Universität Regensburg
Mathematik

A penalty approach to optimal control of Allen-Cahn variational inequalities: MPEC-view

M. Hassan Farshbaf-Shaker

Preprint Nr. 06/2011

A penalty approach to optimal control of Allen-Cahn variational inequalities: MPEC-view

M.Hassan Farshbaf-Shaker*

Abstract

A scalar Allen-Cahn-MPEC problem is considered and a penalization technique is applied to show the existence of an optimal control. We show that the stationary points of the penalized problems converge to weak stationary points of the limit problem.

Key words. Allen-Cahn system, parabolic obstacle problems, MPECs, mathematical programs with complementarity constraints, optimality conditions.

AMS subject classification. 34G25, 35K86, 35R35, 49J20, 65K10

1 Introduction

In a Mini-Workshop *Control of Free Boundaries* in 2007 in Oberwolfach, see [16], the following paradigm optimal control problem involving free boundaries was formulated. Control the interface evolution law

$$V = -H + u, \tag{1.1}$$

where V is the normal velocity and H is the mean curvature of the interface. The space and time dependent quantity u can be used to control the interface. The above formulation is a sharp interface description of the interface. As this is well-known, one drawback of such a description is that it is difficult to handle topological changes, specially if one is interested in numerical simulations. One way to omit these difficulties is to use suitable

*Fakultät für Mathematik, Universität Regensburg, 93040 Regensburg, Germany

approximations of (1.1). Such approximations like diffuse interface models and specially Allen-Cahn models

$$\varepsilon\partial_t y = \varepsilon\Delta y - \frac{1}{\varepsilon}\psi'(y) + u, \tag{1.2}$$

with the smooth double well potential $\psi(u) = \frac{9}{32}(1-u^2)^2$ are used extensively in the phase field community, see [4, 5] and references therein. The approximative models (1.2) are constructed in such a way that they converge to the evolution law (1.1) as $\varepsilon \searrow 0$ and have the advantage that topology changes can be dealt with implicity, see [9]. Here an interface in which a phase field or order parameter rapidly changes its value, is modeled to have a thickness of order ε where $\varepsilon > 0$ is a small parameter. The model is based on a non-convex energy E which has the form $E(y) := E^1(y) + E^2(y)$ and

$$E^1(y) := \int_\Omega \left(\frac{\varepsilon}{2}|\nabla y|^2 + \frac{1}{\varepsilon}\psi(y)\right) dx, \qquad E^2(y) := -\int_\Omega yudx,$$

where $\Omega \subset \mathbb{R}^d$ is an open and bounded domain and $y : \Omega \to \mathbb{R}$ is the phase field, also called order parameter. The potential function ψ is assumed to have two global minima at the points ± 1 and the values ± 1 describe the pure phases. In order to have the Ginzburg-Landau energy $E^1(y)$ of moderate size y favors the values ± 1 due to the potential function. On the other hand given the gradient term $\int_\Omega |\nabla y|^2$ oscillations between the values ± 1 are energetically not favorable. Given an initial distribution the interface motion can be modeled by the steepest decent of E with respect to the $L^2 -$ norm which results then in (1.2). An approach according to the above formulated paradigm problem is now as follows:

$$\min J(y,u) := \int_\Omega \frac{\nu_T}{2}(y(T,x) - y_T(x))^2 dx + \int_{\Omega_T} \frac{\nu_d}{2}(y(t,x) - y_d(t,x))^2 dxdt$$
$$+ \int_{\Omega_T} \frac{\nu_u}{2\varepsilon} u^2 dxdt, \qquad \text{where } \nu_T, \nu_d, \nu_u > 0,$$

such that (1.2) and suitable initial and boundary conditions hold. Here the goal is to transform an initial phase distribution $y_0 : \Omega \to \mathbb{R}$ to some desired phase pattern $y_T : \Omega \to \mathbb{R}$ at a given final time T. Moreover throughout the entire time interval the distribution additionally remains close to y_d. In the formulation (1.2) the potential ψ is a smooth polynomial. Hence, y attains values different from ± 1 in the whole domain Ω and this is a disadvantage from the numerical point of view, where the solution has to be computed on the whole domain instead on the interface. Thus, to overcome this drawback we plan to use an Allen-Cahn variational inequality instead, i.e. using the obstacle potential

$$\psi(y) = \begin{cases} \frac{1}{2}(1-y^2) & \text{if } |y| \le 1, \\ \infty & \text{if } |y| > 1. \end{cases}$$

Introducing $\psi_0(y) := \frac{1}{2}(1-y^2)$ and the indicator function

$$I_{[-1,1]}(y) := \begin{cases} 0 & \text{if } |y| \le 1, \\ \infty & \text{if } |y| > 1, \end{cases}$$

we obtain

$$\psi(y) = \psi_0(y) + I_{[-1,1]}(y).$$

Then the object is given by values identical to 1. The interface $|y| < 1$ now has a small finite thickness proportional to ε. An additional advantage will be that as a consequence one only has to compute the solution in a narrow band around the interface.

Notations and general assumptions In the sequel we always denote by $\Omega \subset R^d$ an open, bounded domain (with spatial dimension d) with boundary $\Gamma = \partial\Omega$. The outer unit normal on Γ is denoted by n. We denote by $L^p(\Omega), W^{k,p}(\Omega)$ for $1 \le p \le \infty$ 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 $H^k(\Omega) = W^{k,2}(\Omega)$, see [1]. 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 the 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 $\|u(\cdot)\|_X \in L^p(0,T)$. We set $\Omega_T := (0,T) \times \Omega$, $\Gamma_T := (0,T) \times \Gamma$. "Generic" positive constants are denoted by C. Furthermore we define following time dependent Sobolev spaces by

$$\begin{aligned} W(0,T) &:= L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^1(\Omega)^*), \\ \mathcal{V} &:= L^2(0,T;H^2(\Omega)) \cap L^\infty(0,T;H^1(\Omega)) \cap H^1(\Omega_T). \end{aligned}$$

Moreover specially for dim $\Omega \le 3$ we will use following Sobolev embeddings

$$H^1(\Omega) \hookrightarrow L^{p*}(\Omega),\ p* \in [1,6], \tag{1.3}$$

$$H^2(\Omega) \hookrightarrow C(\overline{\Omega}), \tag{1.4}$$

and

$$W_2^{\frac{3}{2}}(\Omega) \hookrightarrow W_q^1(\Omega),\ q \in [1,3]. \tag{1.5}$$

Besides we also will use following embedding

$$H^1(0,T;H^1(\Omega)) \cap L^2(0,T;H^2(\Omega)) \hookrightarrow C([0,T];W_2^{\frac{3}{2}}(\Omega)). \tag{1.6}$$

For the rest of the paper we make following assumptions:

(H0) $E^1(y_0) < \infty$.

(H1) Assume $\Omega \subset \mathbb{R}^d$ is bounded and either convex or has a $C^{1,1}-$boundary and let $T > 0$ be a positive time.

Hence, given an initial phase distribution $y(0,\cdot) = y_0 : \Omega \to [-1,1]$ at time $t = 0$ the interface motion can be modeled by the steepest descent of E with respect to the L^2-norm which results, after suitable rescaling of time, in the following Allen-Cahn equation

$$\varepsilon\partial_t y = -\mathrm{grad}_{L^2} E(y) = \varepsilon\Delta y + \frac{1}{\varepsilon}(\psi_0'(y) - \zeta^*) + u,$$

where $\zeta^* \in \partial I_{[-1,1]}$ and $\partial I_{[-1,1]}$ denotes the subdifferential of $I_{[-1,1]}$. This equation leads to the following variational inequality

$$\varepsilon(\partial_t y, \eta - y)_{L^2(\Omega)} + \varepsilon(\nabla y, \nabla(\eta - y))_{L^2(\Omega)} + (\frac{1}{\varepsilon}\psi_0'(y) - u, \eta - y)_{L^2(\Omega)} \geq 0, \tag{1.7}$$

which has to hold for almost all t and all $\eta \in H^1(\Omega)$ with $|\eta| \leq 1$ a.e. in Ω. Our overall optimization problem is now stated as

$$(\mathcal{P}) \quad \begin{cases} \min & J(y,u), \\ \text{over} & y : [0,T] \times \Omega \to [-1,1];\ u : [0,T] \times \Omega \to \mathbb{R}, \\ \text{s.t.} & \varepsilon(\partial_t y, \eta - y) + \varepsilon(\nabla y, \nabla(\eta - y)) \geq (\frac{1}{\varepsilon}y + u, \eta - y), \\ & y(0) = y_0 : \Omega \to [-1,1], \\ & \text{ for almost all } t \text{ and all } \eta : \Omega \to [-1,1]. \end{cases}$$

The resulting optimization problem $(\mathcal{P})$ belongs to the problem class of so-called MPECs (Mathematical Programs with Equilibrium Constraints) which are hard to handle for several reasons. Indeed, it is well known that the variational inequality condition (or equivalently in 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. Approaches for the optimal control of variational inequalities in the classical literature typically introduce a regularization and show that in the limit of a vanishing regularization parameter certain weak generalized first order necessary conditions of optimality are derived, see e.g. [2]. Recently two different approaches are used to obtain weak generalized first order necessary conditions, see [13, 10]. On one hand there are penalization approaches [10, 15], which mostly and exclusively are used for elliptic problems. With such approaches, after getting the necessary optimality conditions by penalization, one tries to show that in the limit of the vanishing penalization parameter certain weak optimality conditions are derived. On the other hand there are relaxation approaches, see [3, 13, 11, 12, 8], which try to relax the complementarity conditions and to regularize the objective functional of the MPCC problem. Also here in the limit of the vanishing relaxation and regularization parameters certain weak optimality conditions are derived. It has to be said that these two approaches are well suited for dealing with elliptic problems. But in the case of parabolic problems additional technical difficulties arise, which lead in the limit to "very" weak optimality conditions (for different notions of stationarity for MPECs we refer to [13]).
In the present work we are interested in applying the penalization approach to our problem $(\mathcal{P})$. Our work is organized as follows. In section 2 we analyse our state equation. Most of the results of this section can be found in different papers, see e.g. [4], so the results are not new. But the penalization functions are different from the ones used in [4]. So we decided to keep our work self-contained and for convenience of the reader, we proved once again well-known results for our special penalization functions. In section 3 we introduce the penalized optimal control problem, prove the existence of minimizers and establish for the case when the spatial dimension is less than three the first order optimality system. In the last section 4 we show that in the limit of the vanishing penalization parameter certain weak optimality conditions are derived.

2 Allen-Cahn variational inequality

In this section we collect and extend known results about the Allen-Cahn variational inequality. All known results, which we will use without proof can be found in the literature, see e.g. [4] and references therein.The Allen-

Cahn variational inequality is given by:
(**ACVI**) *Let be given an initial data* $y_0 \in H^1(\Omega)$ *with* $|y_0| \le 1$ *a.e. in* Ω *and* $E^1(y_0) < \infty$. *Then for a given* $u \in L^2(\Omega_T)$ *find* $y \in H^1(\Omega_T)$ *such that* $y(0) = y_0$, $|y| \le 1$ *a.e. in* Ω_T *and*

$$\varepsilon(\partial_t y, \eta - y)_{L^2(\Omega)} + \varepsilon(\nabla y, \nabla(\eta - y))_{L^2(\Omega)} + \\ + \frac{1}{\varepsilon}(\psi_0'(y), \eta - y)_{L^2(\Omega)} \ge (u, \eta - y)_{L^2(\Omega)},$$

which has to hold for almost all t *and all* $\eta \in H^1(\Omega)$ *with* $|\eta| \le 1$ *a.e. in* Ω.
Due to [4] the problem (**ACVI**) can be reformulated with the help of Lagrange multipliers $\mu^\oplus$ and $\mu^\ominus$ corresponding to the inequality constraints $y \le 1$ and $y \ge -1$.

Lemma 1. *Assume **(H0)** and **(H1)** hold. Let* $u \in L^2(\Omega_T)$ *be given. A function* $y \in \mathcal{V}$ *solves **(ACVI)** if there exist* $\mu^\oplus, \mu^\ominus \in L^2(\Omega_T)$ *such that*

$$\varepsilon\partial_t y - \gamma\varepsilon\Delta y + \frac{1}{\varepsilon}\psi_0'(y) + \frac{1}{\varepsilon}\mu^\oplus - \frac{1}{\varepsilon}\mu^\ominus = u \quad a.e.\ in\ \Omega_T, \tag{2.1}$$

$$y(0) = y_0 \ \ a.e.\ in\ \Omega, \quad n \cdot \nabla y = 0 \quad a.e.\ on\ \Gamma_T, \tag{2.2}$$

$$|y| \le 1 \quad a.e.\ in\ \Omega_T, \tag{2.3}$$

$$\mu^\oplus(y - 1) = 0, \mu^\ominus(y + 1) = 0 \quad a.e.\ in\ \Omega_T, \tag{2.4}$$

$$\mu^\oplus \ge 0, \mu^\ominus \ge 0 \quad a.e.\ in\ \Omega_T. \tag{2.5}$$

The proof of Lemma 1 for $u \equiv 0$ can be found in [4]. The extension of the proof to our case $u \not\equiv 0$ is straightforward. We show the existence of a solution y together with unique Lagrange multipliers $\mu^\oplus$ and $\mu^\ominus$ by a penalty approach for the inequality constraint $|y| \le 1$. In particular, we replace the indicator function in ψ by terms penalizing deviations of y from the interval $[-1, 1]$. Motivated by [15, 6] for arbitrary but fixed and bounded $\gamma \in (0, \infty)$ we define convex functions $\psi_\oplus^\gamma, \psi_\ominus^\gamma \in C^2(\mathbb{R})$ by

$$\psi_\oplus^\gamma(r) := \begin{cases} \frac{1}{2}\left(r - \left(1 + \frac{\gamma}{2}\right)\right)^2 + \frac{\gamma^2}{24} & \text{for} \quad r \ge 1 + \gamma, \\ \frac{1}{6\gamma}(r - 1)^3 & \text{for} \quad 1 < r < 1 + \gamma, \\ 0 & \text{for} \quad r \le 1, \end{cases}$$

$$\psi_\ominus^\gamma(r) := \begin{cases} 0 & \text{for} \quad r \ge -1, \\ -\frac{1}{6\gamma}(r + 1)^3 & \text{for} \quad -1 - \gamma < r < -1, \\ \frac{1}{2}\left(r + \left(1 + \frac{\gamma}{2}\right)\right)^2 + \frac{\gamma^2}{24} & \text{for} \quad r \le -1 - \gamma. \end{cases}$$

We note that $(\psi_\oplus^\gamma)'$ and $(\psi_\ominus^\gamma)'$ are Lipschitz continuous functions where

$$0 \le (\psi_l^\gamma)'' \le 1, \quad l \in \{\oplus, \ominus\}. \tag{2.6}$$

Introducing now the penalized potential function

$$\psi_\sigma^\gamma(r) := \psi_0(r) + \frac{1}{\sigma}\left(\psi_\oplus^\gamma(r) + \psi_\ominus^\gamma(r)\right), \qquad \sigma > 0,$$

we get the penalized Energy

$$E_\sigma^1(y) := \int\limits_\Omega \left(\frac{\varepsilon}{2}|\nabla y|^2 + \frac{1}{\varepsilon}\psi_\sigma^\gamma(y)\right) dx.$$

Steepest decent of E_σ with respect to the L^2- norm gives the following penalized problem:

$$\begin{aligned} \varepsilon\partial_t y_\sigma - \varepsilon\Delta y_\sigma + \frac{1}{\varepsilon}(\psi_\sigma^\gamma)'(y_\sigma) = u_\sigma \quad &\text{in } \Omega_T, \\ y_\sigma(0) = y_0 \ \text{ in } \Omega, \qquad n\cdot\nabla y_\sigma = 0 \quad &\text{on } \Gamma_T. \end{aligned}$$

Defining

$$\mu_\sigma^\oplus := \frac{1}{\sigma}(\psi_\oplus^\gamma)'(y_\sigma) \quad \text{and} \quad \mu_\sigma^\ominus := -\frac{1}{\sigma}(\psi_\ominus^\gamma)'(y_\sigma),$$

we have to solve following semi-linear parabolic equation

$$\varepsilon\partial_t y_\sigma - \varepsilon\Delta y_\sigma + \frac{1}{\varepsilon}\psi_0'(y_\sigma) + \frac{1}{\varepsilon}\mu_\sigma^\oplus - \frac{1}{\varepsilon}\mu_\sigma^\ominus = u_\sigma \quad \text{in } \Omega_T, \tag{2.7}$$

$$y_\sigma(0) = y_0 \ \text{ in } \Omega, \qquad n\cdot\nabla y_\sigma = 0 \quad \text{on } \Gamma_T. \tag{2.8}$$

Theorem 1. *Assume* ***(H0)*** *and* ***(H1)*** *hold. Furthermore let* $u \in L^2(\Omega_T)$*. Then there exists a unique solution* $(y, \mu^\oplus, \mu^\ominus) \in \mathcal{V} \times L^2(\Omega_T) \times L^2(\Omega_T)$ *of (2.1)-(2.5).*

The proof in [4] can be carried out after easy modifications to our problem. However, to be self-contained we will give important aspects of the proof, which are treated in the following two separate Lemmas.

Lemma 2. *Assume* ***(H0)*** *and* ***(H1)*** *hold. Furthermore for* $\sigma > 0$, $u_\sigma \in L^2(\Omega_T)$*. Then there exists a unique solution* $y_\sigma \in \mathcal{V}$ *of (2.7)-(2.8). Moreover for a sequence* $\{u_\sigma\}$ *uniformly bounded in* $L^2(\Omega_T)$ *we have*

$$\begin{array}{lll} y_\sigma & \textit{uniformly bounded} & \textit{in} \ \ \mathcal{V}, \\ \mu_\sigma^\oplus & \textit{uniformly bounded} & \textit{in} \ \ L^2(\Omega_T), \\ \mu_\sigma^\ominus & \textit{uniformly bounded} & \textit{in} \ \ L^2(\Omega_T). \end{array}$$

Proof. The existence of a solution to (2.7)-(2.8) follows by using a standard Galerkin approximation and then passing to the limit, see [4]. The a priori estimates (uniformly in σ) are derived by testing (2.7) by suitable testfunctions like $y_\sigma, \partial_t y_\sigma, -\Delta y_\sigma, \mu_\sigma^\oplus$ and $\mu_\sigma^\ominus$. The key a priori estimate is the energy estimate, which we get by testing (2.7) by $\partial_t y_\sigma$ and carry out partial integration

$$\frac{1}{2}\|\partial_t y_\sigma\|^2_{L^2(\Omega_T)} + E^1_\sigma(y_\sigma(T)) \le E^1(y_0) + \frac{1}{2}\|u_\sigma\|^2_{L^2(\Omega_T)},$$

where we used Young's inequality for the last integral. Using **(H0)** and that $\{u_\sigma\}$ is uniformly bounded in $L^2(\Omega_T)$, we get a $C > 0$ independent of σ and the energy estimate

$$\frac{1}{2}\|\partial_t y_\sigma\|^2_{L^2(\Omega_T)} + E^1_\sigma(y_\sigma(T)) \le C. \tag{2.9}$$

Furthermore we test (2.7) by y_σ and note that $(\mu_\sigma^\oplus - \mu_\sigma^\ominus)y_\sigma \ge 0$, hence we get by standard calculations

$$\frac{\varepsilon}{2}\frac{d}{dt}\|y_\sigma\|^2_{L^2(\Omega)} + \varepsilon\|\nabla y_\sigma\|^2_{L^2(\Omega)} \le \frac{1}{2}\|u_\sigma\|^2_{L^2(\Omega)} + \left(\frac{1}{2}+\frac{1}{\varepsilon}\right)\|y_\sigma\|^2_{L^2(\Omega)}.$$

A Gronwall argument gives that $(y_\sigma)_{\sigma>0}$ is uniformly bounded in $L^\infty(0,T;L^2(\Omega))$. Hence $(y_\sigma)_{\sigma>0}$ is uniformly bounded in $L^\infty(0,T;H^1(\Omega)) \cap H^1(\Omega_T)$. Moreover we multiply (2.7) by $-\Delta y_\sigma$ and integrate. After integration by parts we obtain

$$\frac{d}{dt}\frac{1}{2}\|\nabla y_\sigma\|^2_{L^2(\Omega)} + \|\Delta y_\sigma\|^2_{L^2(\Omega)} + \int_\Omega \frac{1}{\sigma}(\psi^\gamma_\oplus(y_\sigma) + \psi^\gamma_\ominus(y_\sigma))''|\nabla y_\sigma|^2 dx = \|\nabla y_\sigma\|^2_{L^2(\Omega)}.$$

By virtue of $(\psi^\gamma_\oplus(y_\sigma) + \psi^\gamma_\ominus(y_\sigma))'' \ge 0$, a Gronwall argument and elliptic regularity theory we obtain that $(y_\sigma)_{\sigma>0}$ is uniformly bounded in $L^2(0,T;H^2(\Omega))$. Hence, $(y_\sigma)_{\sigma>0}$ is uniformly bounded in $\mathcal{V}$. For details, see e.g. [4]. Moreover since $\mu_\sigma^\oplus \cdot \mu_\sigma^\ominus = 0$ we obtain from (2.7) and the a priori estimates on y_σ that

$$\|\mu_\sigma^\oplus\|_{L^2(\Omega_T)} + \|\mu_\sigma^\ominus\|_{L^2(\Omega_T)} \le C. \tag{2.10}$$

□

As a direct consequence of Lemma 2 we get:

Lemma 3. *Let the assumption of Lemma 2 hold and let $\{u_\sigma\}$ be a sequence in $L^2(\Omega_T)$, $u \in L^2(\Omega_T)$ such that $u_\sigma \to u$ weakly in $L^2(\Omega_T)$. Furthermore* *let $y_\sigma \in \mathcal{V}$ denote the solution of (2.7)-(2.8). Then there exist $y \in \mathcal{V}$ and a subsequence still denoted by $\{y_\sigma\}$ such that as $\sigma \searrow 0$ we have*

$$\begin{array}{lllll} y_\sigma & \longrightarrow & y & \textit{weakly} & \textit{in} \;\; L^2(0,T;H^2(\Omega)), \\ y_\sigma & \longrightarrow & y & \textit{weakly} & \textit{in} \;\; H^1(\Omega_T), \\ y_\sigma & \longrightarrow & y & \textit{weakly-star} & \textit{in} \;\; L^\infty(0,T;H^1(\Omega)), \end{array}$$

The limit element (y,u) satisfies (2.1)-(2.5).

Proof. The convergence results are direct consequences of the estimates given by Lemma 2. Moreover we get from the above estimates

$$\begin{array}{lllll} y_\sigma & \longrightarrow & y & \text{strongly} & \text{in} \;\; L^2(\Omega_T), \\ y_\sigma & \longrightarrow & y & \text{a.e.} & \text{in} \;\; \Omega_T. \end{array}$$

Because of (2.10) there exist $\mu^\oplus, \mu^\ominus \in L^2(\Omega_T)$ such that for a subsequence (still denoted by $\mu_\sigma^\oplus$ and $\mu_\sigma^\ominus$)

$$\mu_\sigma^l \quad \longrightarrow \quad \mu^l \quad \text{weakly in} \quad L^2(\Omega_T) \quad \text{as } \sigma \searrow 0.$$

for $l \in \{\oplus, \ominus\}$. The set $\{\mu_\sigma^\oplus \in L^2(\Omega_T) \mid \mu_\sigma^\oplus \geq 0 \text{ a.e. in } \Omega_T\}$ is convex and closed and hence weakly closed and we obtain $\mu^\oplus \geq 0$ a.e. in Ω_T. An analogue argumentation gives $\mu^\ominus \geq 0$ a.e. in Ω_T. Furthermore the energy estimate (2.9) gives

$$\int_\Omega (\psi_\oplus^\gamma(y_\sigma) + \psi_\ominus^\gamma(y_\sigma))\, dx \leq C\sigma, \tag{2.11}$$

for almost all $t \in [0,T]$. Since $y_\sigma \to y$ a.e. in Ω_T we obtain from Fatou's Lemma

$$\begin{aligned} \int_\Omega (\psi_\oplus^\gamma(y) + \psi_\ominus^\gamma(y))\, dx &= \int_\Omega \liminf_{\sigma\to 0} (\psi_\oplus^\gamma(y_\sigma) + \psi_\ominus^\gamma(y_\sigma))\, dx \\ &\leq \liminf_{\sigma\to 0} \int_\Omega (\psi_\oplus^\gamma(y_\sigma) + \psi_\ominus^\gamma(y_\sigma))\, dx \\ &\leq \lim_{\sigma\to 0} C\sigma = 0, \end{aligned}$$

and we obtain $(\psi_\oplus^\gamma(y) + \psi_\ominus^\gamma(y)) = 0$ a.e. in Ω_T and hence $|y| \leq 1$ a.e. in Ω_T. In addition using the monotonicity of $(\psi_\oplus^\gamma)'$ and $(\psi_\oplus^\gamma)'(1) = 0$ we obtain

$$\mu_\sigma^\oplus(y_\sigma - 1) = \frac{1}{\sigma}(\psi_\oplus^\gamma)'(y_\sigma)(y_\sigma - 1) = \frac{1}{\sigma}[(\psi_\oplus^\gamma)'(y_\sigma) - (\psi_\oplus^\gamma)'(1)](y_\sigma - 1) \geq 0.$$

Since $y_\sigma \to y$ strongly in $L^2(\Omega_T)$ and $\mu_\sigma^\oplus \to \mu^\oplus$ weakly in $L^2(\Omega_T)$ we get

$$\int\limits_{\Omega_T} \mu^\oplus (y-1)dxdt = \lim_{\sigma\to 0} \int\limits_{\Omega_T} \mu_\sigma^\oplus (y_\sigma - 1)dxdt \geq 0.$$

Since $(y-1) \leq 0$ a.e. in Ω_T and $\mu^\oplus \geq 0$ a.e. in Ω_T we hence deduce

$$\mu^\oplus (y-1) = 0 \text{ a.e. in } \Omega_T.$$

An analogue argumentation gives $\mu^\ominus(y+1) = 0$ a.e. in Ω_T. It remains to show uniqueness. Assume that there are two solutions $(y_i, \mu_i^\oplus, \mu_i^\ominus)$, $i = 1, 2$. Defining $\overline{y} := y_1 - y_2, \overline{\mu^l} := \mu_1^l - \mu_2^l$ for $l \in \{\oplus, \ominus\}$ and multiplying the difference of the equation (2.1) for y_1 and y_2 with $\overline{y}$ gives after integration

$$\varepsilon \frac{d}{dt} \|\overline{y}\|_{L^2(\Omega)}^2 + \varepsilon \|\nabla \overline{y}\|_{L^2(\Omega)}^2 + \frac{1}{\varepsilon} \int\limits_\Omega \overline{\mu^\oplus}\,\overline{y} dx - \frac{1}{\varepsilon} \int\limits_\Omega \overline{\mu^\ominus}\,\overline{y} dx = \frac{1}{\varepsilon} \|\overline{y}\|_{L^2(\Omega)}^2.$$

The complementary conditions (2.4)-(2.5) imply that the terms $\overline{\mu^\oplus}\,\overline{y}$ and $-\overline{\mu^\ominus}\,\overline{y}$ are non-negative. We hence deduce

$$\varepsilon \frac{d}{dt} \|\overline{y}\|_{L^2(\Omega)}^2 + \varepsilon \|\nabla \overline{y}\|_{L^2(\Omega)}^2 \leq \frac{1}{\varepsilon} \|\overline{y}\|_{L^2(\Omega)}^2.$$

A Gronwall argument now gives uniqueness of y. □

By virtue of Lemma 2 and Lemma 3 we can reformulate our overall optimization problem $(\mathcal{P})$ as a mathematical program with complementarity constraints (MPCC).

$$(\mathcal{CP}) \quad \begin{cases} \min & J(y,u), \\ \text{over} & (y,u) \in \mathcal{V} \times L^2(\Omega_T), \\ \text{s.t.} & \varepsilon \partial_t y - \varepsilon \Delta y + \frac{1}{\varepsilon}\psi_0'(y) + \frac{1}{\varepsilon}\mu^\oplus - \frac{1}{\varepsilon}\mu^\ominus = u \quad \text{a.e. in } \Omega_T, \\ & y(0) = y_0 \text{ a.e. in } \Omega, \quad n \cdot \nabla y = 0 \quad \text{a.e. on } \Gamma_T, \\ & |y| \leq 1 \quad \text{a.e. in } \Omega_T, \\ & \mu^\oplus (y-1) = 0, \mu^\ominus (y+1) = 0 \quad \text{a.e. in } \Omega_T, \\ & \mu^\oplus \geq 0, \mu^\ominus \geq 0 \quad \text{a.e. in } \Omega_T. \end{cases}$$

3 Penalized optimal control problem

For every $\sigma > 0$ we define the penalized optimal control problem by

$$(\mathcal{CP})_\sigma \qquad \begin{cases} \min & J(y,u), \\ \text{over} & (y,u) \in \mathcal{V} \times L^2(\Omega_T), \\ \text{s.t.} & (2.7) - (2.8). \end{cases}$$

3.1 Existence of an optimal control

Definition 1. *Based on Lemma 2, we introduce the control-to-state operator* $S_\sigma : L^2(\Omega_T) \to \mathcal{V}$*, where* $y_\sigma := S_\sigma(u_\sigma)$ *denotes the solution of (2.7)-(2.8) associated to* u_σ*.*

Lemma 4. *Let* $u_\sigma^i \in L^2(\Omega_T)$ *and* $y_\sigma^i = S_\sigma(u_\sigma^i) \in \mathcal{V}$ $(i = 1, 2)$*, where* $\sigma > 0$*. The following stability estimate holds:*

$$\|y_\sigma^1 - y_\sigma^2\|_{\mathcal{V}} \leq C\|u_\sigma^1 - u_\sigma^2\|_{L^2(\Omega_T)}. \tag{3.1}$$

Proof. First we remark that $\tilde{y}_\sigma := y_\sigma^1 - y_\sigma^2$ satisfies the following initial-boundary value problem:

$$\varepsilon\partial_t\tilde{y}_\sigma - \varepsilon\Delta\tilde{y}_\sigma - \frac{1}{\varepsilon}\tilde{y}_\sigma + \frac{1}{\varepsilon\sigma}\sum_{l=\oplus}^{\ominus}[(\psi_l^\gamma)'(y_\sigma^1) - (\psi_l^\gamma)'(y_\sigma^2)] = \tilde{u}_\sigma \quad \text{in } \Omega_T,$$
$$\tilde{y}_\sigma(0) = 0 \ \text{ in } \Omega, \qquad n \cdot \nabla\tilde{y}_\sigma = 0 \quad \text{on } \Gamma_T.$$

Testing the differential equation by $\tilde{y}_\sigma, \partial_t\tilde{y}_\sigma$ and $-\Delta\tilde{y}_\sigma$ and using the Lipschitz continuity of $(\psi_l^\gamma)', l \in \{\oplus, \ominus\}$, and applying analogue techniques like in the proof of Lemma 2 we get the desired result. □

Theorem 2. *The penalized optimal control problem* $(\mathcal{CP})_\sigma$ *has at least a minimizer.*

Proof. For every $\sigma > 0$ let

$$\mathcal{D}_\sigma := \{(y_\sigma, u_\sigma) \in \mathcal{V} \times L^2(\Omega_T) : (y_\sigma, u_\sigma) \text{ satisfy } (2.7) - (2.8)\}$$

denote the feasible set of $(\mathcal{CP})_\sigma$. Let $\widetilde{u_\sigma} \in L^2(\Omega_T)$ be arbitrary but fixed and $y_\sigma(\widetilde{u_\sigma}) \in \mathcal{V}$ be the solution of (2.7)-(2.8) given by Lemma 2. Then $(y_\sigma(\widetilde{u_\sigma}), \widetilde{u_\sigma}) \in \mathcal{D}_\sigma$. Hence the feasible set is nonempty. Furthermore, the cost functional J is bounded from below. Now let $\{(y_{\sigma,k}, u_{\sigma,k})\} \subset \mathcal{D}_\sigma$ be a minimizing sequence such that

$$\lim_{k\to\infty} J(y_{\sigma,k}, u_{\sigma,k}) = \inf_{(y_\sigma,u_\sigma)\in\mathcal{D}_\sigma} J(y_\sigma, u_\sigma) := d < \infty.$$

Then, we get

$$\begin{array}{lllll} u_{\sigma,k} & \text{bounded} & \text{in} & L^2(\Omega_T) & \text{uniformly in} & k, \\ y_{\sigma,k} & \text{bounded} & \text{in} & L^2(\Omega_T) & \text{uniformly in} & k, \\ y_{\sigma,k}(T) & \text{bounded} & \text{in} & L^2(\Omega) & \text{uniformly in} & k. \end{array}$$

Moreover by using Lemma 2 it follows that $\{y_{\sigma,k}\}$ is bounded in $\mathcal{V}$ uniformly in k. Hence, there exist

$$(\overline{y_\sigma}, \overline{y_\sigma(T)}, \overline{u_\sigma}) \in \mathcal{V} \times L^2(\Omega) \times L^2(\Omega_T)$$

such that on a subsequence (denoted the same) $u_{\sigma,k} \to \overline{u_\sigma}$ weakly in $L^2(\Omega_T)$ and as $k \nearrow \infty$

$$\begin{array}{llll}
y_{\sigma,k} & \longrightarrow \overline{y_\sigma} & \text{weakly} & \text{in } L^2(0,T;H^2(\Omega)), \\
y_{\sigma,k} & \longrightarrow \overline{y_\sigma} & \text{weakly} & \text{in } H^1(\Omega_T), \\
y_{\sigma,k} & \longrightarrow \overline{y_\sigma} & \text{strongly} & \text{in } L^2(\Omega_T), \\
y_{\sigma,k}(T) & \longrightarrow \overline{y_\sigma(T)} & \text{weakly} & \text{in } L^2(\Omega), \\
y_{\sigma,k} & \longrightarrow \overline{y_\sigma} & \text{weakly-star} & \text{in } L^\infty(0,T;H^1(\Omega)), \\
y_{\sigma,k} & \longrightarrow \overline{y_\sigma} & \text{a.e.} & \text{in } \Omega_T.
\end{array}$$

Because of the Lipschitz continuity of $(\psi_l^\gamma)', l \in \{\oplus, \ominus\}$, we have as $k \nearrow \infty$

$$\mu^l_{\sigma,k} \longrightarrow \overline{\mu^l_\sigma} \quad \text{strongly} \quad \text{in } L^2(\Omega_T),$$

for $l \in \{\oplus, \ominus\}$. Therefore,

$$\begin{aligned}
\varepsilon\partial_t\overline{y_\sigma} - \varepsilon\Delta\overline{y_\sigma} + \frac{1}{\varepsilon}\psi_0'(\overline{y_\sigma}) + \frac{1}{\varepsilon}\overline{\mu^\oplus_\sigma} - \frac{1}{\varepsilon}\overline{\mu^\ominus_\sigma} = \overline{u_\sigma} \quad & \text{in } \Omega_T, \\
\overline{y_\sigma(0)} = y_0, \quad n \cdot \nabla\overline{y_\sigma} = 0 \quad & \text{a.e. on } \Gamma_T.
\end{aligned}$$

The weakly lower semi-continuity of J finally yields

$$J(\overline{y_\sigma}, \overline{u_\sigma}) \le \lim_{k\to\infty} J(y_{\sigma,k}, u_{\sigma,k}) = d.$$

Hence $(\overline{y_\sigma}, \overline{u_\sigma})$ is a minimizer of $(\mathcal{CP})_\sigma$. □

As far as globally optimal points are concerned, we find that solutions of the penalized optimal control problem $(\mathcal{CP})_\sigma$ converge to a solution of the problem $(\mathcal{CP})$, as the following theorem shows.

Theorem 3. *Denote by (y_σ, u_σ) the minimizers of the penalized optimal control problems $(\mathcal{CP})_\sigma$. Then there exists a minimizer $(\overline{y}, \overline{u}) \in \mathcal{V} \times L^2(\Omega_T)$ for the problem $(\mathcal{CP})$ such that on a subsequence of minimizers (still denoted by (y_σ, u_σ)) as $\sigma \searrow 0$*

$$\begin{array}{llll}
u_\sigma & \longrightarrow \overline{u} & \textit{strongly} & \textit{in } L^2(\Omega_T), \\
y_\sigma & \longrightarrow \overline{y} & \textit{weakly} & \textit{in } L^2(0,T;H^2(\Omega)), \\
y_\sigma & \longrightarrow \overline{y} & \textit{weakly} & \textit{in } H^1(\Omega_T), \\
y_\sigma & \longrightarrow \overline{y} & \textit{strongly} & \textit{in } L^2(\Omega_T), \\
y_\sigma & \longrightarrow \overline{y} & \textit{weakly-star} & \textit{in } L^\infty(0,T;H^1(\Omega)), \\
y_\sigma & \longrightarrow \overline{y} & \textit{a.e.} & \textit{in } \Omega_T.
\end{array} \tag{3.2}$$

Furthermore we have

$$y_\sigma(T) \longrightarrow \overline{y(T)} \quad \textit{strongly} \quad \textit{in} \quad L^2(\Omega). \tag{3.3}$$

Proof. Let $\tilde{u} \in L^2(\Omega_T)$ be fixed, and denote by $y_\sigma(\tilde{u}) \in \mathcal{V}$ the solution to (2.7)-(2.8). Hence, the estimate

$$J(y_\sigma, u_\sigma) \leq J(y_\sigma(\tilde{u}), \tilde{u}) \tag{3.4}$$

holds true for every $\sigma > 0$. The boundedness of $y_\sigma(\tilde{u})$ given by Lemma 2 implies the boundedness of $\{J(y_\sigma(\tilde{u}), \tilde{u})\}$. Using (3.4), we conclude that also $\{u_\sigma\}$ is uniformly bounded in $L^2(\Omega_T)$, and there exists $\overline{u} \in L^2(\Omega_T)$ such that on a subsequence (also denoted by $\{u_\sigma\}$) as $\sigma \searrow 0$

$$u_\sigma \longrightarrow \overline{u} \quad \text{weakly in} \quad L^2(\Omega_T).$$

Then by Lemma 3 there exists $\overline{y} \in \mathcal{V}$ and a subsequence still denoted by $\{y_\sigma\}$ such that (3.2) holds. Moreover applying interpolation arguments, it can be shown that $L^2(0,T;H^2(\Omega)) \cap H^1(0,T;L^2(\Omega))$ continuously embeds into $C([0,T];H^1(\Omega))$. By Rellich-Kondrachov theorem it follows that $H^1(\Omega)$ is compactly embedded in $L^2(\Omega)$. Hence (3.3) follows. Because of Lemma 3 the limit element $(\overline{y}, \overline{u})$ is feasible for $(\mathcal{CP})$. Now let $(y^*, u^*) \in \mathcal{V} \times L^2(\Omega_T)$ be a minimizer of $(\mathcal{CP})$. Due to the lower semi-continuity of the norm, (3.4) and Lemma 3, we find that

$$\begin{aligned} J(y^*, u^*) \leq J(\overline{y}, \overline{u}) &\leq \liminf_{\sigma \searrow 0} J(y_\sigma, u_\sigma) \leq \limsup_{\sigma \searrow 0} J(y_\sigma, u_\sigma) \\ &\leq \limsup_{\sigma \searrow 0} J(y_\sigma(u^*), u^*) = J(y^*, u^*). \end{aligned}$$

Therefore, $(\overline{y}, \overline{u})$ is optimal for $(\mathcal{CP})$. Furthermore, we see that as $\sigma \searrow 0$

$$J(y_\sigma, u_\sigma) \to J(\overline{y}, \overline{u}),$$

hence $\|u_\sigma\|_{L^2} \to \|\overline{u}\|_{L^2}$, which together with the weak convergence of $\{u_\sigma\}$ implies strong convergence of $\{u_\sigma\}$ in $L^2(\Omega_T)$. □

3.2 Analysis of the linearized state system

For the derivation of first-order optimality conditions, it is essential to show the Fréchet-differentiability of the control-to-state operator, mapping u_σ to

y_σ (see Subsection 3.3.1 below). Suppose $u_\sigma \in L^2(\Omega_T)$ and consider a perturbation $\delta u_\sigma \in L^2(\Omega_T)$. In preparation of the corresponding theorem, we now consider the following linearized version of (2.7)-(2.8):

$$\varepsilon \partial_t y_\sigma^* - \varepsilon \Delta y_\sigma^* + \frac{1}{\varepsilon}(\psi_\sigma^\gamma)''(y_\sigma) y_\sigma^* = \delta u_\sigma \quad \text{in } \Omega_T, \tag{3.5}$$

$$y_\sigma^*(0) = 0 \text{ in } \Omega, \qquad n \cdot \nabla y_\sigma^* = 0 \quad \text{on } \Gamma_T, \tag{3.6}$$

with given functions $y_\sigma, \delta u_\sigma$. Later on $y_\sigma = S_\sigma(u_\sigma)$ will be the solution of the nonlinear state system (2.7)-(2.8) associated to reference control u_σ. In the following we will show that (3.5)-(3.6) admits a solution $y_\sigma^* \in W(0,T)$. This result is then used to establish Fréchet-differentiability of the solution operator S_σ associated to (2.7)-(2.8).

Lemma 5. *Problem (3.5)-(3.6) admits a unique solution $y_\sigma^* \in W(0,T)$.*

Proof. Since for every $\sigma > 0$ which is arbitrary but fixed $(\psi_\sigma^\gamma)''(y_\sigma) \in L^\infty(\Omega_T)$, see (2.6), the existence of a unique weak solution $y_\sigma^* \in W(0,T)$ to (2.7)-(2.8) is a classical result (see [14], Chapter 3, Theorem 5.1). □

3.3 First-order necessary optimality conditions

We start the derivation of first-order conditions with the Fréchet-differentiability of the control-to-state operator S_σ, which is one of the crucial points of the first-order analysis for $(\mathcal{CP})_\sigma$. However, using the analysis for the linearized equation, presented in the previous subsection, yields the desired differentiability of S_σ. Afterwards, we reformulate the derivative of the objective functional by introducing an adjoint PDE system which leads to the first-order necessary optimality conditions in form of a Karush-Kuhn-Tucker (KKT) type optimality system.

3.3.1 Differentiability of the control-to-state mapping

Theorem 4. *Let dim $\Omega \leq 3$. The control-to-state operator S_σ is Fréchet-differentiable from $L^2(\Omega_T)$ to $W(0,T)$. The derivative has the form*

$$S_\sigma'(u_\sigma)\delta u_\sigma = y_\sigma^*,$$

where $y_\sigma^* \in W(0,T)$ is the weak solution of the linearized problem (3.5)-(3.6) in $y_\sigma := S_\sigma(u_\sigma)$.

Proof. We have to prove

$$S_\sigma(u_\sigma + \delta u_\sigma) - S_\sigma(u_\sigma) = D_\sigma\, \delta u_\sigma + r(u_\sigma, \delta u_\sigma),$$

where $D_\sigma : L^2(\Omega_T) \to W(0,T)$ is a linear and continuous operator and

$$\frac{\|r(u_\sigma, \delta u_\sigma)\|_{W(0,T)}}{\|\delta u_\sigma\|_{L^2(Q_T)}} \longrightarrow 0 \quad \text{if } \|\delta u_\sigma\|_{L^2(Q_T)} \longrightarrow 0.$$

Hence, we have $S'_\sigma(u_\sigma) = D_\sigma$. By (1.4) we have $y_\sigma \in L^2(0,T; L^\infty(\Omega))$. Due to [17], §4.3, the Nemytskii-operator (still denoted by $(\psi_\sigma^\gamma)'$) associated to $(\psi_\sigma^\gamma)'$ is Fréchet differentiable from $L^2(0,T;L^\infty(\Omega))$ to $L^2(0,T;L^\infty(\Omega))$. It's derivative is given by

$$(\psi_\sigma^\gamma)''(y_\sigma) : L^2(0,T;L^\infty(\Omega)) \to L^\infty(\Omega_T).$$

Hence, we get

$$(\psi_\sigma^\gamma)'(y_{\sigma,\delta}) - (\psi_\sigma^\gamma)'(y_\sigma) = (\psi_\sigma^\gamma)''(y_\sigma)(y_{\sigma,\delta} - y_\sigma) + r_{y_{\sigma,\delta},y_\sigma},$$

where $y_{\sigma,\delta} = S_\sigma(u_\sigma + \delta u_\sigma)$ and $r_{y_{\sigma,\delta},y_\sigma}$ is the remainder with the form

$$r_{y_{\sigma,\delta},y_\sigma} = \int_0^1 ((\psi_\sigma^\gamma)''(y_\sigma + s(y_{\sigma,\delta} - y_\sigma)) - (\psi_\sigma^\gamma)''(y_\sigma))\, ds\, (y_{\sigma,\delta} - y_\sigma).$$

We estimate $r_{y_{\sigma,\delta},y_\sigma}$ by

$$|r_{y_{\sigma,\delta},y_\sigma}(t,x)| \le C \int_0^1 s|y_{\sigma,\delta} - y_\sigma| ds\, |y_{\sigma,\delta} - y_\sigma| \le C\|(y_{\sigma,\delta} - y_\sigma)(t,\cdot)\|^2_{L^\infty(\Omega)}.$$

Hence, we have

$$\frac{\|r_{y_{\sigma,\delta},y_\sigma}\|_{L^2(0,T;L^\infty(\Omega))}}{\|y_{\sigma,\delta} - y_\sigma\|_{L^2(0,T;L^\infty(\Omega))}} \to 0 \text{ if } \|y_{\sigma,\delta} - y_\sigma\|_{L^2(0,T;L^\infty(\Omega))} \to 0.$$

Therefore we have $y_{\sigma,\delta} - y_\sigma = y_\sigma^* + \hat{y}_\sigma$ with a solution $y_\sigma^* \in W(0,T)$ of (3.5)-(3.6) and a remainder $\hat{y}_\sigma \in W(0,T)$, which satisfy

$$\varepsilon \partial_t \hat{y}_\sigma - \gamma\varepsilon\Delta\hat{y}_\sigma + \frac{1}{\varepsilon}(\psi_\sigma^\gamma)''(y_\sigma)\hat{y}_\sigma = -\frac{1}{\varepsilon} r_{y_{\sigma,\delta},y_\sigma} \quad \text{in } \Omega_T, \tag{3.7}$$

$$\hat{y}_\sigma(0) = 0 \quad \text{in } \Omega, \quad n \cdot \nabla \hat{y}_\sigma = 0 \quad \text{on } \Gamma_T. \tag{3.8}$$

The existence of a weak solution $\hat{y}_\sigma \in W(0,T)$ can be proven in an analogue way as for the system (3.5)-(3.6). By Lemma 4 and (1.3) we have

$$\|y_{\sigma,\delta} - y_\sigma\|_{L^2(0,T;L^\infty(\Omega))} \leq \|y_{\sigma,\delta} - y_\sigma\|_{\mathcal{V}} \leq L\|\delta u_\sigma\|_{L^2(\Omega_T)}.$$

Besides we have

$$\begin{aligned}\frac{\|r_{y_{\sigma,\delta},y_\sigma}\|_{L^2(0,T;L^\infty(\Omega))}}{\|\delta u_\sigma\|_{L^2(\Omega_T)}} &= \frac{\|r_{y_{\sigma,\delta},y_\sigma}\|_{L^2(0,T;L^\infty(\Omega))}}{\|y_{\sigma,\delta} - y_\sigma\|_{L^2(0,T;L^p(\Omega))}} \frac{\|y_{\sigma,\delta} - y_\sigma\|_{L^2(0,T;L^\infty(\Omega))}}{\|\delta u_\sigma\|_{L^2(\Omega_T)}} \\ &\leq \frac{\|r_{y_{\sigma,\delta},y_\sigma}\|_{L^2(0,T;L^\infty(\Omega))}}{\|y_{\sigma,\delta} - y_\sigma\|_{L^2(0,T;L^\infty(\Omega))}} L.\end{aligned}$$

Hence we have $\|r_{y_{\sigma,\delta},y_\sigma}\|_{L^2(0,T;L^\infty(\Omega))} = o(\|\delta u_\sigma\|_{L^2(\Omega_T)})$. By virtue of existence of a solution $\hat{y}_\sigma \in W(0,T)$ to (3.7)-(3.8) we get

$$\|\hat{y}_\sigma\|_{W(0,T)} = o(\|\delta u_\sigma\|_{L^2(\Omega_T)}).$$

We denote the map $\delta u_\sigma \to y_\sigma^*$ by D_σ, which is linear and continuous. Finally we end up with

$$S_\sigma(u_\sigma + \delta u_\sigma) - S_\sigma(u_\sigma) = y_{\sigma,\delta} - y_\sigma = D_\sigma\, \delta u_\sigma + r(u_\sigma, \delta u_\sigma),$$

where $r(u_\sigma, \delta u_\sigma) = \hat{y}_\sigma$ provides the claimed properties. □

3.3.2 Optimality conditions

Now we are in the position to state the first-order necessary optimality conditions for $(\mathcal{CP})_\sigma$. Defining

$$\lambda_\sigma^\oplus := \frac{1}{\sigma}(\psi_\oplus^\gamma)''(y_\sigma)p_\sigma, \quad \lambda_\sigma^\ominus := \frac{1}{\sigma}(\psi_\ominus^\gamma)''(y_\sigma)p_\sigma,$$

we have:

Theorem 5. *Let $\sigma > 0$, $n \leq 3$,* ***(H0)*** *and* ***(H1)*** *hold. Then there exist functions $(y_\sigma, u_\sigma, p_\sigma) \in \mathcal{V} \times L^2(\Omega_T) \times W(0,T)$ such that the following first order optimality system holds*

$$\varepsilon\partial_t y_\sigma - \varepsilon\Delta y_\sigma + \frac{1}{\varepsilon}\psi_0'(y_\sigma) + \frac{1}{\varepsilon}\mu_\sigma^\oplus - \frac{1}{\varepsilon}\mu_\sigma^\ominus = u_\sigma \quad \text{in } \Omega_T, \tag{3.9}$$

$$y_\sigma(0) = y_0 \ \text{in } \Omega, \qquad n \cdot \nabla y_\sigma = 0 \quad \text{on } \Gamma_T, \tag{3.10}$$

$$\frac{\nu_u}{\varepsilon}u_\sigma - p_\sigma = 0 \quad \text{in } \Omega_T, \tag{3.11}$$

$$-\varepsilon\partial_t p_\sigma - \varepsilon\Delta p_\sigma + \frac{1}{\varepsilon}\psi_0''(y_\sigma)p_\sigma + \frac{1}{\varepsilon}\lambda_\sigma^\oplus + \frac{1}{\varepsilon}\lambda_\sigma^\ominus = \nu_d(y_\sigma - y_d) \quad \text{in } \Omega_T, \tag{3.12}$$

$$p_\sigma(T,\cdot) = \nu_T(y_\sigma(T,\cdot) - y_T) \ \text{in } \Omega, \qquad n \cdot \nabla p_\sigma = 0 \quad \text{on } \Gamma_T. \tag{3.13}$$

Proof. Let (u_σ, y_σ) be an optimal solution of $(\mathcal{CP})_\sigma$. From Theorem 4 we know that S_σ is Fréchet-differentiable from $L^2(\Omega_T)$ to $W(0,T)$. Therefore

$$\frac{d}{d\theta} J(S_\sigma(u_\sigma + \theta\,\delta u_\sigma), u_\sigma + \theta\,\delta u_\sigma)|_{\theta=0} = \\ = \nu_T \int_\Omega (y_\sigma(T,\cdot) - y_T) y_\sigma^*(T,\cdot) dx + \nu_d \int_{\Omega_T} (y_\sigma - y_d) y_\sigma^* dxdt + \frac{\nu_u}{\varepsilon} \int_{\Omega_T} u_\sigma \delta u_\sigma dxdt, \tag{3.14}$$

where $y_\sigma^* = S'_\sigma(u_\sigma)\delta u_\sigma$ is the weak solution of the linearized problem (3.5)-(3.6) in $y_\sigma := S_\sigma(u_\sigma)$, see Theorem 4.

We transform (3.14) into another form by introducing the formally adjoint system to (3.5)-(3.6). The adjoint variable p_σ is the solution of the following adjoint problem:

$$-\varepsilon\partial_t p_\sigma - \varepsilon\Delta p_\sigma + \frac{1}{\varepsilon}(\psi_\sigma^\gamma)''(y_\sigma)p_\sigma = \nu_d(y_\sigma - y_d) \quad \text{in } \Omega_T, \tag{3.15}$$

$$n \cdot \nabla p_\sigma = 0 \quad \text{on } \Gamma_T, \tag{3.16}$$

$$p_\sigma(T,\cdot) = \nu_T(y_\sigma(T,\cdot) - y_T) \quad \text{in } \Omega. \tag{3.17}$$

We apply Lemma 5 to prove existence of solutions to (3.15)-(3.17). We introduce the transformation $\tau := T - t$ and $p_\sigma(t) := \tilde{p}_\sigma(\tau)$. Hence, we get the following system

$$\varepsilon\partial_\tau \tilde{p}_\sigma - \varepsilon\Delta\tilde{p}_\sigma + \frac{1}{\varepsilon}\psi_\sigma''(y_\sigma)\tilde{p}_\sigma = \nu_d(y_\sigma - y_d) \quad \text{in } \Omega_T \tag{3.18}$$

$$n \cdot \nabla\tilde{p}_\sigma = 0, \quad \text{on } \Gamma_T \tag{3.19}$$

$$\tilde{p}_\sigma(0,\cdot) = \nu_T(y_\sigma(T,\cdot) - y_T) \quad \text{in } \Omega. \tag{3.20}$$

Arguing as in the proof of Lemma 5 we get a solution $\tilde{p}_\sigma \in W(0,T)$, hence $p_\sigma \in W(0,T)$. To prove (3.11) we test (3.12) by y_σ^*, which is the solution of the linearized problem (3.5)-(3.6) in $y_\sigma := S_\sigma(u_\sigma)$. Integration by parts gives

$$\int_{\Omega_T} p_\sigma \delta u_\sigma dxdt = \frac{\nu_u}{\varepsilon} \int_{\Omega_T} u_\sigma \delta u_\sigma dxdt.$$

□

4 Optimality conditions for the limit problem

For the rest of the paper we make use of the following assumptions:

(OA) Let $\{u_\sigma\}$ be bounded in $L^2(\Omega_T) \cap H^1(0,T;L^2(\Omega))$ uniformly in $\sigma > 0$ and $u_0 \in L^2(\Omega)$

Lemma 6. *Let dim $\Omega \leq 3$ and **(OA)** hold. Furthermore $y_0 \in H^2(\Omega)$ with $|y_0| \leq 1$ a.e. in Ω and for every $\sigma > 0$, let $(y_\sigma, u_\sigma, p_\sigma) \in \mathcal{V} \times L^2(\Omega_T) \times W(0,T)$ be a solution of the optimality system (3.9)-(3.13). Then the following estimates hold*

$$\begin{array}{llll} 1.) & y_\sigma & \textit{uniformly bounded} & \textit{in} \ \ \mathcal{V}, \\ 2.) & y_\sigma & \textit{uniformly bounded} & \textit{in} \ \ H^1(0,T;H^1(\Omega)) \cap W^{1,\infty}(0,T;L^2(\Omega)), \\ 3.) & \mu_\sigma^\ominus & \textit{uniformly bounded} & \textit{in} \ \ L^2(\Omega_T), \\ 4.) & \mu_\sigma^\oplus & \textit{uniformly bounded} & \textit{in} \ \ L^2(\Omega_T), \\ 5.) & p_\sigma & \textit{uniformly bounded} & \textit{in} \ \ L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega)), \\ 6.) & \partial_t p_\sigma & \textit{uniformly bounded} & \textit{in} \ \ W(0,T)^*, \\ 7.) & \lambda_\sigma^\ominus + \lambda_\sigma^\oplus & \textit{uniformly bounded} & \textit{in} \ \ W(0,T)^*, \\ 8.) & \lambda_\sigma^\ominus & \textit{uniformly bounded} & \textit{in} \ \ W(0,T)^*, \\ 9.) & \lambda_\sigma^\oplus & \textit{uniformly bounded} & \textit{in} \ \ W(0,T)^*. \end{array} \tag{4.1}$$

Proof. 1.), 3.) and 4.) are direct consequences of Lemma 2. To prove 2.) we formally differentiate (3.9) with respect to time and obtain

$$\varepsilon\partial_{tt}y_\sigma - \varepsilon\Delta(\partial_t y_\sigma) + \frac{1}{\varepsilon\sigma}(\Psi_\oplus^\gamma + \Psi_\ominus^\gamma)''(y_\sigma)\partial_t y_\sigma = \frac{1}{\varepsilon}\partial_t y_\sigma + \partial_t u_\sigma \quad \text{in } \Omega_T, \tag{4.2}$$

$$y_\sigma(0) = y_0 \ \text{ in } \Omega, \qquad n \cdot \nabla(\partial_t y_\sigma) = 0 \quad \text{on } \Gamma_T. \tag{4.3}$$

Now formally testing (4.2) by $\partial_t y_\sigma$ and noting that $(\Psi_\oplus^\gamma + \Psi_\ominus^\gamma)''(y_\sigma) \geq 0$ it follows

$$\frac{\varepsilon}{2}\frac{d}{dt}\|\partial_t y_\sigma\|^2_{L^2(\Omega)} + \varepsilon\|\nabla(\partial_t y_\sigma)\|^2_{L^2(\Omega)} \leq C(\varepsilon)(\|\partial_t y_\sigma\|^2_{L^2(\Omega)} + \|\partial_t u_\sigma\|^2_{L^2(\Omega)}). \tag{4.4}$$

Integrating with respect to t, using **(OA)** and 1.) we get

$$\frac{\varepsilon}{2}\|\partial_t y_\sigma(t)\|^2_{L^2(\Omega)} + \|\nabla(\partial_t y_\sigma)\|^2_{L^2(\Omega)} \leq \frac{C(\varepsilon)}{2}\|\partial_t y_0\|^2_{L^2(\Omega)}. \tag{4.5}$$

Using (3.9)-(3.10) and noting that $(\psi_\oplus^\gamma)'(y_0) = (\psi_\ominus^\gamma)'(y_0) = 0$ we can estimate the right hand side of (4.5) by

$$\|\partial_t y_0\|^2_{L^2(\Omega)} \leq C(\|\Delta y_0\|^2_{L^2(\Omega)} + \frac{1}{\varepsilon^2}\|y_0\|^2_{L^2(\Omega)} + \|u_0\|^2_{L^2(\Omega)}). \tag{4.6}$$

Inserting (4.6) into (4.5) and using **(OA)**, $y_0 \in H^2(\Omega)$ with $|y_0| \leq 1$ a.e. in Ω we get 2.) We have to remark here that the previous calculations can be done rigorously be using standard Galerkin technique, see e.g. [7].

Now we prove 5.). We introduce the transformation $\tau := T - t$ and $p_\sigma(t) := e^{\alpha\tau}\tilde{p}_\sigma(\tau)$. Hence, we get the following system

$$\varepsilon\partial_\tau\tilde{p}_\sigma - \varepsilon\Delta\tilde{p}_\sigma + \frac{1}{\varepsilon}[(\psi_\sigma^\gamma)''(y_\sigma) + \alpha\varepsilon^2]\tilde{p}_\sigma = \nu_d e^{-\alpha\tau}(y_\sigma - y_d) \quad \text{in } \Omega_T, \tag{4.7}$$

$$n \cdot \nabla\tilde{p}_\sigma = 0 \quad \text{on } \Gamma_T, \tag{4.8}$$

$$\tilde{p}_\sigma(0,\cdot) = \nu_T(y_\sigma(T,\cdot) - y_T) \quad \text{in } \Omega. \tag{4.9}$$

Now testing (4.7) by $\tilde{p}_\sigma$ and choosing $\alpha > 0$ so that $(\psi_\sigma^\gamma)''(y_\sigma) + \alpha\varepsilon^2 \geq C_0 > 0$ we get by standard calculations the existence of a constant $C(\tau) > 0$, independent of σ, such that

$$\frac{\varepsilon}{2}\frac{d}{dt}\|\tilde{p}_\sigma\|^2_{L^2(\Omega)} + \varepsilon\|\nabla\tilde{p}_\sigma\|^2_{L^2(\Omega)} \leq C(\tau)\|\tilde{p}_\sigma\|^2_{L^2(\Omega)}.$$

Now by a Gronwall argument we get $\|\tilde{p}_\sigma\|_{L^2(0,T;H^1(\Omega))} \leq C(\tau)$, hence

$$\|p_\sigma\|_{L^2(0,T;H^1(\Omega))\cap L^\infty(0,T;L^2(\Omega))} \leq C.$$

To prove 6.) let $v \in W(0,T)$. Using integration by parts we obtain

$$\langle\partial_t p_\sigma, v\rangle = -\langle\partial_t v, p_\sigma\rangle + \nu_T(y_\sigma(T) - y_T, v(T))_{L^2(\Omega)} - (p_\sigma(0), v(0))_{L^2(\Omega)}.$$

The continuous injection of $W(0,T)$ into $C([0,T];L^2(\Omega))$ yields

$$|\langle\partial_t p_\sigma, v\rangle| \leq \left(\|p_\sigma\|_{L^2(0,T;H^1(\Omega))} + \nu_T\|y_\sigma(T) - y_T\|_{L^2(\Omega)} + \|p_\sigma(0)\|_{L^2(\Omega)}\right)\|v\|_{W(0,T)}.$$

Hence from 1.), 2.) and 5.) we deduce 6.).

The boundedness of $\lambda_\sigma^\oplus + \lambda_\sigma^\ominus$ in $W(0,T)^*$ follows from the adjoint equation (3.12) and 6.). To prove 9.) we define $\Phi^\oplus \in C^\infty(\mathbb{R})$, $0 \leq \Phi^\oplus(r) \leq 1$, $r \in \mathbb{R}$, $\Phi^\oplus \equiv 1$ on $\{r \geq 1\}$, $\Phi^\oplus \equiv 0$ on $\{r \leq 0\}$ and $|(\Phi^\oplus)'| \leq 2$ and get for a $v \in W(0,T)$

$$\|\Phi^\oplus(y_\sigma)v\|_{W(0,T)} \leq C\|v\|_{W(0,T)}. \tag{4.10}$$

We want to prove (4.10). First we have

$$\|\nabla[\Phi^\oplus(y_\sigma)v]\|_{L^2(\Omega_T)} \leq \|(\Phi^\oplus)'(y_\sigma)\nabla y_\sigma\, v\|_{L^2(\Omega_T)} + \|\Phi^\oplus(y_\sigma)\nabla v\|_{L^2(\Omega_T)}.$$

For the first summand on the right hand side of the above inequality we have using the Hölder inequality

$$\|(\Phi^\oplus)'(y_\sigma)\nabla y_\sigma\, v\|_{L^2(\Omega_T)} \leq C\|\nabla y_\sigma\|_{L^\infty(0,T;L^3(\Omega))}\|v\|_{L^2(0,T;L^6(\Omega))}.$$

By (1.3), (1.5), (1.6) and 2.) we get

$$\|(\Phi^{\oplus})'(y_\sigma)\nabla y_\sigma\, v\|_{L^2(\Omega_T)} \leq C\|v\|_{L^2(0,T;H^1(\Omega))}.$$

Furthermore we have

$$\begin{aligned}\|\partial_t[\Phi^{\oplus}(y_\sigma)v]\|_{L^2(0,T;(H^1(\Omega))^*)} \leq \|(\Phi^{\oplus})'(y_\sigma)\partial_t y_\sigma\, v\|_{L^2(0,T;(H^1(\Omega))^*)}+ \\ + \|\Phi^{\oplus}(y_\sigma)\partial_t v\|_{L^2(0,T;(H^1(\Omega))^*)}.\end{aligned}$$

To estimate the first summand on the right hand side of the above inequality we have to use the Hölder inequality for a $\phi \in L^2(0,T;H^1(\Omega))$

$$\int_0^T \int_\Omega (\Phi^{\oplus})'(y_\sigma)\partial_t y_\sigma\, v\, \phi\, dxdt \leq C\|\partial_t y_\sigma\|_{L^\infty(0,T;L^2(\Omega))}\, \|v\|_{L^2(0,T;L^6(\Omega))}\, \|\phi\|_{L^2(0,T;L^3(\Omega))}.$$

By (1.3) and 2.) we estimate

$$\int_0^T \int_\Omega (\Phi^{\oplus})'(y_\sigma)\partial_t y_\sigma\, v\, \phi\, dxdt \leq C\|v\|_{L^2(0,T;H^1(\Omega))}\, \|\phi\|_{L^2(0,T;H^1(\Omega))}$$

and get

$$\begin{aligned}\|(\Phi^{\oplus})'(y_\sigma)\partial_t y_\sigma v\|_{L^2(0,T;(H^1(\Omega))^*)} &= \sup_{\phi\in L^2(0,T;H^1(\Omega))} \frac{\int_0^T \langle(\Phi^{\oplus})'(y_\sigma)\partial_t y_\sigma\, v, \phi\rangle}{\|\phi\|_{L^2(0,T;H^1(\Omega))}} \\ &\leq C\|v\|_{L^2(0,T;H^1(\Omega))}.\end{aligned}$$

In conclusion, the assertion (4.10) is proved. To complete the proof of 9.) we have for a $v \in W(0,T)$

$$\begin{aligned}|\langle\lambda_\sigma^{\oplus}, v\rangle_{W(0,T)^*,W(0,T)}| &= |\langle\lambda_\sigma^{\oplus}, \Phi^{\oplus}(y_\sigma)v\rangle_{W(0,T)^*,W(0,T)}| = \\ &= |\langle\lambda_\sigma^{\oplus} + \lambda_\sigma^{\ominus}, \Phi^{\oplus}(y_\sigma)v\rangle_{W(0,T)^*,W(0,T)}| \leq C\|v\|_{W(0,T)},\end{aligned}$$

where we used 7.) and (4.10) for the last inequality. Hence, we get

$$\|\lambda_\sigma^{\oplus}\|_{W(0,T)^*} \leq C.$$

Analogously by

$$\|\Phi^{\ominus}(y_\sigma)v\|_{W(0,T)} \leq C\|v\|_{W(0,T)},$$

where we have $\Phi^{\ominus} \in C^\infty(\mathbb{R})$, $-1 \leq \Phi^{\ominus}(r) \leq 0$, $r \in \mathbb{R}$, $\Phi^{\ominus} \equiv -1$ on $\{r \leq -1\}$, $\Phi^{\ominus} \equiv 0$ on $\{r \geq 0\}$ and $|(\Phi^{\ominus})'| \leq 2$, we get $\|\lambda_\sigma^{\ominus}\|_{W(0,T)^*} \leq C$. □

Now we can state the main result of this section. Defining the functions

$$[y_\sigma + 1]^{\oplus} := \max(y_\sigma + 1, 0), \qquad [y_\sigma - 1]^{\ominus} := \min(y_\sigma - 1, 0),$$

and the space $\tilde{\mathcal{V}} := \mathcal{V} \cap H^1(0,T;H^1(\Omega)) \cap W^{1,\infty}(0,T;L^2(\Omega))$ we have:

Theorem 6. *Let dim $\Omega \leq 3$ and let $\{y_\sigma, u_\sigma, p_\sigma\}$ be a sequence of solutions of the optimality system (3.9)-(3.13). If **(OA)** holds, then there exist*

$$\{y^*, u^*, p^*\} \in \tilde{\mathcal{V}} \times L^2(\Omega_T) \times L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$$

and a subsequence still denoted by $\{y_\sigma, u_\sigma, p_\sigma\}$ such that as $\sigma \searrow 0$

$$\begin{array}{lllll} y_\sigma & \longrightarrow & y^* & \textit{weakly} & \textit{in} \;\; L^2(0,T;H^2(\Omega)) \cap H^1(\Omega_T), \\ y_\sigma & \longrightarrow & y^* & \textit{weakly} & \textit{in} \;\; H^1(0,T;H^1(\Omega)), \\ y_\sigma & \longrightarrow & y^* & \textit{weakly-star} & \textit{in} \;\; L^\infty(0,T;H^1(\Omega)), \\ y_\sigma & \longrightarrow & y^* & \textit{weakly-star} & \textit{in} \;\; W^{1,\infty}(0,T;L^2(\Omega)), \\ \mu_\sigma^\oplus & \longrightarrow & \mu_*^\oplus & \textit{weakly} & \textit{in} \;\; L^2(\Omega_T), \\ \mu_\sigma^\ominus & \longrightarrow & \mu_*^\ominus & \textit{weakly} & \textit{in} \;\; L^2(\Omega_T), \\ u_\sigma & \longrightarrow & u^* & \textit{weakly} & \textit{in} \;\; L^2(\Omega_T), \\ p_\sigma & \longrightarrow & p^* & \textit{weakly} & \textit{in} \;\; L^2(0,T;H^1(\Omega)), \\ p_\sigma & \longrightarrow & p^* & \textit{weakly-star} & \textit{in} \;\; L^\infty(0,T;L^2(\Omega)), \\ \partial_t p_\sigma & \longrightarrow & \partial_t p^* & \textit{weakly} & \textit{in} \;\; W(0,T)^*, \\ \lambda_\sigma^\oplus & \longrightarrow & \lambda_*^\oplus & \textit{weakly} & \textit{in} \;\; W(0,T)^*, \\ \lambda_\sigma^\ominus & \longrightarrow & \lambda_*^\ominus & \textit{weakly} & \textit{in} \;\; W(0,T)^*. \end{array} \tag{4.11}$$

The limit element $\{y^*, u^*, p^*\}$ satisfies the following optimality system

$$\frac{1}{\varepsilon}\left\langle \lambda_*^\oplus + \lambda_*^\ominus, v \right\rangle_{W(0,T)^*, W(0,T)} + \varepsilon \left\langle p^*, \partial_t v \right\rangle_{L^2(0,T;H^1(\Omega)), L^2(0,T;H^1(\Omega)^*)} +$$
$$+ \varepsilon(\nabla p^*, \nabla v)_{L^2(\Omega_T)} + \nu_d(y^* - y_d, v)_{L^2(\Omega_T)} + \nu_T(y^*(T,\cdot) - y_T, v(T,\cdot))_{L^2(\Omega)} -$$
$$- \frac{1}{\varepsilon}(p^*, v)_{L^2(\Omega_T)} = 0, \qquad \forall v \in W_0(0,T) \tag{4.12}$$

$$\frac{\nu_u}{\varepsilon} u^* - p^* = 0 \quad \textit{in } \Omega_T, \tag{4.13}$$

$$\varepsilon \partial_t y^* - \varepsilon \Delta y^* - \frac{1}{\varepsilon} y^* + \frac{1}{\varepsilon}\mu_*^\oplus - \frac{1}{\varepsilon}\mu_*^\ominus = u^* \quad \textit{a.e. in } \Omega_T, \tag{4.14}$$

$$y^*(0) = y_0 \quad \textit{in } \Omega, \qquad n \cdot \nabla y^* = 0 \quad \textit{a.e. on } \Gamma_T, \tag{4.15}$$

with the complementarity conditions

$$|y^*| < 1 \quad \textit{a.e. in } \Omega_T, \tag{4.16}$$

$$\mu_*^\oplus(y^* - 1) = 0, \mu_*^\ominus(y^* + 1) = 0 \quad \textit{a.e. in } \Omega_T, \tag{4.17}$$

$$\mu_*^\oplus \geq 0, \mu_*^\ominus \geq 0 \quad \textit{a.e. in } \Omega_T, \tag{4.18}$$

$$\lim_{\sigma \searrow 0} (\lambda_\sigma^\ominus, [y_\sigma + 1]^\oplus)_{L^2(\Omega_T)} = 0, \tag{4.19}$$

$$\lim_{\sigma \searrow 0} (\lambda_\sigma^\oplus, [y_\sigma - 1]^\ominus)_{L^2(\Omega_T)} = 0, \tag{4.20}$$

$$\liminf_{\sigma \searrow 0} (\lambda_\sigma^\ominus, p_\sigma)_{L^2(\Omega_T)} \geq 0, \tag{4.21}$$

$$\liminf_{\sigma \searrow 0} (\lambda_\sigma^\oplus, p_\sigma)_{L^2(\Omega_T)} \geq 0, \tag{4.22}$$

where $W_0(0,T) := \{v \in W(0,T) : v(0,\cdot) = 0\}$*. Furthermore, for every* $\omega > 0$*, there exists a subset* $Q_\omega \subset \{(t,x) \in \Omega_T : |y^*(t,x)| < 1\}$ *with* $meas(\{(t,x) \in \Omega_T : |y^*(t,x)| < 1\} \setminus Q_\omega) \leq \omega$*, such that*

$$\lambda_\sigma^\oplus + \lambda_\sigma^\ominus \quad \longrightarrow \quad 0 \quad \textit{uniformly} \quad \textit{in} \quad Q_\omega. \tag{4.23}$$

Proof. The convergence results are direct consequences of the estimates given by Lemma 6. To show (4.14)-(4.18) we proceed like in the proof of Lemma 2. Now let $v \in W(0,T)$ be chosen such that $v(0) \equiv 0$ in $L^2(\Omega)$. We multiply the adjoint equation (3.9) by v and use integration by parts. Passing to the limit $\sigma \searrow 0$, then yields the weak formulation of the adjoint equation as given in (4.12). Because of $(\psi_\ominus^\gamma)''(y_\sigma)[y_\sigma + 1]^\oplus = 0$ and $(\psi_\oplus^\gamma)''(y_\sigma)[y_\sigma - 1]^\ominus = 0$ we easily obtain

$$\lim_{\sigma \to 0} (\lambda_\sigma^\ominus, [y_\sigma + 1]^\oplus)_{L^2(\Omega_T)} = 0 \quad \text{and} \quad \lim_{\sigma \to 0} (\lambda_\sigma^\oplus, [y_\sigma - 1]^\ominus)_{L^2(\Omega_T)} = 0.$$

Furthermore we have

$$(\lambda_\sigma^l, p_\sigma)_{L^2(\Omega_T)} = \int_{\Omega_T} \frac{1}{\sigma} (\psi_l^\gamma)''(y_\sigma) |p_\sigma|^2 dxdt \geq 0$$

for $l \in \{\oplus, \ominus\}$ and $\sigma > 0$. Hence, we obtain (4.21)-(4.22).
By Theorem 6 we know that there exists a subsequence (denoted the same) such that $y_\sigma \longrightarrow y^*$ a.e. in Ω_T. Hence for almost every $\{(t,x) \in \Omega_T : |y^*(t,x)| < 1\}$ we have that $|y_\sigma(t,x)| < 1$ for σ sufficiently small. Therefore,

$$\lambda_\sigma^\oplus + \lambda_\sigma^\ominus \quad \longrightarrow \quad 0 \quad \text{a.e.} \quad \text{in} \quad \{(t,x) \in \Omega_T : |y^*(t,x)| < 1\}.$$

Due to Egorov's theorem, the quantity $(\lambda_\sigma^\oplus + \lambda_\sigma^\ominus)|_{\{(t,x) \in \Omega_T : |y^*(t,x)| < 1\}}$ then converges uniformly with respect to the underlying measure to zero, i.e., for every $\omega > 0$, there exists a subset $Q_\omega \subset \{(t,x) \in \Omega_T : |y^*(t,x)| < 1\}$ with $meas(\{(t,x) \in \Omega_T : |y^*(t,x)| < 1\} \setminus Q_\omega) \leq \omega$, such that

$$\lambda_\sigma^\oplus + \lambda_\sigma^\ominus \quad \longrightarrow \quad 0 \quad \text{uniformly} \quad \text{in} \quad Q_\omega.$$

Hence, (4.23) is proven. □

Remark 1. *Our convergence results are based on the assumption that* u_σ *stays inside some uniformly bounded set as* $\sigma \searrow 0$*. The optimality conditions are hence derived for accumulation points of stationarity point of the penalized subproblems, only.*

The optimality conditions (4.12)-(4.23) of Theorem 6 define a "very" weak form of stationarity points of the Allen-Cahn problem (see [13] for different definitions of stationarity). The results of Theorem 6 can be interpreted in the following way: The accumulation points of stationary points of the penalized subproblems satisfy "very" weak optimality conditions.
The weakness of the result is due to the low regularity of $\lambda^\oplus + \lambda^\ominus$. In fact, if $\lambda^\oplus + \lambda^\ominus$ is bounded in $L^2(0,T;H^1(\Omega)^*)$, then the results can be strengthened as the following corollary states

Corollary 1. *Let the assumptions of Theorem 6 be satisfied. Furthermore, we assume that $\{\lambda^\oplus_\sigma + \lambda^\ominus_\sigma\}$ is bounded in $L^2(0,T;H^1(\Omega)^*)$ uniformly with respect to $\sigma > 0$. Then there exist*

$$\{y^*, u^*, p^*\} \in \tilde{\mathcal{V}} \times L^2(\Omega_T) \times W(0,T)$$

and a subsequence still denoted by $\{y_\sigma, u_\sigma, p_\sigma\}$ such that as $\sigma \searrow 0$

$$\begin{array}{lllll}
y_\sigma & \longrightarrow & y^* & \textit{weakly} & \textit{in} \;\; L^2(0,T;H^2(\Omega)) \cap H^1(\Omega_T), \\
y_\sigma & \longrightarrow & y^* & \textit{weakly} & \textit{in} \;\; H^1(0,T;H^1(\Omega)), \\
y_\sigma & \longrightarrow & y^* & \textit{weakly-star} & \textit{in} \;\; L^\infty(0,T;H^1(\Omega)), \\
y_\sigma & \longrightarrow & y^* & \textit{weakly-star} & \textit{in} \;\; W^{1,\infty}(0,T;L^2(\Omega)), \\
\mu^\oplus_\sigma & \longrightarrow & \mu^\oplus_* & \textit{weakly} & \textit{in} \;\; L^2(\Omega_T), \\
\mu^\ominus_\sigma & \longrightarrow & \mu^\ominus_* & \textit{weakly} & \textit{in} \;\; L^2(\Omega_T), \\
u_\sigma & \longrightarrow & u^* & \textit{weakly} & \textit{in} \;\; L^2(\Omega_T), \\
p_\sigma & \longrightarrow & p^* & \textit{weakly} & \textit{in} \;\; W(0,T), \\
\lambda^\oplus_\sigma & \longrightarrow & \lambda^\oplus_* & \textit{weakly} & \textit{in} \;\; L^2(0,T;H^1(\Omega)^*), \\
\lambda^\ominus_\sigma & \longrightarrow & \lambda^\ominus_* & \textit{weakly} & \textit{in} \;\; L^2(0,T;H^1(\Omega)^*).
\end{array} \tag{4.24}$$

The limit element $\{y^*, u^*, p^*\}$ satisfies the following optimality system

$$\begin{aligned}
&\left\langle \frac{1}{\varepsilon}(\lambda^\oplus_* + \lambda^\ominus_*) - \varepsilon\,\partial_t p^*, v \right\rangle_{L^2(0,T;H^1(\Omega)^*),L^2(0,T;H^1(\Omega))} \\
&+ \varepsilon(\nabla p^*, \nabla v)_{L^2(\Omega_T)} + \nu_d(y^* - y_d, v)_{L^2(\Omega_T)} - \\
&- \frac{1}{\varepsilon}(p^*, v)_{L^2(\Omega_T)} = 0 \qquad \forall v \in L^2(0,T;H^1(\Omega)), && (4.25)\\
&p^*(T,\cdot) = \nu_T(y^*(T,\cdot) - y_T) \quad \textit{in } \Omega, && (4.26)\\
&\frac{\nu_u}{\varepsilon}u^* - p^* = 0 \quad \textit{in } \Omega_T, && (4.27)\\
&\varepsilon\partial_t y^* - \varepsilon\Delta y^* - \frac{1}{\varepsilon}y^* + \frac{1}{\varepsilon}\mu^\oplus_* - \frac{1}{\varepsilon}\mu^\ominus_* = u^* \quad \textit{a.e. in } \Omega_T, && (4.28)\\
&y^*(0) = y_0 \quad \textit{in } \Omega, \qquad n \cdot \nabla y^* = 0 \quad \textit{a.e. on } \Gamma_T, && (4.29)
\end{aligned}$$

with the complementarity conditions

$$|y^*| < 1 \quad a.e. \ in \ \Omega_T, \tag{4.30}$$

$$\mu_*^{\oplus}(y^* - 1) = 0, \mu_*^{\ominus}(y^* + 1) = 0 \quad a.e. \ in \ \Omega_T, \tag{4.31}$$

$$\mu_*^{\oplus} \geq 0, \mu_*^{\ominus} \geq 0 \quad a.e. \ in \ \Omega_T, \tag{4.32}$$

$$\left\langle \lambda_*^{\ominus}, [y^* + 1] \right\rangle_{L^2(0,T;H^1(\Omega)^*),L^2(0,T;H^1(\Omega))} = 0, \tag{4.33}$$

$$\left\langle \lambda_*^{\oplus}, [y^* - 1] \right\rangle_{L^2(0,T;H^1(\Omega)^*),L^2(0,T;H^1(\Omega))} = 0. \tag{4.34}$$

$$\liminf_{\sigma \searrow 0} (\lambda_\sigma^{\ominus}, p_\sigma)_{L^2(\Omega_T)} \geq 0, \tag{4.35}$$

$$\liminf_{\sigma \searrow 0} (\lambda_\sigma^{\oplus}, p_\sigma)_{L^2(\Omega_T)} \geq 0, \tag{4.36}$$

Furthermore, for every $\omega > 0$*, there exists a subset* $Q_\omega \subset \{(t,x) \in \Omega_T : |y^*(t,x)| < 1\}$ *with* $meas(\{(t,x) \in \Omega_T : |y^*(t,x)| < 1\} \setminus Q_\omega) \leq \omega$*, such that*

$$\lambda_\sigma^{\oplus} + \lambda_\sigma^{\ominus} \longrightarrow 0 \quad uniformly \quad in \quad Q_\omega. \tag{4.37}$$

Proof. If $\{\lambda_\sigma^{\oplus} + \lambda_\sigma^{\ominus}\}$ *is bounded in* $L^2(0,T;H^1(\Omega)^*)$ *uniformly with respect to* $\sigma > 0$*, then the adjoint equation (3.12) immediately yields uniform boundedness of* $\{\partial_t p_\sigma\}$ *in* $L^2(0,T;H^1(\Omega)^*)$ *with respect to* $\sigma > 0$ *and hence uniform boundedness of* $\{p_\sigma\}$ *in* $W(0,T)$ *with respect to* $\sigma > 0$*. Consequently*

$$\begin{array}{llll} p_\sigma & \longrightarrow & p^* & weakly \quad in \quad W(0,T), \\ \lambda_\sigma^{\oplus} & \longrightarrow & \lambda_*^{\oplus} & weakly \quad in \quad L^2(0,T;H^1(\Omega)^*), \\ \lambda_\sigma^{\ominus} & \longrightarrow & \lambda_*^{\ominus} & weakly \quad in \quad L^2(0,T;H^1(\Omega)^*). \end{array} \tag{4.38}$$

Further, (4.19) and (4.20) imply (4.33) and (4.34).

Acknowledgment. The author wishes to thank Luise Blank and Harald Garcke for many fruiteful hints and discussions. This work was supported by a DFG grant within the Priority Program SPP 1253 (*Optimization with Differential Equations*), which is gratefully acknowledged.

References

[1] R. ADAMS, *Sobolev spaces* Academic Press, Boston, 1975.

[2] V. BARBU, *Optimal control of variational inequalities*, Research Notes in Mathematics 100 (1984), Pitman Advanced Publishing Program, Boston London Melbourne.

[3] M. Bergounioux, *Optimal control of an Obstacle Problem*, Appl. Math. Optim., 36 (1997), pp.147-172

[4] L. Blank, H. Garcke, L. Sarbu, V. Styles, *Primal-dual active set methods for Allen-Cahn variational inequalities with non-local constraints*, Preprint-Nr.: SPP1253-09-01 (2009).

[5] L. Blank, H. Garcke, L. Sarbu, V. Styles, *Non-local Allen-Cahn systems: Analysis and a primal dual active set method*, Universität Regensburg, Mathematik, Preprint Nr. 02/2011

[6] J.F. Blowey, C.M. Elliott, *The Cahn-Hillard gradient theory for phase separation with non-smooth free energy Part I: Mathematical analysis*, Euro. Jnl of Applied Mathematics (1991), vol. 2, pp.233-279

[7] L.C. Evans, *Partial differential equations*, Graduate Studies in Mathematics 19. American Mathematical Society, Providence, RI (1998).

[8] M.H. Farshbaf-Shaker, *A Relaxation approach to vector-valued Allen-Cahn MPEC problems*, to appear

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

[10] M. Hintermüller, I. Kopacka, *A smooth penalty approach and a nonlinear multigrid algorithm for elliptic mpecs*, Computational Optimization and Applications, Springer Netherlands, DOI 10.1007/s10589-009-9307-9

[11] M. Hintermüller, I. Kopacka, *Mathematical Programs with Complementarity Constraints in Function Space: C- and Strong Stationarity and a Path-Following Algorithm* , SIAM J. Optim. Volume 20, Issue 2, pp. 868-902 (2009).

[12] M. Hintermüller, 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. on Control and Optimization, Vol. 48 (2010)

[13] I. Kopacka, *MPECs/MPCCs in Function Space: First Order Optimality Concepts, Path-Following, and Multilevel Algorithms* Dissertation, Karl-Franzens-Universität Graz

[14] O.A. Ladyzenskaja, V.A. Solonikov and N.N. Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type*, Translations of Mathematical Monographs 23, American Mathematical Society, 1968.

[15] F. Mignot, J.P. Puel, *Optimal control in some variational inequalities*, SIAM J. on Control and Optimization, Vol. 22, No. 3, May 1984

[16] Mini-Workshop:*Control of Free Boundaries.* Abstracts from the mini-workshop held February 11-17. 2007. Organized by Charles M. Elliott, Michael Hinze and Vanessa Styles. Oberwolfach reports. Vol. 4, no.1.

[17] F. Tröltzsch, *Optimal Control of Partial Differential Equations, Theory, Methods and Applications*, Graduate Studies in Mathematics Vol. 112, AMS 2010.