. We therefore skip here the proof, and for details we refer the interested reader to [14].

Lemma 1 (CC$_\sigma$) *Let $\sigma \in (0, \frac{1}{4})$ be given. For any $(\boldsymbol{c}_0, \boldsymbol{g}) \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega) \times L^2(0,T; L^2(\Gamma_g, \mathbb{R}^d))$ there exist unique functions*

$$(\boldsymbol{c}_\sigma, \boldsymbol{u}_\sigma) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T; H^1_D(\Omega, \mathbb{R}^d))$$

such that $\boldsymbol{c}_\sigma(0,\cdot) = \boldsymbol{c}_0(\cdot)$ a.e. in Ω and

$$\int_0^T \int_\Omega \partial_t \boldsymbol{c}_\sigma \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_\Omega \nabla \boldsymbol{c}_\sigma \cdot \nabla \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t + \\ - \int_0^T \int_\Omega (\boldsymbol{P}_{\boldsymbol{\Sigma}}(\boldsymbol{c}_\sigma + \boldsymbol{\xi}_\sigma - D_{\boldsymbol{c}} W(\boldsymbol{c}_\sigma, \mathcal{E}(\boldsymbol{u}_\sigma)))) \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t = 0, \tag{2.1}$$

$$\int_0^T \langle \mathcal{E}(\boldsymbol{u}_\sigma) - \mathcal{E}^*(\boldsymbol{c}_\sigma), \mathcal{E}(\boldsymbol{\eta}) \rangle_{\mathcal{C}} \,\mathrm{d}t = \int_0^T \int_{\Gamma_g} \boldsymbol{g} \cdot \boldsymbol{\eta} \,\mathrm{d}s\,\mathrm{d}t, \tag{2.2}$$

which has to hold for all $\boldsymbol{\chi} \in L^2(0,T; \boldsymbol{H}^1(\Omega))$ and $\boldsymbol{\eta} \in L^2(0,T; H^1_D(\Omega, \mathbb{R}^d))$.

Remark 1 *It follows from Lemma 1, in particular, that the control-to-state operator* $\boldsymbol{S}_\sigma : L^2(0,T; L^2(\Gamma_g,\mathbb{R}^d)) \to \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$ *given by*

$$\boldsymbol{g} \mapsto \boldsymbol{S}_\sigma(\boldsymbol{g}) := (\boldsymbol{S}_{\sigma|1}(\boldsymbol{g}), \boldsymbol{S}_{\sigma|2}(\boldsymbol{g}), \boldsymbol{S}_{\sigma|3}(\boldsymbol{g})) := (\boldsymbol{c}_\sigma, \boldsymbol{u}_\sigma, \boldsymbol{\xi}_\sigma)$$

is well-defined.

The next step is to prove a priori estimates uniformly in $\sigma \in (0, \frac{1}{4})$ for the solution $(\boldsymbol{c}_\sigma, \boldsymbol{u}_\sigma, \boldsymbol{\xi}_\sigma) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$ to (2.1)–(2.2). We have the following result:

Lemma 2 *There exists a positive constant* K_1 *independent of* $\sigma \in (0,\frac{1}{4})$ *such that we have: whenever* $(\boldsymbol{c}_\sigma, \boldsymbol{u}_\sigma, \boldsymbol{\xi}_\sigma) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$ *is the solution to (2.1)–(2.2) for some* $\boldsymbol{g} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ *and some* $\sigma \in (0,\frac{1}{4})$*, then it holds:*

$$\|\boldsymbol{c}_\sigma\|_{\boldsymbol{\mathcal{V}}} + \|\boldsymbol{u}_\sigma\|_{L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))} + \|\boldsymbol{\xi}_\sigma\|_{\boldsymbol{L}^2(\Omega_T)} \le K_1(1 + \|\boldsymbol{g}\|_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))}). \tag{2.3}$$

Proof. Suppose that $\sigma \in (0,\frac{1}{4})$ and $\boldsymbol{g} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ are arbitrarily chosen, and let $(\boldsymbol{c}_\sigma, \boldsymbol{u}_\sigma, \boldsymbol{\xi}_\sigma) = \boldsymbol{S}_\sigma(\boldsymbol{g})$. The result will be established in a series of a priori estimates. To this end, we will in the following denote by C_i, $i \in \mathbb{N}$, positive constants which do not depend on σ:

<u>*First a priori estimate:*</u>
Applying in (2.2) the testfunction $\boldsymbol{\eta} := \boldsymbol{u}_\sigma \in L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ and using **(A1.2)** we get

$$\theta\|\mathcal{E}(\boldsymbol{u}_\sigma)\|^2_{\boldsymbol{L}^2(\Omega_T)} \le \int_0^T \langle \mathcal{E}^*(\boldsymbol{c}_\sigma), \mathcal{E}(\boldsymbol{u}_\sigma)\rangle_{\mathcal{C}}\,\mathrm{d}t + \int_0^T \int_{\Gamma_g} \boldsymbol{g}\cdot\boldsymbol{u}_\sigma\,\mathrm{d}s\,\mathrm{d}t.$$

Using the inequalities of Korn and Young, the trace theorem and (1.5) we obtain

$$\|\boldsymbol{u}_\sigma\|_{L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))} \le C_1(\|\boldsymbol{c}_\sigma\|_{\boldsymbol{L}^2(\Omega_T)} + \|\boldsymbol{g}\|_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))}). \tag{2.4}$$

<u>*Second a priori estimate:*</u>
We add $2\boldsymbol{P}_{\boldsymbol{\Sigma}}\boldsymbol{c}_\sigma$ on both sides of (2.1) and test the resulting equation by $\boldsymbol{\chi} := \chi_{(0,\tau)}\partial_t\boldsymbol{c}_\sigma \in L^2(0,T;\boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega))$ for some arbitrary $\tau \in (0,T]$, where $\chi_{(0,\tau)}$ is the characteristic function of the interval $(0,\tau)$, to find the estimate

$$\begin{aligned} &\|\partial_t\boldsymbol{c}_\sigma\|^2_{\boldsymbol{L}^2(\Omega_\tau)} + \frac{1}{2}\|\boldsymbol{c}_\sigma(\tau)\|^2_{\boldsymbol{H}^1(\Omega)} + \frac{1}{\sigma}\int_0^\tau\int_\Omega \hat{\boldsymbol{\Phi}}(\boldsymbol{c}_\sigma)\cdot\partial_t\boldsymbol{c}_\sigma\,\mathrm{d}x\,\mathrm{d}t \\ &\le \frac{1}{2}\|\boldsymbol{c}_0\|^2_{\boldsymbol{H}^1(\Omega)} + \int_0^\tau\int_\Omega |\boldsymbol{c}_\sigma|\,|\partial_t\boldsymbol{c}_\sigma|\,\mathrm{d}x\,\mathrm{d}t + \int_0^\tau\int_\Omega |D_cW(\boldsymbol{c}_\sigma,\mathcal{E}(\boldsymbol{u}_\sigma))|\,|\partial_t\boldsymbol{c}_\sigma|\,\mathrm{d}x\,\mathrm{d}t. \end{aligned}$$

Note that $\frac{1}{\sigma}\int_0^\tau\int_\Omega \hat{\boldsymbol{\Phi}}(\boldsymbol{c}_\sigma)\cdot\partial_t\boldsymbol{c}_\sigma\,\mathrm{d}x\,\mathrm{d}t = \int_\Omega \frac{1}{\sigma}\hat{\Psi}^{\ominus}(\boldsymbol{c}(\tau))\,\mathrm{d}x \ge 0$. Moreover, applying Young's inequality, (1.7), (1.5) and **(A2.2)** we have

$$\int_0^\tau\int_\Omega |D_cW(\boldsymbol{c}_\sigma,\mathcal{E}(\boldsymbol{u}_\sigma))|^2\,\mathrm{d}x\,\mathrm{d}t \le C_2\|\mathcal{E}(\boldsymbol{u}_\sigma)\|^2_{\boldsymbol{L}^2(\Omega_\tau)} + C_3\|\boldsymbol{c}_\sigma\|^2_{\boldsymbol{L}^2(\Omega_\tau)}.$$

By (2.4) and Gronwall's inequality we end up with

$$\|\partial_t\boldsymbol{c}_\sigma\|_{\boldsymbol{L}^2(\Omega_T)} + \|\boldsymbol{c}_\sigma\|_{L^\infty(0,T;\boldsymbol{H}^1(\Omega))} \le C_4(1 + \|\boldsymbol{g}\|_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))}). \tag{2.5}$$

Third a priori estimate:
We test (2.1) by $\boldsymbol{\chi} := -\chi_{(0,\tau)}\Delta \boldsymbol{c}_\sigma \in L^2(0,T;\boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega))$ and integrate over Ω_T and by parts, using the boundary conditions, to obtain

$$\begin{aligned}\frac{1}{2}\|\nabla \boldsymbol{c}_\sigma(\tau)\|^2_{\boldsymbol{L}^2(\Omega)} + \|\Delta \boldsymbol{c}_\sigma\|^2_{\boldsymbol{L}^2(\Omega_\tau)} - \frac{1}{\sigma}\int_0^\tau\int_\Omega \hat{\boldsymbol{\Phi}}(\boldsymbol{c}_\sigma)\cdot\Delta \boldsymbol{c}_\sigma \,\mathrm{d}x\,\mathrm{d}t \\ = \frac{1}{2}\|\nabla \boldsymbol{c}_0\|^2_{\boldsymbol{L}^2(\Omega)} + \|\nabla \boldsymbol{c}_\sigma\|^2_{\boldsymbol{L}^2(\Omega_\tau)} + \int_0^\tau\int_\Omega |D_c W(\boldsymbol{c}_\sigma,\mathcal{E}(\boldsymbol{u}_\sigma))|\,|\Delta \boldsymbol{c}_\sigma|\,\mathrm{d}x\,\mathrm{d}t\end{aligned}$$

Note that $-\frac{1}{\sigma}\int_0^\tau\int_\Omega \hat{\boldsymbol{\Phi}}(\boldsymbol{c}_\sigma)\cdot\Delta\boldsymbol{c}_\sigma\,\mathrm{d}x\,\mathrm{d}t = \frac{1}{\sigma}\int_0^\tau\int_\Omega D_c\hat{\boldsymbol{\Phi}}(\boldsymbol{c}_\sigma)\nabla\boldsymbol{c}_\sigma\cdot\nabla\boldsymbol{c}_\sigma\,\mathrm{d}x\,\mathrm{d}t \geq 0$. Now from (2.4), (2.5) and Young's inequality we infer

$$\|\Delta \boldsymbol{c}_\sigma\|_{\boldsymbol{L}^2(\Omega_T)} \leq C_5(1+\|\boldsymbol{g}\|_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))}).$$

Elliptic regularity theory gives

$$\|\boldsymbol{c}_\sigma\|_{L^2(0,T;\boldsymbol{H}^2(\Omega))} \leq C_6(1+\|\boldsymbol{g}\|_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))}). \tag{2.6}$$

Fourth a priori estimate:
Following the lines of [3] we also get the estimate

$$\frac{1}{\sigma}\|\hat{\boldsymbol{\Phi}}(\boldsymbol{c}_\sigma)\|_{\boldsymbol{L}^2(\Omega_T)} \leq C_7. \tag{2.7}$$

and the assertion of the lemma is finally proved. □

Invoking the results of Lemma 1 and Lemma 2 we can prove the existence and uniqueness of a solution to the elastic vector-valued Allen–Cahn variational inequality (**CC**):

Proof of Theorem 1:
By virtue of Lemma 2 there exists a sequence $\{\sigma_n\} \subset (0,\frac{1}{4})$ with $\sigma_n \searrow 0$ as $n \to \infty$ and limit elements $(\boldsymbol{c},\boldsymbol{u},\boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$, such that, as $n\to\infty$,

$$\begin{array}{lllll}\boldsymbol{c}_{\sigma_n} & \longrightarrow & \boldsymbol{c} & \text{weakly} & \text{in } H^1(0,T;\boldsymbol{L}^2(\Omega))\cap L^2(0,T;\boldsymbol{H}^2(\Omega)),\\ \boldsymbol{u}_{\sigma_n} & \longrightarrow & \boldsymbol{u} & \text{weakly} & \text{in } L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)),\\ \boldsymbol{\xi}_{\sigma_n} & \longrightarrow & \boldsymbol{\xi} & \text{weakly} & \text{in } \boldsymbol{L}^2(\Omega_T).\end{array} \tag{2.8}$$

Passing to the limit in (2.1)-(2.2), written for σ_n, $n\in\mathbb{N}$, and using (2.8) and (1.7) we obtain that $(\boldsymbol{c},\boldsymbol{u},\boldsymbol{\xi})$ solve (1.10)–(1.11). Because the set $\{\boldsymbol{\xi}\in\boldsymbol{L}^2(\Omega_T) : \boldsymbol{\xi}\geq\boldsymbol{0}$ a.e. in $\Omega_T\}$ is weakly closed we obtain $\boldsymbol{\xi}\geq\boldsymbol{0}$ a.e. in Ω_T. The same is true for the subset $\boldsymbol{L}^2_{\boldsymbol{\Sigma}}(\Omega_T)$ and we get $\boldsymbol{c}\in\boldsymbol{L}^2_{\boldsymbol{\Sigma}}(\Omega_T)$. To prove (1.12) we make use of the Lipschitz continuity of $\hat{\boldsymbol{\Phi}}$:

$$\begin{aligned}\|\hat{\boldsymbol{\Phi}}(\boldsymbol{c})\|_{\boldsymbol{L}^2(\Omega_T)} &\leq \|\hat{\boldsymbol{\Phi}}(\boldsymbol{c})-\hat{\boldsymbol{\Phi}}(\boldsymbol{c}_{\sigma_n})\|_{\boldsymbol{L}^2(\Omega_T)} + \|\hat{\boldsymbol{\Phi}}(\boldsymbol{c}_{\sigma_n})\|_{\boldsymbol{L}^2(\Omega_T)}\\ &\leq C_{Lip}\|\boldsymbol{c}-\boldsymbol{c}_{\sigma_n}\|_{\boldsymbol{L}^2(\Omega_T)} + \|\hat{\boldsymbol{\Phi}}(\boldsymbol{c}_{\sigma_n})\|_{\boldsymbol{L}^2(\Omega_T)} \quad \forall n\in\mathbb{N}.\end{aligned}$$

Because of (2.8) and (2.7) we get $\|\hat{\boldsymbol{\Phi}}(\boldsymbol{c})\|_{\boldsymbol{L}^2} = 0$ and thus, $\boldsymbol{c}\geq\boldsymbol{0}$ a.e. in Ω_T. Moreover as $n\to\infty$

$$(\boldsymbol{\xi},\boldsymbol{c})_{\boldsymbol{L}^2(\Omega_T)} \longleftarrow (\boldsymbol{\xi}_{\sigma_n},\boldsymbol{c}_{\sigma_n})_{\boldsymbol{L}^2(\Omega_T)} = -\frac{1}{\sigma_n}(\hat{\boldsymbol{\Phi}}(\boldsymbol{c}_{\sigma_n}),\boldsymbol{c}_{\sigma_n})_{\boldsymbol{L}^2(\Omega_T)} \leq 0, \tag{2.9}$$

and hence $(\boldsymbol{\xi},\boldsymbol{c})_{\boldsymbol{L}^2(\Omega_T)} \leq 0$. However, since $\boldsymbol{\xi}\geq\boldsymbol{0}$ and $\boldsymbol{c}\geq\boldsymbol{0}$ we have that $(\boldsymbol{\xi},\boldsymbol{c})_{\boldsymbol{L}^2(\Omega_T)} = 0$. Therefore, the existence assertion of the theorem is proven. For uniqueness we follow the lines of [3]. This needs no repetition here and the reader is referred to the mentioned paper. □

Remark 2 *It follows from Theorem 1, in particular, that the control-to-state operator* $\boldsymbol{S}_0 : L^2(0,T; L^2(\Gamma_g,\mathbb{R}^d)) \to \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$ *defined by*

$$\boldsymbol{g} \mapsto \boldsymbol{S}_0(\boldsymbol{g}) := (\boldsymbol{S}_{0|1}(\boldsymbol{g}), \boldsymbol{S}_{0|2}(\boldsymbol{g}), \boldsymbol{S}_{0|3}(\boldsymbol{g})) := (\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{\xi}), \tag{2.10}$$

where $(\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{\xi})$ *denotes the solution to (***CC***) associated to* $\boldsymbol{g}$*, is well-defined.*

Remark 3 *By the same arguments as in the proof of Theorem 1 we can conclude that for any sequence* $\{\sigma_n\} \subset (0, \frac{1}{4})$ *with* $\lim_{n\to\infty} \sigma_n = 0$ *it follows:*

(i) Whenever the sequence $\{\boldsymbol{g}_{\sigma_n}\} \subset L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ *converges to* $\boldsymbol{g}$ *weakly in* $L^2(0,T; L^2(\Gamma_g,\mathbb{R}^d))$ *as* $n \to \infty$ *then there is some subsequence, which is again indexed by* n*, such that* $\{\boldsymbol{S}_{\sigma_n|1}(\boldsymbol{g}_{\sigma_n})\}$ *converges to* $\boldsymbol{S}|_{0|1}(\boldsymbol{g})$ *weakly in* $L^2(0,T;\boldsymbol{H}^2(\Omega)) \cap H^1(0,T;\boldsymbol{L}^2(\Omega))$ *as* $n \to \infty$*.*

(ii) Due to the continuous embedding of $L^2(0,T;\boldsymbol{H}^2(\Omega)) \cap H^1(0,T;\boldsymbol{L}^2(\Omega))$ *into* $C([0,T]; \boldsymbol{H}^1(\Omega))$ *and the compact embedding of* $\boldsymbol{H}^1(\Omega)$ *into* $\boldsymbol{L}^2(\Omega)$ *(due to Rellich-Kondrachov) we obtain the strong convergence of the sequence* $\{\boldsymbol{S}_{\sigma_n|1}(\boldsymbol{g}_{\sigma_n})(T)\}$ *in* $\boldsymbol{L}^2(\Omega)$*. Furthermore, Aubin-Lions' lemma provides the strong convergence of* $\{\boldsymbol{S}_{\sigma_n|1}(\boldsymbol{g}_{\sigma_n})\}$ *in* $\boldsymbol{L}^2(\Omega_T)$*.*

(iii) Moreover, we have

$$\lim_{n\to\infty} J(\boldsymbol{S}_{\sigma_n|1}(\boldsymbol{h}), \boldsymbol{h}) = J(\boldsymbol{S}_{0|1}(\boldsymbol{h}), \boldsymbol{h}) \quad \forall \boldsymbol{h} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d)).$$

3 Existence and approximation of optimal controls

Our first aim in this section is to prove the following existence result:

Theorem 2 *The optimal control problem* $(\mathcal{P}_0)$ *admits a solution.*

Before proving Theorem 2, we introduce a family of auxiliary optimal control problems $(\mathcal{P}_\sigma)$ parametrized by $\sigma \in (0, \frac{1}{4})$. We define

$$(\mathcal{P}_\sigma) \qquad \begin{cases} \min & J(\boldsymbol{c}, \boldsymbol{g}), \\ \text{over} & (\boldsymbol{c}, \boldsymbol{g}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d)), \\ \text{s.t.} & (\mathbf{CC}_\sigma) \text{ holds .} \end{cases}$$

The following lemma can be shown by the direct method in the calculus of variations, while making use of Lemma 2:

Lemma 3 *Let* $\sigma \in (0, \frac{1}{4})$ *be given. Then the optimal control problem* $(\mathcal{P}_\sigma)$ *admits a solution.*

Proof of Theorem 2: By virtue of Lemma 3, for any $\sigma \in (0, \frac{1}{4})$, we may pick a solution $(\boldsymbol{c}_\sigma, \boldsymbol{u}_\sigma, \boldsymbol{\xi}_\sigma)$ for the optimal control problem $(\mathcal{P}_\sigma)$. Obviously, we have

$$(\boldsymbol{c}_\sigma, \boldsymbol{u}_\sigma, \boldsymbol{\xi}_\sigma) = \boldsymbol{S}_\sigma(\boldsymbol{g}_\sigma) \quad \forall \sigma \in (0, \frac{1}{4}).$$

For an arbitrary chosen element $\widehat{\boldsymbol{g}} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ we have

$$J(\boldsymbol{c}_\sigma,\boldsymbol{g}_\sigma) \le J(\boldsymbol{S}_\sigma(\widehat{\boldsymbol{g}}),\widehat{\boldsymbol{g}}) \quad \forall \sigma \in (0,\frac{1}{4}).$$

Hence, there exists a subsequence $\{\boldsymbol{g}_{\sigma_n}\}$ with $\sigma_n \searrow 0$ as $n \to \infty$ and a limit element $\boldsymbol{g} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ such that as $n \to \infty$

$$\boldsymbol{g}_{\sigma_n} \longrightarrow \boldsymbol{g} \quad \text{weakly in} \quad L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d)). \tag{3.1}$$

Using arguments as in Theorem 1 we find from Lemma 2 that there exist limit elements $(\boldsymbol{c},\boldsymbol{u},\boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$ such that the convergence properties (2.8) are satisfied and $(\boldsymbol{c},\boldsymbol{u},\boldsymbol{\xi}) = \boldsymbol{S}_0(\boldsymbol{g})$, i.e. the element $((\boldsymbol{c},\boldsymbol{u},\boldsymbol{\xi}),\boldsymbol{g})$ is admissible for $(\mathcal{P}_0)$. It remains to show, that $((\boldsymbol{c},\boldsymbol{u},\boldsymbol{\xi}),\boldsymbol{g})$ is in fact optimal for $(\mathcal{P}_0)$. To this end, let $\widehat{\boldsymbol{g}} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ be arbitrary. Invoking the convergence properties in (2.8) and using the weak sequential lower semicontinuity of the cost functional (1.9), we obtain

$$\begin{aligned} J(\boldsymbol{c},\boldsymbol{g}) &= J(\boldsymbol{S}_{0|1}(\boldsymbol{g}),\boldsymbol{g}) \le \liminf_{n\to\infty} J(\boldsymbol{S}_{\sigma_n|1}(\boldsymbol{g}_{\sigma_n}),\boldsymbol{g}_{\sigma_n}) \\ &\le \liminf_{n\to\infty} J(\boldsymbol{S}_{\sigma_n|1}(\widehat{\boldsymbol{g}}),\widehat{\boldsymbol{g}}) \le \lim_{n\to\infty} J(\boldsymbol{S}_{\sigma_n|1}(\widehat{\boldsymbol{g}}),\widehat{\boldsymbol{g}}) = J(\boldsymbol{S}_{0|1}(\widehat{\boldsymbol{g}}),\widehat{\boldsymbol{g}}), \end{aligned} \tag{3.2}$$

where for the last equality the continuity of the cost functional with respect to the first variable was used, see Remark 3. With this, the assertion is completely proved. □

Theorem 2 does not yield any information on whether every solution to the optimal control problem $(\mathcal{P}_0)$ can be approximated by a sequence of solutions of $(\mathcal{P}_\sigma)$. As already announced in the introduction, we are not able to prove such a general "global" result. Instead, we can only give a "local" answer for every individual optimizer of $(\mathcal{P}_0)$. For this purpose, we employ a trick due to Barbu [2]. To this end, let $((\bar{\boldsymbol{c}},\bar{\boldsymbol{u}},\bar{\boldsymbol{\xi}}),\bar{\boldsymbol{g}}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T) \times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$, where $(\bar{\boldsymbol{c}},\bar{\boldsymbol{u}},\bar{\boldsymbol{\xi}}) = \boldsymbol{S}_0(\bar{\boldsymbol{g}})$, be an arbitrary but fixed solution to $(\mathcal{P}_0)$. We associate with this solution the "*adapted cost functional*"

$$\widetilde{J}(\boldsymbol{c},\boldsymbol{g}) = J(\boldsymbol{c},\boldsymbol{g}) + \frac{1}{2}\|\boldsymbol{g}-\bar{\boldsymbol{g}}\|^2_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))}$$

and a corresponding "*adapted optimal control problem*"

$$(\widetilde{\mathcal{P}}_\sigma) \qquad \begin{cases} \min & \widetilde{J}(\boldsymbol{c},\boldsymbol{g}), \\ \text{over} & (\boldsymbol{c},\boldsymbol{g}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d)), \\ \text{s.t.} & (\mathbf{CC}_\sigma) \text{ holds .} \end{cases}$$

With a proof that resembles that of Lemma 3 and needs no repetition here, we can show the following result:

Lemma 4 *Let $\sigma \in (0,\frac{1}{4})$ be given. Then the optimal control problem $(\widetilde{\mathcal{P}}_\sigma)$ admits a solution.*

We are now in the position to give a partial answer to the question raised above. We have the following result:

Theorem 3 *Suppose that $((\bar{\boldsymbol{c}},\bar{\boldsymbol{u}},\bar{\boldsymbol{\xi}}),\bar{\boldsymbol{g}}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T) \times L^2(0,T;$ $L^2(\Gamma_g,\mathbb{R}^d))$ is any fixed solution to the optimal control problem $(\mathcal{P}_0)$. Then there exists a sequence $\{\sigma_n\} \subset (0,\frac{1}{4})$ with $\sigma_n \searrow 0$ as $n \to \infty$, and for any $n \in \mathbb{N}$, there exists a solution pair*

$((\bar{\boldsymbol{c}}_{\sigma_n}, \bar{\boldsymbol{u}}_{\sigma_n}, \bar{\boldsymbol{\xi}}_{\sigma_n}), \bar{\boldsymbol{g}}_{\sigma_n}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T; H^1_D(\Omega, \mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T) \times L^2(0,T; L^2(\Gamma_g, \mathbb{R}^d))$ *solving the adapted problem* $(\widetilde{\mathcal{P}}_{\sigma_n})$ *and such that, as* $n \to \infty$*,*

$$\begin{array}{lll} \bar{\boldsymbol{g}}_{\sigma_n} & \longrightarrow \bar{\boldsymbol{g}} & \textit{strongly in } L^2(0,T; L^2(\Gamma_g, \mathbb{R}^d)), \\ \bar{\boldsymbol{c}}_{\sigma_n} & \longrightarrow \bar{\boldsymbol{c}} & \textit{weakly in } L^2(0,T; \boldsymbol{H}^2(\Omega)) \cap H^1(0,T; \boldsymbol{L}^2(\Omega)), \\ \bar{\boldsymbol{\xi}}_{\sigma_n} & \longrightarrow \bar{\boldsymbol{\xi}} & \textit{weakly in } \boldsymbol{L}^2(\Omega_T), \\ \bar{\boldsymbol{u}}_{\sigma_n} & \longrightarrow \bar{\boldsymbol{u}} & \textit{weakly in } L^2(0,T; H^1_D(\Omega, \mathbb{R}^d)), \\ \widetilde{J}(\bar{\boldsymbol{c}}_{\sigma_n}, \bar{\boldsymbol{g}}_{\sigma_n}) & \longrightarrow J(\bar{\boldsymbol{c}}, \bar{\boldsymbol{g}}). & \end{array} \tag{3.3}$$

Proof. For every $\sigma \in (0, \frac{1}{4})$, we pick an optimal pair $((\bar{\boldsymbol{c}}_\sigma, \bar{\boldsymbol{u}}_\sigma, \bar{\boldsymbol{\xi}}_\sigma), \bar{\boldsymbol{g}}_\sigma) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T; H^1_D(\Omega, \mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T) \times L^2(0,T; L^2(\Gamma_g, \mathbb{R}^d))$ for the adapted problem $(\widetilde{\mathcal{P}}_\sigma)$. Moreover, for any $\sigma \in (0, \frac{1}{4})$ we have

$$\widetilde{J}(\bar{\boldsymbol{c}}_\sigma, \bar{\boldsymbol{g}}_\sigma) \leq \widetilde{J}(\boldsymbol{S}_{\sigma|1}(\bar{\boldsymbol{g}}), \bar{\boldsymbol{g}}) = J(\boldsymbol{S}_{\sigma|1}(\bar{\boldsymbol{g}}), \bar{\boldsymbol{g}}). \tag{3.4}$$

Now, from Remark 3 we can infer that there exists some subsequence $\{\sigma_n\} \subset (0, \frac{1}{4})$ with $\sigma_n \searrow 0$ as $n \to \infty$ and a $\boldsymbol{g} \in L^2(0,T; L^2(\Gamma_g, \mathbb{R}^d))$ satisfying

$$\bar{\boldsymbol{g}}_{\sigma_n} \longrightarrow \boldsymbol{g} \quad \text{weakly in} L^2(0,T; L^2(\Gamma_g, \mathbb{R}^d)) \quad \text{as } n \to \infty. \tag{3.5}$$

Moreover, owing to Lemma 2, we may without loss of generality assume that there is some limit element $(\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T; H^1_D(\Omega, \mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$ such that the second, third and fourth line of (3.3) are satisfied with $(\bar{\boldsymbol{c}}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{\xi}})$ replaced by $(\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{\xi})$. Following the arguments of the proof of Theorem 1 we can show that actually $(\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{\xi}) = \boldsymbol{S}_0(\boldsymbol{g})$, which implies, in particular, that $((\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{\xi}), \boldsymbol{g})$ is admissible for $(\mathcal{P}_0)$.
We now aim to prove $\boldsymbol{g} = \bar{\boldsymbol{g}}$. Once this will be shown, we can deduce from the unique solvability of the state system (**CC**), see Theorem 1, that also $(\boldsymbol{c}, \boldsymbol{\xi}, \boldsymbol{u}) = (\bar{\boldsymbol{c}}, \bar{\boldsymbol{\xi}}, \bar{\boldsymbol{u}})$.
Indeed, we have, owing to the weakly sequential lower semicontinuity of $\widetilde{J}$, and in view of the optimality property of $((\bar{\boldsymbol{c}}, \bar{\boldsymbol{\xi}}, \bar{\boldsymbol{u}}), \bar{\boldsymbol{g}})$ for problem $(\mathcal{P}_0)$

$$\begin{aligned} \liminf_{n\to\infty} \widetilde{J}(\bar{\boldsymbol{c}}_{\sigma_n}, \bar{\boldsymbol{g}}_{\sigma_n}) &\geq J(\boldsymbol{c}, \boldsymbol{g}) + \frac{1}{2}\|\boldsymbol{g} - \bar{\boldsymbol{g}}\|^2_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))} \\ &\geq J(\bar{\boldsymbol{c}}, \bar{\boldsymbol{g}}) + \frac{1}{2}\|\boldsymbol{g} - \bar{\boldsymbol{g}}\|^2_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))}. \end{aligned} \tag{3.6}$$

On the other hand, taking the limit superior as $n \to \infty$ on both side of (3.4) and invoking Remark 3 we have

$$\limsup_{n\to\infty} \widetilde{J}(\bar{\boldsymbol{c}}_{\sigma_n}, \bar{\boldsymbol{g}}_{\sigma_n}) \leq J(\boldsymbol{S}_{0|1}(\bar{\boldsymbol{g}}), \bar{\boldsymbol{g}}) = J(\bar{\boldsymbol{c}}, \bar{\boldsymbol{g}}). \tag{3.7}$$

Combining (3.6) with (3.7), we have thus shown that $\|\boldsymbol{g} - \bar{\boldsymbol{g}}\|^2_{L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))} = 0$, so that $\boldsymbol{g} = \bar{\boldsymbol{g}}$. Moreover, (3.6) and (3.7) also imply that

$$J(\bar{\boldsymbol{c}}, \bar{\boldsymbol{g}}) = \widetilde{J}(\bar{\boldsymbol{c}}, \bar{\boldsymbol{g}}) = \lim_{n\to\infty} \widetilde{J}(\bar{\boldsymbol{c}}_{\sigma_n}, \bar{\boldsymbol{g}}_{\sigma_n})$$

which proves the last line of (3.3), and, at the same time, also the first line of (3.3). The assertion is thus completely proven. □

4 The optimality system

In this section our aim is to derive first-order necessary optimality conditions for the optimal control problem $(\mathcal{P}_0)$. This will be achieved by deriving first-order necessary optimality conditions for the adapted optimal control problems $(\widetilde{\mathcal{P}}_\sigma)$ and passing to the limit as $\sigma \searrow 0$. We will finally show that in the limit certain generalized first-order necessary conditions hold.

4.1 The linearized system

For the derivation of first-order optimality conditions it is essential to show the Fréchet-differentiability of the control-to-state operator. In view of the occurrence of the indicator function in (1.3), this is impossible for the control-to-state operator $\boldsymbol{S}_0$ of the state system (1.10)–(1.11). It is, however, possible for the control-to-state operators $\boldsymbol{S}_\sigma$ of the approximating systems (2.1)–(2.2), see Section 4.2. In preparation of a corresponding theorem, we now consider for given $\boldsymbol{h} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ the linearized version of (2.1)–(2.2):

$$\int_0^T \int_\Omega \partial_t \dot{\boldsymbol{c}}_\sigma \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_\Omega \nabla \dot{\boldsymbol{c}}_\sigma \cdot \nabla \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t - \int_0^T \int_\Omega \dot{\boldsymbol{c}}_\sigma \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t + \\ - \int_0^T \int_\Omega (\boldsymbol{P_\Sigma}(\mathrm{D}(-\frac{1}{\sigma}\hat{\boldsymbol{\Phi}})(\overline{\boldsymbol{c}}_\sigma)\dot{\boldsymbol{c}}_\sigma - D_c W(\dot{\boldsymbol{c}}_\sigma, \mathcal{E}(\dot{\boldsymbol{u}}_\sigma))) \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t = 0, \tag{4.1}$$

$$\int_0^T \langle \mathcal{E}(\dot{\boldsymbol{u}}_\sigma) - \mathcal{E}^*(\dot{\boldsymbol{c}}_\sigma), \mathcal{E}(\boldsymbol{\eta}) \rangle_{\mathcal{C}} \,\mathrm{d}t = \int_0^T \int_{\Gamma_g} \boldsymbol{h} \cdot \boldsymbol{\eta} \,\mathrm{d}s\,\mathrm{d}t, \tag{4.2}$$

which has to hold for all $\boldsymbol{\chi} \in L^2(0,T;\boldsymbol{H}^1(\Omega))$ and $\boldsymbol{\eta} \in L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ with $\dot{\boldsymbol{c}}_\sigma(0,\cdot) = \boldsymbol{0}$ a.e. in Ω and $\overline{\boldsymbol{c}}_\sigma = \boldsymbol{S}_{\sigma|1}(\overline{\boldsymbol{g}}_\sigma)$.

Existence and uniqueness of a solution $(\dot{\boldsymbol{c}}_\sigma, \dot{\boldsymbol{u}}_\sigma) \in \boldsymbol{\mathcal{V}_{T\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ to the system (4.1)–(4.2) follow by the subsequent Theorem 4.

4.2 Differentiability of the control-to-state operator $\boldsymbol{S}_\sigma$

We have the following differentiability result:

Theorem 4 *Let $\sigma \in (0,\frac{1}{4})$ be given and $\boldsymbol{g} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ be arbitrary. Then the control-to-state mapping $(\boldsymbol{S}_{\sigma|1}, \boldsymbol{S}_{\sigma|2})$, viewed as a mapping from $L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ into $\boldsymbol{\mathcal{V}_\Sigma} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$, is Fréchet-differentiable at $\boldsymbol{g} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$, and the Fréchet derivative is given by*

$$(\mathrm{D}\boldsymbol{S}_{\sigma|1}(\boldsymbol{g})(\boldsymbol{h}), \mathrm{D}\boldsymbol{S}_{\sigma|2}(\boldsymbol{g})(\boldsymbol{h})) = (\dot{\boldsymbol{c}}_\sigma, \dot{\boldsymbol{u}}_\sigma)$$

where for any given $\boldsymbol{h} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ the pair $(\dot{\boldsymbol{c}}_\sigma, \dot{\boldsymbol{u}}_\sigma)$ denotes the solution to the linearized system (4.1)–(4.2).

In preparation for proving the abovementioned theorem we discuss some preparatory lemmata introducing some auxiliary problems:

Lemma 5 *For given $(\boldsymbol{r},\boldsymbol{g},\boldsymbol{c}_0) \in \boldsymbol{L}^2(\Omega_T) \times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d)) \times \boldsymbol{H}^1(\Omega)$ there exists a unique pair $(\boldsymbol{c},\boldsymbol{u}) \in \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ such that $\boldsymbol{c}(\cdot,0) = \boldsymbol{c}_0(\cdot)$ and*

$$\int_0^T \int_\Omega \partial_t \boldsymbol{c} \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_\Omega \nabla \boldsymbol{c} \cdot \nabla \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t - \int_0^T \int_\Omega \boldsymbol{P}_\Sigma(\boldsymbol{c}) \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t + \\ + \int_0^T \int_\Omega \boldsymbol{P}_\Sigma(\mathrm{D}_c W(\boldsymbol{c}, \mathcal{E}(\boldsymbol{u}))) \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int_\Omega \boldsymbol{r} \cdot \boldsymbol{\chi} \,\mathrm{d}x\,\mathrm{d}t, \tag{4.3}$$

$$\int_0^T \langle \mathcal{E}(\boldsymbol{u}) - \mathcal{E}^*(\boldsymbol{c}), \mathcal{E}(\boldsymbol{\eta}) \rangle_{\mathcal{C}} \,\mathrm{d}t = \int_0^T \int_{\Gamma_g} \boldsymbol{g} \cdot \boldsymbol{\eta} \,\mathrm{d}s\,\mathrm{d}t, \tag{4.4}$$

which has to hold for all $\boldsymbol{\chi} \in L^2\left(0,T;\boldsymbol{H}^1(\Omega)\right)$ and $\boldsymbol{\eta} \in L^2\left(0,T;H^1_D\left(\Omega,\mathbb{R}^d\right)\right)$.

Remark 4 *Standard theory for linear parabolic equations, see e.g. [8], provide for any given* $\boldsymbol{u} \in L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ *the existence of a solution* $\boldsymbol{c} \in \boldsymbol{\mathcal{V}}$ *to* (4.3)*. Moreover, Lax-Milgram's theorem gives the existence of a solution* $\boldsymbol{u} \in L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ *to* (4.4) *for any given* $\boldsymbol{c} \in \boldsymbol{\mathcal{V}}$*. Hence, applying Banach's fixed point theorem establishes the existence of a solution* $(\boldsymbol{c},\boldsymbol{u})$ *to* (4.3) $-$ (4.4)*. Uniqueness follows by Korn's inequality and a standard Gronwall argument. For details, see e.g. [14].*

The assertion of Lemma 5 motivates to define the operator

$$\begin{aligned} \mathcal{L}: \boldsymbol{L}^2(\Omega_T)\times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))\times \boldsymbol{H}^1(\Omega) &\;\to\; \boldsymbol{\mathcal{V}}\times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)),\\ (\boldsymbol{r},\boldsymbol{g},\boldsymbol{c}_0) &\;\mapsto\; \mathcal{L}(\boldsymbol{r},\boldsymbol{g},\boldsymbol{c}_0):=(\boldsymbol{c},\boldsymbol{u}), \end{aligned} \tag{4.5}$$

where $(\boldsymbol{c},\boldsymbol{u}) \in \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ is defined as the solution to (4.3)-(4.4). Moreover, we introduce the operators $\mathcal{R}(\cdot):=\mathcal{L}(\cdot,\boldsymbol{0},\boldsymbol{0})$, $\mathcal{G}(\cdot):=\mathcal{L}(\boldsymbol{0},\cdot,\boldsymbol{0})$ and $\mathcal{I}(\cdot):=\mathcal{L}(\boldsymbol{0},\boldsymbol{0},\cdot)$. Using similar a priori estimates as in the proof of Lemma 2 it follows that the mapping $\mathcal{R}$ and $\mathcal{G}$ are continuous, see e.g. [14].

Lemma 6 *Let* $\sigma \in (0,\frac{1}{4})$ *be given and assume that* $\overline{\boldsymbol{c}} \in \boldsymbol{\mathcal{V}}$ *is fixed. Then the operator*

$$\widetilde{\mathcal{R}}: \boldsymbol{\mathcal{V}}\times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \to \boldsymbol{\mathcal{V}}\times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$$

defined by

$$\widetilde{\mathcal{R}}\,(\boldsymbol{c},\boldsymbol{u}) := -\mathcal{R}(\boldsymbol{P}_\Sigma(\mathrm{D}(-\frac{1}{\sigma}\hat{\boldsymbol{\Phi}})(\overline{\boldsymbol{c}})\boldsymbol{c})) + (\boldsymbol{c},\boldsymbol{u})$$

admits a linear and continuous inverse mapping.

Proof. We exploit the bounded inverse theorem. To this end, let $\bar{\boldsymbol{c}} \in \boldsymbol{\mathcal{V}}$ be given. Due to the continuous embedding $\boldsymbol{\mathcal{V}} \subset \boldsymbol{L}^4(\Omega_T)$ it follows from [1] that the Nemytskii-operator $-\frac{1}{\sigma}\hat{\boldsymbol{\Phi}} : \boldsymbol{L}^4(\Omega_T) \to \boldsymbol{L}^2(\Omega_T)$ is Fréchet-differentiable. Using this and Lemma 5 the continuity of $\widetilde{\mathcal{R}}$ is shown. To prove bijectivity we have to establish the existence and uniqueness of an element $(\widetilde{\boldsymbol{c}},\widetilde{\boldsymbol{u}}) \in \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ for given $(\boldsymbol{c},\boldsymbol{u}) \in \boldsymbol{\mathcal{V}}\times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ such that the condition $\widetilde{\mathcal{R}}\,(\widetilde{\boldsymbol{c}},\widetilde{\boldsymbol{u}}) = (\boldsymbol{c},\boldsymbol{u})$ is fulfilled, which is equivalent to

$$(\boldsymbol{c}-\widetilde{\boldsymbol{c}},\boldsymbol{u}-\widetilde{\boldsymbol{u}}) = -\mathcal{R}(\boldsymbol{P}_\Sigma(\mathrm{D}(-\frac{1}{\sigma}\hat{\boldsymbol{\Phi}})(\overline{\boldsymbol{c}})\widetilde{\boldsymbol{c}})). \tag{4.6}$$

Using (4.3) $-$ (4.4) the expression (4.6) reads as

$$\begin{aligned} &\int_0^T\!\!\int_\Omega \partial_t\widetilde{\boldsymbol{c}}\cdot\boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\!\!\int_\Omega \nabla\widetilde{\boldsymbol{c}}\cdot\nabla\boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t - \int_0^T\!\!\int_\Omega \boldsymbol{P}_\Sigma(\widetilde{\boldsymbol{c}})\cdot\boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t + \\ &\qquad - \int_0^T\!\!\int_\Omega (\boldsymbol{P}_\Sigma(\mathrm{D}(-\frac{1}{\sigma}\hat{\boldsymbol{\Phi}})(\bar{\boldsymbol{c}})\widetilde{\boldsymbol{c}} - D_cW(\widetilde{\boldsymbol{c}},\mathcal{E}(\widetilde{\boldsymbol{u}})))) \cdot\boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_0^T\!\!\int_\Omega \partial_t\boldsymbol{c}\cdot\boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\!\!\int_\Omega \nabla\boldsymbol{c}\cdot\nabla\boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t + \\ &\qquad - \int_0^T\!\!\int_\Omega (\boldsymbol{P}_\Sigma(\boldsymbol{c} - D_cW(\boldsymbol{c},\mathcal{E}(\boldsymbol{u})))) \cdot\boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t, \end{aligned} \tag{4.7}$$

$$\int_0^T \langle \mathcal{E}(\widetilde{\boldsymbol{u}}) - \mathcal{E}^*(\widetilde{\boldsymbol{c}}), \mathcal{E}(\boldsymbol{\eta})\rangle_{\mathcal{C}}\;\mathrm{d}t = 0, \tag{4.8}$$

which has to hold for all $\boldsymbol{\chi} \in L^2(0,T;\boldsymbol{H}^1(\Omega))$ and $\boldsymbol{\eta} \in L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ with $\widetilde{\boldsymbol{c}}\,(\cdot,0) = \boldsymbol{c}\,(\cdot,0)$.

The existence and uniqueness of a solution to (4.7) − (4.8) follow by using similar arguments as in Lemma 5 which then provides that $\widetilde{\mathcal{R}}$ is bijective. The statement of the lemma follows then by the bounded inverse theorem. □

Proof of Theorem 4: We will utilize the implicit function theorem to prove Fréchet-differentiability of $\boldsymbol{S}_\sigma$. To this end let us introduce the mapping

$$F_\sigma : \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d)) \to \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$$

defined by

$$F_\sigma(\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma,\boldsymbol{g}) := (\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma) - \mathcal{R}(\boldsymbol{P}_\Sigma((-\frac{1}{\sigma}\hat{\boldsymbol{\Phi}})(\boldsymbol{c}_\sigma))) - \mathcal{G}(\boldsymbol{g}) - \mathcal{I}(\boldsymbol{c}_0).$$

First, Lemma 1 implies that for every $(\boldsymbol{c}_0,\boldsymbol{g}) \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)\times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ there is a $(\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ such that

$$(\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma) = \mathcal{R}(\boldsymbol{P}_\Sigma(-\frac{1}{\sigma}\hat{\boldsymbol{\Phi}}(\boldsymbol{c}_\sigma))) + \mathcal{G}(\boldsymbol{g}) + \mathcal{I}(\boldsymbol{c}_0)$$

and consequently $F_\sigma(\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma,\boldsymbol{g}) = \boldsymbol{0}$. Moreover, by virtue of Lemma 6 the mapping

$$\mathrm{D}_{(c,u)}F_\sigma(\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma,\boldsymbol{g}) : \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \to \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$$

$$\mathrm{D}_{(c,u)}F_\sigma(\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma,\boldsymbol{g})(\dot{\boldsymbol{c}}_\sigma,\dot{\boldsymbol{u}}_\sigma) = (\dot{\boldsymbol{c}}_\sigma,\dot{\boldsymbol{u}}_\sigma) - \mathcal{R}(\boldsymbol{P}_\Sigma(\mathrm{D}(-\frac{1}{\sigma}\hat{\boldsymbol{\Phi}})(\boldsymbol{c}_\sigma)\dot{\boldsymbol{c}}_\sigma))$$

is invertible for all $(\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma,\boldsymbol{g}) \in \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times L^2\left(0,T;L^2\left(\Gamma_g,\mathbb{R}^d\right)\right)$. The implicit function theorem implies the existence and uniqueness of a differentiable operator $\boldsymbol{B} : U(\boldsymbol{g}) \to \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$, where $U(\boldsymbol{g})$ is some open neighborhood of $\boldsymbol{g}$ in $L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$, such that $F\left(\boldsymbol{B}(\boldsymbol{h}),\boldsymbol{h}\right) = \boldsymbol{0}$ for all $\boldsymbol{h} \in U(\boldsymbol{g})$. So we can conclude, that $\boldsymbol{B} \equiv \left(\boldsymbol{S}_{\sigma|1}|_{U(\boldsymbol{g})},\boldsymbol{S}_{\sigma|2}|_{U(\boldsymbol{g})}\right)$. The particular form of $\mathrm{D}\boldsymbol{S}_\sigma$ immediately follows from

$$(\mathrm{D}\boldsymbol{S}_{\sigma|1}(\boldsymbol{g}),\mathrm{D}\boldsymbol{S}_{\sigma|2}(\boldsymbol{g})) = \mathrm{D}\boldsymbol{B}(\boldsymbol{g}) = -\left(\mathrm{D}_{(c,u)}F_\sigma(\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma,\boldsymbol{g})\right)^{-1} \circ \mathrm{D}_g F_\sigma(\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma,\boldsymbol{g})$$

where $(\boldsymbol{c}_\sigma,\boldsymbol{u}_\sigma) = (\boldsymbol{S}_{\sigma|1}(\boldsymbol{g}),\boldsymbol{S}_{\sigma|2}(\boldsymbol{g}))$, see Remark 1. This means, that $(\dot{\boldsymbol{c}}_\sigma,\dot{\boldsymbol{u}}_\sigma) \in \boldsymbol{\mathcal{V}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$, where $(\dot{\boldsymbol{c}}_\sigma,\dot{\boldsymbol{u}}_\sigma) = (\mathrm{D}\boldsymbol{S}_{\sigma|1}(\boldsymbol{g})(\boldsymbol{h}),\mathrm{D}\boldsymbol{S}_{\sigma|2}(\boldsymbol{g})(\boldsymbol{h}))$ solves the linearized system (4.1) – (4.2). Uniqueness can be shown in a standard way using Gronwall's inequality. □

Remark 5 *From Theorem 4 it easily follows, using the quadratic form of $\widetilde{J}$ and chain rule, that for any $\sigma \in (0,\frac{1}{4})$ the reduced cost functional $j_\sigma(\boldsymbol{g}) := \widetilde{J}(\boldsymbol{S}_{\sigma|1}(\boldsymbol{g}),\boldsymbol{g})$ is Fréchet differentiable and the derivative has the form*

$$Dj_\sigma(\boldsymbol{g}) = D_c\widetilde{J}(\boldsymbol{S}_{\sigma|1}(\boldsymbol{g}),\boldsymbol{g}) \circ D\boldsymbol{S}_{\sigma|1}(\boldsymbol{g}) + D_g\widetilde{J}(\boldsymbol{S}_{\sigma|1}(\boldsymbol{g}),\boldsymbol{g}). \tag{4.9}$$

4.3 First-order necessary optimality conditions for $(\widetilde{\mathcal{P}}_\sigma)$

Suppose now that $\bar{\boldsymbol{g}} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ is any local minimizer for $(\mathcal{P}_0)$ with associated state $(\bar{\boldsymbol{c}},\bar{\boldsymbol{u}},\bar{\boldsymbol{\xi}}) = \boldsymbol{S}_0(\bar{\boldsymbol{g}}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$, see (2.10). Inserting (1.17) and (1.9) in (4.9) and applying Theorem 4 yields the following result:

Corollary 1 *Let $\sigma \in (0, \frac{1}{4})$ be given. If $\bar{\boldsymbol{g}}_\sigma \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ is an optimal control for the adapted control problem $(\widetilde{\mathcal{P}}_\sigma)$ with associated state $(\bar{\boldsymbol{c}}_\sigma, \bar{\boldsymbol{u}}_\sigma) = (\boldsymbol{S}_{\sigma|1}(\bar{\boldsymbol{g}}_\sigma), \boldsymbol{S}_{\sigma|2}(\bar{\boldsymbol{g}}_\sigma)) \in \boldsymbol{\mathcal{V}}_\Sigma \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ then we have for every $\boldsymbol{h} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$*

$$\nu_T \int_\Omega (\bar{\boldsymbol{c}}_\sigma(T,\cdot) - \boldsymbol{c}_T)\cdot \dot{\boldsymbol{c}}_\sigma(T,\cdot)\,\mathrm{d}x + \nu_d \int_0^T \int_\Omega (\bar{\boldsymbol{c}}_\sigma - \boldsymbol{c}_d)\cdot \dot{\boldsymbol{c}}_\sigma \,\mathrm{d}x\,\mathrm{d}t + \\ + \int_0^T \int_{\Gamma_g} (\nu_g \bar{\boldsymbol{g}}_\sigma + (\bar{\boldsymbol{g}}_\sigma - \bar{\boldsymbol{g}}))\cdot \boldsymbol{h}\,\mathrm{d}s\,\mathrm{d}t = 0, \tag{4.10}$$

where $\dot{\boldsymbol{c}}_\sigma = \mathrm{D}\boldsymbol{S}_{\sigma|1}(\bar{\boldsymbol{g}}_\sigma)(\boldsymbol{h})$, see Theorem 4.

We are now in the position to derive the first-order necessary optimality conditions for the control problem $(\widetilde{\mathcal{P}}_\sigma)$:

Theorem 5 *Let $\sigma \in (0, \frac{1}{4})$ be given and define $\boldsymbol{\zeta}_\sigma := \mathrm{D}\left(-\frac{1}{\sigma}\hat{\boldsymbol{\Phi}}\right)(\bar{\boldsymbol{c}}_\sigma)\bar{\boldsymbol{p}}_\sigma$. Moreover, assume that $\bar{\boldsymbol{g}}_\sigma \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ is an optimal control for the control problem $(\widetilde{\mathcal{P}}_\sigma)$ with associated state $(\bar{\boldsymbol{c}}_\sigma, \bar{\boldsymbol{u}}_\sigma) = (\boldsymbol{S}_{\sigma|1}(\bar{\boldsymbol{g}}_\sigma), \boldsymbol{S}_{\sigma|2}(\bar{\boldsymbol{g}}_\sigma)) \in \boldsymbol{\mathcal{V}}_\Sigma \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$. Then the adjoint state system*

$$- \int_0^T \int_\Omega \partial_t \bar{\boldsymbol{p}}_\sigma \cdot \boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_\Omega \nabla \bar{\boldsymbol{p}}_\sigma \cdot \nabla \boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t - \int_0^T \int_\Omega \bar{\boldsymbol{p}}_\sigma \cdot \boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t + \\ - \int_0^T \int_\Omega \boldsymbol{P_\Sigma}(\boldsymbol{\zeta}_\sigma - D_p W(\bar{\boldsymbol{p}}_\sigma, \mathcal{E}(\bar{\boldsymbol{q}}_\sigma)))\cdot \boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t \\ = \nu_d \int_0^T \int_\Omega (\bar{\boldsymbol{c}}_\sigma - \boldsymbol{c}_d)\cdot \boldsymbol{\chi}\,\mathrm{d}x\,\mathrm{d}t, \tag{4.11}$$

$$\int_0^T \langle \mathcal{E}(\bar{\boldsymbol{q}}_\sigma) - \mathcal{E}^*(\bar{\boldsymbol{p}}_\sigma), \mathcal{E}(\boldsymbol{\eta})\rangle_{\mathcal{C}}\,\mathrm{d}t = 0, \tag{4.12}$$

which has to hold for all $\boldsymbol{\chi} \in L^2(0,T;\boldsymbol{H}^1(\Omega))$ and $\boldsymbol{\eta} \in L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ with

$$\bar{\boldsymbol{p}}_\sigma(T,\cdot) = \nu_T(\bar{\boldsymbol{c}}_\sigma(T,\cdot) - \boldsymbol{c}_T) \text{ a.e. in } \Omega, \tag{4.13}$$

has a unique solution $(\bar{\boldsymbol{p}}_\sigma, \bar{\boldsymbol{q}}_\sigma) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{T\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$, and we have

$$\bar{\boldsymbol{q}}_\sigma + \nu_g \bar{\boldsymbol{g}}_\sigma + (\bar{\boldsymbol{g}}_\sigma - \bar{\boldsymbol{g}}) = \boldsymbol{0} \text{ a.e. on } (0,T)\times\Gamma_g. \tag{4.14}$$

Proof. First observe that system (4.11)-(4.13) is a linear backward-in-time parabolic boundary value problem, which after the time transformation $t \mapsto T - t$ takes the form of a standard parabolic initial value problem. The well-posedness of a solution follows as indicated in Remark 4. At this point, we may perform standard calculation, using repeated integration by parts of the systems (4.1)-(4.2) and (4.11)-(4.13), which provides

$$\nu_T \int_\Omega (\bar{\boldsymbol{c}}_\sigma(T,\cdot) - \boldsymbol{c}_T)\cdot \dot{\boldsymbol{c}}_\sigma(T,\cdot)\,\mathrm{d}x + \nu_d \int_0^T \int_\Omega (\bar{\boldsymbol{c}}_\sigma - \boldsymbol{c}_d)\cdot \dot{\boldsymbol{c}}_\sigma \,\mathrm{d}x\,\mathrm{d}t \\ = \int_0^T \int_{\Gamma_g} \bar{\boldsymbol{q}}_\sigma \cdot \boldsymbol{h}\,\mathrm{d}s\,\mathrm{d}t \tag{4.15}$$

for all $\boldsymbol{h} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$, so that (4.14) follows from (4.10). For details, see e.g. [14]. □

4.4 The optimality conditions for $(\mathcal{P}_0)$

Suppose now that $\bar{\boldsymbol{g}} \in L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ is a local minimizer for $(\mathcal{P}_0)$ with associated state $(\bar{\boldsymbol{c}},\bar{\boldsymbol{u}},\bar{\boldsymbol{\xi}}) = \boldsymbol{S}_0(\bar{\boldsymbol{g}}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T)$, see (2.10). Then, by Theorem 3 we can find a sequence $\{\sigma_n\} \subset (0,\frac{1}{4})$ with $\sigma_n \searrow 0$ as $n \to \infty$ and, for any $n \in \mathbb{N}$, an optimal pair $((\bar{\boldsymbol{c}}_{\sigma_n},\bar{\boldsymbol{u}}_{\sigma_n},\bar{\boldsymbol{\xi}}_{\sigma_n}),\bar{\boldsymbol{g}}_{\sigma_n}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T) \times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ of the adapted optimal control problem $(\widetilde{\mathcal{P}}_{\sigma_n})$, such that the convergences (3.3) hold true. Moreover, by Theorem 5 there exist for any $n \in \mathbb{N}$ the corresponding adjoint variables $(\bar{\boldsymbol{p}}_{\sigma_n},\bar{\boldsymbol{q}}_{\sigma_n}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{T\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$ to problem $(\widetilde{\mathcal{P}}_{\sigma_n})$. We now derive some a priori estimates for the adjoint state variables $(\bar{\boldsymbol{p}}_{\sigma_n},\bar{\boldsymbol{q}}_{\sigma_n})$. To this end, we define $\boldsymbol{\mathcal{W}}_0(0,T) = \{\boldsymbol{v} \in \boldsymbol{\mathcal{W}}(0,T) : \boldsymbol{v}(0,\cdot) = \boldsymbol{0}\}$ which is a Banach space as a subspace of $\boldsymbol{\mathcal{W}}(0,T)$. Thus, we can define the dual space $\boldsymbol{\mathcal{W}}_0(0,T)^*$, where the dual pairing between elements $\boldsymbol{z} \in \boldsymbol{\mathcal{W}}_0(0,T)^*$ and $\boldsymbol{v} \in \boldsymbol{\mathcal{W}}_0(0,T)$ is denoted by $\langle\langle \boldsymbol{z},\boldsymbol{v}\rangle\rangle$.

Lemma 7 *There is some constant $K_2 > 0$ independent of $n \in \mathbb{N}$ such that*

$$\begin{aligned}&\|\bar{\boldsymbol{p}}_{\sigma_n}\|_{L^\infty(0,T;\boldsymbol{L}^2(\Omega))\cap L^2(0,T;\boldsymbol{H}^1(\Omega))}+\\ &\quad+\|\bar{\boldsymbol{q}}_{\sigma_n}\|_{L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))}+\|\partial_t\bar{\boldsymbol{p}}_{\sigma_n}\|_{\boldsymbol{\mathcal{W}}_0(0,T)^*}+\|\boldsymbol{P}_{\boldsymbol{\Sigma}}(\boldsymbol{\zeta}_{\sigma_n})\|_{\boldsymbol{\mathcal{W}}_0(0,T)^*}\le K_2.\end{aligned} \tag{4.16}$$

Proof. In the following, C_i, $i \in \mathbb{N}$, denote positive constants which are independent of $n \in \mathbb{N}$. Using the testfunction $\bar{\boldsymbol{q}}_{\sigma_n}$ in (4.12), written for σ_n, $n \in \mathbb{N}$, we get by the same arguments as in the proof of Lemma 2, see (2.4), that

$$\|\bar{\boldsymbol{q}}_{\sigma_n}\|_{L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))} \le C_1\|\bar{\boldsymbol{p}}_{\sigma_n}\|_{\boldsymbol{L}^2(\Omega_T)} \quad \forall n \in \mathbb{N}. \tag{4.17}$$

Moreover applying for any $\tau \in (0,T)$ the testfunction $\chi_{(\tau,T)}\bar{\boldsymbol{p}}_{\sigma_n}$ in (4.11), written for σ_n, $n \in \mathbb{N}$, we obtain for all $n \in \mathbb{N}$

$$\begin{aligned}&\frac{1}{2}\|\bar{\boldsymbol{p}}_{\sigma_n}(\tau)\|^2_{\boldsymbol{L}^2(\Omega)}+\|\nabla\bar{\boldsymbol{p}}_{\sigma_n}\|^2_{L^2(\tau,T;\boldsymbol{L}^2(\Omega))}+\frac{1}{\sigma_n}\int_\tau^T\int_\Omega D\hat{\boldsymbol{\Phi}}(\bar{\boldsymbol{c}}_{\sigma_n})|\bar{\boldsymbol{p}}_{\sigma_n}|^2\,\mathrm{d}x\,\mathrm{d}t\\ &\le\frac{1}{2}\|\nu_T(\bar{\boldsymbol{c}}_{\sigma_n}(T)-\boldsymbol{c}_T)\|^2_{\boldsymbol{L}^2(\Omega)}+\|\bar{\boldsymbol{p}}_{\sigma_n}\|^2_{L^2(\tau,T;\boldsymbol{L}^2(\Omega))}+\\ &\quad+\int_\tau^T\int_\Omega|D_pW(\bar{\boldsymbol{p}}_{\sigma_n},\mathcal{E}(\bar{\boldsymbol{q}}_{\sigma_n}))|\,|\bar{\boldsymbol{p}}_{\sigma_n}|\,\mathrm{d}x\,\mathrm{d}t+ \qquad (4.18)\\ &\quad+\int_\tau^T\int_\Omega|\nu_d(\bar{\boldsymbol{c}}_{\sigma_n}-\boldsymbol{c}_d)||\bar{\boldsymbol{p}}_{\sigma_n}|\,\mathrm{d}x\,\mathrm{d}t. \qquad (4.19)\end{aligned}$$

First, we observe that the last term in the first line of (4.19) is nonnegative. Now we recall that by Lemma 2 the sequence $\{\|\bar{\boldsymbol{c}}_{\sigma_n}\|_{\boldsymbol{\mathcal{V}}}\}$ is bounded. Therefore, using the final time condition in (4.13), applying Young's inequality and invoking Gronwall's inequality, we find the estimate

$$\|\bar{\boldsymbol{p}}_{\sigma_n}\|_{L^\infty(0,T;\boldsymbol{L}^2(\Omega))\cap L^2(0,T;\boldsymbol{H}^1(\Omega))} \le C_2 \quad \forall n \in \mathbb{N}. \tag{4.20}$$

Next, we derive a bound for the time derivative of $\bar{\boldsymbol{p}}_{\sigma_n}$. To this end, let $\boldsymbol{v} \in \boldsymbol{\mathcal{W}}_0(0,T)$ be arbitrary. Using integration by parts, we then have for all $n \in \mathbb{N}$

$$\langle\langle\partial_t\bar{\boldsymbol{p}}_{\sigma_n},\boldsymbol{v}\rangle\rangle=-\int_0^T\langle\partial_t\boldsymbol{v},\bar{\boldsymbol{p}}_{\sigma_n}\rangle\,\mathrm{d}t+\int_\Omega\bar{\boldsymbol{p}}_{\sigma_n}(T,\cdot)\,\boldsymbol{v}(T,\cdot)\,\mathrm{d}x.$$

Now observe that $\boldsymbol{\mathcal{W}}(0,T)$ is continuously embedded in $C([0,T];\boldsymbol{L}^2(\Omega))$. We thus obtain

$$\begin{aligned}&|\langle\langle\partial_t\bar{\boldsymbol{p}}_{\sigma_n},\boldsymbol{v}\rangle\rangle|\\ &\quad\le\left(\|\bar{\boldsymbol{p}}_{\sigma_n}\|_{L^2(0,T;\boldsymbol{H}^1(\Omega))}+\|(\bar{\boldsymbol{p}}_{\sigma_n}(T,\cdot)\|_{\boldsymbol{L}^2(\Omega)}\right)\|\boldsymbol{v}\|_{\boldsymbol{\mathcal{W}}_0(0,T)},\end{aligned}$$

and by using (4.20), (4.13) and the uniform bound on $\{\|\bar{\boldsymbol{c}}_{\sigma_n}\|_{\boldsymbol{\mathcal{V}}}\}$ this gives

$$\|\partial_t \bar{\boldsymbol{p}}_{\sigma_n}\|_{\boldsymbol{\mathcal{W}}_0(0,T)^*} \leq C_3 \quad \forall n \in \mathbb{N}. \tag{4.21}$$

Finally, comparison in (4.11), invoking the estimates (4.20) and (4.21), yields that

$$\|\boldsymbol{P}_{\boldsymbol{\Sigma}}(\boldsymbol{\zeta}_{\sigma_n})\|_{\boldsymbol{\mathcal{W}}_0(0,T)^*} \leq C_4 \quad \forall n \in \mathbb{N}.$$

and the assertion is proved. ☐ We draw some consequences from Lemma 7. At first, it follows from (4.16) that there is some subsequence, which is again indexed by n, such that, as $n \to \infty$,

$$\begin{array}{llll} \bar{\boldsymbol{p}}_{\sigma_n} & \longrightarrow \boldsymbol{p} & \text{weakly} & \text{in} \quad L^2(0,T;\boldsymbol{H}^1(\Omega)), \\ \bar{\boldsymbol{q}}_{\sigma_n} & \longrightarrow \boldsymbol{q} & \text{weakly} & \text{in} \quad L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)), \\ \boldsymbol{P}_{\boldsymbol{\Sigma}}(\boldsymbol{\zeta}_{\sigma_n}) & \longrightarrow \boldsymbol{\zeta} & \text{weakly-star} & \text{in} \quad \boldsymbol{\mathcal{W}}_0(0,T)^*, \end{array} \tag{4.22}$$

for suitable limits $(\boldsymbol{p},\boldsymbol{q},\boldsymbol{\zeta})$. Therefore, passing to the limit as $n \to \infty$ in (4.14) and (4.11) – (4.13) we obtain

$$\boldsymbol{q} + \nu_g \bar{\boldsymbol{g}} = \boldsymbol{0} \text{ a.e. on } (0,T) \times \Gamma_g \tag{4.23}$$

together with the adjoint state system:

$$\begin{aligned} &-\langle\langle \boldsymbol{\zeta}, \boldsymbol{v}\rangle\rangle + \int_0^T \langle \partial_t \boldsymbol{v}, \boldsymbol{p}\rangle \,\mathrm{d}t + \int_0^T \int_\Omega \nabla \boldsymbol{p} \cdot \nabla \boldsymbol{v} \,\mathrm{d}x\,\mathrm{d}t + \\ &- \int_0^T \int_\Omega \boldsymbol{p}\cdot \boldsymbol{v} \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_\Omega \boldsymbol{P}_{\boldsymbol{\Sigma}}(D_p W(\boldsymbol{p}, \mathcal{E}(\boldsymbol{q}))) \cdot \boldsymbol{v} \,\mathrm{d}x\,\mathrm{d}t + \\ &- \int_0^T \int_\Omega \nu_d (\bar{\boldsymbol{c}} - \boldsymbol{c}_d)\cdot \boldsymbol{v} \,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \nu_T (\bar{\boldsymbol{c}}(T,\cdot) - \boldsymbol{c}_T)\cdot \boldsymbol{v}(T)\,\mathrm{d}x = 0, \end{aligned} \tag{4.24}$$

$$\int_0^T \langle \mathcal{E}(\boldsymbol{q}) - \mathcal{E}(\boldsymbol{p}), \mathcal{E}(\boldsymbol{\eta})\rangle_{\mathcal{C}} \,\mathrm{d}t = 0, \tag{4.25}$$

which has to hold for all $\boldsymbol{v} \in \boldsymbol{\mathcal{W}}_0(0,T)$ and all $\boldsymbol{\eta} \in L^2(0,T;H^1_D(\Omega,\mathbb{R}^d))$. Next, we show that the limit elements $(\boldsymbol{p},\boldsymbol{q},\boldsymbol{\zeta})$ satisfy some sort of complementarity slackness condition. To this end, observe that we obviously have

$$(-\boldsymbol{\zeta}_{\sigma_n}, \bar{\boldsymbol{p}}_{\sigma_n})_{\boldsymbol{L}^2(\Omega_T)} = \frac{1}{\sigma_n}\int_0^T \int_\Omega D\hat{\boldsymbol{\Phi}}(\bar{\boldsymbol{c}}_{\sigma_n})|\bar{\boldsymbol{p}}_{\sigma_n}|^2 \,\mathrm{d}x\,\mathrm{d}t \geq 0 \quad \forall n \in \mathbb{N}.$$

We thus obtain

$$\lim_{n\to\infty} (\boldsymbol{\zeta}_{\sigma_n}, \bar{\boldsymbol{p}}_{\sigma_n})_{\boldsymbol{L}^2(\Omega_T)} \leq 0. \tag{4.26}$$

Moreover, we have that $(\boldsymbol{\zeta}_{\sigma_n}, \max(\boldsymbol{0}, \bar{\boldsymbol{c}}_{\sigma_n}))_{\boldsymbol{L}^2(\Omega_T)} = 0$ for all $n \in \mathbb{N}$ and hence

$$\lim_{n\to\infty} (\boldsymbol{\zeta}_{\sigma_n}, \max(\boldsymbol{0}, \bar{\boldsymbol{c}}_{\sigma_n}))_{\boldsymbol{L}^2(\Omega_T)} = 0. \tag{4.27}$$

Similarly as in [19, Theorem 4.3.4] we infer for all $n \in \mathbb{N}$

$$\begin{aligned} &(\bar{\boldsymbol{\xi}}_{\sigma_n}, \bar{\boldsymbol{c}}_{\sigma_n})_{\boldsymbol{L}^2(\Omega_T)} \\ &= \sum_{i=1}^N \left(\int_{\{-\sigma_n < \bar{c}_{\sigma_n,i} < 0\}} \frac{1}{2\sigma_n^2} \bar{c}^3_{\sigma_n,i} + \int_{\{\bar{c}_{\sigma_n,i} \leq -\sigma_n\}} \left(-\frac{1}{\sigma_n}\bar{c}_{\sigma_n,i} - \frac{1}{2}\right)\bar{c}_{\sigma_n,i} \right), \end{aligned} \tag{4.28}$$

where $\{-\sigma_n < \bar{c}_{\sigma_n,i} < 0\} := \{(t,x) \in \Omega_T \mid -\sigma_n < \bar{c}_{\sigma_n,i} < 0$ a.e. in $\Omega_T\}$ and $\{\bar{c}_{\sigma_n,i} \le -\sigma_n\} := \{(t,x) \in \Omega_T \mid \bar{c}_{\sigma_n,i} \le -\sigma_n$ a.e. in $\Omega_T\}$. From $\lim_{n\to\infty}(\bar{\boldsymbol{\xi}}_{\sigma_n}, \bar{\boldsymbol{c}}_{\sigma_n})_{\boldsymbol{L}^2(\Omega_T)} = 0$, see (2.9), and the fact that both summands in (4.28) are non-positive, we deduce that all terms in (4.28) tend individually to zero and so we get as $n \to \infty$

$$\left\| \chi_{\{-\sigma_n < \bar{c}_{\sigma_n,i} < 0\}} \frac{1}{\sqrt{2}\sigma_n} \left|\bar{c}_{\sigma_n,i}\right|^{\frac{3}{2}} \right\|_{L^2(\Omega_T)} \to 0 \quad \forall i = 1, \dots, N \tag{4.29}$$

and

$$\left\| \chi_{\{\bar{c}_{\sigma_n,i} \le -\sigma_n\}} \sqrt{\sigma_n} \left(\frac{1}{\sigma_n} \bar{c}_{\sigma_n,i} + \frac{1}{2} \right) \right\|_{L^2(\Omega_T)} \to 0 \quad \forall i = 1, \dots, N, \tag{4.30}$$

where for the last convergence result we use that on $\{\bar{c}_{\sigma_n,i} \le -\sigma_n\}$ we have

$$\left| \frac{1}{\sigma_n} \bar{c}_{\sigma_n,i} + \frac{1}{2} \right| = -\frac{1}{\sigma_n} \bar{c}_{\sigma_n,i} - \frac{1}{2} \le -\frac{1}{\sigma_n} \bar{c}_{\sigma_n,i} = \frac{1}{\sigma_n} \left|\bar{c}_{\sigma_n,i}\right| \quad \forall n \in \mathbb{N}.$$

Using the a priori estimate (4.19) once more, we have for all $n \in \mathbb{N}$

$$\left|(\boldsymbol{\zeta}_{\sigma_n}, \bar{\boldsymbol{p}}_{\sigma_n})_{\boldsymbol{L}^2(\Omega_T)}\right| \le C \left(\|\bar{\boldsymbol{p}}_{\sigma_n}\|^2_{L^\infty(0,T;\boldsymbol{L}^2(\Omega)) \cap L^2(0,T;\boldsymbol{H}^1(\Omega))} + 1 \right).$$

Invoking Lemma 7 gives then that

$$\left|(\boldsymbol{\zeta}_{\sigma_n}, \bar{\boldsymbol{p}}_{\sigma_n})_{\boldsymbol{L}^2(\Omega_T)}\right| \le C_5 \quad \forall n \in \mathbb{N}. \tag{4.31}$$

Moreover, we have for all $n \in \mathbb{N}$

$$\begin{aligned} &\left|(\boldsymbol{\zeta}_{\sigma_n}, \bar{\boldsymbol{p}}_{\sigma_n})_{\boldsymbol{L}^2(\Omega_T)}\right| \\ &= \sum_{i=1}^N \left[\int_{\{-\sigma_n < \bar{c}_{\sigma_n,i} < 0\}} \frac{1}{\sigma_n^2} \left(-\bar{c}_{\sigma_n,i}\right) \left(\bar{p}_{\sigma_n,i}\right)^2 + \int_{\{\bar{c}_{\sigma_n,i} \le -\sigma_n\}} \frac{1}{\sigma_n} \left(\bar{p}_{\sigma_n,i}\right)^2 \right]. \end{aligned}$$

Since again both terms are non-negative, using (4.31), we obtain for all $i = 1, \dots, N$

$$\left\| \chi_{\{-\sigma_n < \bar{c}_{\sigma_n,i} < 0\}} \frac{1}{\sigma_n} \left|\bar{c}_{\sigma_n,i}\right|^{\frac{1}{2}} \bar{p}_{\sigma_n,i} \right\|_{L^2(\Omega_T)} \le C_6 \quad \forall n \in \mathbb{N}, \tag{4.32}$$

and

$$\left\| \chi_{\{c_{\sigma,i} \le -\sigma_n\}} \frac{1}{\sqrt{\sigma_n}} \bar{p}_{\sigma_n,i} \right\|_{L^2(\Omega_T)} \le C_7 \quad \forall n \in \mathbb{N}. \tag{4.33}$$

We have for all $n \in \mathbb{N}$

$$\begin{aligned} &\left(\bar{\boldsymbol{\xi}}_{\sigma_n}, \bar{\boldsymbol{p}}_{\sigma_n}\right)_{\boldsymbol{L}^2(\Omega_T)} = \\ &= \sum_{i=1}^N \left[\frac{1}{\sqrt{2}} \left(\chi_{\{-\sigma_n < \bar{c}_{\sigma_n,i} < 0\}} \frac{1}{\sqrt{2}\sigma_n} \left|\bar{c}_{\sigma_n,i}\right|^{\frac{3}{2}}, \frac{1}{\sigma_n} \left|\bar{c}_{\sigma_n,i}\right|^{\frac{1}{2}} \bar{p}_{\sigma_n,i} \right)_{L^2(\Omega_T)} \right] + \\ &+ \sum_{i=1}^N \left[\left(\chi_{\{\bar{c}_{\sigma_n,i} \le -\sigma_n\}} \sqrt{\sigma_n} \left(\frac{1}{\sigma_n} \bar{c}_{\sigma_n,i} + \frac{1}{2} \right), \frac{1}{\sqrt{\sigma_n}} \bar{p}_{\sigma_n,i} \right)_{L^2(\Omega_T)} \right] \end{aligned}$$

and applying now the estimates (4.29), (4.30), (4.32) and (4.33) gives

$$\lim_{n\to\infty} (\bar{\boldsymbol{p}}_{\sigma_n}, \bar{\boldsymbol{\xi}}_{\sigma_n})_{\boldsymbol{L}^2(\Omega_T)} = 0. \tag{4.34}$$

We now combine the results established above with Theorem 3 to obtain the following result:

Theorem 6 *Let* $((\bar{\boldsymbol{c}}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{\xi}}), \bar{\boldsymbol{g}}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T) \times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$, *where* $(\bar{\boldsymbol{c}}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{\xi}}) = \boldsymbol{S}_0(\bar{\boldsymbol{g}})$, *be an optimal pair for* $(\mathcal{P}_0)$. *Then the following assertions hold true:*

(i) There exists a sequence $\{\sigma_n\} \subset (0, \frac{1}{4})$ *with* $\sigma_n \searrow 0$ *as* $n \to \infty$, *and for any* $n \in \mathbb{N}$ *a solution pair* $((\bar{\boldsymbol{c}}_{\sigma_n}, \bar{\boldsymbol{u}}_{\sigma_n}, \bar{\boldsymbol{\xi}}_{\sigma_n}), \bar{\boldsymbol{g}}_{\sigma_n}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \times L^2(0,T;H^1_D(\Omega,\mathbb{R}^d)) \times \boldsymbol{L}^2(\Omega_T) \times L^2(0,T;L^2(\Gamma_g,\mathbb{R}^d))$ *to the adapted optimal control problem* $(\widetilde{\mathcal{P}}_{\sigma_n})$, *such that (3.3) holds as* $n \to \infty$.

(ii) Whenever sequences $\{\sigma_n\} \subset (0, \frac{1}{4})$ *and* $((\bar{\boldsymbol{c}}_{\sigma_n}, \bar{\boldsymbol{u}}_{\sigma_n}, \bar{\boldsymbol{\xi}}_{\sigma_n}), \bar{\boldsymbol{g}}_{\sigma_n})$ *having the properties described in (i) are given, the following holds true: to any subsequence* $\{n_k\}_{k\in\mathbb{N}}$ *of* $\mathbb{N}$ *there are a subsequence* $\{n_{k_l}\}_{l\in\mathbb{N}}$ *and some* $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\zeta})$ *such that*

- *the relations (4.26),(4.27), and (4.34) hold (where the sequences are indexed by* n_{k_l} *and the limits are taken for* $l \to \infty$*), and*
- *the gradient equation (4.23) and the adjoint system (4.24) are satisfied.*

References

[1] J. Appell and P.P. Zabrejko. *Nonlinear superposition operators*. Cambridge University Press, 1990.

[2] V. Barbu. *Optimal control of variational inequalities*. Research notes in mathematics. Pitman Advanced Pub. Program, 1984.

[3] L. Blank, H. Garcke, L. Sarbu, and V. Styles. Non-local Allen-Cahn systems: Analysis and a primal dual active set method. *IMA J. Numer. Anal.*, 2013.

[4] T. Blesgen and U. Weikard. Multi-component Allen-Cahn equation for elastically stressed solids. *Electron. J. Differential Equations*, 2005(89):1–17, 2005.

[5] J. F. Blowey and C. M. Elliott. The Cahn-Hilliard gradient theory for phase separation with non-smooth free energy Part I: Mathematical analysis. *European J. Appl. Math.*, 2:233–280, 8 1991.

[6] A. Bossavit and A. Damlamian. *Free boundary problems: applications and theory*, volume 1. Pitman Advanced Publishing Program, 1985.

[7] C.M. Elliott, , Y. Giga, M. Hinze, and V. Styles. Workshop: Interfaces and Free Boundaries: Analysis, Control and Simulation. *Oberwolfach Reports*, 15, 2013.

[8] L.C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics. American Mathematical Society, 1998.

[9] M. H. Farshbaf-Shaker. A penalty approach to optimal control of Allen-Cahn variational inequalities: MPEC-view. *Numer. Funct. Anal. Optim.*, 33(11):1321–1349, 2012.

[10] A. Fasano and M. Primicerio. *Free boundary problems: theory and applications*. Research notes in mathematics. Pitman Advanced Pub. Program, 1983.

[11] H. Garcke. *On mathematical models for phase separation in elastically stressed solids*. Habilitation thesis, University of Bonn, 2000.

[12] H. Garcke. On a Cahn-Hilliard model for phase separation with elastic misfit. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 22(2):165 – 185, 2005.

[13] H. Garcke and B. Nestler. A mathematical model for grain growth in thin metallic films. *Math. Models Methods Appl. Sci.*, 10(06):895–921, 2000.

[14] C. Hecht. Existence theory and necessary optimality conditions for the control of the elastic Allen-Cahn system. Diplomarbeit, Universität Regensburg, 2011.

[15] M. Hintermüller and I. Kopacka. Mathematical programs with complementarity constraints in function space: C-and strong stationarity and a path-following algorithm. *SIAM J. Optim.*, 20(2):868–902, 2009.

[16] M. Hintermüller and T. Surowiec. First-order optimality conditions for elliptic mathematical programs with equilibrium constraints via variational analysis. *SIAM J. Optim.*, 21(4):1561–1593, 2011.

[17] M. Hintermüller and M.H. Tber. An inverse problem in American options as a mathematical program with equilibrium constraints: C-stationarity and an active-set-Newton solver. *SIAM J. Control Optim.*, 48(7):4419–4452, 2010.

[18] M. Hintermüller and D. Wegner. Distributed optimal control of the Cahn-Hilliard system including the case of a double-obstacle homogeneous free energy density. *SIAM J. Control Optim.*, 50(1):388–418, 2012.

[19] I. Kopacka. *MPECs/MPCCs in function space: first order optimality concepts, path-following, and multilevel algorithms*. PhD thesis, Department of Mathematics and Scientific Computing, University of Graz, 2009.

[20] T. Merkle. *The Cahn-Larché system: a model for spinodal decomposition in eutectic solder; modelling, analysis and simulation*. PhD thesis, Universität Stuttgart, 2005.

[21] F. Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*. Graduate Studies in Mathematics. American Mathematical Society, 2010.