Universität Regensburg
Mathematik

A relaxation approach to vector-valued Allen-Cahn MPEC problems

M. Hassan Farshbaf-Shaker

Preprint Nr. 27/2011

A relaxation approach to vector-valued Allen-Cahn MPEC problems

M.Hassan Farshbaf-Shaker*

Abstract

In this paper we consider a vector-valued Allen-Cahn MPEC problem. To derive optimality conditions we exploit a regularization-relaxation technique. The optimality system of the regularized-relaxed subproblems are investigated by applying the classical result of Zowe and Kurcyusz. Finally we show that the stationary points of the regularized-relaxed subproblems converge to weak stationary points of the limit problem.

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

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

1 Introduction

The field of the mathematical and numerical analysis of systems of nonlinear PDE's involving interfaces and free boundaries is a burgeoning area of research. Many such systems arise from mathematical models in material science and fluid dynamics such as phase separation in alloys, crystal growth, dynamics of multi-phase fluids and epitaxial growth. In applications of these mathematical models, suitable performance indices and appropriate control actions have to be specified. Mathematically this leads to optimization problems with PDE constraints including free boundaries, see [16]. Surveys and articles concerning the mathematical and numerical approaches to optimal control of free boundary problems may be found in [10, 5]. In this paper we consider an Allen-Cahn model as a phase-field model to describe the interface

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

evolution. Phase-field methods provide a natural method for dealing with the complex topological changes that occur, see [6]. The interface between the phases is replaced by a thin transitional layer of width $O(\varepsilon)$ where ε is a small parameter. The underlying non-convex energy functional is based on the Ginzburg-Landau energy

$$E(\boldsymbol{y}) := \int_\Omega \left(\frac{\varepsilon}{2} |\nabla \boldsymbol{y}|^2 + \frac{1}{\varepsilon} \Psi(\boldsymbol{y}) \right) dx, \qquad \varepsilon > 0, \tag{1.1}$$

where $\Omega \subset \mathbb{R}^d$ is an open and bounded domain, $\boldsymbol{y} : (0,T) \times \Omega \to \mathbb{R}^N$ is the phase field vector (in our setting the state variable) and Ψ is the bulk potential. Since each component of $\boldsymbol{y} := (y_1, \dots, y_N)^T$ stands for the fraction of one phase, the phase space for the order parameter $\boldsymbol{y}$ is the Gibbs simplex

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

Here $\boldsymbol{v} \geq \boldsymbol{0}$ means $v_i \geq 0$ for all $i \in \{1, \dots, N\}$, $\boldsymbol{1} = (1, \dots, 1)^T$. For the bulk potential $\Psi : \mathbb{R}^N \to \mathbb{R}_0^+ \cup \{\infty\}$ we consider the multi obstacle potential

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

where $I_{\boldsymbol{G}}$ is the indicator function of the Gibbs simplex. We are interested in phase kinetics, so the next procedure is to minimize (1.1) under the constraint (1.2). For details, see [11, 12].

Notations. In the sequel we always denote by $\Omega \subset R^d$ a bounded domain (with spatial dimension d) with boundary $\Gamma = \partial\Omega$. The outer unit normal on Γ is denoted by n. Vectors are defined by boldface letters. Moreover we define $\mathbb{R}_+^N := \{\boldsymbol{v} \in \mathbb{R}^N \mid \boldsymbol{v} \geq \boldsymbol{0}\}$ and the affine hyperplane

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

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

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

is a 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 $1 \leq p \leq \infty$ 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 $H^k(\Omega) = W^{k,2}(\Omega)$. 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 vector-valued function spaces by boldface letters, $\boldsymbol{L}^2(\Omega) := L^2(\Omega;\mathbb{R})^N$. Moreover we define $\boldsymbol{L}^2_+(\Omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \mid \boldsymbol{v} \in \mathbb{R}^N_+ \text{ a.e. in } \Omega\}$ which is a convex cone in $\boldsymbol{L}^2(\Omega)$; $\boldsymbol{L}^2_{\boldsymbol{\Sigma}}(\Omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \mid \boldsymbol{v} \in \boldsymbol{\Sigma} \text{ a.e. in } \Omega\}$ which is a convex subset of $\boldsymbol{L}^2(\Omega)$ and $\boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \mid \boldsymbol{v} \in \boldsymbol{T\Sigma} \text{ a.e. in } \Omega\}$ which is a subspace of $\boldsymbol{L}^2(\Omega)$ and hence also a Hilbert space. Furthermore we have $\boldsymbol{L}^2_{\boldsymbol{G}}(\Omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \mid \boldsymbol{v} \in \boldsymbol{G} \text{ a.e. in } \Omega\}$ 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 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

$$\boldsymbol{W}(0,T) := L^2(0,T;\boldsymbol{H}^1(\Omega)) \cap H^1(0,T;\boldsymbol{H}^1(\Omega)^*).$$

Moreover we use $\boldsymbol{L}^2_i(\Omega_T) := L^2(0,T;\boldsymbol{L}^2_i(\Omega))$, where $i \in \{+, \boldsymbol{\Sigma}, \boldsymbol{T\Sigma}\}$, $\boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} := \boldsymbol{\mathcal{V}} \cap \boldsymbol{L}^2_{\boldsymbol{\Sigma}}(\Omega_T)$ and $\boldsymbol{W}(0,T)_i := \boldsymbol{W}(0,T) \cap \boldsymbol{L}^2_i(\Omega_T)$ where $i \in \{\boldsymbol{\Sigma}, \boldsymbol{T\Sigma}\}$. We also have $\boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} := \{\boldsymbol{u} \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}} \mid n \cdot \nabla \boldsymbol{u} = 0 \text{ a.e. in } \Gamma_T\}$. Here for vector-valued functions we define the $\boldsymbol{L}^2$ inner product by

$$(\boldsymbol{\xi}, \boldsymbol{y})_{\boldsymbol{L}^2} := \sum_{i=1}^N (\xi_i, y_i)_{L^2}, \tag{1.3}$$

For the rest of the paper we make the following assumption

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

Hence, given an initial phase distribution $\boldsymbol{y}(0,\cdot) = \boldsymbol{y}_0 : \Omega \to \boldsymbol{G}$ 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

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

where $\boldsymbol{\zeta}^* \in \partial I_G$ and ∂I_G denotes the subdifferential of I_G. As for the scalar case, see e.g [3, 8], this equation leads to the following Allen-Cahn variational inequality

*Let **(H1)** hold. For given initial data* $\boldsymbol{y}_0 \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)$ *find* $\boldsymbol{y} \in L^2(0,T;\boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)) \cap H^1(0,T;\boldsymbol{L}^2(\Omega))$ *such that* $\boldsymbol{y}(0) = \boldsymbol{y}_0$ *and*

$$\varepsilon(\partial_t \boldsymbol{y}, \boldsymbol{\chi} - \boldsymbol{y})_{\boldsymbol{L}^2(\Omega)} + \varepsilon(\nabla \boldsymbol{y}, \nabla(\boldsymbol{\chi} - \boldsymbol{y}))_{\boldsymbol{L}^2(\Omega)} \geq (\frac{1}{\varepsilon}\boldsymbol{y}, \boldsymbol{\chi} - \boldsymbol{y})_{\boldsymbol{L}^2(\Omega)},$$

which has to hold for almost all $t \in [0,T]$ *and all* $\boldsymbol{\chi} \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)$.

1.1 Allen-Cahn MPEC problem

Now we introduce our overall optimization problem. Our goal is to transform an initial phase distribution $\boldsymbol{y}_0 : \Omega \to \mathbb{R}$ with minimal cost of control to some desired phase pattern $\boldsymbol{y}_T : \Omega \to \mathbb{R}$ at a given final time T, where furthermore the distribution remains throughout the entire time interval close to a given distribution $\boldsymbol{y}_d$.

Our upper level problem is

$$\begin{cases} \text{min} & J(\boldsymbol{y},\boldsymbol{u}) := \frac{\nu_d}{2}\|\boldsymbol{y} - \boldsymbol{y}_d\|^2_{\boldsymbol{L}^2(\Omega_T)} + \frac{\nu_T}{2}\|\boldsymbol{y}(T,\cdot) - \boldsymbol{y}_T\|^2_{\boldsymbol{L}^2(\Omega)} + \frac{\nu_u}{2\varepsilon}\|\boldsymbol{u}\|^2_{\boldsymbol{L}^2(\Omega_T)} \\ \text{over} & (\boldsymbol{y},\boldsymbol{u}) \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{G}} \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \\ \text{s.t.} & (\textbf{ACVI}) \text{ holds .} \end{cases}$$

Our lower level problem (**ACVI**) is:

*Let **(H1)** hold. For given initial data* $\boldsymbol{y}_0 \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)$ *and given control* $\boldsymbol{u} \in \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ *find* $\boldsymbol{y} \in L^2(0,T;\boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)) \cap H^1(0,T;\boldsymbol{L}^2(\Omega))$ *such that* $\boldsymbol{y}(0) = \boldsymbol{y}_0$ *and*

$$\varepsilon(\partial_t \boldsymbol{y}, \boldsymbol{\chi} - \boldsymbol{y})_{\boldsymbol{L}^2(\Omega)} + \varepsilon(\nabla \boldsymbol{y}, \nabla(\boldsymbol{\chi} - \boldsymbol{y}))_{\boldsymbol{L}^2(\Omega)} \geq (\frac{1}{\varepsilon}\boldsymbol{y} + \boldsymbol{u}, \boldsymbol{\chi} - \boldsymbol{y})_{\boldsymbol{L}^2(\Omega)}, \quad (1.4)$$

which has to hold for almost all $t \in [0,T]$ *and all* $\boldsymbol{\chi} \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)$.

Here, ν_d, ν_T, ν_u are positive constants. The resulting optimization problem belongs to the problem class of the so-called MPECs (Mathematical Programs with Equilibrium Constraints) which are hard to handle for several reasons. Indeed, we note that due to the structure of the feasible set classical constraint qualifications such as the Mangasarian-Fromovitz constraint qualifications do not hold true. As a result the existence of Lagrange multipliers of the upper level problem for characterizing first order optimality cannot be derived from standard KKT theory. These kinds of problems have

been extensively studied by many authors, as for example V. Barbu [1], M. Bergounioux [2] or more recently M. Hintermüller and I. Kopacka [13].
In this work our aim is to derive first order optimality conditions of C-stationarity-type (for different notions of stationarity for MPECs we refer to [15]). In contrast to [8] our approach in this paper consists of using first a relaxation technique to extend the feasible set of the resulting MPEC and secondly a Moreau-Yosida based regularization to avoid the lower regularity of the Lagrange multiplier of the upper level problem corresponding to the state constraint in the relaxed problem. We derive first order optimality conditions of the regularized-relaxed subproblems using the classical result of Zowe and Kurcyusz [17] and we study the limit for vanishing relaxation parameter and regularization parameter $\gamma \uparrow +\infty$. We derive the limit optimality system without considering global solutions (minimizers) of the regularized-relaxed subproblems. The approach reflects the typical situation for nonlinear and non-convex minimization problems, where solution procedures guarantee stationarity points only rather than global minimizers.
The rest of the paper is organized as follows. In section 2 we analyze the vector-valued Allen-Cahn inequality as the lower level problem; the existence of a solution to the inequality is proven by a penalization technique, see for similar results in [4]. Furthermore the complementarity formulation for the Allen-Cahn inequality is given. In section 3 the MPCC (Mathematical programming with complementarity constraints) problem is formulated, which is a special case of an MPEC. To derive the optimality system for the MPCC we use a regularization relaxation technique in section 4. Furthermore we investigate the convergence behavior of minimizers with respect to the relaxation and regularization parameters. We also derive first order optimality systems for the regularized-relaxed subproblems. In section 5 we investigate the convergence behavior of stationarity points to the original problem.

2 Lower level problem: Allen-Cahn variational inequality

We begin with defining the operator $\boldsymbol{A} : \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}}^{hN} \to \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)^*$ by

$$(\boldsymbol{A}\boldsymbol{y}, \boldsymbol{\chi}) := (-\Delta \boldsymbol{y}, \boldsymbol{\chi})_{\boldsymbol{L}^2(\Omega_T)} \text{ for all } \boldsymbol{\chi} \in \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T).$$

Following [4] the problem (**ACVI**) can be reformulated with the help of the slack variable (Lagrange multiplier of the lower level problem) $\boldsymbol{\xi}$ corresponding to the inequality constraint $\boldsymbol{y} \geq 0$, which results in the following complementarity-problem (**CCP**):

*Let **(H1)** hold. For given initial data* $\boldsymbol{y}_0 \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)$ *and* $\boldsymbol{u} \in \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ *find* $\boldsymbol{y} \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}}$ *such that* $\boldsymbol{y}(0) = \boldsymbol{y}_0$ *and*

$$(\varepsilon \partial_t \boldsymbol{y} + \varepsilon \boldsymbol{A}\boldsymbol{y} - \frac{1}{\varepsilon}(\boldsymbol{\xi} + \boldsymbol{y}) - \boldsymbol{u}, \boldsymbol{\chi})_{\boldsymbol{L}^2(\Omega_T)} = 0, \tag{2.1}$$

which has to hold for all $\boldsymbol{\chi} \in \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$*. Moreover we have the complementarity conditions*

$$(CC) \quad \begin{cases} \boldsymbol{y} \geq 0 \ \textit{a.e. in } \Omega_T, \\ \boldsymbol{\xi} \geq 0 \ \textit{a.e. in } \Omega_T, \\ (\boldsymbol{\xi}, \boldsymbol{y})_{\boldsymbol{L}^2(\Omega_T)} = 0, \end{cases} \tag{2.2}$$

By Riesz representation theorem we indentify $\boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)^*$ with $\boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ and rewrite (2.1) as an operator equation

$$(LLP) \quad \begin{cases} (\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \times \boldsymbol{L}^2(\Omega_T) \\ \varepsilon \partial_t \boldsymbol{y} + \varepsilon \boldsymbol{A}\boldsymbol{y} - \frac{1}{\varepsilon}(\boldsymbol{y} + \boldsymbol{\xi}) = \boldsymbol{u} \text{ in } \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \\ \boldsymbol{y}(0) = \boldsymbol{y}_0 \text{ a.e. in } \Omega. \end{cases}$$

Lemma 1. *Let **(H1)** hold and* $(\boldsymbol{y}_0, \boldsymbol{u}) \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega) \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ *be given. A function* $\boldsymbol{y} \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}}$ *solves **(ACVI)** if there exists* $\boldsymbol{\xi} \in \boldsymbol{L}^2(\Omega_T)$ *such that (LLP) and (CC) are fulfilled.*

Proof. Let $\boldsymbol{y} \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}}$ be the solution to (LLP) and (CC). For $\boldsymbol{\chi} \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)$, the function $(\boldsymbol{\chi} - \boldsymbol{y}) \in \boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega) \subset \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega)$ is an admissible testfunction in (2.1). After partial integration we get

$$(\varepsilon \partial_t \boldsymbol{y} - \frac{1}{\varepsilon}(\boldsymbol{\mu} + \boldsymbol{y}) - \boldsymbol{u}, \boldsymbol{\chi} - \boldsymbol{y})_{\boldsymbol{L}^2(\Omega)} + \varepsilon(\nabla \boldsymbol{y}, \nabla(\boldsymbol{\chi} - \boldsymbol{y}))_{\boldsymbol{L}^2(\Omega)} = 0,$$

for a.e. $t \in [0, T]$. Using the property $\boldsymbol{\chi} \geq \boldsymbol{0}$ and (CC) gives

$$(\boldsymbol{\xi}, \boldsymbol{\chi} - \boldsymbol{y})_{\boldsymbol{L}^2(\Omega_T)} \geq 0.$$

Hence we obtain for all $\boldsymbol{\chi} \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)$ and almost all $t \in [0, T]$

$$(\varepsilon \partial_t \boldsymbol{y} - \frac{1}{\varepsilon}\boldsymbol{y} - \boldsymbol{u}, \boldsymbol{\chi} - \boldsymbol{y})_{\boldsymbol{L}^2(\Omega)} + \varepsilon(\nabla \boldsymbol{y}, \nabla(\boldsymbol{\chi} - \boldsymbol{y}))_{\boldsymbol{L}^2(\Omega)} \geq 0,$$

and hence $\boldsymbol{y}$ solves (**ACVI**). □

Theorem 1. *Let **(H1)** hold. Given* $(\boldsymbol{y}_0, \boldsymbol{u}) \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega) \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ *there exists a unique solution* $(\boldsymbol{y}, \boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \times \boldsymbol{L}^2(\Omega_T)$ *to **(CCP)**.*

Proof. We will give here a sketch of the main steps of the proof. For detailed calculations we refer to a similar proof in [4].
1. Step: Regularized problems We introduce the following regularization of the obstacle potential $\Psi(\boldsymbol{y})$:

$$\Psi^\delta(\boldsymbol{y}) = \Psi_0(\boldsymbol{y}) + \frac{1}{\delta}\hat{\Psi}(\boldsymbol{y}),$$

where

$$\hat{\Psi}(\boldsymbol{y}) = \sum_{i=1}^{N} \min(y_i, 0)^2.$$

Define the function $\hat{\Phi}(r) = 2\min(r, 0)$ for all $r \in \mathbb{R}$ and note that $\hat{\Psi}'_{\boldsymbol{y}}(\boldsymbol{y}) := \hat{\boldsymbol{\Phi}}(\boldsymbol{y}) = \{\hat{\Phi}(y_i)\}_{i=1}^N$.
We now solve the following regularized Allen-Cahn equation $(\mathbf{ACVI})_\delta$: *Let **(H1)** hold. Given $\boldsymbol{y}_0 \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)$ and $\boldsymbol{u} \in \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ find $\boldsymbol{y}_\delta \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}}$ such that $\boldsymbol{y}_\delta(0) = \boldsymbol{y}_0$ and*

$$\varepsilon(\partial_t \boldsymbol{y}_\delta, \boldsymbol{\chi})_{\boldsymbol{L}^2(\Omega)} + \varepsilon(\nabla \boldsymbol{y}_\delta, \nabla \boldsymbol{\chi})_{\boldsymbol{L}^2(\Omega)} + (\frac{1}{\varepsilon}\Psi^\delta_{,\boldsymbol{y}}(\boldsymbol{y}_\delta) - \boldsymbol{u}_\delta, \boldsymbol{\chi})_{\boldsymbol{L}^2(\Omega)} = 0, \quad (2.3)$$

which has to hold for almost all $t \in [0, T]$ and all $\boldsymbol{\chi} \in \boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega)$.
For every $\delta \in (0, 1]$ one can show the unique solvability of (2.3) by classical theory of parabolic partial differential equations and then pass to the limit. Following [4] we reformulate (2.3) by using $\Psi^\delta_{,\boldsymbol{y}}(\boldsymbol{y}_\delta) = \frac{1}{\delta}\hat{\boldsymbol{\Phi}}(\boldsymbol{y}_\delta) - \boldsymbol{y}_\delta$ and defining $\boldsymbol{\xi}_\delta := -\frac{1}{\delta}\hat{\boldsymbol{\Phi}}(\boldsymbol{y}_\delta)$. Hence, we have

$$\varepsilon(\partial_t \boldsymbol{y}_\delta, \boldsymbol{\chi})_{\boldsymbol{L}^2(\Omega)} + \varepsilon(\nabla \boldsymbol{y}_\delta, \nabla \boldsymbol{\chi})_{\boldsymbol{L}^2(\Omega)} - (\frac{1}{\varepsilon}(\boldsymbol{y}_\delta + \boldsymbol{\xi}_\delta) + \boldsymbol{u}_\delta, \boldsymbol{\chi})_{\boldsymbol{L}^2(\Omega)} = 0, \quad (2.4)$$

for all $\boldsymbol{\chi} \in \boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega)$.

2. Step: A priori estimates Let **(H1)** hold and $\boldsymbol{y}_0 \in \boldsymbol{H}^1_{\boldsymbol{G}}(\Omega)$. For a sequence $\{\boldsymbol{u}_\delta\}_{\delta\in(0,1]}$ uniformly bounded in $\boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ it is shown in [4] that

$$\begin{array}{lllll} \boldsymbol{y}_\delta & \text{bounded} & \text{in} & \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} & \text{uniformly in} \quad \delta \in (0,1], \\ \boldsymbol{\xi}_\delta & \text{bounded} & \text{in} & \boldsymbol{L}^2(\Omega_T) & \text{uniformly in} \quad \delta \in (0,1]. \end{array} \quad (2.5)$$

3. Step: Passing to the limit From Step 2 we get the convergence results as $\delta \searrow 0$

$$\begin{array}{lllll} \boldsymbol{y}_\delta & \longrightarrow & \boldsymbol{y} & \text{weakly} & \text{in} \quad L^2(0,T;\boldsymbol{H}^2(\Omega)), \\ \boldsymbol{y}_\delta & \longrightarrow & \boldsymbol{y} & \text{weakly} & \text{in} \quad H^1(0,T;\boldsymbol{L}^2(\Omega)), \\ \boldsymbol{y}_\delta & \longrightarrow & \boldsymbol{y} & \text{weak-star} & \text{in} \quad L^\infty(0,T;\boldsymbol{H}^1(\Omega)), \\ \boldsymbol{\xi}_\delta & \longrightarrow & \boldsymbol{\xi} & \text{weakly} & \text{in} \quad \boldsymbol{L}^2(\Omega_T), \end{array} \quad (2.6)$$

The set $\{\boldsymbol{\xi}_\delta \in \boldsymbol{L}^2(\Omega_T) : \boldsymbol{\xi}_\delta \geq \mathbf{0}$ a.e. in $\Omega_T\}$ is convex and closed and hence weakly closed and we obtain $\boldsymbol{\xi} \geq \mathbf{0}$ a.e. in Ω_T. Furthermore, the convex and closed subset $\boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}}^{hN}$ is weakly closed and we obtain that $\boldsymbol{y} \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}}^{hN}$. For proving $\boldsymbol{y} \geq \mathbf{0}$ we refer the reader to [4, 9]. We get moreover as $\delta \searrow 0$

$$(\boldsymbol{\xi}, \boldsymbol{y})_{\boldsymbol{L}^2(\Omega)} \longleftarrow (\boldsymbol{\xi}_\delta, \boldsymbol{y}_\delta)_{\boldsymbol{L}^2(\Omega)} = -\frac{1}{\delta}(\hat{\boldsymbol{\Phi}}(\boldsymbol{y}_\delta), \boldsymbol{y}_\delta)_{\boldsymbol{L}^2(\Omega)} \leq 0,$$

and hence $(\boldsymbol{\xi}, \boldsymbol{y})_{\boldsymbol{L}^2(\Omega)} \leq 0$. However, since $\boldsymbol{\xi} \geq \mathbf{0}$ and $\boldsymbol{y} \geq \mathbf{0}$ we have that $(\boldsymbol{\xi}, \boldsymbol{y})_{\boldsymbol{L}^2(\Omega)} = 0$ a.e. in $(0, T)$. Hence, $(\boldsymbol{y}, \boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}}^{hN} \times \boldsymbol{L}^2(\Omega_T)$ is the solution to (**CCP**). For uniqueness we refer the reader to [4]. □

The following proposition will be useful for establishing the next results.

Proposition 1. *Let $(\boldsymbol{u}_k)_{k\geq 1}$ be an uniformly bounded sequence in $\boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ and $(\boldsymbol{y}_k, \boldsymbol{\xi}_k)_{k\geq 1}$ the corresponding solutions of (**CCP**). Then there exits $(\boldsymbol{y}, \boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}}^{hN} \times \boldsymbol{L}^2(\Omega_T)$ and a subsequence still denoted by $(\boldsymbol{y}_k, \boldsymbol{u}_k, \boldsymbol{\xi}_k)_{k\geq 0}$ such that as $k \uparrow \infty$*

$$\begin{array}{lllll} \boldsymbol{y}_k & \longrightarrow & \boldsymbol{y} & \textit{weakly} & \textit{in} \;\; L^2(0,T;\boldsymbol{H}^2(\Omega)), \\ \boldsymbol{y}_k & \longrightarrow & \boldsymbol{y} & \textit{weakly} & \textit{in} \;\; H^1(0,T;\boldsymbol{L}^2(\Omega)), \\ \boldsymbol{y}_k & \longrightarrow & \boldsymbol{y} & \textit{weak-star} & \textit{in} \;\; L^\infty(0,T;\boldsymbol{H}^1(\Omega)), \\ \boldsymbol{\xi}_k & \longrightarrow & \boldsymbol{\xi} & \textit{weakly} & \textit{in} \;\; \boldsymbol{L}^2(\Omega_T), \end{array} \tag{2.7}$$

*and $(\boldsymbol{y}, \boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}}^{hN} \times \boldsymbol{L}^2(\Omega_T)$ fulfil (**CCP**).*

Proof. For every $\boldsymbol{u}_k$ the corresponding solutions to (2.4) are given by $(\boldsymbol{y}_{\delta,k}, \boldsymbol{\xi}_{\delta,k})_{k\geq 1}$. By (2.5) we have

$$(\boldsymbol{y}_{\delta,k}, \boldsymbol{\xi}_{\delta,k}) \quad \text{bounded} \quad \text{in} \quad \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}}^{hN} \times L^2(\Omega_T) \quad \text{uniformly in} \quad \delta \quad \text{and} \quad k.$$

By virtue of the lower semi-continuity of the norm we get

$$(\boldsymbol{y}_k, \boldsymbol{\xi}_k) \quad \text{bounded} \quad \text{in} \quad \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}}^{hN} \times L^2(\Omega_T) \quad \text{uniformly in} \quad k.$$

Continuing as in the proof of Theorem 1 we get (2.1). We get furthermore as $\delta \searrow 0$

$$(\boldsymbol{\xi}, \boldsymbol{y})_{\boldsymbol{L}^2(\Omega)} \longleftarrow (\boldsymbol{\xi}_k, \boldsymbol{y}_k)_{\boldsymbol{L}^2(\Omega)} = 0,$$

because of the strong and weak convergence of $\boldsymbol{y}_k$ and $\boldsymbol{\xi}_k$ in $\boldsymbol{L}^2(\Omega)$. The rest of the proof is similar to the proof of Theorem 1. □

3 Upper level problem: Optimal control problem

We consider the time dependent vectorial Allen-Cahn-MPCC problem:

$$(\mathcal{P}_0) \qquad \min\{J(\boldsymbol{y},\boldsymbol{u}) \mid (\boldsymbol{y},\boldsymbol{u},\boldsymbol{\xi}) \in \mathcal{D}_0\}$$

where $\mathcal{D}_0$ is the feasible set given by

$$(\mathcal{D}_0) \qquad \begin{cases} (\boldsymbol{y},\boldsymbol{u},\boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \times \boldsymbol{L}^2(\Omega_T) \\ \varepsilon\partial_t \boldsymbol{y} + \varepsilon \boldsymbol{A}\boldsymbol{y} - \frac{1}{\varepsilon}(\boldsymbol{y}+\boldsymbol{\xi}) = \boldsymbol{u} \text{ in } \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \\ \boldsymbol{y}(0) = \boldsymbol{y}_0 \text{ a.e. in } \Omega \\ \boldsymbol{y} \geq 0 \text{ a.e. in } \Omega_T,\ \boldsymbol{\xi} \geq 0 \text{ a.e. in } \Omega_T,\ (\boldsymbol{\xi},\boldsymbol{y})_{\boldsymbol{L}^2(\Omega_T)} = 0. \end{cases}$$

Theorem 2. *The problem $(\mathcal{P}_0)$ has at least one solution.*

Proof. Let $(\boldsymbol{y}_k,\boldsymbol{u}_k,\boldsymbol{\xi}_k)_{k\geq 0}$ be a minimizing sequence for $(\mathcal{P}_0)$ such that

$$\inf(\mathcal{P}_0) \leq J(\boldsymbol{y}_k,\boldsymbol{u}_k) \leq \inf(\mathcal{P}_0) + \frac{1}{k}.$$

Then $(\boldsymbol{u}_k)_{k\geq 0}$ is bounded in $\boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ and by Proposition 1 there exists $(\boldsymbol{y},\boldsymbol{u},\boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \times \boldsymbol{L}^2(\Omega_T)$ and a subsequence still denoted by $(\boldsymbol{y}_k,\boldsymbol{u}_k,\boldsymbol{\xi}_k)_{k\geq 0}$ such that (2.7) holds. Moreover we easily can check by the same proof-techniques as in the proof of Theorem 1 that $(\boldsymbol{y},\boldsymbol{u},\boldsymbol{\xi}) \in \mathcal{D}_0$ which implies that $(\boldsymbol{y},\boldsymbol{u},\boldsymbol{\xi})$ is a feasible point for $(\mathcal{P}_0)$. On the other hand (2.7) and the weak lower semi-continuity of norms yield

$$J(\boldsymbol{y},\boldsymbol{u}) \leq \liminf_{k\uparrow\infty} J(\boldsymbol{y}_k,\boldsymbol{u}_k) \leq \inf(\mathcal{P}_0).$$

Consequently $(\boldsymbol{y},\boldsymbol{u},\boldsymbol{\xi})$ is an optimal solution of $(\mathcal{P}_0)$. □

Following [2], we add from now on an explicit constraint to $(\mathcal{P}_0)$ involving the multiplier $\boldsymbol{\xi}$ in $\boldsymbol{L}^2(\Omega_T)$. The new time dependent vectorial Allen-Cahn-MPEC problem reads

$$(\mathcal{P}) \qquad \min\{J(\boldsymbol{y},\boldsymbol{u}) \mid (\boldsymbol{y},\boldsymbol{u},\boldsymbol{\xi}) \in \mathcal{D}\},$$

where

$$(\mathcal{D}) \qquad \begin{cases} (\boldsymbol{y},\boldsymbol{u},\boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \times \boldsymbol{L}^2(\Omega_T) \\ \varepsilon\partial_t \boldsymbol{y} + \varepsilon \boldsymbol{A}\boldsymbol{y} - \frac{1}{\varepsilon}(\boldsymbol{y}+\boldsymbol{\xi}) = \boldsymbol{u} \text{ in } \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \\ \boldsymbol{y}(0) = \boldsymbol{y}_0 \text{ a.e. in } \Omega \\ \boldsymbol{y} \geq 0 \text{ a.e. in } \Omega_T,\ \boldsymbol{\xi} \geq 0 \text{ a.e. in } \Omega_T,\ (\boldsymbol{\xi},\boldsymbol{y})_{\boldsymbol{L}^2(\Omega_T)} = 0, \\ \frac{1}{2}\|\boldsymbol{\xi}\|^2_{\boldsymbol{L}^2(\Omega_T)} \leq R, \end{cases}$$

where R is a sufficiently large positive constant. For instance, R may be the largest positive number that may be computed by the machine, see [2]. However, as $(\mathcal{P})$ lacks constraint regularity, for deriving stationarity conditions for $(\mathcal{P}_0)$ in the next section we relax the constraints of $(\mathcal{P})$ such that the relaxed version of $(\mathcal{P})$ satisfies well-known constraint qualifications of mathematical programming in Banach spaces [17]. In this context, it turns out that the well posedness of the relaxed version of $(\mathcal{P})$ depends on the new constraint for $\boldsymbol{\xi}$, see [2].

4 Regularized-relaxed upper level problems

In this section we introduce and study a regularized-relaxed version of the optimal control problem $(\mathcal{P})$. Following the approaches in [13], [14], our objective is to characterize some type of C-stationarity of critical points of $(\mathcal{P})$. This is achieved by passing to the limit with respect to the regularization and relaxation parameters. The regularized-relaxed problems are defined as follows:

$$(\mathcal{P}_\gamma) \qquad \min\{J_\gamma(\boldsymbol{y},\boldsymbol{u}) \mid (\boldsymbol{y},\boldsymbol{u},\boldsymbol{\xi}) \in \mathcal{D}_\gamma\},$$

where $J_\gamma(\boldsymbol{y},\boldsymbol{u}) := J(\boldsymbol{y},\boldsymbol{u}) + \frac{1}{2\gamma}\sum_{i=1}^{N} \|\max(0,\overline{\lambda}-\gamma y_i)\|^2_{L^2(\Omega_T)}$ and

$$(\mathcal{D}_\gamma) \quad \begin{cases} (\boldsymbol{y},\boldsymbol{u},\boldsymbol{\xi}) \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \times \boldsymbol{L}^2(\Omega_T) \\ \varepsilon\partial_t \boldsymbol{y} + \varepsilon \boldsymbol{A}\boldsymbol{y} - \frac{1}{\varepsilon}(\boldsymbol{y}+\boldsymbol{\xi}) = \boldsymbol{u} \text{ in } \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \\ \boldsymbol{y}(0) = \boldsymbol{y}_0 \text{ a.e. in } \Omega \\ \boldsymbol{\xi} \geq 0 \text{ a.e. in } \Omega_T, \\ \alpha_\gamma \geq (\boldsymbol{\xi},\boldsymbol{y})_{\boldsymbol{L}^2(\Omega_T)}, \\ R \geq \frac{1}{2}\|\boldsymbol{\xi}\|^2_{\boldsymbol{L}^2(\Omega_T)}, \end{cases}$$

where $\overline{\lambda} \in L^2(\Omega_T)$, which mimics a regular version of the multiplier associated to $\boldsymbol{y} \geq 0$, is arbitrary fixed with $\overline{\lambda} \geq 0$ a.e. in Ω_T. Note that we add a regularization term $\frac{1}{2\gamma}\sum_{i=1}^{N} \|\max(0,\overline{\lambda}+\gamma y_i)\|^2_{L^2(\Omega_T)}$ to $J(\boldsymbol{y},\boldsymbol{u})$ with γ denoting the associated regularization parameter. This step relaxes the pointwise state constraint $\boldsymbol{y} \geq 0$ a.e. in Ω_T. The derivative of the regularization-term serves as a regular (i.e, $L^2(\Omega_T)-$) approximation of the multiplier associated with $\boldsymbol{y} \geq 0$ a.e. in Ω_T. Further we relax $(\boldsymbol{\xi},\boldsymbol{y})_{\boldsymbol{L}^2(\Omega_T)} = 0$ by allowing $(\boldsymbol{\xi},\boldsymbol{y})_{\boldsymbol{L}^2(\Omega_T)} \leq \alpha_\gamma$ for some $\alpha_\gamma > 0$. These modifications motivate the description of $(\mathcal{P}_\gamma)$ as the regularized-relaxed version of $(\mathcal{P})$. Subsequently we are interested in $\gamma \uparrow \infty$ and $\alpha_\gamma \downarrow 0$ as $\gamma \uparrow \infty$. Let $\mathcal{D}_\gamma$ and $\mathcal{D}$ denote the feasible sets of $(\mathcal{P}_\gamma)$ and $(\mathcal{P})$, respectively. Observe that we have

$$\mathcal{D}_\gamma \supseteq \mathcal{D} \neq \emptyset. \tag{4.1}$$

4.1 Minimizers of the upper level problems

Theorem 3. *For every $\gamma > 0$, the regularized-relaxed problem $(\mathcal{P}_\gamma)$ admits at least one minimizer (globally optimal solution) which is denoted by $(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma, \boldsymbol{\xi}_\gamma)$.*

Proof. For the proof let $\gamma > 0$ be arbitrary but fixed. Since $J_\gamma \geq 0$ and because of (4.1) $\mathcal{D}_\gamma \neq \emptyset$ the infimum $d := J_\gamma(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma)$ in $\mathcal{D}_\gamma$ exists and hence we find a minimizing sequence $(\boldsymbol{y}_\gamma^k, \boldsymbol{u}_\gamma^k, \boldsymbol{\xi}_\gamma^k)_{k\geq 1} \subset \mathcal{D}_\gamma$ with

$$\lim_{k\uparrow\infty} J_\gamma(\boldsymbol{y}_\gamma^k, \boldsymbol{u}_\gamma^k) = d.$$

As $\{J_\gamma(\boldsymbol{y}_\gamma^k, \boldsymbol{u}_\gamma^k)\}$ is bounded, $\{\boldsymbol{u}_\gamma^k\}$ is bounded in $\boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$. Then by virtue of Proposition 1 there exits $(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{\xi}_\gamma}) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}}^{hN} \times \boldsymbol{L}^2(\Omega_T)$ and a subsequence still denoted by $(\boldsymbol{y}_\gamma^k, \boldsymbol{u}_\gamma^k, \boldsymbol{\xi}_\gamma^k)_{k\geq 1}$ such that

$$\begin{array}{lllll} \boldsymbol{y}_\gamma^k & \longrightarrow & \overline{\boldsymbol{y}_\gamma} & \text{weakly} & \text{in } L^2(0,T;\boldsymbol{H}^2(\Omega)), \\ \boldsymbol{y}_\gamma^k & \longrightarrow & \overline{\boldsymbol{y}_\gamma} & \text{weakly} & \text{in } H^1(0,T;\boldsymbol{L}^2(\Omega)), \\ \boldsymbol{y}_\gamma^k & \longrightarrow & \overline{\boldsymbol{y}_\gamma} & \text{weak-star} & \text{in } L^\infty(0,T;\boldsymbol{H}^1(\Omega)), \\ \boldsymbol{\xi}_\gamma^k & \longrightarrow & \overline{\boldsymbol{\xi}_\gamma} & \text{weakly} & \text{in } \boldsymbol{L}^2(\Omega_T). \end{array} \tag{4.2}$$

We next show that the limit point $(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{\xi}_\gamma}) \in \mathcal{D}_\gamma$. It is clear that

$$\alpha_\gamma \geq (\boldsymbol{y}_k, \boldsymbol{\xi}_k)_{\boldsymbol{L}^2(\Omega_T)} \to (\overline{\boldsymbol{y}}, \overline{\boldsymbol{\xi}})_{\boldsymbol{L}^2(\Omega_T)}.$$

The rest of the proof is similar to the proof of Theorem 1. The weak convergence of $(\boldsymbol{y}_\gamma^k, \boldsymbol{u}_\gamma^k, \boldsymbol{\xi}_\gamma^k)$ as $k \uparrow \infty$, the feasibility of $(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}, \overline{\boldsymbol{\xi}_\gamma})$ and the lower semi-continuity of J_γ give

$$d = \liminf_{k\uparrow\infty} J_\gamma(\boldsymbol{y}_\gamma^k, \boldsymbol{u}_\gamma^k) \geq J_\gamma(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}, \overline{\boldsymbol{\xi}_\gamma}) \geq d.$$

Therefore $(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}, \overline{\boldsymbol{\xi}_\gamma}) \in \mathcal{D}_\gamma$ is an optimal solution of $(\mathcal{P}_\gamma)$ for every $\gamma > 0$. □

Next we are interested in the convergence behavior of optimal solutions with respect to the regularization and relaxation parameters. For each $\gamma > 0$, let α_γ satisfy $\alpha_\gamma \downarrow 0$ as $\gamma \uparrow \infty$. We now show that the minimizers of the relaxed-regularized problems $(\mathcal{P}_\gamma)$ converge to a minimizer of $(\mathcal{P})$.

Theorem 4. *For every $\gamma > 0$, let $(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma, \boldsymbol{\xi}_\gamma)$ be a solution of $(\mathcal{P}_\gamma)$. Then there exist*

$$(\boldsymbol{y}^*, \boldsymbol{u}^*, \boldsymbol{\xi}^*) \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \times \boldsymbol{L}^2(\Omega_T)$$

and a subsequence still denoted by $(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma, \boldsymbol{\xi}_\gamma)_{\gamma>0}$ such that as $\gamma \uparrow \infty$

$$\begin{array}{lllll} \boldsymbol{y}_\gamma & \longrightarrow & \boldsymbol{y}^* & \textit{weakly} & \textit{in} \;\; L^2(0,T;\boldsymbol{H}^2(\Omega)), \\ \boldsymbol{y}_\gamma & \longrightarrow & \boldsymbol{y}^* & \textit{weakly} & \textit{in} \;\; H^1(0,T;\boldsymbol{L}^2(\Omega)), \\ \boldsymbol{y}_\gamma & \longrightarrow & \boldsymbol{y}^* & \textit{weak-star} & \textit{in} \;\; L^\infty(0,T;\boldsymbol{H}^1(\Omega)), \\ \boldsymbol{u}_\gamma & \longrightarrow & \boldsymbol{u}^* & \textit{strongly} & \textit{in} \;\; \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T), \\ \boldsymbol{\xi}_\gamma & \longrightarrow & \boldsymbol{\xi}^* & \textit{weakly} & \textit{in} \;\; \boldsymbol{L}^2(\Omega_T). \end{array}$$

Furthermore $\frac{1}{2\gamma}\sum_{i=1}^{N} \| \max(0, \overline{\lambda} + \gamma y_i)\|^2_{L^2(\Omega_T)} \to 0$ *as* $\gamma \uparrow \infty$ *and* $(\boldsymbol{y}^*, \boldsymbol{u}^*, \boldsymbol{\xi}^*)$ *is a solution of* $(\mathcal{P})$.

Proof. We consider the point $(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}, \overline{\boldsymbol{\xi}_\gamma})$ that is a solution to the problem $(\mathcal{P}_\gamma)$. Then $(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}, \overline{\boldsymbol{\xi}_\gamma}) \in \mathcal{D}_\gamma$ for all $\gamma > 0$. Hence for each $\gamma \geq 1$ we can estimate

$$\begin{aligned} J_\gamma(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}) &\leq J_\gamma(\boldsymbol{0}, \boldsymbol{0}) \\ &\leq \frac{1}{2}\|\boldsymbol{y}_T\|^2_{\boldsymbol{L}^2(\Omega)} + \frac{1}{2}\|\boldsymbol{y}_d\|^2_{\boldsymbol{L}^2(\Omega_T)} + \frac{1}{2}\sum_{i=1}^{N} \| \max(0, \overline{\lambda})\|^2_{L^2(\Omega_T)}. \end{aligned}$$

Hence

$$\overline{\boldsymbol{u}_\gamma} \quad \text{is bounded in} \quad \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \quad \text{uniformly in} \quad \gamma \in (0,\infty) \tag{4.3}$$

and for all $1 \leq i \leq N$

$$\frac{1}{\sqrt{2\gamma}} \max(0, \overline{\lambda} - \gamma \overline{y^i_\gamma}) \quad \text{is bounded in} \quad L^2(\Omega_T) \quad \text{uniformly in } \gamma \in (0,\infty). \tag{4.4}$$

By virtue of (4.3) and Proposition 1 there exist $(\boldsymbol{y}^*, \boldsymbol{\xi}^*) \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \times \boldsymbol{L}^2(\Omega_T)$ and a subsequence still denoted by $(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}, \overline{\boldsymbol{\xi}_\gamma})_{\gamma \geq 0}$ such that as $k \uparrow \infty$

$$\begin{array}{lllll} \overline{\boldsymbol{y}_\gamma} & \longrightarrow & \boldsymbol{y}^* & \text{weakly} & \text{in} \;\; L^2(0,T;\boldsymbol{H}^2(\Omega)), \\ \overline{\boldsymbol{y}_\gamma} & \longrightarrow & \boldsymbol{y}^* & \text{weakly} & \text{in} \;\; H^1(0,T;\boldsymbol{L}^2(\Omega)), \\ \overline{\boldsymbol{y}_\gamma} & \longrightarrow & \boldsymbol{y}^* & \text{weak-star} & \text{in} \;\; L^\infty(0,T;\boldsymbol{H}^1(\Omega)), \\ \overline{\boldsymbol{\xi}_\gamma} & \longrightarrow & \boldsymbol{\xi}^* & \text{weakly} & \text{in} \;\; \boldsymbol{L}^2(\Omega_T), \end{array} \tag{4.5}$$

and $(\boldsymbol{y}^*, \boldsymbol{\xi}^*) \in \boldsymbol{\mathcal{V}}_{\boldsymbol{\Sigma}}^{hN} \times \boldsymbol{L}^2(\Omega_T) \in \mathcal{D}$. The set $\{\overline{\boldsymbol{\xi}_\gamma} \in \boldsymbol{L}^2(\Omega_T) : \overline{\boldsymbol{\xi}_\gamma} \geq \mathbf{0}$ a.e. in $\Omega_T\}$ is weakly closed and we obtain

$$\boldsymbol{\xi}^* \geq \mathbf{0} \text{ a.e. in } \Omega_T. \tag{4.6}$$

Furthermore, we have

$$(\boldsymbol{y}^*, \boldsymbol{\xi}^*)_{\boldsymbol{L}^2(\Omega_T)} = \lim_{\gamma \uparrow \infty} (\overline{\boldsymbol{\xi}_\gamma}, \overline{\boldsymbol{y}_\gamma})_{\boldsymbol{L}^2(\Omega_T)} \leq \lim_{\gamma \uparrow \infty} \alpha_\gamma = 0. \tag{4.7}$$

and

$$R \geq \frac{1}{2} \|\boldsymbol{\xi}^*\|^2_{\boldsymbol{L}^2(\Omega_T)}.$$

Moreover from (4.4) we obtain

$$\| \max(0, \frac{\overline{\lambda}}{\gamma} - \gamma \overline{y^i_\gamma})\|_{L^2(\Omega_T)} \to 0, \quad \text{ as } \gamma \uparrow \infty \quad \forall 1 \leq i \leq N.$$

Since $\overline{\boldsymbol{y}_\gamma}$ converges strongly in $\boldsymbol{L}^2(\Omega_T)$, without loss of generality we may assume that $\overline{\boldsymbol{y}_\gamma}$ converges to $\boldsymbol{y}^*$ a.e. in Ω_T. Taking the limit and applying Fatou's lemma we conclude that

$$\begin{aligned} \| \max(0, -(y^i)^*)\|^2_{L^2(\Omega_T)} &= \| \liminf_{\gamma \uparrow \infty} \max(0, \frac{\overline{\lambda}}{\gamma} - \gamma \overline{y^i_\gamma})\|^2_{L^2(\Omega_T)} \\ &\leq \liminf_{\gamma \uparrow \infty} \| \max(0, \frac{\overline{\lambda}}{\gamma} - \gamma \overline{y^i_\gamma})\|^2_{L^2(\Omega_T)} \leq \lim_{\gamma \uparrow \infty} \frac{2c}{\gamma} = 0. \end{aligned}$$

Consequently

$$\| \max(0, -(y^i)^*)\|^2_{L^2(\Omega_T)} = 0 \quad \forall 1 \leq i \leq N.$$

and

$$\boldsymbol{y}^* \geq \mathbf{0} \text{ a.e. in } \Omega_T. \tag{4.8}$$

This with (4.6) and (4.7) implies

$$(\boldsymbol{y}^*, \boldsymbol{\xi}^*)_{\boldsymbol{L}^2(\Omega_T)} = 0.$$

Now let $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{\xi}})$ be an optimal control of $(\mathcal{P})$. Note that by (4.1) $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{\xi}}) \in \mathcal{D}_\gamma$ and $(\boldsymbol{y}^*, \boldsymbol{u}^*, \boldsymbol{\xi}^*) \in \mathcal{D}$. We therefore conclude

$$\begin{aligned} J(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}) &\leq J(\boldsymbol{y}^*, \boldsymbol{u}^*), \\ J_\gamma(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}) &\leq J_\gamma(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}) \quad \forall \gamma > 0. \end{aligned}$$

Using the lower semi-continuity of J, the definition of J_γ and the non-negativity of $\tilde{\boldsymbol{y}}$ it follows that

$$\begin{aligned} J(\boldsymbol{y}^*, \boldsymbol{u}^*) &\leq \liminf_{\gamma\uparrow\infty} J(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}) \\ &\leq \liminf_{\gamma\uparrow\infty} J_\gamma(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}) \leq \limsup_{\gamma\uparrow\infty} J_\gamma(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}) \leq \limsup_{\gamma\uparrow\infty} J_\gamma(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}) = J(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}) \\ &\leq J(\boldsymbol{y}^*, \boldsymbol{u}^*). \end{aligned}$$

Therefore

$$\lim_{\gamma\uparrow\infty} J_\gamma(\overline{\boldsymbol{y}_\gamma}, \overline{\boldsymbol{u}_\gamma}) = J(\boldsymbol{y}^*, \boldsymbol{u}^*) = J(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}).$$

and $(\boldsymbol{y}^*, \boldsymbol{u}^*, \boldsymbol{\xi}^*) \in \mathcal{D}$ is optimal for $(\mathcal{P})$. The convergence of the objective function values yields as $\gamma \uparrow \infty$

$$\frac{1}{2\gamma} \sum_{i=1}^{N} \| \max(0, \overline{\lambda} - \gamma \overline{y^i_\gamma}) \|^2_{L^2(\Omega_T)} \to 0 \quad \wedge \quad \|\overline{\boldsymbol{u}_\gamma}\|^2_{\boldsymbol{L}^2(\Omega_T)} \to \|\boldsymbol{u}^*\|^2_{\boldsymbol{L}^2(\Omega_T)}.$$

As weak convergence together with norm-convergence in $\boldsymbol{L}^2(\Omega_T)$ imply strong convergence, this yields the strong convergence of $\{\boldsymbol{u}_\gamma\}$ in $\boldsymbol{L}^2(\Omega_T)$. □

4.2 First order optimality conditions

In the previous section, our analysis required minimizers (global solutions) of the regularized-relaxed problems. However, finding globally optimal solutions (in particular by means of numerical algorithms) is difficult in practice. Often, one rather has to rely on stationary points, i.e. points satisfying first order optimality conditions, or on local solutions. In this subsection we derive the first order optimality system for the regularized-relaxed problems $(\mathcal{P}_\gamma)_{\gamma>0}$ using the mathematical programming approach in Banach spaces due to Zowe and Kurcyusz [17]. Let $\mathcal{X}$ and $\mathcal{Z}$ be real Banach spaces. For

$$\begin{array}{llll} F: & \mathcal{X} & \longrightarrow & \mathbb{R} & \text{Frechét-differentiable functional ,} \\ g: & \mathcal{X} & \longrightarrow & \mathcal{Z} & \text{continuously Frechét-differentiable ,} \end{array}$$

we consider the following mathematical program:

$$\min\{F(x) \mid \ g(x) \in M, \ x \in C\}, \tag{4.9}$$

where C is a convex closed subset of $\mathcal{X}$ and M a closed cone in $\mathcal{Z}$ with vertex at 0. We define the notion of local optimality

Definition 1. *We call $\hat{x}$ a local solution of (4.9) if there is some $\sigma > 0$ such that*

$$F(\hat{x}) \leq F(x)$$

for all $x \in C$ with $g(x) \in M$ and $\|\hat{x} - x\|_{\mathcal{X}} \leq \sigma$.

Now we suppose that the problem (4.9) has an local optimal solution $\hat{x}$, and we introduce the conical hulls of $C - \{\hat{x}\}$ and $M - \{z\}$, respectively, by

$$\begin{array}{rcl} C(\hat{x}) & = & \{x \in \mathcal{X} \mid \exists \beta \geq 0,\ \exists c \in C,\ x = \beta(c - \hat{x})\}, \\ M(z) & = & \{\zeta \in \mathcal{Z} \mid \exists \lambda \geq 0,\ \exists k \in M,\ \zeta = k - \lambda z\}. \end{array}$$

The main result in [17] on the existence of a Lagrange multiplier for (4.9) is stated next.

Theorem 5. *Let $\hat{x}$ be an optimal solution of the problem (4.9) satisfying the following constraints qualification*

$$g'(\hat{x}) \cdot C(\hat{x}) - M(g(\hat{x})) = \mathcal{Z}. \tag{4.10}$$

Then there exists a Lagrange multiplier $z^* \in \mathcal{Z}^*$ such that

$$\langle z^*, \zeta \rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq 0 \quad \forall \zeta \in M, \tag{4.11}$$

$$\langle z^*, g(\hat{x}) \rangle_{\mathcal{Z}^*,\mathcal{Z}} = 0, \tag{4.12}$$

$$F'(\hat{x}) - z^* \circ g'(\hat{x}) \in C(\hat{x})_+, \tag{4.13}$$

where $A_+ = \{x^* \in \mathcal{X}^* : \langle x^*, a \rangle_{\mathcal{X}^*,\mathcal{X}} \geq 0\, \forall a \in A\}$, $\mathcal{Z}^*$ and $\mathcal{X}^*$ are the topological dual spaces of $\mathcal{Z}$ and $\mathcal{X}$, respectively, and $(z^* \circ g'(\hat{x}))d = \langle z^*, g'(\hat{x})d \rangle_{\mathcal{Z}^*,\mathcal{Z}}$ $\forall d \in \mathcal{X}$.

We apply Theorem 5 to $(\mathcal{P}_\gamma)$. For this purpose we set

$$\begin{aligned} &\mathcal{X} = \boldsymbol{\mathcal{V}}^{hN} \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \times \boldsymbol{L}^2(\Omega_T), \\ &C = \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \times \boldsymbol{L}^2(\Omega_T), \\ &\mathcal{Z} = \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \times \boldsymbol{L}^2(\Omega) \times \boldsymbol{L}^2(\Omega_T) \times \mathbb{R} \times \mathbb{R}, \\ &M = \{\boldsymbol{0}\} \times \{\boldsymbol{0}\} \times \boldsymbol{L}^2_+(\Omega_T) \times \mathbb{R}_+ \times \mathbb{R}_+, \\ &\hat{x} = (\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma, \boldsymbol{\xi}_\gamma), \\ &F(\hat{x}) = J_\gamma(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma), \\ &g(x) = \begin{cases} \varepsilon \partial_t \boldsymbol{y}_\gamma + \varepsilon \boldsymbol{A} \boldsymbol{y}_\gamma - \frac{1}{\varepsilon} \boldsymbol{y}_\gamma - \boldsymbol{u}_\gamma - \frac{1}{\varepsilon} \boldsymbol{\xi}_\gamma, \\ \boldsymbol{y}_\gamma(0) - \boldsymbol{y}_0, \\ \boldsymbol{\xi}_\gamma, \\ \alpha_\gamma - (\boldsymbol{\xi}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)}, \\ R - \frac{1}{2} \|\boldsymbol{\xi}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)}. \end{cases} \end{aligned}$$

Then we have for the convex hull of $\boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \times \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T) \times \boldsymbol{L}^2(\Omega_T) - \{(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma, \boldsymbol{\xi}_\gamma)^T\}$

$$C\begin{pmatrix}\boldsymbol{y}_\gamma\\ \boldsymbol{u}_\gamma\\ \boldsymbol{\xi}_\gamma\end{pmatrix} = \left\{\begin{pmatrix}\boldsymbol{c}\\ \boldsymbol{d}\\ \boldsymbol{e}\end{pmatrix} \in \mathcal{X} \mid \exists\beta \geq 0,\ \exists \begin{pmatrix}\tilde{\boldsymbol{c}}\\ \tilde{\boldsymbol{d}}\\ \tilde{\boldsymbol{e}}\end{pmatrix} \in C,\ \begin{pmatrix}\boldsymbol{c}\\ \boldsymbol{d}\\ \boldsymbol{e}\end{pmatrix} = \beta \begin{pmatrix}\tilde{\boldsymbol{c}} - \boldsymbol{y}_\gamma\\ \tilde{\boldsymbol{d}} - \boldsymbol{u}_\gamma\\ \tilde{\boldsymbol{e}} - \boldsymbol{\xi}_\gamma\end{pmatrix}\right\}.$$

The constraint qualification (4.10) in our setting requires the existence of

$$\begin{array}{rcll}
\boldsymbol{c} & := & (c_1, \cdots, c_N)^T & \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{T\Sigma}},\\
\boldsymbol{d} & := & (d_1, \cdots, d_N)^T & \in \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T),\\
\boldsymbol{e} & := & (e_1, \cdots, e_N)^T & \in \boldsymbol{L}^2(\Omega_T),\\
\boldsymbol{k} & := & (k_1, \cdots, k_N)^T & \in \boldsymbol{L}^2_+(\Omega_T),
\end{array}$$

and $(k_{N+1}, k_{N+2}, \lambda)^T \in \mathbb{R}^3_+$ such that for arbitrary given

$$\begin{array}{rcll}
\boldsymbol{z}_1 & := & (z_1, \cdots, z_N)^T & \in \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T),\\
\boldsymbol{z}_2 & := & (z_{N+1}, \cdots, z_{2N})^T & \in \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega),\\
\boldsymbol{z}_3 & := & (z_{2N+1}, \cdots, z_{3N})^T & \in \boldsymbol{L}^2(\Omega_T),
\end{array}$$

and $(z_{3N+1}, z_{3N+2}) \in \mathbb{R}^2$ the following system holds

$$\boldsymbol{z}_1 = \varepsilon\partial_t \boldsymbol{c} + \varepsilon \boldsymbol{A}\boldsymbol{c} - \frac{1}{\varepsilon}\boldsymbol{c} - \boldsymbol{d} - \frac{1}{\varepsilon}\boldsymbol{e} \quad \text{in } \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T), \tag{4.14}$$

$$\boldsymbol{z}_2 = \boldsymbol{c}(0) \quad \text{in } \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega), \tag{4.15}$$

$$\boldsymbol{z}_3 = \boldsymbol{e} - (\boldsymbol{k} - \lambda\boldsymbol{\xi}_\gamma) \quad \text{in } \boldsymbol{L}^2(\Omega_T), \tag{4.16}$$

$$z_{3N+1} = -(\boldsymbol{c}, \boldsymbol{\xi}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} - (\boldsymbol{e}, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} - (k_{N+1} - \lambda(\alpha_\gamma - (\boldsymbol{\xi}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)})), \tag{4.17}$$

$$z_{3N+2} = -(\boldsymbol{\xi}_\gamma, \boldsymbol{e})_{\boldsymbol{L}^2(\Omega_T)} - (k_{N+2} - \lambda(R - \frac{1}{2}\|\boldsymbol{\xi}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)})). \tag{4.18}$$

By virtue of **(H1)** and by the classical theory of parabolic partial differential equations (see [7], for example), the system

$$\varepsilon\partial_t \boldsymbol{c} + \varepsilon \boldsymbol{A}\boldsymbol{c} - \frac{1}{\varepsilon}\boldsymbol{c} = \boldsymbol{z}_1 - \frac{1}{\varepsilon}\boldsymbol{e} \quad \text{in } \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T), \tag{4.19}$$

$$\boldsymbol{z}_2 = \boldsymbol{c}(0) \quad \text{in } \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega), \tag{4.20}$$

admits a unique solution $\boldsymbol{c} \in \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{T\Sigma}}$ for every $\boldsymbol{z}_1 \in \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ and $\boldsymbol{e} \in \boldsymbol{L}^2(\Omega_T)$.

Therefore, a solution of (4.14)-(4.18) is obtained by choosing

$$
\begin{aligned}
\boldsymbol{d} &= \boldsymbol{0},\\
\lambda &= \rho^3, \quad \boldsymbol{e} = \rho \boldsymbol{f} - \rho^2 \boldsymbol{\xi}_\gamma\\
\boldsymbol{c} &\text{ solution of } (4.19) - (4.20),\\
\boldsymbol{k} &= (\rho^3 - \rho^2)\boldsymbol{\xi}_\gamma + \rho \boldsymbol{f} - \boldsymbol{z}_3,\\
k_{N+1} &= \rho^3(\alpha_\gamma - (\boldsymbol{\xi}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)}) + \rho^2(\boldsymbol{\xi}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} - \rho(\boldsymbol{y}_\gamma, \boldsymbol{f})_{\boldsymbol{L}^2(\Omega_T)}\\
&\quad - (\boldsymbol{c}, \boldsymbol{\xi}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} - z_{3N+1},\\
k_{N+2} &= \rho^3(R - \frac{1}{2}\|\boldsymbol{\xi}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)}) + \rho^2\|\boldsymbol{\xi}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} - \rho(\boldsymbol{\xi}_\gamma, \boldsymbol{f})_{\boldsymbol{L}^2(\Omega_T)} - z_{3N+2}
\end{aligned}
$$

for some $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega_T)$ with $\boldsymbol{f} > \boldsymbol{0}$ a.e. in Ω_T, and $\rho > 0$ large enough such that $\boldsymbol{k}, k_{N+1}$ and k_{N+2} are nonnegative. □

Consequently problem $(\mathcal{P}_\gamma)$ satisfies the constraint qualification (4.10). Hence, according to Theorem 5, the set of Lagrange multipliers is nonempty and bounded, i.e. introducing

$$\lambda^i_\gamma := \max(0, \overline{\lambda} - \gamma y^i_\gamma), \qquad \boldsymbol{\lambda}_\gamma := (\lambda^1_\gamma, \dots, \lambda^N_\gamma)^T,$$

we have the following

Proposition 2. *Let $(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma, \boldsymbol{\xi}_\gamma)$ be a solution for the problem $(\mathcal{P}_\gamma)$. Then there exists a Lagrange multiplier vector $(\boldsymbol{p}_\gamma, \boldsymbol{\mu}_\gamma, r_\gamma, \kappa_\gamma)$ in $\boldsymbol{W}(0,T)_{\boldsymbol{T\Sigma}} \times \boldsymbol{L}^2(\Omega_T) \times \mathbb{R} \times \mathbb{R}$ such that the following first order optimality system holds*

$$
\begin{aligned}
&- \varepsilon \partial_t \boldsymbol{p}_\gamma + \varepsilon \boldsymbol{A}^* \boldsymbol{p}_\gamma - \frac{1}{\varepsilon}\boldsymbol{p}_\gamma - r_\gamma \boldsymbol{\xi}_\gamma +\\
&\qquad\qquad + \boldsymbol{\lambda}_\gamma = \nu_d(\boldsymbol{y}_\gamma - \boldsymbol{y}_d) \quad \text{in} \quad L^2(0,T;\boldsymbol{H}^1(\Omega)^*), && (4.21)\\
&\boldsymbol{p}_\gamma(T,\cdot) = \nu_T(\boldsymbol{y}_\gamma(T,\cdot) - \boldsymbol{y}_T), \ \text{a.e. in } \Omega, && (4.22)\\
&\boldsymbol{p}_\gamma + \frac{\nu_u}{\varepsilon}\boldsymbol{u}_\gamma = \boldsymbol{0} \ \text{a.e. in } \Omega_T, && (4.23)\\
&\kappa_\gamma \boldsymbol{\xi}_\gamma + \frac{1}{\varepsilon}\boldsymbol{p}_\gamma - \boldsymbol{\mu}_\gamma + r_\gamma \boldsymbol{y}_\gamma = \boldsymbol{0} \ \text{a.e. in } \Omega_T, && (4.24)\\
&\boldsymbol{\xi}_\gamma \geq \boldsymbol{0} \ \text{a.e. in } \Omega_T, \quad \boldsymbol{\mu}_\gamma \geq \boldsymbol{0} \ \text{a.e. in } \Omega_T, \quad (\boldsymbol{\xi}_\gamma, \boldsymbol{\mu}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} = 0, && (4.25)\\
&\kappa_\gamma \geq 0, \quad \frac{1}{2}\|\boldsymbol{\xi}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} \leq R, \quad \frac{\kappa_\gamma}{2}\|\boldsymbol{\xi}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} = \kappa_\gamma R, && (4.26)\\
&r_\gamma \geq 0, \quad (\boldsymbol{y}_\gamma, \boldsymbol{\xi}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \leq \alpha_\gamma, \quad r_\gamma(\boldsymbol{y}_\gamma, \boldsymbol{\xi}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} = r_\gamma \alpha_\gamma, && (4.27)\\
&\varepsilon \partial_t \boldsymbol{y}_\gamma + \varepsilon \boldsymbol{A} \boldsymbol{y}_\gamma - \frac{1}{\varepsilon}\boldsymbol{y}_\gamma - \boldsymbol{u}_\gamma - \frac{1}{\varepsilon}\boldsymbol{\xi}_\gamma = \boldsymbol{0} \quad \text{in} \quad \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T), && (4.28)\\
&\boldsymbol{y}_\gamma(0) = \boldsymbol{y}_0 \ \text{a.e. in } \Omega, && (4.29)
\end{aligned}
$$

Proof. For every fixed $0 < \gamma$ we know by virtue of Theorem 5 that $\boldsymbol{p}_\gamma \in \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$. Furthermore we obtain $\boldsymbol{p}_\gamma \in \boldsymbol{W}(0,T)_{\boldsymbol{T\Sigma}}$ by the classical theory of parabolic partial differential equations, see for example [7]. □

5 Optimality for the limit problem ($\mathcal{P}$)

In this section we investigate the convergence of a sequence $(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma, \boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma, \boldsymbol{\mu}_\gamma, r_\gamma, \kappa_\gamma)_{\gamma>0}$ satisfying the optimality conditions (4.21)-(4.29). For this purpose we make the following assumptions:

- **(O1)** Let $\{\boldsymbol{u}_\gamma\}$ be bounded in $\boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T)$ uniformly in $\gamma \in (0,\infty)$,
- **(O2)** we choose α_γ such that $\frac{1}{\alpha_\gamma\sqrt{\gamma}} \leq C$,
- **(O3)** we assume that $\kappa_\gamma\gamma \leq C$.

Here and in what follows, C denotes a generic positive constant that may take different values at different occurrences but not depending on γ. We also introduce the notations

$$\begin{aligned}
&\vartheta^i_\gamma = r_\gamma\xi^i_\gamma - \lambda^i_\gamma, \quad \boldsymbol{\vartheta}_\gamma := (\vartheta^i_\gamma)^N_{i=1},\\
&N^i_\gamma = \{(t,x) \in \Omega_T : y^i_\gamma < 0\},\\
&P^i_\gamma = \Omega_T \setminus N^i_\gamma,\\
&\Pi^i_\gamma = \{(t,x) \in \Omega_T : \overline{\lambda} - \gamma y^i_\gamma \geq 0\}.
\end{aligned}$$

Lemma 2. *Let $\gamma > 0$, **(O1)-(O3)** hold and let $(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma, \boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma, \boldsymbol{\mu}_\gamma, r_\gamma, \kappa_\gamma)$ be a solution of the optimality system (4.21)-(4.29). Then we have*

1.)	$\boldsymbol{y}_\gamma$	*is bounded*	*in*	$\boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}}$	*uniformly in*	$\gamma \in (0,\infty)$,
2.)	$\boldsymbol{y}_\gamma(T)$	*is bounded*	*in*	$\boldsymbol{L}^2_{\boldsymbol{\Sigma}}(\Omega)$	*uniformly in*	$\gamma \in (0,\infty)$,
3.)	$\boldsymbol{p}_\gamma$	*is bounded*	*in*	$L^2(0,T;\boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega))$	*uniformly in*	$\gamma \in (0,\infty)$,
4.)	$\boldsymbol{p}_\gamma(0)$	*is bounded*	*in*	$\boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega)$	*uniformly in*	$\gamma \in (0,\infty)$,
5.)	$\boldsymbol{u}_\gamma$	*is bounded*	*in*	$L^2(0,T;\boldsymbol{H}^1(\Omega))$	*uniformly in*	$\gamma \in (0,\infty)$,
6.)	$\frac{1}{\sqrt{\gamma}}\boldsymbol{\lambda}_\gamma$	*is bounded*	*in*	$\boldsymbol{L}^2(\Omega_T)$	*uniformly in*	$\gamma \in (1,\infty)$,
7.)	$\partial_t\boldsymbol{p}_\gamma$	*is bounded*	*in*	$\boldsymbol{W}(0,T)^*$	*uniformly in*	$\gamma \in (0,\infty)$,
8.)	$\boldsymbol{\vartheta}_\gamma$	*is bounded*	*in*	$\boldsymbol{W}(0,T)^*$	*uniformly in*	$\gamma \in (0,\infty)$,

Proof. By virtue of **(O1)** Proposition 1 gives the estimates 1) and 2).

3) Testing each component of (4.21) by p^i_γ and summing over $i = 1, .., N$ we get:

$$\begin{aligned}&\frac{\varepsilon}{2}\|\boldsymbol{p}_\gamma(0)\|^2_{\boldsymbol{L}^2(\Omega)} + \varepsilon\|\nabla \boldsymbol{p}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} = \\ &= \left(\frac{\varepsilon}{2}\|\boldsymbol{p}_\gamma(T)\|^2_{\boldsymbol{L}^2(\Omega)} + \frac{1}{\varepsilon}\|\boldsymbol{p}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} + r_\gamma(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)}\right) + \\ &\quad \left(-(\boldsymbol{\lambda}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} + (\boldsymbol{y}_\gamma - \boldsymbol{y}_d, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)}\right).\end{aligned}$$

For continuing the proof we need two claims.
Claim 1: $(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \le 0$.
Proof of Claim 1 Multiplying (4.24) by $\boldsymbol{\xi}_\gamma$ and taking into account that $(\boldsymbol{\xi}_\gamma, \boldsymbol{\mu}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} = 0$ from (4.25), we obtain

$$(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} = -\kappa_\gamma\|\boldsymbol{\xi}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} - \alpha_\gamma r_\gamma \le 0.$$

Claim 2 :$-(\boldsymbol{\lambda}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \le \|\overline{\lambda}\|^2_{L^2(\Omega_T)}\|\boldsymbol{p}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} + C$.
Proof of Claim 2 From the definition of λ^i_γ and (1.3) it follows

$$-(\boldsymbol{\lambda}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} = \sum_{i=1}^N \left\{-(\overline{\lambda}, p^i_\gamma)_{L^2(\Pi^i_\gamma)} + \gamma(y^i_\gamma, p^i_\gamma)_{L^2(\Pi^i_\gamma \cap P^i_\gamma)} + \gamma(y^i_\gamma, p^i_\gamma)_{L^2(\Pi^i_\gamma \cap N^i_\gamma)}\right\}.$$

On $\Pi^i_\gamma \cap P^i_\gamma$ we have $0 \le \gamma y^i_\gamma \le \overline{\lambda}$. Then

$$-(\boldsymbol{\lambda}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \le \|\overline{\lambda}\|_{L^2(\Omega_T)}\|\boldsymbol{p}_\gamma\|_{\boldsymbol{L}^2(\Omega_T)} + \sum_{i=1}^N \gamma(y^i_\gamma, p^i_\gamma)_{L^2(\Pi^i_\gamma \cap N^i_\gamma)}.$$

Multiplying (4.24) componentwise by $\gamma y^i_\gamma \chi_{N^i_\gamma}$ where $\chi_{N^i_\gamma}$ is the characteristic function of N^i_γ, we get

$$\begin{aligned}\frac{\gamma}{\varepsilon}(y^i_\gamma, p^i_\gamma)_{L^2(N^i_\gamma)} &= -\gamma\kappa_\gamma(y^i_\gamma, \xi^i_\gamma)_{L^2(N^i_\gamma)} + \gamma(y^i_\gamma, \mu^i_\gamma)_{L^2(N^i_\gamma)} - \gamma r_\gamma(y^i_\gamma, y^i_\gamma)_{L^2(N^i_\gamma)} \\ &\le -\gamma\kappa_\gamma(y^i_\gamma, \xi^i_\gamma)_{L^2(N^i_\gamma)}.\end{aligned}$$

By virtue of **(O3)**, the boundedness of $\boldsymbol{\xi}_\gamma$ in $\boldsymbol{L}^2(\Omega_T)$, and 1) we obtain

$$-(\boldsymbol{\lambda}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \le \|\overline{\lambda}\|^2_{L^2(\Omega_T)}\|\boldsymbol{p}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} + C.$$

Now using (4.22), 1), *Claim 1* and *Claim 2* we get

$$\frac{1}{2}\|\boldsymbol{p}_\gamma(0)\|^2_{\boldsymbol{L}^2(\Omega)} + \|\nabla \boldsymbol{p}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} \le C(\varepsilon)\|\boldsymbol{p}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)}.$$

A Grönwall argument finally provides 3) and 4).
5) is obtained by (4.23) and 3).
6) By testing (4.21) componentwise against y_γ^i and (4.28) componentwise against p_γ^i and finally summation over $i = 1, .., N$ we get after standard calculations

$$\begin{aligned}(\boldsymbol{\lambda}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} =& r_\gamma(\boldsymbol{\xi}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} - \frac{1}{\varepsilon}(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} + \\ &+ \nu_T(\boldsymbol{y}_\gamma(T), \boldsymbol{y}_\gamma(T) - \boldsymbol{y}_T)_{L^2(\Omega)} + \nu_d(\boldsymbol{y}_\gamma, \boldsymbol{y}_\gamma - \boldsymbol{y}_d)_{\boldsymbol{L}^2(\Omega_T)}.\end{aligned}$$

For continuing the proof we need a further claim.
Claim 3: $r_\gamma \alpha_\gamma \le C$.
Proof of Claim 3 Multiplying (4.24) by $\boldsymbol{\xi}_\gamma$ and using (4.25) and (4.27) we get

$$\begin{aligned} r_\gamma \alpha_\gamma &= r_\gamma(\boldsymbol{\xi}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \\ &= -(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} + (\boldsymbol{\xi}_\gamma, \boldsymbol{\mu}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} - \kappa_\gamma \|\boldsymbol{\xi}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} \\ &= -(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} - \kappa_\gamma \|\boldsymbol{\xi}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} \\ &\le -(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \\ &\le \|\boldsymbol{\xi}_\gamma\|_{\boldsymbol{L}^2(\Omega_T)} \|\boldsymbol{p}_\gamma\|_{\boldsymbol{L}^2(\Omega_T)}. \end{aligned} \tag{5.1}$$

Then from 3) and the boundness of $\boldsymbol{\xi}_\gamma$ in $\boldsymbol{L}^2(\Omega_T)$ we deduce *Claim 3*.
Therefore from 1), 3) and *Claim 3* we deduce that

$$\left|(\boldsymbol{\lambda}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)}\right| \le C. \tag{5.2}$$

The definition of $\boldsymbol{\lambda}_\gamma$ yields

$$\|\boldsymbol{\lambda}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} = \sum_{i=1}^N (\lambda_\gamma^i, \lambda_\gamma^i)_{L^2(\Omega_T)} = \sum_{i=1}^N \left\{ (\lambda_\gamma^i, \overline{\lambda})_{L^2(\Omega_T)} - \gamma(\lambda_\gamma^i, y_\gamma^i)_{\Pi_\gamma^i} \right\}.$$

Then

$$\frac{1}{\gamma}\|\boldsymbol{\lambda}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} \le \frac{1}{\gamma}\|\boldsymbol{\lambda}_\gamma\|_{\boldsymbol{L}^2(\Omega_T)} \|\overline{\lambda}\|_{L^2(\Omega_T)} + \left|(\boldsymbol{\lambda}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)}\right|.$$

Using (5.2) we obtain

$$\frac{1}{\gamma}\|\boldsymbol{\lambda}_\gamma\|^2_{\boldsymbol{L}^2(\Omega_T)} \le \frac{1}{\sqrt{\gamma}}\|\boldsymbol{\lambda}_\gamma\|_{\boldsymbol{L}^2(\Omega_T)} \|\overline{\lambda}\|_{L^2(\Omega_T)} + C \quad \forall \gamma \ge 1.$$

In particular we infer

$$\frac{1}{\sqrt{\gamma}}\|\boldsymbol{\lambda}_\gamma\|_{\boldsymbol{L}^2(\Omega_T)} \le C \quad \forall \gamma \ge 1.$$

7) Let $\boldsymbol{v} \in \boldsymbol{W}(0,T)$. Using integration by parts we obtain

$$\langle \partial_t \boldsymbol{p}_\gamma, \boldsymbol{v}\rangle = -\langle \partial_t \boldsymbol{v}, \boldsymbol{p}_\gamma\rangle + \nu_T(\boldsymbol{y}_\gamma(T) - \boldsymbol{y}_T, \boldsymbol{v}(T))_{\boldsymbol{L}^2(\Omega)} - (\boldsymbol{p}_\gamma(0), \boldsymbol{v}(0))_{\boldsymbol{L}^2(\Omega)}.$$

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

$$|\langle \partial_t \boldsymbol{p}_\gamma, \boldsymbol{v}\rangle| \le (\|\boldsymbol{p}_\gamma\|_{L^2(0,T;\boldsymbol{H}^1(\Omega))} + \|\nu_T(\boldsymbol{y}_\gamma(T) - \boldsymbol{y}_T)\|_{\boldsymbol{L}^2(\Omega)} + \|\boldsymbol{p}_\gamma(0)\|_{\boldsymbol{L}^2(\Omega)})\|\boldsymbol{v}\|_{\boldsymbol{W}(0,T)}.$$

From 1) and 4) we deduce that

$$\|\partial_t \boldsymbol{p}_\gamma\|_{\boldsymbol{W}(0,T)^*} \le C.$$

8) The boundedness of $\boldsymbol{\vartheta}_\gamma$ in $\boldsymbol{W}(0,T)^*$ follows from (4.14). □

Theorem 6. *Let **(O1)-(O3)** hold and let* $(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma, \boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma, \boldsymbol{\mu}_\gamma, r_\gamma, \kappa_\gamma)_{\gamma>0}$ *be a sequence of solutions of the optimality system (4.21)-(4.29). Then there exists*

$$\begin{pmatrix} \boldsymbol{y}^\star \\ \boldsymbol{u}^\star \\ \boldsymbol{\xi}^\star \\ \boldsymbol{p}^\star \\ \boldsymbol{\vartheta}^\star \end{pmatrix} \in \begin{pmatrix} \boldsymbol{\mathcal{V}}^{hN}_{\boldsymbol{\Sigma}} \\ L^2(0,T;\boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega)) \\ \boldsymbol{L}^2(\Omega_T) \\ L^2(0,T;\boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega)) \\ \boldsymbol{W}(0,T)^* \end{pmatrix}$$

and a subsequence still denoted by $(\boldsymbol{y}_\gamma, \boldsymbol{u}_\gamma, \boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma, \boldsymbol{\mu}_\gamma, r_\gamma, \kappa_\gamma)_{\gamma>0}$ *such that*

$$\begin{array}{llll}
\boldsymbol{y}_\gamma \longrightarrow \boldsymbol{y}^\star & \textit{weakly} & \textit{in} & L^2(0,T;\boldsymbol{H}^2(\Omega)), \\
\boldsymbol{y}_\gamma \longrightarrow \boldsymbol{y}^\star & \textit{weakly} & \textit{in} & H^1(0,T;\boldsymbol{L}^2(\Omega)), \\
\boldsymbol{y}_\gamma \longrightarrow \boldsymbol{y}^\star & \textit{weak-star} & \textit{in} & L^\infty(0,T;\boldsymbol{H}^1(\Omega)), \\
\boldsymbol{u}_\gamma \longrightarrow \boldsymbol{u}^\star & \textit{weakly} & \textit{in} & L^2(0,T;\boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega)), \\
\boldsymbol{\xi}_\gamma \longrightarrow \boldsymbol{\xi}^\star & \textit{weakly} & \textit{in} & \boldsymbol{L}^2(\Omega_T), \\
\boldsymbol{p}_\gamma \longrightarrow \boldsymbol{p}^\star & \textit{weakly} & \textit{in} & L^2(0,T;\boldsymbol{H}^1_{\boldsymbol{T\Sigma}}(\Omega)), \\
\partial_t\boldsymbol{p}_\gamma \longrightarrow \partial_t\boldsymbol{p}^\star & \textit{weakly} & \textit{in} & \boldsymbol{W}(0,T)^*, \\
\boldsymbol{\vartheta}_\gamma \longrightarrow \boldsymbol{\vartheta}^\star & \textit{weakly} & \textit{in} & \boldsymbol{W}(0,T)^*.
\end{array}$$

The limit element $(\boldsymbol{y}^\star, \boldsymbol{u}^\star, \boldsymbol{\xi}^\star, \boldsymbol{p}^\star, \boldsymbol{\vartheta}^\star)$ *satisfies the following optimality system:*

$$\begin{aligned}
&\varepsilon\langle \boldsymbol{p}^\star, \partial_t \boldsymbol{\varphi}\rangle_{L^2(0,T;\boldsymbol{H}^1(\Omega)),L^2(0,T;\boldsymbol{H}^1(\Omega)^*)} + \varepsilon(\nabla \boldsymbol{p}^\star, \nabla\boldsymbol{\varphi})_{\boldsymbol{L}^2(\Omega_T)}+\\
&+(\frac{1}{\varepsilon}\boldsymbol{p}^\star + \nu_d(\boldsymbol{y}^\star - \boldsymbol{y}_d), \boldsymbol{\varphi})_{\boldsymbol{L}^2(\Omega_T)} - \langle \boldsymbol{\vartheta}^\star, \boldsymbol{\varphi}\rangle_{\boldsymbol{W}(0,T)^*,\boldsymbol{W}(0,T)} +\\
&+\nu_T(\boldsymbol{y}^\star(T) - \boldsymbol{y}_T, \boldsymbol{\varphi}(T))_{\boldsymbol{L}^2(\Omega)} = \boldsymbol{0},\\
&\forall \boldsymbol{\varphi} \in \boldsymbol{Z} = \{\boldsymbol{z} \in \boldsymbol{W}(0,T), \boldsymbol{z}(0,\cdot) = \boldsymbol{0}\}, \qquad (5.3)
\end{aligned}$$

$$\boldsymbol{p}^\star + \frac{\nu_u}{\varepsilon}\boldsymbol{u}^\star = \boldsymbol{0} \text{ a.e. in } \Omega_T, \qquad (5.4)$$

$$\boldsymbol{y}^\star \geq 0 \text{ a.e. in } \Omega_T,\ \boldsymbol{\xi}^\star \geq 0 \text{ a.e. in } \Omega_T,\ (\boldsymbol{\xi}^\star, \boldsymbol{y}^\star)_{\boldsymbol{L}^2(\Omega_T)} = 0, \qquad (5.5)$$

$$\varepsilon\partial_t \boldsymbol{y}^\star + \varepsilon \boldsymbol{A}\boldsymbol{y}^\star - \frac{1}{\varepsilon}\boldsymbol{y}^\star - \boldsymbol{u}^\star - \frac{1}{\varepsilon}\boldsymbol{\xi}^\star = \boldsymbol{0} \text{ in } \boldsymbol{L}^2_{\boldsymbol{T\Sigma}}(\Omega_T), \qquad (5.6)$$

$$\boldsymbol{y}^\star(0) = \boldsymbol{y}_0 \text{ a.e. in } \Omega, \qquad (5.7)$$

$$R \geq \frac{1}{2}\|\boldsymbol{\xi}^\star\|^2_{\boldsymbol{L}^2(\Omega_T)}, \qquad (5.8)$$

$$\lim_{\gamma\uparrow\infty}(\boldsymbol{p}_\gamma, \boldsymbol{\xi}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} = 0, \qquad (5.9)$$

$$\lim_{\gamma\uparrow\infty}(\boldsymbol{\vartheta}_\gamma, \boldsymbol{y}^+_\gamma)_{\boldsymbol{L}^2(\Omega_T)} = 0, \qquad (5.10)$$

$$\lim_{\gamma\uparrow\infty}(\boldsymbol{\vartheta}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \leq 0. \qquad (5.11)$$

Proof. The convergence results are consequences of the estimates 1)-8) established in Lemma 2, and (5.3), (5.4), (5.6) and (5.7) are consequences of the convergence results. To show (5.5) we use 6) of Lemma 2 and proceed like in the proof of Theorem 4.

Next we prove (5.9). If (r_γ) is bounded, then $\lim_{\gamma\uparrow\infty}(\boldsymbol{p}_\gamma, \boldsymbol{\xi}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} = 0$ follows immediately from (5.1), **(O2)**, **(O3)** and the boundedness of $\boldsymbol{\xi}_\gamma$ in $\boldsymbol{L}^2(\Omega_T)$. In the case where (r_γ) is unbounded we take $\boldsymbol{p}_\gamma$ as a testfunction in the adjoint system (4.21)-(4.22) and estimate

$$\begin{aligned}
r_\gamma(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \geq & \varepsilon\|\nabla \boldsymbol{y}_\gamma\|_{\boldsymbol{L}^2(\Omega_T)} + \frac{\varepsilon}{2}\|\boldsymbol{p}_\gamma(0)\|^2_{\boldsymbol{L}^2(\Omega)} - \frac{\varepsilon}{2}\|\boldsymbol{y}_\gamma(T) - \boldsymbol{y}_T\|^2_{\boldsymbol{L}^2(\Omega)}+\\
&+ (\boldsymbol{\lambda}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} + \nu_d(\boldsymbol{y}_\gamma - \boldsymbol{y}_d, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)}-\\
&- \frac{1}{\varepsilon}\|\boldsymbol{y}_\gamma\|_{\boldsymbol{L}^2(\Omega_T)}.
\end{aligned}$$

This together with 1), 2), 3), 4) of Lemma 2, *Claim 2* in the proof of Lemma 2 and (5.2) yields

$$r_\gamma(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \geq -C,$$

which implies

$$\lim_{\gamma\uparrow\infty}(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \geq 0.$$

From *Claim 1* in the proof of Lemma 2 we consequently get

$$\lim_{\gamma\uparrow\infty}(\boldsymbol{\xi}_\gamma, \boldsymbol{p}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} = 0.$$

We prove (5.10): Due to (5.1) we find

$$\lim_{\gamma\uparrow\infty} r_\gamma \alpha_\gamma = 0. \tag{5.12}$$

Moreover we have

$$\begin{aligned}
0 \leq (\boldsymbol{\lambda}_\gamma, \boldsymbol{y}_\gamma^+)_{\boldsymbol{L}^2(\Omega_T)} &= \sum_{i=1}^N (\lambda_\gamma^i, (y_\gamma^i)^+)_{L^2(\Omega_T)} \\
&\leq \sum_{i=1}^N \left\{ (\overline{\lambda}, y_\gamma^i)_{L^2(\Pi_\gamma^i \cap P_\gamma^i)} - \gamma (y_\gamma^i, y_\gamma^i)_{L^2(\Pi_\gamma^i \cap P_\gamma^i)} \right\} \\
&\leq \frac{N}{\gamma} \|\overline{\lambda}\|_{L^2(\Omega_T)}
\end{aligned} \tag{5.13}$$

and using the uniform boundedness of $\boldsymbol{\xi}_\gamma$ in $\boldsymbol{L}^2(\Omega_T)$ and **(O2)** we obtain

$$\begin{aligned}
r_\gamma(\boldsymbol{\xi}_\gamma, \boldsymbol{y}_\gamma^+)_{\boldsymbol{L}^2(\Omega_T)} &= r_\gamma(\boldsymbol{\xi}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} - r_\gamma \sum_{i=1}^N (\xi_\gamma^i, y_\gamma^i)_{L^2(N_\gamma^i)} \\
&= r_\gamma \alpha_\gamma - r_\gamma \sum_{i=1}^N (\xi_\gamma^i, y_\gamma^i)_{L^2(N_\gamma^i)} \\
&\leq r_\gamma \alpha_\gamma + r_\gamma \alpha_\gamma \sum_{i=1}^N \frac{1}{\alpha_\gamma \sqrt{\gamma}} \|\xi_\gamma^i\|_{L^2(\Omega_T)} \|\sqrt{\gamma} y_\gamma^i\|_{L^2(N_\gamma^i)} \\
&\leq C r_\gamma \alpha_\gamma \sum_{i=1}^N \|\sqrt{\gamma} y_\gamma^i\|_{L^2(N_\gamma^i)}.
\end{aligned} \tag{5.14}$$

From (5.13) and (5.14) it follows that

$$\begin{aligned}
(\boldsymbol{\vartheta}_\gamma, \boldsymbol{y}_\gamma^+)_{\boldsymbol{L}^2(\Omega_T)} &= (r_\gamma \boldsymbol{\xi}_\gamma - \boldsymbol{\lambda}_\gamma, \boldsymbol{y}_\gamma^+)_{\boldsymbol{L}^2(\Omega_T)} \\
&\leq C r_\gamma \alpha_\gamma \sum_{i=1}^N \|\sqrt{\gamma} y_\gamma^i\|_{L^2(N_\gamma^i)} + \frac{N}{\gamma} \|\overline{\lambda}\|_{L^2(\Omega_T)}.
\end{aligned}$$

Hence, because of (5.12), it suffices to show that $\sum_{i=1}^{N} \|\sqrt{\gamma} y_\gamma^i\|_{L^2(N_\gamma^i)}$ is bounded for proving (5.10). For this purpose we use (5.2) and find

$$
\begin{aligned}
-C \leq (\boldsymbol{\lambda}_\gamma, \boldsymbol{y}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} &= \sum_{i=1}^{N} (\overline{\lambda} - \gamma y_\gamma^i, y_\gamma^i)_{L^2(\Pi_\gamma^i)} \\
&= \sum_{i=1}^{N} \left\{ (\overline{\lambda}, y_\gamma^i)_{L^2(\Pi_\gamma^i)} - \gamma (y_\gamma^i, y_\gamma^i)_{L^2(\Pi_\gamma^i \cap P_\gamma^i)} - \gamma (y_\gamma^i, y_\gamma^i)_{L^2(N_\gamma^i)} \right\}
\end{aligned}
$$

and further

$$
\begin{aligned}
\sum_{i=1}^{N} \gamma (y_\gamma^i, y_\gamma^i)_{L^2(N_\gamma^i)} &\leq \sum_{i=1}^{N} (\overline{\lambda}, y_\gamma^i)_{L^2(\Pi_\gamma^i)} + C \\
&\leq \sum_{i=1}^{N} \|\overline{\lambda}\|_{L^2(\Omega_T)} \|\boldsymbol{y}_\gamma\|_{L^2(\Omega_T)} + C \leq C.
\end{aligned}
$$

Consequently

$$
\sum_{i=1}^{N} \|\sqrt{\gamma} y_\gamma^i\|_{L^2(N_\gamma^i)} \leq C. \tag{5.15}
$$

and (5.10) holds true.
Finally we prove (5.11): We multiply (4.24) by $\boldsymbol{\lambda}_\gamma$ and estimate

$$
\begin{aligned}
-(\boldsymbol{p}_\gamma, \boldsymbol{\lambda}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} =& \kappa_\gamma (\boldsymbol{\xi}_\gamma, \boldsymbol{\lambda}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} - (\boldsymbol{\mu}_\gamma, \boldsymbol{\lambda}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} + r_\gamma (\boldsymbol{y}_\gamma, \boldsymbol{\lambda}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \\
\leq& \kappa_\gamma (\boldsymbol{\xi}_\gamma, \boldsymbol{\lambda}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} + r_\gamma (\boldsymbol{y}_\gamma, \boldsymbol{\lambda}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \\
\leq& \kappa_\gamma \sum_{i=1}^{N} (\xi_\gamma^i, \overline{\lambda})_{L^2(\Pi_\gamma^i)} - \kappa_\gamma \gamma \sum_{i=1}^{N} (\xi_\gamma^i, y_\gamma^i)_{L^2(N_\gamma^i)} \\
& + r_\gamma \sum_{i=1}^{N} (y_\gamma^i, \overline{\lambda})_{L^2(\Pi_\gamma^i \cap P_\gamma^i)} - \gamma r_\gamma \sum_{i=1}^{N} (y_\gamma^i, y_\gamma^i)_{L^2(\Pi_\gamma^i \cap P_\gamma^i)}.
\end{aligned}
$$

Since $\frac{\overline{\lambda}}{\gamma} \geq y_\gamma^i \geq 0$ on $\Pi_\gamma^i \cap P_\gamma^i$ we obtain

$$
\begin{aligned}
-(\boldsymbol{p}_\gamma, \boldsymbol{\lambda}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \leq& \kappa_\gamma \sum_{i=1}^{N} (\xi_\gamma^i, \overline{\lambda})_{L^2(\Pi_\gamma^i)} - \kappa_\gamma \sqrt{\gamma} \sum_{i=1}^{N} (\xi_\gamma^i, \sqrt{\gamma} y_\gamma^i)_{L^2(N_\gamma^i)} + \\
& + \frac{r_\gamma}{\gamma} \sum_{i=1}^{N} (\overline{\lambda}, \overline{\lambda})_{L^2(\Pi_\gamma^i \cap P_\gamma^i)}.
\end{aligned}
$$

From **(O2)**, **(O3)** and (5.12) we infer

$$\lim_{\gamma\uparrow\infty} \kappa_\gamma \sqrt{\gamma} = 0 \text{ and } \lim_{\gamma\uparrow\infty} \frac{r_\gamma}{\gamma} = \lim_{\gamma\uparrow\infty} \frac{1}{\sqrt{\gamma}} \frac{r_\gamma \alpha_\gamma}{\sqrt{\gamma}\alpha_\gamma} = 0.$$

Hence, using the boundedness of $\boldsymbol{\xi}_\gamma$ in $\boldsymbol{L}^2(\Omega_T)$ and (5.15) we get

$$\limsup_{\gamma\uparrow\infty} -(\boldsymbol{p}_\gamma, \boldsymbol{\lambda}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \leq 0. \tag{5.16}$$

On the other hand from *Claim 1* in the proof of Lemma 2 we have

$$\begin{aligned}(\boldsymbol{p}_\gamma, \boldsymbol{\vartheta}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} &= r_\gamma (\boldsymbol{p}_\gamma, \boldsymbol{\xi}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} - (\boldsymbol{p}_\gamma, \boldsymbol{\lambda}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \\ &\leq -(\boldsymbol{p}_\gamma, \boldsymbol{\lambda}_\gamma)_{\boldsymbol{L}^2(\Omega_T)}.\end{aligned} \tag{5.17}$$

Therefore, from (5.16) and (5.17), we deduce

$$\limsup_{\gamma\uparrow\infty} (\boldsymbol{p}_\gamma, \boldsymbol{\vartheta}_\gamma)_{\boldsymbol{L}^2(\Omega_T)} \leq 0,$$

which completes the proof. □

The optimality conditions (5.3)-(5.11) of Theorem 6 define a weak form of C-stationarity for the Allen-Cahn optimization problem. The results of Theorem 6 can be interpreted in the following way: The accumulation points of stationary points of the regularized-relaxed subproblems satisfy optimality conditions of W-stationarity-type. The product conditions, necessary for a C-stationarity-type condition, are satisfied in the sence of limits of pairings of weakly convergent sequences. The weak result is due to the low regularity of $\boldsymbol{\vartheta}_\gamma$.

Acknowledgment. The author wishes to thank Luise Blank and Harald Garcke for many fruitful hints and discussions. This work was supported by the SPP 1253 “Optimization with Partial Differential Equations” of the German Science Foundation (DFG) through the grant GA 695/5-2.

References

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

[2] M. Bergounioux, Optimal control of an obstacle problem, Appl. Math. Optim., 36(1997), pp. 147-172.

[3] L. Blank, H. Garcke, L. Sarbu, V. Styles, Primal-dual active set methods for Allen-Cahn variational inequalities with non-local constraints, DFG priority program "Optimization with PDEs", Preprint SPP1253-09-01.

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

[5] A. Bossavit, A. Damlamian and M. Fremond (editors) Free Boundary Problems: applications and theory Pitman Research Notes in Mathematics 121 Pitman Boston London Melbourne Pitman, London (1985).

[6] M. Brokate and J. Sprekels, Hysteresis and Phase Transitions, Applied Mathematical Sciences 121, Springer (1996).

[7] L.C. Evans, Partial Differential Equations. American Mathematical Society, Providence (1998).

[8] M. H. Farshbaf-Shaker, A penalty approach to optimal control of Allen-Cahn variational inequalities: MPEC-view, Universität Regensburg, Preprint Nr.06/2011

[9] M. H. Farshbaf-Shaker, Optimal control of vector-valued elastic Allen-Cahn MPECs, in preparation

[10] A. Fasano and M. Primicerio (editors) Free Boundary Problems, Theory Applications Pitman Research Notes in Mathematics 79 Pitman Boston London Melbourne Pitman, London (1983).

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

[12] H. Garcke, On a Cahn-Hilliard Model for Phase Separation with Elastic Misfit, Ann. de l'IHP - Analyse non linéaire 22 (2) (2005), 165-185.

[13] 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)

[14] 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. Control Optim., Vol. 48, No. 7, pp. 4419-4452.

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

[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] J. Zowe, S. Kurcyusz, Regularity and stability for the mathematical programming problem in Banach spaces, Applied Mathematics and Optimization 5 (1979) 49-62.