**Weierstraß-Institut**
**für Angewandte Analysis und Stochastik**

**Leibniz-Institut im Forschungsverbund Berlin e. V.**

Preprint ISSN 2198-5855

# Optimal control for shape memory alloys of the one-dimensional Frémond model

Pierluigi Colli[1], M. Hassan Farshbaf-Shaker[2], Ken Shirakawa[3], Noriaki Yamazaki[4]

submitted: July 6, 2020

[1] Dipartimento di Matematica
Università degli Studi di Pavia
Via Ferrata 5
27100 Pavia
Italy
E-Mail: pierluigi.colli@unipv.it

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: shaker@wias-berlin.de

[3] Department of Mathematics
Faculty of Education
Chiba University
1-33 Yayoi-cho, Inage-ku
Chiba, 263-8522
Japan
E-Mail: sirakawa@faculty.chiba-u.jp

[4] Department of Mathematics
Faculty of Engineering
Kanagawa University
3-27-1 Rokkakubashi, Kanagawa-ku
Yokohama, 221-8686
Japan
E-Mail: noriaki@kanagawa-u.ac.jp

No. 2737
Berlin 2020

2010 *Mathematics Subject Classification.* 49J20, 35K55, 35R35.

*Key words and phrases.* Optimal control problem, one-dimensional Frémond model, shape memory alloys, Mosco convergence, subdifferentials.

Pierluigi Colli gratefully acknowledges some support from the Italian Ministry of Education, University and Research (MIUR): Dipartimenti di Eccellenza Program (2018–2022) – Dept. of Mathematics "F. Casorati", University of Pavia; the GNAMPA (Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni) of INdAM (Istituto Nazionale di Alta Matematica); the IMATI – C.N.R. Pavia, Italy. In addition, this work was supported by Grant-in-Aid for Scientific Research (C) No. 16K05224 (Ken Shirakawa), JSPS. The second author MHFS acknowledges the financial support of the Berlin Mathematics Research Center MATH+ through Project AA2-4.

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

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

# Optimal control for shape memory alloys of the one-dimensional Frémond model

Pierluigi Colli, M. Hassan Farshbaf-Shaker, Ken Shirakawa, Noriaki Yamazaki

**Abstract**

In this paper, we consider optimal control problems for the one-dimensional Frémond model for shape memory alloys. This model is constructed in terms of basic functionals like free energy and pseudo-potential of dissipation. The state problem is expressed by a system of partial differential equations involving the balance equations for energy and momentum. We prove the existence of an optimal control that minimizes the cost functional for a nonlinear and nonsmooth state problem. Moreover, we show the necessary condition of the optimal pair by using optimal control problems for approximating systems.

## 1 Introduction

This paper is concerned with optimal control problems for a simplified version of the mathematical model proposed by Michel Frémond to describe the thermomechanical evolution of a shape memory alloy. In the one-dimensional setting, one can think to a metallic wire, which has the surprising property that it could be permanently deformed and then be forced to recover its original shape just by thermal means. In the microscopic scale, such phenomenon has been ascribed to (solid-solid) phase transitions between different configurations of the metallic lattice, known as austenite and martensite from the metallurgical terminology.

The Frémond model is a macroscopic model which is constructed in terms of basic functionals like free energy and pseudo-potential of dissipation, and it turns out to be consistent with the fundamental laws of Thermodynamics (cf. [26, Chapter 13]). The model leads to the system of partial differential equations and related conditions that is stated below. The balance equations for energy and momentum are coupled with the partial differential inclusion governing the evolution of the pointwise phase variables $\chi_1, \chi_2$ that are related to the volumetric fractions of austenite and martensite phases. The other unknown variable of the system is the absolute temperature $\theta$ and, in the fixed one-dimensional bounded interval $\Omega := (0,1)$, the following system is considered:

$$(L_0\theta - L_1\chi_1)_t - h\theta_{xx} = a_0 f(t,x) \ \text{ in } Q := (0,T)\times\Omega, \tag{1.1}$$

$$\mu_0 \begin{pmatrix}\chi_1\\ \chi_2\end{pmatrix}_t - \mu_1 \begin{pmatrix}\chi_1\\ \chi_2\end{pmatrix}_{xx} + \partial I_K(\chi_1,\chi_2) \ni \begin{pmatrix} l(\theta^c - \theta) \\ -\beta a_1 g(t)\alpha(\theta) + \beta\alpha(\theta)^2\chi_2\end{pmatrix} \ \text{ in } Q, \tag{1.2}$$

$$-h\theta_x(t,0) + k(\theta(t,0) - a_2\gamma_0(t)) = h\theta_x(t,1) + k(\theta(t,1) - a_3\gamma_1(t)) = 0, \quad t\in(0,T), \tag{1.3}$$

$$(\chi_i)_x(t,0) = (\chi_i)_x(t,1) = 0, \quad t\in(0,T),\ i=1,2, \tag{1.4}$$

DOI 10.20347/WIAS.PREPRINT.2737 Berlin 2020

$$\theta(0,x) = \theta_0(x), \quad \chi_i(0,x) = \chi_{i,0}(x), \quad x \in \Omega,\ i = 1,2. \tag{1.5}$$

The initial boundary value problem (SMA):={(1.1), (1.2), (1.3), (1.4), (1.5)} is based on energy balance and phase dynamics, while the (longitudinal) displacement $u$, which plays the major role in the momentum balance, does not appear explicitly in (SMA). Indeed, the momentum balance equation in the quasi-stationary form reads $\sigma_x = 0$, where the stress $\sigma$ is related to the strain $u_x$, the temperature $\theta$, and the phase proportions by the following constitutive relation:

$$\sigma = u_x + \beta\alpha(\theta)\chi_2.$$

As the boundary conditions $u(t,0) = 0$ (one end of the wire fixed), $\sigma(1,t) = \beta g(t)$ (external traction prescribed) are usually considered, it turns out that in the one-dimensional case $\sigma$ and $u$ can be completely determined in terms of data and other unknowns (see especially the papers [12,21] dealing with the one dimensional problem). That is the reason why in the dynamics of $\chi_i$, $i = 1,2$, written above, one finds the complicate expression $-\beta a_1 g(t)\alpha(\theta) + \beta\alpha(\theta)^2\chi_2$ instead of $-\beta\alpha(\theta)u_x$. In fact, it is time to point out that, in our system, $L_0$, $L_1$, $h$, $k$, $\mu_0$ $\mu_1$, $l$, $\beta$, and $\theta^c$ are positive coefficients with proper physical meaning; in particular, $\theta^c$ represents a critical temperature. In addition, $a_0$, $a_1$, $a_2$, and $a_3$ are fixed real numbers. Furthermore, $f : Q \to \mathbb{R}$ stands for a known source term, while $g$, $\gamma_0$, $\gamma_1$ are given functions defined on the finite time interval $[0,T]$. The nonlinearity $\alpha$ acting on temperature values is a smooth nonnegative decreasing function, vanishing on the interval $[\theta^{Cu}, +\infty)$ for a certain fixed temperature (the so-called Curie point) $\theta^{Cu} > \theta^c$: see, for instance, [13, assumptions (2.12), (2.13)]. Actually, among the properties of $\alpha$, in our analysis we just use the fact that $\alpha \in W^{2,\infty}(\mathbb{R})$. As the Frémond model assumes a nondifferentiable free energy, in (SMA) we meet the maximal monotone graph $\partial I_K$, representing the subdifferential of the indicator function $I_K$ of the plane triangle $K$ (cf. Figure 1 below):

$$K := \left\{(\xi,\eta) \in \mathbb{R}^2\ ;\ 0 \le \xi \le 1,\ |\eta| \le \xi\right\}. \tag{1.6}$$

The set $K$ is convex, and contains the admissible phase proportions. We also notice that

$$I_K(\chi_1,\chi_2) := \begin{cases} 0, & \text{if } (\chi_1,\chi_2) \in K, \\ \infty, & \text{otherwise.} \end{cases} \tag{1.7}$$

An updated and detailed presentation of the Frémond model and related system of equations and conditions, applying to the multidimensional case as well, is provided in [8, 9], [26, Chapter 13], and [27]. We also point out [8, 9] for existence and uniqueness results in the three-dimensional situation: here, the various nonlinear terms arising in the derivation of the model are taken into account. For a list of related references as well as for a survey of previous mathematical work, we address the reader to [7, 13, 21]. The large time behavior of solutions is investigated in [17, 21, 22] in connection with the convergence to steady-state solutions, global attractors, and so on. However, the study of the optimal control problem for the Frémond model has not been reported to date, up to our knowledge. The reason for that is, in our opinion, due to the difficulties created by the presence of the plane triangle set $K$.

In this paper, we deal with the optimal control problem (OP) of (SMA) defined as follows:

**Problem (OP):** Find a quadruplet of control functions $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$ such that

$$J(f^*, g^*, \gamma_0^*, \gamma_1^*) = \inf_{(f,g,\gamma_0,\gamma_1) \in \mathcal{U}_{ad}^M} J(f, g, \gamma_0, \gamma_1),$$

where $(f^*, g^*, \gamma_0^*, \gamma_1^*)$ is called an optimal control for (OP); here, putting

$$\mathcal{U} := L^2(0,T;L^2(\Omega)) \times H^1(0,T) \times H^1(0,T) \times H^1(0,T),$$

$\mathcal{U}_{ad}^M$ is the control space specified by

$$\mathcal{U}_{ad}^M := \left\{ (f,g,\gamma_0,\gamma_1) \in \mathcal{U} \;\middle|\; \begin{array}{c} |g|_{H^1(0,1)} \le M, \\ |\gamma_i|_{H^1(0,1)} \le M,\ i = 0,1 \end{array} \right\} \tag{1.8}$$

for some fixed positive number $M$, and $J(f,g,\gamma_0,\gamma_1)$ is the cost functional defined by

$$\begin{aligned} J(f,g,\gamma_0,\gamma_1) \ &:= \frac{c_0}{2}\int_0^T |(\theta-\theta_d)(t)|^2_{L^2(\Omega)}dt + \frac{c_1}{2}\int_0^T |(\chi_1-\chi_{1,d})(t)|^2_{L^2(\Omega)}dt \\ &\quad + \frac{c_2}{2}\int_0^T |(\chi_2-\chi_{2,d})(t)|^2_{L^2(\Omega)}dt \\ &\quad + \frac{m_0}{2}\int_0^T a_0^2|f(t)|^2_{L^2(\Omega)}dt + \frac{m_1}{2}\int_0^T a_1^2|g(t)|^2dt \\ &\quad + \frac{m_2}{2}\int_0^T a_2^2|\gamma_0(t)|^2dt + \frac{m_3}{2}\int_0^T a_3^2|\gamma_1(t)|^2dt, \end{aligned} \tag{1.9}$$

where $(f,g,\gamma_0,\gamma_1) \in \mathcal{U}_{ad}^M$ denotes the generic control and the triplet of functions $(\theta,\chi_1,\chi_2)$ yields the unique solution to the state problem (SMA) with the source term $(f,g,\gamma_0,\gamma_1)$. We also point out that the given constants $c_0$, $c_1$, $c_2$, $m_0$, $m_1$, $m_2$, $m_3$ are nonnegative, and $\theta_d \in L^2(0,T;L^2(\Omega))$, $\chi_{1,d} \in L^2(0,T;L^2(\Omega))$, $\chi_{2,d} \in L^2(0,T;L^2(\Omega))$ represent the known desired target profiles.

Note that if the constant $a_0$ is equal to $0$, then (OP) is a boundary control problem. Similarly, if $a_1 = a_2 = a_3 = 0$, then (OP) reduces to a distributed control problem with the heat source as control. In addition, we remark that $\gamma_0$ (resp. $\gamma_1$) denotes the outside temperature control function at $x = 0$ (resp. $x = 1$).

There is a vast amount of literature on optimal control problems for variational inequalities, phase transitions problems and so on. In particular, we refer to the contributions [1, 4, 6, 11, 14–16, 23–25, 28–31, 35, 36, 38, 42, 43, 45, 46]. However, to the best of our knowledge, no result is available for the optimal control analyisis of problems like (SMA), probably because of the triangular shape of $K$ and the non-smooth nonlinearity of the two-components constraints $\partial I_K(\chi_1,\chi_2)$ in (1.2).

The novelties of this work are as follows:

**(a)** We show the existence of an optimal control for (OP).

**(b)** We propose an approximation procedure for (SMA) and (OP). Then, we show the existence of approximating solutions to (SMA). In addition, we investigate the approximating control problems of (OP).

**(c)** We show the relationship between the limits ($\omega$-limit points) of sequences of approximating optimal controls and the optimal controls of the limiting problem (OP).

**(d)** We show the necessary conditions for the approximating optimal control problems.

**(e)** We derive a weak formula of the necessary conditions for the original problem (OP), through the limiting observation of approximating situations.

Consequently, an effective approximating approach to the optimal controls of our control problem (OP) will be presented as a further conclusion derived from the main results. Also, it is worth considering the approximating optimal control problems from the view-point of numerical analysis, since the triangle convexity of $K$ and the full nonlinearity of the constraint $\partial I_K(\cdot,\cdot)$ in (SMA) cause us the difficulty to set the numerical experiments for (OP).

The plan of this paper is as follows. In Section 2, the main theorems, denoted by Theorems 2.1–2.5, are stated. In Section 3, we check the well-posedness of the state problem $(\mathrm{SMA})$ and this will help us to prove Theorem 2.1 concerned with the existence of an optimal control for (OP). The following Sections 4–5 are devoted to the proofs of Theorems 2.2, 2.3, and 2.4, corresponding to items (b), (c), and (d), respectively. The final Section 6 contains the proof of Theorem 2.5, which corresponds to item (e).

## 1.1 Notations and basic assumptions

We first state the notations that are used throughout this paper.

For any reflexive Banach space $B$, we denote by $|\cdot|_B$ the norm of $B$, and denote by $B'$ the dual space of $B$. Additionally, we denote by $\langle\cdot,\cdot\rangle_{B',B}$ the duality pairing between $B'$ and $B$. Furthermore, for a positive integer $m \in \mathbb{N}$, we use the product space $B^m$:

$$B^m := \prod_{i=1}^{m} B = \overbrace{B \times B \times \cdots \times B}^{m\text{-factors}}$$

with the norm:

$$|z|_{B^m} := \sum_{i=1}^{m} |z_i|_B \text{ for } z = (z_1, z_2, \cdots, z_m) \in B^m.$$

In particular, we put $H := L^2(\Omega)$ with the usual real Hilbert structure, and denote by $(\cdot,\cdot)_H$ the inner product in $H$, for simplicity.

In addition, let $V$ be the Sobolev space $H^1(\Omega)$ with the inner product and norm:

$$(z, w)_V := (z_x, w_x)_H + \frac{k}{h}\left(z(0)w(0) + z(1)w(1)\right) \text{ for any } z, w \in V,$$

and

$$|z|_V := \left\{ |z_x|_H^2 + \frac{k}{h}\left(|z(0)|^2 + |z(1)|^2\right) \right\}^{1/2} \text{ for any } z \in V,$$

which are equivalent to the standard inner product and norm of $H^1(\Omega)$.

We now list some notation and definitions of subdifferentials of convex functions. For a proper (i.e., not identically equal to infinity), l.s.c. (lower semi-continuous), and convex function $\phi : H \to \mathbb{R} \cup \{\infty\}$, the effective domain $D(\phi)$ of $\phi$ is defined by $D(\phi) := \{z \in H;\ \phi(z) < \infty\}$. We denote by $\partial\phi$ the subdifferential of $\phi$ in the topology of $H$. In general, the subdifferential is a possibly multi-valued operator from $H$ into itself, and for any $z \in H$, the value $\partial\phi(z)$ is defined as:

$$\partial\phi(z) := \{z^* \in H\ ;\ (z^*, y - z)_H \le \phi(y) - \phi(z) \text{ for all } y \in H\}.$$

Then, a set $D(\partial\phi) := \{z \in H\ ;\ \partial\phi(z) \ne \emptyset\}$ is called the domain of $\partial\phi$. For various properties and related notions of a proper, l.s.c., convex function $\phi$ and its subdifferential $\partial\phi$, we refer to the monograph by Brézis [10]. In particular, for those in Banach spaces, we quote the books by Barbu [3,5].

We also recall a notion of convergence for convex functions, developed by Mosco [37].

**Definition 1.1** (cf. [37]). *Let $\phi$, $\phi_n$ ($n \in \mathbb{N}$) be proper, l.s.c., and convex functions on $H$. Then, we say that $\phi_n$ converges to $\phi$ on $H$ in the sense of Mosco [37] as $n \to \infty$ if the following two conditions are satisfied:*

(i) *for any subsequence $\{\phi_{n_k}\}_{k\in\mathbb{N}} \subset \{\phi_n\}_{n\in\mathbb{N}}$, if $z_k \to z$ weakly in $H$ as $k \to \infty$, then*

$$\liminf_{k\to\infty} \phi_{n_k}(z_k) \geq \phi(z);$$

(ii) *for any $z \in D(\phi)$, there is a sequence $\{z_n\}_{n\in\mathbb{N}}$ in $H$ such that*

$$z_n \to z \text{ in } H \text{ as } n \to \infty \quad \textit{and} \quad \lim_{n\to\infty} \phi_n(z_n) = \phi(z).$$

Next, we give some assumptions on data. Throughout this paper, we assume the following conditions (A1)–(A4).

**(A1)** $T > 0$, $L_0 > 0$, $L_1 > 0$, $h > 0$, $k > 0$, $\mu_0 > 0$, $\mu_1 > 0$, $l > 0$, $\beta > 0$, $\theta^c > 0$, $c_0 \geq 0$, $c_1 \geq 0$, $c_2 \geq 0$, $m_0 \geq 0$, $m_1 \geq 0$, $m_2 \geq 0$, $m_3 \geq 0$, and $M > 0$ are fixed constants. In addition, $a_0$, $a_1$, $a_2$, $a_3$ are fixed real numbers.

**(A2)** $\alpha \in W^{2,\infty}(\mathbb{R})$.

**(A3)** $\theta_0 \in V$, and $\chi_{i,0} \in V$ $(i = 1, 2)$ with $(\chi_{1,0}, \chi_{2,0}) \in K$, a.e. in $\Omega$.

**(A4)** $\theta_d \in L^2(0, T; H)$, $\chi_{1,d} \in L^2(0, T; H)$, $\chi_{2,d} \in L^2(0, T; H)$ are the given desired target profiles.

Finally, throughout this paper, $N_i$ and $\nu_i$, $i = 1, 2, 3, \cdots$, denote positive (or nonnegative) constants depending only on their argument(s).

# 2 Main results

We begin by defining the notion of solutions to (SMA). To this end, given $\gamma_i \in H^1(0, T)$, $i = 1, 2$, we define $\gamma$ by putting (cf. [22, (2.4)]):

$$\gamma(t, x) := \frac{k}{2h + k}(a_3\gamma_1(t) - a_2\gamma_0(t))x + \frac{ha_3\gamma_1(t) + (h + k)a_2\gamma_0(t)}{2h + k}, \quad (t, x) \in Q. \tag{2.1}$$

It is easy to check that $\gamma \in W^{1,2}(0, T; H^2(\Omega))$ solves the boundary value problem

$$\begin{cases} \gamma_{xx}(t, x) = 0 \text{ for any } (t, x) \in Q, \\ -h\gamma_x(t, 0) + k(\gamma(t, 0) - a_2\gamma_0(t)) = h\gamma_x(t, 1) + k(\gamma(t, 1) - a_3\gamma_1(t)) = 0, \ t \in (0, T). \end{cases} \tag{2.2}$$

**Definition 2.1.** *Let $\theta_0 \in V$ and $\chi_{i,0} \in V$ ($i = 1, 2$). Then, a triplet of functions $(\theta, \chi_1, \chi_2)$ is called a solution to* (SMA), *or* $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)$ *when the data are specified, on $[0, T]$, if the following conditions are satisfied:*

**(S1)** $\theta \in W^{1,2}(0,T;H) \cap L^\infty(0,T;V)$.

**(S2)** $\chi_i \in W^{1,2}(0,T;H) \cap L^\infty(0,T;V)$, $i = 1,2$.

**(S3)** *For all* $z \in V$ *and a.a.* $t \in (0,T)$,

$$(L_0\theta_t(t) - L_1(\chi_1)_t(t), z)_H + h(\theta(t) - \gamma(t), z)_V = (a_0 f(t), z)_H.$$

**(S4)** *There is a pair of functions* $(\xi_1, \xi_2) \in (L^2(0,T;H))^2$ *such that*

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \in \partial I_K(\chi_1, \chi_2), \quad \text{a.e. in } Q$$

*and*

$$\sum_{i=1}^{2} \{\mu_0((\chi_i)_t(t), z_i)_H + \mu_1((\chi_i)_x(t), (z_i)_x)_H + (\xi_i(t), z_i)_H\}$$
$$= l\,(\theta^c - \theta(t), z_1)_H + \left(-\beta a_1 g(t)\alpha(\theta(t)) + \beta\alpha(\theta(t))^2\chi_2(t), z_2\right)_H$$
$$\text{for any } (z_1, z_2) \in V \times V \text{ and a.a. } t \in (0,T).$$

**(S5)** $\theta(0) = \theta_0$ *in* $H$, *and* $\chi_i(0) = \chi_{i,0}$ *in* $H$, $i = 1,2$.

Here, we recall the known result of the existence-uniqueness and boundedness of solutions to (SMA).

**Proposition 2.1.** *[22, Theorems 2.1 and 2.2] Suppose that assumptions* (A1), (A2), *and* (A3) *hold. Let* $f \in L^2(0,T;H)$, $g \in H^1(0,T)$, $\gamma_0 \in H^1(0,T)$, *and* $\gamma_1 \in H^1(0,T)$. *Then, there is a unique solution* $(\theta, \chi_1, \chi_2)$ *to* (SMA;$\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1$) *on* $[0,T]$ *in the sense of Definition 2.1. In addition, there is a positive constant* $N_1$, *independent of* $f$, $g$, $\gamma_0$, $\gamma_1$, $\theta_0$, $\chi_{1,0}$, *and* $\chi_{2,0}$, *such that the following estimate holds:*

$$\begin{aligned} &\sup_{t\in[0,T]} |\theta(t)|_V^2 + \sum_{i=1}^{2} \sup_{t\in[0,T]} |\chi_i(t)|_V^2 + \int_0^T |\theta_t(t)|_H^2 dt + \sum_{i=1}^{2} \int_0^T |(\chi_i)_t(t)|_H^2 dt \\ \le\ & N_1\Big( |\theta_0|_V^2 + |\chi_{1,0}|_V^2 + |\chi_{2,0}|_V^2 + a_2^2|\gamma_0(0)|^2 + a_3^2|\gamma_1(0)|^2 \\ &\qquad + a_0^2|f|_{L^2(0,T;H)}^2 + a_1^2|g|_{L^2(0,T)}^2 + a_2^2|\gamma_0|_{W^{1,2}(0,T)}^2 + a_3^2|\gamma_1|_{W^{1,2}(0,T)}^2 + 1\Big) \end{aligned} \tag{2.3}$$

In the next Section 3, we give a sketch of the proof of Proposition 2.1. For the other results of (SMA), we refer to [22], for instance.

We now state the first main result of this paper, which is concerned with the existence of an optimal control for (OP).

**Theorem 2.1.** *Suppose that assumptions* (A1), (A2), (A3), *and* (A4) *hold. Then, the problem* (OP) *has at least one optimal control* $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$, *namely,*

$$J(f^*, g^*, \gamma_0^*, \gamma_1^*) = \inf_{(f,g,\gamma_0,\gamma_1)\in\mathcal{U}_{ad}^M} J(f, g, \gamma_0, \gamma_1).$$

**Remark 2.1.** *Note that Theorem 2.1 do not cover the uniqueness of optimal controls for* (OP)*. Although Hoffmann–Jiang [30] reported the uniqueness of optimal controls for a regular Fix–Caginalp system, their technique is not applicable to our problem* (OP)*, because of the constraint* $\partial I_K(\chi_1, \chi_2)$ *in* (1.2)*. Therefore, the uniqueness question of optimal controls for* (OP) *remains open.*

In Section 3, we prove Theorem 2.1 by the quite standard method. In fact, by using the result of convergence of solutions to (SMA), we give the proof of Theorem 2.1.

Note that it is very difficult to show the necessary conditions for (OP) directly, because of the constraint $\partial I_K(\chi_1, \chi_2)$ in (1.2) (cf. Remark 2.1). Therefore, by investigating approximating problems for (SMA) and (OP), we show a limiting optimality system for (OP). To this end, we consider the following smooth function $\widehat{K}^\varepsilon$ on $\mathbb{R}^2$ for each $\varepsilon \in (0, 1]$:

**(A5)** For each $\varepsilon \in (0, 1]$, the function $\widehat{K}^\varepsilon$ is convex and non-negative on $\mathbb{R}^2$ such that $\widehat{K}^\varepsilon \in C^2(\mathbb{R}^2)$, $\partial_i \partial_j \widehat{K}^\varepsilon \in W^{1,\infty}(\mathbb{R}^2)$ $(i, j = 1, 2)$,

$$K^\varepsilon := \left\{ (z_1, z_2) \in \mathbb{R}^2 \ ; \ \widehat{K}^\varepsilon(z_1, z_2) = 0 \right\} \supset K,$$

$$|\partial_i \partial_j \widehat{K}^\varepsilon(z_1, z_2)| \le \frac{1}{\varepsilon} \text{ for any } i, j = 1, 2, \text{ and any } (z_1, z_2) \in \mathbb{R}^2, \tag{2.4}$$

and

$$\widehat{K}^\varepsilon \text{ converges to } I_K \text{ on } \mathbb{R}^2 \text{ in the sense of Mosco [37] as } \varepsilon \to 0,$$

where $\partial_i \widehat{K}^\varepsilon(z_1, z_2)$ is the partial derivative of $\widehat{K}^\varepsilon(z_1, z_2)$ with respect to the variable $z_i$ $(i = 1, 2)$, namely, $\partial_i := \partial / \partial z_i$.

**Remark 2.2.** *A function with properties as in* (A5) *has already been used in [6, 40, 41]. Indeed, for each* $\varepsilon \in (0, 1]$*, a non-decreasing function* $F^\varepsilon$ *is defined by:*

$$F^\varepsilon(r) := \textit{sign}(r) \int_0^{|r|} \min\left\{ \frac{1}{\varepsilon}, \frac{[s-1]^+}{\varepsilon^2} \right\} ds \quad \textit{for } r \in \mathbb{R},$$

*where* $[\,\cdot\,]^+$ *denotes the positive part of functions, and sign*$(\cdot)$ *is a signum function so that sign*$(0) = 0$*. In addition, let* $\widehat{F}^\varepsilon$ *be a primitive of* $F^\varepsilon$ *such that*

$$\widehat{F}^\varepsilon(0) = 0 \quad \textit{and} \quad \widehat{F}^\varepsilon(r) \ge 0 \quad \textit{for all } r \in \mathbb{R}.$$

*Then, we observe that* $F^\varepsilon$ *is a* $C^1$*-function with derivative* $(F^\varepsilon)' \in W^{1,\infty}(\mathbb{R})$*,*

$$0 \le (F^\varepsilon)'(r) \le \frac{1}{\varepsilon} \ \textit{for any } r \in \mathbb{R},$$

*and* $\widehat{F}^\varepsilon$ *converges to* $I_{[-1,1]}$ *on* $\mathbb{R}$ *in the sense of Mosco [37] as* $\varepsilon \to 0$*, where* $I_{[-1,1]}$ *is the indicator function of the closed interval* $[-1, 1]$*, that is defined by*

$$I_{[-1,1]}(\tau) := \begin{cases} 0, & \textit{if } \tau \in [-1, 1], \\ \infty, & \textit{otherwise.} \end{cases}$$

Note that the function $\widehat{K}^\varepsilon$ in assumption (A5) can be easily defined for each $\varepsilon \in (0,1]$. We here give a typical example of $\widehat{K}^\varepsilon$.

**Example 2.1.** *For each $\varepsilon \in (0,1]$, let $K^\varepsilon$ be a smooth closed convex set in $\mathbb{R}^2$ such that $K^\varepsilon$ includes the convex set $K$, the boundary of $K^\varepsilon$ is described by the combination of linear functions and cubic functions (cf. $|z_1|^3 + |z_2|^3 = constant$), and $K^\varepsilon$ converges to $K$ in the sense of Hausdorff distance as $\varepsilon \to 0$. More precisely, the pictures of $K$ and its approximating set $K^\varepsilon$ are illustrated in Figure 1. Then, by arguing similarly as in Remark 2.2, we can define the smooth convex function $\widehat{K}^\varepsilon$ satisfying assumption $(\mathrm{A5})$. The typical graph of $\widehat{K}^\varepsilon$ is illustrated in Figure 2, which is described by the combination of a cubic surface, a smooth surface, and so on (cf. $(|z_1|^3+|z_2|^3)/\varepsilon^2$, $\sqrt{|z_1|^3+|z_2|^3}/\sqrt{\varepsilon}$, and so on). For such constructions, we refer to the Appendix.*

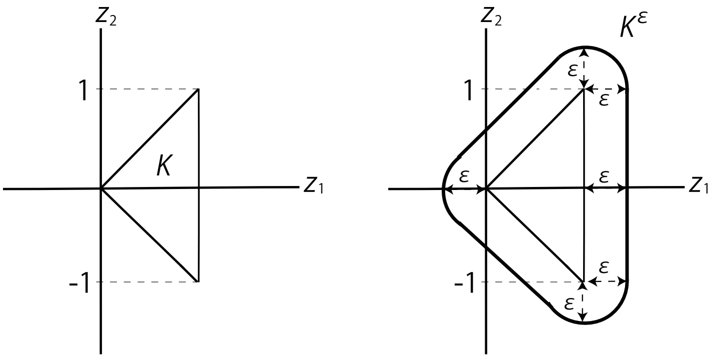


Figure 1: Convex set $K$ and its approximating set

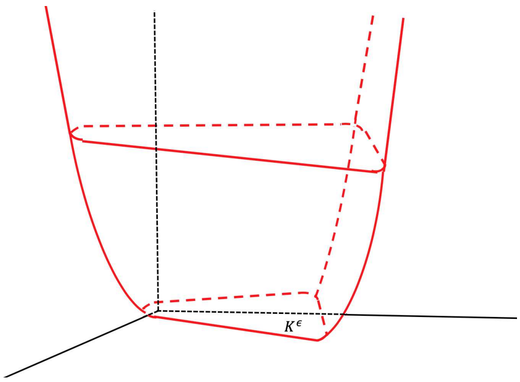


Figure 2: The typical graph of $\widehat{K}^\varepsilon$

Now, for each $\varepsilon \in (0,1]$, we present the following approximating state system for (SMA), denoted by (SMA)$^\varepsilon$:

**Problem (SMA)$^\varepsilon$.**

$$(L_0\theta^\varepsilon - L_1\chi_1^\varepsilon)_t - h\theta^\varepsilon_{xx} = a_0 f(t,x) \quad \text{in } Q = (0,T)\times\Omega, \tag{2.5}$$

$$\mu_0 \begin{pmatrix} \chi_1^\varepsilon \\ \chi_2^\varepsilon \end{pmatrix}_t - \mu_1 \begin{pmatrix} \chi_1^\varepsilon \\ \chi_2^\varepsilon \end{pmatrix}_{xx} + \nabla\widehat{K}^\varepsilon(\chi_1^\varepsilon,\chi_2^\varepsilon) = \begin{pmatrix} l(\theta^c-\theta^\varepsilon) \\ -\beta a_1 g(t)\alpha(\theta^\varepsilon) + \beta\alpha(\theta^\varepsilon)^2\chi_2^\varepsilon \end{pmatrix} \quad \text{in } Q, \tag{2.6}$$

$$-h\theta^\varepsilon_x(t,0)+k(\theta^\varepsilon(t,0)-a_2\gamma_0(t))=h\theta^\varepsilon_x(t,1)+k(\theta^\varepsilon(t,1)-a_3\gamma_1(t))=0,\ \ t\in(0,T), \tag{2.7}$$

$$(\chi^\varepsilon_i)_x(t,0)=(\chi^\varepsilon_i)_x(t,1)=0,\quad t\in(0,T),\ i=1,2, \tag{2.8}$$

$$\theta^\varepsilon(0,x)=\theta_0(x),\quad \chi^\varepsilon_i(0,x)=\chi_{i,0}(x),\quad x\in\Omega,\ i=1,2. \tag{2.9}$$

In the rest, we denote (SMA)$^\varepsilon$ by (SMA;$\theta_0,\chi_{1,0},\chi_{2,0},f,g,\gamma_0,\gamma_1)^\varepsilon$ when the data of the initial value $\theta_0,\chi_{1,0},\chi_{2,0}$ and the control functions $f,g,\gamma_0,\gamma_1$ are specified. Note that for each $\varepsilon\in(0,1]$, the constraint $\partial I_K(\chi_1,\chi_2)$ as in (1.2) is approximated by $\nabla\widehat{K}^\varepsilon(\chi^\varepsilon_1,\chi^\varepsilon_2)(=\partial\widehat{K}^\varepsilon(\chi^\varepsilon_1,\chi^\varepsilon_2))$. In a similar way to Proposition 2.1, we immediately get the following proposition, concerned with the solvability of the approximating state problem (SMA)$^\varepsilon$ ($\varepsilon\in(0,1]$).

**Proposition 2.2.** *Suppose that assumptions* (A1), (A2), (A3), *and* (A5) *hold. Let* $f\in L^2(0,T;H)$, $g\in H^1(0,T)$, $\gamma_0\in H^1(0,T)$, *and* $\gamma_1\in H^1(0,T)$. *Then, for each* $\varepsilon\in(0,1]$, *there is a unique solution* $(\theta^\varepsilon,\chi^\varepsilon_1,\chi^\varepsilon_2)$ *to* (SMA;$\theta_0,\chi_{1,0},\chi_{2,0},f,g,\gamma_0,\gamma_1)^\varepsilon$ *on* $[0,T]$, *which solves the equations* (2.5)–(2.9) *in the following sense:*

**(i)** $\theta^\varepsilon\in W^{1,2}(0,T;H)\cap L^\infty(0,T;V)$.

**(ii)** $\chi^\varepsilon_i\in W^{1,2}(0,T;H)\cap L^\infty(0,T;V)$, $i=1,2$.

**(iii)** *For all* $z\in V$ *and a.a.* $t\in(0,T)$,

$$(L_0\theta^\varepsilon_t(t)-L_1(\chi^\varepsilon_1)_t(t),z)_H+h(\theta^\varepsilon(t)-\gamma(t),z)_V=(a_0f(t),z)_H,$$

*where* $\gamma$ *is the function defined in* (2.1).

**(iv)** *For all* $(z_1,z_2)\in V\times V$ *and a.a.* $t\in(0,T)$,

$$\begin{aligned}&\sum_{i=1}^{2}\left\{\mu_0((\chi^\varepsilon_i)_t(t),z_i)_H+\mu_1((\chi^\varepsilon_i)_x(t),(z^\varepsilon_i)_x)_H+(\partial_i\widehat{K}^\varepsilon(\chi^\varepsilon_1(t),\chi^\varepsilon_2(t)),z_i)_H\right\}\\&\quad=l\left(\theta^c-\theta^\varepsilon(t),z_1\right)_H+\left(-\beta a_1g(t)\alpha(\theta^\varepsilon(t))+\beta\alpha(\theta^\varepsilon(t))^2\chi^\varepsilon_2(t),z_2\right)_H.\end{aligned}$$

**(v)** $\theta^\varepsilon(0)=\theta_0$ *in* $H$, *and* $\chi^\varepsilon_i(0)=\chi_{i,0}$ *in* $H$, $i=1,2$.

*In addition, there is a positive constant* $N_2$, *independent of* $\varepsilon$, $f$, $g$, $\gamma_0$, $\gamma_1$, $\theta_0$, $\chi_{1,0}$, *and* $\chi_{2,0}$, *such that the following estimate holds:*

$$\begin{aligned}&\sup_{t\in[0,T]}|\theta^\varepsilon(t)|^2_V+\sum_{i=1}^{2}\sup_{t\in[0,T]}|\chi^\varepsilon_i(t)|^2_V+\int_0^T|\theta^\varepsilon_t(t)|^2_Hdt+\sum_{i=1}^{2}\int_0^T|(\chi^\varepsilon_i)_t(t)|^2_Hdt\\&\le\ N_2\Big(|\theta_0|^2_V+|\chi_{1,0}|^2_V+|\chi_{2,0}|^2_V+a_2^2|\gamma_0(0)|^2+a_3^2|\gamma_1(0)|^2\\&\qquad+a_0^2|f|^2_{L^2(0,T;H)}+a_1^2|g|^2_{L^2(0,T)}+a_2^2|\gamma_0|^2_{W^{1,2}(0,T)}+a_3^2|\gamma_1|^2_{W^{1,2}(0,T)}+1\Big).\end{aligned} \tag{2.10}$$

In Section 4, we give the sketch of the proof of Proposition 2.2.

Now, for each $\varepsilon \in (0,1]$ and $\delta \geq 0$, we present an approximating optimal control problem for (OP), denoted by (OP)$_\delta^\varepsilon$, as follows:

**Problem (OP)$_\delta^\varepsilon$.** Find an optimal control $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \in \mathcal{U}_{ad}^M$, namely,

$$J_\delta^\varepsilon(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) = \inf_{(f,g,\gamma_0,\gamma_1)\in\mathcal{U}_{ad}^M} J_\delta^\varepsilon(f,g,\gamma_0,\gamma_1),$$

where $J_\delta^\varepsilon(f,g,\gamma_0,\gamma_1)$ is the cost functional defined by

$$\begin{aligned} J_\delta^\varepsilon(f,g,\gamma_0,\gamma_1) \ := & \frac{c_0}{2}\int_0^T |(\theta^\varepsilon - \theta_d)(t)|_H^2 dt + \frac{c_1}{2}\int_0^T |(\chi_1^\varepsilon - \chi_{1,d})(t)|_H^2 dt \\ & + \frac{c_2}{2}\int_0^T |(\chi_2^\varepsilon - \chi_{2,d})(t)|_H^2 dt \\ & + \frac{m_0}{2}\int_0^T a_0^2 |f(t)|_H^2 dt + \frac{m_1}{2}\int_0^T a_1^2 |g(t)|^2 dt \\ & + \frac{m_2}{2}\int_0^T a_2^2 |\gamma_0(t)|^2 dt + \frac{m_3}{2}\int_0^T a_3^2 |\gamma_1(t)|^2 dt \\ & + \frac{\delta}{2}\int_0^T |(f-f^*)(t)|_H^2 dt + \frac{\delta}{2}\int_0^T |(g-g^*)(t)|^2 dt \\ & + \frac{\delta}{2}\int_0^T |(\gamma_0-\gamma_0^*)(t)|^2 dt + \frac{\delta}{2}\int_0^T |(\gamma_1-\gamma_1^*)(t)|^2 dt. \end{aligned} \tag{2.11}$$

Here, $(f,g,\gamma_0,\gamma_1) \in \mathcal{U}_{ad}^M$ is the control and the triplet of functions $(\theta^\varepsilon, \chi_1^\varepsilon, \chi_2^\varepsilon)$ is the unique solution to the state problem $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)^\varepsilon$. Moreover, as in (1.9), $\theta_d \in L^2(0,T;H)$, $\chi_{1,d} \in L^2(0,T;H)$, $\chi_{2,d} \in L^2(0,T;H)$ are the given desired target profiles, while $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$ is any fixed optimal control for (OP) obtained in Theorem 2.1.

We now state the second main result of this paper, which is concerned with the existence of an optimal control for (OP)$_\delta^\varepsilon$ for each $\varepsilon \in (0,1]$ and $\delta \geq 0$.

**Theorem 2.2.** *Suppose that assumptions* (A1)–(A5) *hold. Let* $\varepsilon \in (0,1]$, $\delta \geq 0$, *and let* $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$ *be a chosen optimal control for* (OP) *given by Theorem 2.1. Then, the approximating problem* $(\mathrm{OP})_\delta^\varepsilon$ *has at least one optimal control* $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \in \mathcal{U}_{ad}^M$*, namely,*

$$J_\delta^\varepsilon(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) = \inf_{(f,g,\gamma_0,\gamma_1)\in\mathcal{U}_{ad}^M} J_\delta^\varepsilon(f,g,\gamma_0,\gamma_1).$$

The following third main result of the paper is concerned with the relationship between (OP) and (OP)$_\delta^\varepsilon$.

**Theorem 2.3.** *Suppose that all the assumptions of Theorem 2.2 hold. Then, the following two statements hold.*

**(I)** *Let* $\delta = 0$, $\varepsilon \in (0,1]$, *and let* $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \in \mathcal{U}_{ad}^M$ *be an optimal control for the approximating problem* $(\mathrm{OP})_0^\varepsilon$*. In addition, let* $(\theta^{*,\varepsilon}, \chi_1^{*,\varepsilon}, \chi_2^{*,\varepsilon})$ *be the unique solution to the state problem* $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})^\varepsilon$ *on* $[0,T]$*. Then, there exist a subsequence* $\{\varepsilon_n\}_{n\in\mathbb{N}} \subset \{\varepsilon\}_{\varepsilon\in(0,1]}$*, a quadruplet of functions* $(f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**}) \in \mathcal{U}_{ad}^M$*, and the unique solution* $(\theta^{**}, \chi_1^{**}, \chi_2^{**})$ *to* $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**})$ *on* $[0,T]$ *such that* $(f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**})$ *is an optimal control for* (OP), $\varepsilon_n \to 0$*, and*

$$f^{*,\varepsilon_n} \to f^{**} \quad \text{weakly in } L^2(0,T;H), \tag{2.12}$$

$$g^{*,\varepsilon_n} \to g^{**} \text{ weakly in } H^1(0,T), \text{ and in } L^2(0,T), \tag{2.13}$$

$$\gamma_i^{*,\varepsilon_n} \to \gamma_i^{**} \text{ weakly in } H^1(0,T), \text{ and in } L^2(0,T), \ (i=0,1), \tag{2.14}$$

$$(\theta^{*,\varepsilon_n}, \chi_1^{*,\varepsilon_n}, \chi_2^{*,\varepsilon_n}) \to (\theta^{**}, \chi_1^{**}, \chi_2^{**}) \text{ in } (C([0,T];H))^3 \tag{2.15}$$

*as $n \to \infty$.*

**(II)** *Let $\delta > 0$, $\varepsilon \in (0,1]$, and let $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \in \mathcal{U}_{ad}^M$ be an optimal control for the approximating problem $(\mathrm{OP})_\delta^\varepsilon$. Let $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$ be an optimal control for* (OP) *obtained in Theorem 2.1. In addition, let $(\theta^{*,\varepsilon}, \chi_1^{*,\varepsilon}, \chi_2^{*,\varepsilon})$ and $(\theta^*, \chi_1^*, \chi_2^*)$ be the unique solution to the state problem $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})^\varepsilon$ and $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^*, g^*, \gamma_0^*, \gamma_1^*)$ on $[0,T]$, respectively. Then, there exist a subsequence $\{\varepsilon_n\}_{n\in\mathbb{N}} \subset \{\varepsilon\}_{\varepsilon\in(0,1]}$ such that $\varepsilon_n \to 0$,*

$$f^{*,\varepsilon_n} \to f^* \text{ in } L^2(0,T;H), \tag{2.16}$$

$$g^{*,\varepsilon_n} \to g^* \text{ weakly in } H^1(0,T), \text{ and in } L^2(0,T), \tag{2.17}$$

$$\gamma_i^{*,\varepsilon_n} \to \gamma_i^* \text{ weakly in } H^1(0,T), \text{ and in } L^2(0,T), \ (i=0,1), \tag{2.18}$$

$$(\theta^{*,\varepsilon_n}, \chi_1^{*,\varepsilon_n}, \chi_2^{*,\varepsilon_n}) \to (\theta^*, \chi_1^*, \chi_2^*) \text{ in } (C([0,T];H))^3 \tag{2.19}$$

*as $n \to \infty$.*

The proofs of Theorems 2.2 and 2.3 are given in Section 5. To show Theorem 2.3, we use the fact that the unique solution $(\theta^\varepsilon, \chi_1^\varepsilon, \chi_2^\varepsilon)$ to $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^\varepsilon, g^\varepsilon, \gamma_0^\varepsilon, \gamma_1^\varepsilon)^\varepsilon$ converges to the solution $(\theta, \chi_1, \chi_2)$ to $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)$ in $(C([0,T];H))^3$ as $\varepsilon \to 0$, if the data $(f^\varepsilon, g^\varepsilon, \gamma_0^\varepsilon, \gamma_1^\varepsilon)$ converges to $(f, g, \gamma_0, \gamma_1)$ as $\varepsilon \to 0$ in some appropriate sense.

The fourth main result is concerned with the necessary condition of an optimal sevenfold $(\theta^{*,\varepsilon}, \chi_1^{*,\varepsilon}, \chi_2^{*,\varepsilon}, f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})$ for $(\mathrm{OP})_\delta^\varepsilon$, where $(\theta^{*,\varepsilon}, \chi_1^{*,\varepsilon}, \chi_2^{*,\varepsilon})$ is the unique solution to the state problem $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})^\varepsilon$ on $[0,T]$, while the quadruplet $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \in \mathcal{U}_{ad}^M$ denotes the optimal control for $(\mathrm{OP})_\delta^\varepsilon$ obtained in Theorem 2.2.

**Theorem 2.4.** *Suppose that all the assumptions of Theorem 2.2 hold. Let the quadruplet $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$ be any optimal control for* (OP) *obtained in Theorem 2.1. In addition, for the fixed number $\varepsilon \in (0,1]$ and $\delta \geq 0$, let $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \in \mathcal{U}_{ad}^M$ be an optimal control for the approximating problem $(\mathrm{OP})_\delta^\varepsilon$ obtained in Theorem 2.2, with $(\theta^{*,\varepsilon}, \chi_1^{*,\varepsilon}, \chi_2^{*,\varepsilon})$ being the unique solution to the state problem $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})^\varepsilon$ on $[0,T]$. Then, there exists a unique solution $(p^\varepsilon, q_1^\varepsilon, q_2^\varepsilon)$ to the adjoint equations on $[0,T]$ as follows:*

$$(p^\varepsilon, q_1^\varepsilon, q_2^\varepsilon) \in (W^{1,2}(0,T;H) \cap L^\infty(0,T;V))^3; \tag{2.20}$$

$$\begin{aligned} -L_0 p_t^\varepsilon - h p_{xx}^\varepsilon + l q_1^\varepsilon + \beta a_1 g^{*,\varepsilon}(t)\alpha'(\theta^{*,\varepsilon}) q_2^\varepsilon - 2\beta\alpha'(\theta^{*,\varepsilon})\alpha(\theta^{*,\varepsilon})\chi_2^{*,\varepsilon} q_2^\varepsilon \\ = c_0(\theta^{*,\varepsilon} - \theta_d) \text{ in } Q; \end{aligned} \tag{2.21}$$

$$-h p_x^\varepsilon(t,0) + k p^\varepsilon(t,0) = h p_x^\varepsilon(t,1) + k p^\varepsilon(t,1) = 0, \ \ t \in (0,T), \tag{2.22}$$

$$\begin{aligned} -\mu_0(q_1^\varepsilon)_t + L_1 p_t^\varepsilon - \mu_1(q_1^\varepsilon)_{xx} + \partial_1^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon})q_1^\varepsilon + \partial_1\partial_2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon})q_2^\varepsilon \\ = c_1(\chi_1^{*,\varepsilon}-\chi_{1,d}) \text{ in } Q; \end{aligned} \tag{2.23}$$

$$\begin{aligned} -\mu_0(q_2^\varepsilon)_t - \mu_1(q_2^\varepsilon)_{xx} + \partial_2\partial_1\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon})q_1^\varepsilon + \partial_2^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon})q_2^\varepsilon - \beta\alpha(\theta^{*,\varepsilon})^2 q_2^\varepsilon \\ = c_2(\chi_2^{*,\varepsilon}-\chi_{2,d}) \text{ in } Q; \end{aligned} \tag{2.24}$$

$$(q_1^\varepsilon)_x(t,0) = (q_1^\varepsilon)_x(t,1) = (q_2^\varepsilon)_x(t,0) = (q_1^\varepsilon)_x(t,1) = 0, \ \ t\in(0,T), \tag{2.25}$$

$$p^\varepsilon(T,x) = q_1^\varepsilon(T,x) = q_2^\varepsilon(T,x) = 0, \quad x\in\Omega. \tag{2.26}$$

*In addition, $(p^\varepsilon, q_1^\varepsilon, q_2^\varepsilon)$ satisfies the following inequality:*

$$\begin{aligned} &\int_0^T a_0((a_0 m_0 f^{*,\varepsilon} + p^\varepsilon)(t), (\check{f} - f^{*,\varepsilon})(t))_H dt \\ &\quad + \int_0^T a_1(a_1 m_1 g^{*,\varepsilon}(t) - (\beta\alpha(\theta^{*,\varepsilon}(t)), q_2^\varepsilon(t))_H)(\check{g} - g^{*,\varepsilon})(t)dt \\ &\quad + \int_0^T a_2(a_2 m_2 \gamma_0^{*,\varepsilon}(t) + kp^\varepsilon(t,0))(\check{\gamma}_0 - \gamma_0^{*,\varepsilon})(t)dt \\ &\quad + \int_0^T a_3(a_3 m_3 \gamma_1^{*,\varepsilon}(t) + kp^\varepsilon(t,1))(\check{\gamma}_1 - \gamma_1^{*,\varepsilon})(t)dt \\ &\quad + \delta\int_0^T ((f^{*,\varepsilon} - f^*)(t), (\check{f} - f^{*,\varepsilon})(t))_H dt \\ &\quad + \delta\int_0^T (g^{*,\varepsilon} - g^*)(t)(\check{g} - g^{*,\varepsilon})(t)dt \\ &\quad + \delta\int_0^T (\gamma_0^{*,\varepsilon} - \gamma_0^*)(t)(\check{\gamma}_0 - \gamma_0^{*,\varepsilon})(t)dt \\ &\quad + \delta\int_0^T (\gamma_1^{*,\varepsilon} - \gamma_1^*)(t)(\check{\gamma}_1 - \gamma_1^{*,\varepsilon})(t)dt \\ &\geq 0, \quad \forall(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1) \in \mathcal{U}_{ad}^M. \end{aligned} \tag{2.27}$$

In Section 5, we prove Theorem 2.4 by showing the result of Gâteaux differentiability of the cost functional $J_\delta^\varepsilon(\cdot,\cdot,\cdot,\cdot)$.

In Theorem 2.4, we get the optimality condition for (OP)$_\delta^\varepsilon$. However, in general, it is difficult to show the necessary condition of the optimal control for (OP), since the subdifferential $\partial I_K(\cdot,\cdot)$ in (1.2) is not smooth. Thus, by using the approximating problems (OP)$_\delta^\varepsilon$, we give the optimality condition for (OP).

We now state the final main result of this paper, which is concerned with the necessary condition of the optimal control for (OP).

**Theorem 2.5.** *Suppose that all the assumptions of Theorem 2.2 hold. Let the quadruplet $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$ be any optimal control for* (OP) *obtained in Theorem 2.1. Let $(\theta^*, \chi_1^*, \chi_2^*)$ be the unique solution to the state problem* (SMA;$\theta_0, \chi_{1,0}, \chi_{2,0}, f^*, g^*, \gamma_0^*, \gamma_1^*$) *on $[0,T]$. In addition, let us set:*

$$W := \{\zeta \in H^1(Q)\ ;\ \zeta(0,x) = 0, \text{ a.a. } x\in\Omega\}.$$

*Then, there are the functions*

$$p \in W^{1,2}(0,T;H)\cap L^\infty(0,T;V), \quad q_i \in L^2(0,T;V)\cap L^\infty(0,T;H) \ (i=1,2),$$

*and the elements* $\varsigma_i \in W'$ $(i = 1, 2)$ *satisfying the following:*

$$\begin{aligned} -L_0 p_t - h p_{xx} + l q_1 + \beta a_1 g^*(t) \alpha'(\theta^*) q_2 - 2\beta \alpha'(\theta^*) \alpha(\theta^*) \chi_2^*(t) q_2 \\ = c_0(\theta^* - \theta_d) \text{ in } Q, \end{aligned} \tag{2.28}$$

$$-h p_x(t, 0) + k p(t, 0) = h p_x(t, 1) + k p(t, 1) = 0, \quad t \in (0, T), \tag{2.29}$$

$$\begin{aligned} \int_0^T (\mu_0 q_1(t), \zeta_t(t))_H dt + \int_0^T (L_1 p_t(t), \zeta(t))_H dt + \int_0^T (\mu_1 (q_1)_x(t), \zeta_x(t))_H dt \\ + \langle \varsigma_1, \zeta \rangle_{W', W} = c_1 \int_0^T (\chi_1^*(t) - \chi_{1,d}(t), \zeta(t))_H dt \text{ for all } \zeta \in W, \end{aligned} \tag{2.30}$$

$$\begin{aligned} \int_0^T (\mu_0 q_2(t), \zeta_t(t))_H dt + \int_0^T (\mu_1 (q_2)_x(t), \zeta_x(t))_H dt + \langle \varsigma_2, \zeta \rangle_{W', W} \\ - \int_0^T (\beta \alpha(\theta^*(t))^2 q_2(t), \zeta(t))_H dt \\ = c_2 \int_0^T (\chi_2^*(t) - \chi_{2,d}(t), \zeta(t))_H dt \text{ for all } \zeta \in W, \end{aligned} \tag{2.31}$$

$$p(T, x) = 0, \quad x \in \Omega. \tag{2.32}$$

*In addition,* $(p, q_1, q_2)$ *satisfies the following inequality:*

$$\begin{aligned} &\int_0^T a_0((a_0 m_0 f^* + p)(t), (\check{f} - f^*)(t))_H dt \\ &\quad + \int_0^T a_1(a_1 m_1 g^*(t) - (\beta \alpha(\theta^*(t)), q_2(t))_H)(\check{g} - g^*)(t) dt \\ &\quad + \int_0^T a_2(a_2 m_2 \gamma_0^*(t) + k p(t, 0))(\check{\gamma}_0 - \gamma_0^*)(t) dt \\ &\quad + \int_0^T a_3(a_3 m_3 \gamma_1^*(t) + k p(t, 1))(\check{\gamma}_1 - \gamma_1^*)(t) dt \\ &\geq 0, \quad \forall (\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}_{ad}^M. \end{aligned} \tag{2.33}$$

In Section 6, we prove Theorem 2.5 by letting $\varepsilon \to 0$ in (2.21)–(2.27).

**Remark 2.3.** *The identities* (2.30) *and* (2.31) *can be regarded as some variational forms of*

$$-\mu_0 (q_1)_t + L_1 p_t - \mu_2 (q_1)_{xx} + \varsigma_1 = c_1(\chi_1^* - \chi_{1,d}),$$

*and*

$$-\mu_0 (q_2)_t - \mu_1 (q_2)_{xx} + \varsigma_2 - \beta \alpha(\theta^*)^2 q_2 = c_2(\chi_2^* - \chi_{2,d})$$

*in the distribution sense, respectively.*

# 3 Optimal control for (OP)

In this section, we prove Theorem 2.1, which is concerned with the existence of an optimal control for (OP). Throughout this section, we suppose that all the assumptions of Theorem 2.1 are made.

We begin with the sketch of the proof of Proposition 2.1.

*Proof of Proposition 2.1.* This proposition has already been proved in [22, Theorems 2.1 and 2.2]. However, we give the sketch of the proof of this Proposition 2.1 to make use of a similar idea in the approximating state problem (SMA)$^\varepsilon$ ($\varepsilon \in (0,1]$).

Note that $g \in C([0,T])$, because $H^1(0,T)$ is compactly embedded in $C([0,T])$. Then, by the standard monotone arguments as in [22, Theorem 2.1], we can show the uniqueness of solutions to $(\mathrm{SMA})$. Therefore, we here omit the detailed proof of the uniqueness of solutions to $(\mathrm{SMA})$.

We next give the sketch of the proof of existence of solutions to $(\mathrm{SMA})$. To this end, note that $(\mathrm{SMA})$ can be reformulated to the following system of abstract evolution equations:

$$\theta_t(t) - \frac{L_1}{L_0}(\chi_1)_t(t) + \partial\varphi^t(\theta(t)) \ni \frac{a_0}{L_0} f(t) \text{ in } H \text{ for } t \in (0,T), \tag{3.1}$$

$$\frac{d}{dt}(\chi_1(t),\chi_2(t)) + \partial\psi(\chi_1(t),\chi_2(t)) + G^t_{\theta(t)}(\chi_1(t),\chi_2(t)) \ni (0,0) \text{ in } H\times H \quad \text{for } t \in (0,T), \tag{3.2}$$

$$\theta(0) = \theta_0 \text{ in } H, \text{ and } (\chi_1(0),\chi_2(0)) = (\chi_{1,0},\chi_{2,0}) \text{ in } H\times H. \tag{3.3}$$

Here, $\partial\varphi^t(\cdot)$ is the subdifferential of a time-dependent convex function $\varphi^t(\cdot)$ on $H$ for each $t \in [0,T]$, defined by

$$\varphi^t(z) := \begin{cases} \dfrac{h}{2L_0}|z-\gamma(t)|_V^2, & \text{if } z \in V, \\ \infty, & \text{otherwise,} \end{cases} \tag{3.4}$$

where $\gamma$ is the function defined in (2.1). In addition, $\partial\psi(\cdot,\cdot)$ is the subdifferential of a convex function $\psi$ on $H\times H$, defined by:

$$\psi(z_1,z_2) := \begin{cases} \dfrac{\mu_1}{2\mu_0}\displaystyle\sum_{i=1}^2 \int_\Omega |(z_i)_x(x)|^2 dx \\ \qquad + \dfrac{1}{\mu_0}\displaystyle\int_\Omega I_K(z_1(x),z_2(x))dx, & \text{if } z_i \in V,\ i=1,2, \\ \infty, \quad \text{otherwise.} \end{cases} \tag{3.5}$$

Furthermore, $G^t_{\theta(t)}(\cdot,\cdot)$ is a time-dependent operator on $H\times H$ for each $\theta \in C([0,T];H)$ and $t \in [0,T]$, defined by

$$G^t_{\theta(t)}(z_1,z_2) := \frac{1}{\mu_0}\left(-l(\theta^c-\theta(t)), \beta a_1 g(t)\alpha(\theta(t)) - \beta\alpha(\theta(t))^2 z_2\right) \quad \text{for any } (z_1,z_2) \in H\times H. \tag{3.6}$$

Then, we can show the existence of a solution to (3.1)–(3.3) by employing the fixed point argument for continuous operators in compact convex sets (e.g., Schauder's fixed point theorem). Indeed, by modifying the proof of [22, Theorem 2.1] (cf. [34, Theorem 2.1]), we can construct a solution to (3.1)–(3.3). Hence, we omit the detailed proof.

In addition, from the standard calculations (cf. [22, Theorem 2.2]), we obtain (2.3). Indeed, multiplying (1.1) by $(1/L_1)(\theta(t)-\gamma(t))$ and $\theta_t(t)-\gamma_t(t)$, multiplying (1.2) by $(\mu_1/l\mu_0)(\chi_1(t),\chi_2(t))$ and $(1/l)((\chi_1)_t(t),(\chi_2)_t(t))$, using Young's inequality, and integrating in time, we get the a priori estimate (2.3). For such arguments, we refer to [22, Theorem 2.2].

Thus, the proof of Proposition 2.1 is complete. □

We now state the result of continuous dependence of solutions to $(\mathrm{SMA})$.

**Proposition 3.1.** *Suppose that all the assumptions of Proposition 2.1 hold. In addition, assume* $\{f_n\}_{n\in\mathbb{N}} \subset L^2(0,T;H)$, $\{(g_n,\gamma_{0,n},\gamma_{1,n})\}_{n\in\mathbb{N}} \subset (H^1(0,T))^3$, $f \in L^2(0,T;H)$, $(g,\gamma_0,\gamma_1) \in (H^1(0,T))^3$*, and*

$$f_n \to f \text{ weakly in } L^2(0,T;H), \tag{3.7}$$

$$g_n \to g \text{ weakly in } H^1(0,T), \tag{3.8}$$

$$(\gamma_{0,n},\gamma_{1,n}) \to (\gamma_0,\gamma_1) \text{ in } (C[0,T])^2 \tag{3.9}$$

*as* $n \to \infty$*. Let* $(\theta_n,\chi_{1,n},\chi_{2,n})$ *and* $(\theta,\chi_1,\chi_2)$ *denote the unique solutions to the state problems* $(\mathrm{SMA};\theta_0,\chi_{1,0},\chi_{2,0},f_n,g_n,\gamma_{0,n},\gamma_{1,n})$ *and* $(\mathrm{SMA};\theta_0,\chi_{1,0},\chi_{2,0},f,g,\gamma_0,\gamma_1)$*, respectively, on* $[0,T]$*. Then,* $(\theta_n,\chi_{1,n},\chi_{2,n})$ *converges to* $(\theta,\chi_1,\chi_2)$ *in the sense that*

$$(\theta_n,\chi_{1,n},\chi_{2,n}) \to (\theta,\chi_1,\chi_2) \text{ in } (C([0,T];H))^3 \quad \text{as } n\to\infty. \tag{3.10}$$

*Proof.* By (2.3), there are a subsequence $\{n_k\}_{k\in\mathbb{N}} \subset \{n\}_{n\in\mathbb{N}}$, the triplet of functions $(\theta,\chi_1,\chi_2) \in (W^{1,2}(0,T;H)\cap L^\infty(0,T;V))^3$ such that $n_k\to\infty$,

$$\left.\begin{array}{ll}(\theta_{n_k},\chi_{1,n_k},\chi_{2,n_k}) \to (\theta,\chi_1,\chi_2) & \text{in } (C([0,T];H))^3, \\ & \text{weakly in } (W^{1,2}(0,T;H))^3, \\ & \text{weakly-}* \text{ in } (L^\infty(0,T;V))^3\end{array}\right\} \tag{3.11}$$

as $k\to\infty$.

We now show that $(\theta,\chi_1,\chi_2)$ is a solution to $(\mathrm{SMA};\theta_0,\chi_{1,0},\chi_{2,0},f,g,\gamma_0,\gamma_1)$ on $[0,T]$. To this end, note that the solution $(\theta_{n_k},\chi_{1,n_k},\chi_{2,n_k})$ to $(\mathrm{SMA})$ satisfies the following abstract evolution equations (cf. Proposition 2.1):

$$(\theta_{n_k})_t(t) + \partial\varphi^t_{n_k}(\theta_{n_k}(t)) \ni \frac{a_0}{L_0} f_{n_k}(t) + \frac{L_1}{L_0}(\chi_{1,n_k})_t(t) \text{ in } H \text{ for } t\in(0,T), \tag{3.12}$$

$$\begin{aligned}\frac{d}{dt}(\chi_{1,n_k}(t),\chi_{2,n_k}(t)) + \partial\psi(\chi_{1,n_k}(t),\chi_{2,n_k}(t)) \ni -G^t_{\theta_{n_k}(t)}(\chi_{1,n_k}(t),\chi_{2,n_k}(t)) \\ \text{in } H\times H \text{ for } t\in(0,T),\end{aligned} \tag{3.13}$$

$$\theta_{n_k}(0)=\theta_0 \text{ in } H, \text{ and } (\chi_{1,n_k}(0),\chi_{2,n_k}(0)) = (\chi_{1,0},\chi_{2,0}) \text{ in } H\times H, \tag{3.14}$$

where $\varphi^t_{n_k}(\cdot)$ is the time-dependent convex function defined by (3.4) with $\gamma(t)$ replaced by $\gamma_{n_k}(t)$, and $\psi(\cdot,\cdot)$ is the convex function defined by (3.5). In addition, $G^t_{\theta_{n_k}(t)}(\cdot,\cdot)$ is a time-dependent operator on $H\times H$ defined by (3.6) with $\theta(t)$ and $g(t)$ replaced by $\theta_{n_k}(t)$ and $g_{n_k}(t)$, respectively:

$$\begin{aligned}G^t_{\theta_{n_k}(t)}(z_1,z_2) := \frac{1}{\mu_0}(-l(\theta^c-\theta_{n_k}(t)), \beta a_1 g_{n_k}(t)\alpha(\theta_{n_k}(t)) - \beta\alpha(\theta_{n_k}(t))^2 z_2) \\ \text{for any } (z_1,z_2)\in H\times H.\end{aligned} \tag{3.15}$$

From (A2), (3.11), and Lebesgue's dominated convergence theorem, note that

$$\alpha(\theta_{n_k}) \to \alpha(\theta) \text{ in } L^2(0,T;H) \text{ as } k\to\infty. \tag{3.16}$$

We also note from (3.8) and the compact embedding $H^1(0,T) \hookrightarrow C([0,T])$ that

$$g_{n_k} \to g \text{ in } C([0,T]) \text{ as } k \to \infty,$$

taking a subsequence if necessary. Thus, we observe from (A2), (3.11), and (3.16) that

$$G^{(\cdot)}_{\theta_{n_k}(\cdot)}(\chi_{1,n_k}, \chi_{2,n_k}) \to G^{(\cdot)}_{\theta(\cdot)}(\chi_1, \chi_2) \text{ weakly in } (L^2(0,T;H))^2 \tag{3.17}$$

as $k \to \infty$. In addition, we easily observe from (2.1) and (3.9) that

$$\gamma_{n_k} \to \gamma \text{ in } C([0,T];V) \text{ as } k \to \infty, \tag{3.18}$$

thus,

$$\varphi^t_{n_k} \to \varphi^t \text{ on } H \text{ in the sense of Mosco [37] as } k \to \infty \text{ for all } t \in [0,T]. \tag{3.19}$$

Applying the abstract convergence theorem established in [2, 33] with (3.7), (3.11), (3.17), and (3.19), there is a triplet of functions $(\widetilde{\theta}, \widetilde{\chi}_1, \widetilde{\chi}_2)$ (taking a subsequence if necessary) such that $(\widetilde{\theta}, \widetilde{\chi}_1, \widetilde{\chi}_2) \in (W^{1,2}(0,T;H) \cap L^\infty(0,T;V))^3$,

$$(\theta_{n_k}, \chi_{1,n_k}, \chi_{2,n_k}) \to (\widetilde{\theta}, \widetilde{\chi}_1, \widetilde{\chi}_2) \text{ in } (C([0,T];H))^3 \text{ as } k \to \infty, \tag{3.20}$$

and $(\widetilde{\theta}, \widetilde{\chi}_1, \widetilde{\chi}_2)$ is the unique solution to the following system:

$$\widetilde{\theta}_t(t) + \partial\varphi^t(\widetilde{\theta}(t)) \ni \frac{a_0}{L_0} f(t) + \frac{L_1}{L_0}(\chi_1)_t(t) \text{ in } H \text{ for } t \in (0,T), \tag{3.21}$$

$$\frac{d}{dt}(\widetilde{\chi}_1(t), \widetilde{\chi}_2(t)) + \partial\psi(\widetilde{\chi}_1(t), \widetilde{\chi}_2(t)) \ni -G^t_{\theta(t)}(\chi_1(t), \chi_2(t)) \text{ in } H \times H \text{ for } t \in (0,T), \tag{3.22}$$

$$\widetilde{\theta}(0) = \theta_0 \text{ in } H, \text{ and } (\widetilde{\chi}_1(0), \widetilde{\chi}_2(0)) = (\chi_{1,0}, \chi_{2,0}) \text{ in } H \times H. \tag{3.23}$$

On account of the uniqueness of solutions to (3.21)–(3.23) and to the state problem $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)$ (cf. (3.1)–(3.3)), we conclude from (3.11) and (3.20) that $(\theta, \chi_1, \chi_2) = (\widetilde{\theta}, \widetilde{\chi}_1, \widetilde{\chi}_2)$ is a unique solution to $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)$ on $[0,T]$, and the convergence (3.10) holds without extracting any subsequence from $\{n\}_{n\in\mathbb{N}}$. Thus, the proof of Proposition 3.1 has been completed. □

We now prove the main Theorem 2.1 of this paper, which is concerned with the existence of an optimal solution to (OP).

*Proof of Theorem 2.1.* By the quite standard method, we can prove Theorem 2.1. Indeed, let $\{(f_n, g_n, \gamma_{0,n}, \gamma_{1,n})\}_{n\in\mathbb{N}} \subset \mathcal{U}^M_{ad}$ be a minimizing sequence such that

$$\lim_{n\to\infty} J(f_n, g_n, \gamma_{0,n}, \gamma_{1,n}) = \inf_{(f,g,\gamma_0,\gamma_1)\in\mathcal{U}^M_{ad}} J(f, g, \gamma_0, \gamma_1).$$

Then, from the definition (1.9) of $J(f_n, g_n, \gamma_{0,n}, \gamma_{1,n})$, it follows that $\{f_n\}_{n\in\mathbb{N}}$ is bounded in $L^2(0,T;H)$. In addition, from the definition (1.8) of $\mathcal{U}^M_{ad}$, we see that $\{(g_n, \gamma_{0,n}, \gamma_{1,n})\}_{n\in\mathbb{N}}$ is bounded in $(H^1(0,T))^3$.

Note that $H^1(0,T)$ is compactly embedded in $C([0,T])$. Therefore, there are a subsequence $\{n_k\}_{k\in\mathbb{N}} \subset \{n\}_{n\in\mathbb{N}}$ and the quadruplet of functions $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$ such that $n_k \to \infty$ and

$$f_{n_k} \to f^* \text{ weakly in } L^2(0,T;H), \tag{3.24}$$

$$g_n \to g^* \text{ weakly in } H^1(0,T), \text{ in } C([0,T]), \tag{3.25}$$

$$(\gamma_{0,n}, \gamma_{1,n}) \to (\gamma_0^*, \gamma_1^*) \text{ weakly in } (H^1(0,T))^2, \text{ in } (C([0,T]))^2 \tag{3.26}$$

as $k \to \infty$. Indeed, we get $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$, because $\mathcal{U}_{ad}^M$ is a convex and closed subset, hence is weakly closed.

Let $(\theta_{n_k}, \chi_{1,n_k}, \chi_{2,n_k})$ uniquely solve problem $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f_{n_k}, g_{n_k}, \gamma_{0,n_k}, \gamma_{1,n_k})$ on $[0,T]$. Then, by Proposition 3.1, we observe that

$$(\theta_{n_k}, \chi_{1,n_k}, \chi_{2,n_k}) \to (\theta^*, \chi_1^*, \chi_2^*) \text{ in } (C([0,T];H))^3 \text{ as } k \to \infty, \tag{3.27}$$

where $(\theta^*, \chi_1^*, \chi_2^*)$ is a unique solution to $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^*, g^*, \gamma_0^*, \gamma_1^*)$ on $[0,T]$.

Therefore, it follows from (3.24)–(3.27) and the weak lower semicontinuity of $L^2$-norm that

$$J(f^*, g^*, \gamma_0^*, \gamma_1^*) \le \lim_{k\to\infty} J(f_{n_k}, g_{n_k}, \gamma_{0,n_k}, \gamma_{1,n_k}) = \inf_{(f,g,\gamma_0,\gamma_1)\in\mathcal{U}_{ad}^M} J(f, g, \gamma_0, \gamma_1),$$

which implies that $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$ is an optimal control to (OP). Thus, the proof of Theorem 2.1 is complete. □

## 4 Approximating problems (SMA)$^\varepsilon$ and (OP)$_\delta^\varepsilon$

In this section, we consider the approximating problems $(\mathrm{SMA})^\varepsilon$ and (OP)$_\delta^\varepsilon$ of (SMA) and (OP), respectively, for each $\varepsilon \in (0,1]$ and $\delta \ge 0$. After showing the solvability of $(\mathrm{SMA})^\varepsilon$, we prove Theorems 2.2 and 2.3, which is concerned with the existence of optimal control for (OP)$_\delta^\varepsilon$ and the relationship between (OP) and (OP)$_\delta^\varepsilon$.

We begin by proving Proposition 2.2, which is concerned with the solvability of the approximating system $(\mathrm{SMA})^\varepsilon$.

*Proof of Proposition 2.2.* By a similar argument to (SMA), we can construct the unique solution $(\theta^\varepsilon, \chi_1^\varepsilon, \chi_2^\varepsilon)$ to $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)^\varepsilon$ on $[0,T]$ satisfying the bounded estimate (2.10). Indeed, the approximating problem $(\mathrm{SMA})^\varepsilon$ is reformulated to the following system of abstract evolution equations (cf. (3.1)–(3.3)):

$$\theta_t^\varepsilon(t) - \frac{L_1}{L_0}(\chi_1^\varepsilon)_t(t) + \partial\varphi^t(\theta^\varepsilon(t)) \ni \frac{a_0}{L_0} f(t) \text{ in } H \text{ for } t \in (0,T), \tag{4.1}$$

$$\frac{d}{dt}(\chi_1^\varepsilon(t), \chi_2^\varepsilon(t)) + \partial\psi^\varepsilon(\chi_1^\varepsilon(t), \chi_2^\varepsilon(t)) + G_{\theta^\varepsilon(t)}^t(\chi_1^\varepsilon(t), \chi_2^\varepsilon(t)) = (0,0) \text{ in } H \times H \quad \text{for } t \in (0,T), \tag{4.2}$$

$$\theta^\varepsilon(0) = \theta_0 \text{ in } H, \text{ and } (\chi_1^\varepsilon(0), \chi_2^\varepsilon(0)) = (\chi_{1,0}, \chi_{2,0}) \text{ in } H \times H, \tag{4.3}$$

where $\varphi^t(\cdot)$ and $G^t_{(\cdot)}(\cdot,\cdot)$ are same ones defined by (3.4) and (3.6), respectively. In addition, for each $\varepsilon \in (0,1]$, $\psi^\varepsilon$ is a proper, l.s.c., and convex function on $H \times H$, defined by:

$$\psi^\varepsilon(z_1,z_2) := \begin{cases} \dfrac{\mu_1}{2\mu_0}\displaystyle\sum_{i=1}^{2}\int_\Omega |(z_i)_x(x)|^2 dx \\ \quad +\dfrac{1}{\mu_0}\displaystyle\int_\Omega \widehat{K}^\varepsilon(z_1(x),z_2(x))dx, \quad \text{if } z_i \in V,\ i=1,2, \\ \infty, \quad \text{otherwise,} \end{cases} \tag{4.4}$$

where $\widehat{K}^\varepsilon(\cdot,\cdot)$ is the function in the assumption (A5).

Therefore, by the slight modification of the proof of [22, Theorems 2.1 and 2.2], we can show that $(\mathrm{SMA};\theta_0,\chi_{1,0},\chi_{2,0},f,g,\gamma_0,\gamma_1)^\varepsilon$ has a unique $(\theta^\varepsilon,\chi_1^\varepsilon,\chi_2^\varepsilon)$ fulfilling (i)–(v) and the a priori estimate (2.10) in Proposition 2.2. For the detailed arguments, we refer to [22, Theorems 2.1 and 2.2].

Thus, the proof of Proposition 2.2 is complete. □

We now state the following result of continuous dependence between $(\mathrm{SMA})$ and $(\mathrm{SMA})^\varepsilon$ ($\varepsilon \in (0,1]$).

**Proposition 4.1.** *Suppose that all the assumptions of Proposition 2.2 hold. In addition, assume* $\varepsilon \in (0,1]$, $\{f^\varepsilon\}_{\varepsilon\in(0,1]} \subset L^2(0,T;H)$, $\{(g^\varepsilon,\gamma_0^\varepsilon,\gamma_1^\varepsilon)\}_{\varepsilon\in(0,1]} \subset (H^1(0,T))^3$, $f \in L^2(0,T;H)$, $(g,\gamma_0,\gamma_1) \in (H^1(0,T))^3$*, and*

$$f^\varepsilon \to f \text{ weakly in } L^2(0,T;H), \tag{4.5}$$

$$g^\varepsilon \to g \text{ weakly in } H^1(0,T), \tag{4.6}$$

$$(\gamma_0^\varepsilon,\gamma_1^\varepsilon) \to (\gamma_0,\gamma_1) \text{ in } (C[0,T])^2 \tag{4.7}$$

*as* $\varepsilon \to 0$*. Let* $(\theta^\varepsilon,\chi_1^\varepsilon,\chi_2^\varepsilon)$ *be the unique solution to the approximating state problem* $(\mathrm{SMA};\theta_0,\chi_{1,0},\chi_{2,0},f^\varepsilon,g^\varepsilon,\gamma_0^\varepsilon,\gamma_1^\varepsilon)^\varepsilon$ *on* $[0,T]$*. Then,* $(\theta^\varepsilon,\chi_1^\varepsilon,\chi_2^\varepsilon)$ *converges to the unique solution* $(\theta,\chi_1,\chi_2)$ *to* $(\mathrm{SMA};\theta_0,\chi_{1,0},\chi_{2,0},f,g,\gamma_0,\gamma_1)$ *on* $[0,T]$ *in the sense that*

$$(\theta^\varepsilon,\chi_1^\varepsilon,\chi_2^\varepsilon) \to (\theta,\chi_1,\chi_2) \text{ in } (C([0,T];H))^3 \text{ as } \varepsilon \to 0. \tag{4.8}$$

*Proof.* By (2.10) with (4.5), (4.6), and (4.7), there are a subsequence $\{\varepsilon_n\}_{n\in\mathbb{N}} \subset \{\varepsilon\}_{\varepsilon\in(0,1]}$ and the triplet of functions $(\theta,\chi_1,\chi_2) \in (W^{1,2}(0,T;H)\cap L^\infty(0,T;V))^3$ such that $\varepsilon_n \to 0$,

$$\left.\begin{aligned} (\theta^{\varepsilon_n},\chi_1^{\varepsilon_n},\chi_2^{\varepsilon_n}) \to (\theta,\chi_1,\chi_2) \quad & \text{in } (C([0,T];H))^3, \\ & \text{weakly in } (W^{1,2}(0,T;H))^3, \\ & \text{weakly-}* \text{ in } L^\infty((0,T;V))^3 \end{aligned}\right\} \tag{4.9}$$

as $n \to \infty$.

By similar arguments used in the proof of Proposition 3.1, we can show that $(\theta,\chi_1,\chi_2)$ is a solution to $(\mathrm{SMA};\theta_0,\chi_{1,0},\chi_{2,0},f,g,\gamma_0,\gamma_1)$ on $[0,T]$. Indeed, note that the solution $(\theta^{\varepsilon_n},\chi_1^{\varepsilon_n},\chi_2^{\varepsilon_n})$ to $(\mathrm{SMA})^{\varepsilon_n}$ satisfies the system of abstract evolution equations (4.1)–(4.3) with $\varepsilon$ replaced by $\varepsilon_n$. In addition, we observe from assumption (A5) that

$$\psi^{\varepsilon_n} \to \psi \text{ on } H\times H \text{ in the sense of Mosco [37] as } n \to \infty,$$

where $\psi^{\varepsilon_n}$ and $\psi$ are convex functions defined in (4.4) with $\varepsilon$ replaced by $\varepsilon_n$ and (3.5), respectively.

By a similar manner to the proof of Proposition 3.1, we conclude that $(\theta, \chi_1, \chi_2)$ is a unique solution to $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)$ on $[0, T]$, and the convergence (4.8) holds without extracting any subsequence from $\{\varepsilon\}_{\varepsilon\in(0,1]}$. Thus, the proof of Proposition 4.1 is complete. □

We now prove the second main theorem of this paper (Theorem 2.2), which is concerned with the existence of an optimal control for (OP)$^\varepsilon_\delta$ for each $\varepsilon \in (0, 1]$ and $\delta \geq 0$.

*Proof of Theorem 2.2.* By an argument similar to that as in Proposition 3.1 (cf. Proposition 4.1), we can obtain the result of convergence of solutions to $(\mathrm{SMA})^\varepsilon$. Hence, for each $\varepsilon \in (0, 1]$ and $\delta \geq 0$, the proof of the existence of an optimal control $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \in \mathcal{U}^M_{ad}$ for (OP)$^\varepsilon_\delta$ will be a slight modification of that as in Theorem 2.1. Thus, we omit the detailed proof of this Theorem 2.2. □

We next prove Theorem 2.3 concerning the relationship between the optimal control problems (OP) and (OP)$^\varepsilon_\delta$.

*Proof of Theorem 2.3.* We first show (I). Assume $\delta = 0$. Let $\{(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})\}_{\varepsilon\in(0,1]} \subset \mathcal{U}^M_{ad}$ be a sequence of optimal controls for (OP)$^\varepsilon_0$. Let $(f, g, \gamma_0, \gamma_1)$ be arbitrary function in $\mathcal{U}^M_{ad}$. In addition, let $(\theta^\varepsilon, \chi_1^\varepsilon, \chi_2^\varepsilon)$ be a unique solution to the approximating state problem $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)^\varepsilon$ on $[0, T]$, and let $(\theta, \chi_1, \chi_2)$ be a unique solution to the original state problem $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)$ on $[0, T]$. Then, we observe from Proposition 4.1 that

$$(\theta^\varepsilon, \chi_1^\varepsilon, \chi_2^\varepsilon) \to (\theta, \chi_1, \chi_2) \text{ in } (C([0, T]; H))^3 \text{ as } \varepsilon \to 0. \tag{4.10}$$

Since $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})$ is the optimal control for (OP)$^\varepsilon_0$, we observe that

$$\begin{aligned}
&J_0^\varepsilon(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \\
\leq &J_0^\varepsilon(f, g, \gamma_0, \gamma_1) \\
= &\frac{c_0}{2}\int_0^T |(\theta^\varepsilon - \theta_d)(t)|_H^2 dt + \frac{c_1}{2}\int_0^T |(\chi_1^\varepsilon - \chi_{1,d})(t)|_H^2 dt \\
&+ \frac{c_2}{2}\int_0^T |(\chi_2^\varepsilon - \chi_{2,d})(t)|_H^2 dt \\
&+ \frac{m_0}{2}\int_0^T a_0^2 |f(t)|_H^2 dt + \frac{m_1}{2}\int_0^T a_1^2 |g(t)|^2 dt \\
&+ \frac{m_2}{2}\int_0^T a_2^2 |\gamma_0(t)|^2 dt + \frac{m_3}{2}\int_0^T a_3^2 |\gamma_1(t)|^2 dt.
\end{aligned} \tag{4.11}$$

Clearly, it follows from (1.8), (2.11), (4.10), and (4.11) that $\{(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})\}_{\varepsilon\in(0,1]}$ is bounded in $\mathcal{U}^M_{ad}$ with respect to $\varepsilon \in (0, 1]$. Therefore, taking account of the compact embedding $H^1(0, T) \hookrightarrow C([0, T])$, there are a subsequence $\{\varepsilon_n\}_{n\in\mathbb{N}} \subset \{\varepsilon\}_{\varepsilon\in(0,1]}$ and the quadruplet of functions $(f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**}) \in \mathcal{U}^M_{ad}$ such that $\varepsilon_n \to 0$,

$$f^{*,\varepsilon_n} \to f^{**} \text{ weakly in } L^2(0, T; H), \tag{4.12}$$

$$g^{*,\varepsilon_n} \to g^{**} \text{ weakly in } H^1(0, T), \text{ in } C([0, T]), \tag{4.13}$$

$$(\gamma_0^{*,\varepsilon_n}, \gamma_1^{*,\varepsilon_n}) \to (\gamma_0^{**}, \gamma_1^{**}) \text{ weakly in } (H^1(0,T))^2, \text{ in } (C([0,T]))^2, \tag{4.14}$$

as $n \to \infty$.

Let $(\theta^{*,\varepsilon_n}, \chi_1^{*,\varepsilon_n}, \chi_2^{*,\varepsilon_n})$ be the unique solution to the approximating state problem $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^{*,\varepsilon_n}, g^{*,\varepsilon_n}, \gamma_0^{*,\varepsilon_n}, \gamma_1^{*,\varepsilon_n})^{\varepsilon_n}$ on $[0,T]$. Then, by (4.12)–(4.14) and Proposition 4.1, we observe that $(\theta^{*,\varepsilon_n}, \chi_1^{*,\varepsilon_n}, \chi_2^{*,\varepsilon_n})$ converges to the unique solution $(\theta^{**}, \chi_1^{**}, \chi_2^{**})$ to the original state system $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**})$ on $[0,T]$ in the sense that

$$(\theta^{*,\varepsilon_n}, \chi_1^{*,\varepsilon_n}, \chi_2^{*,\varepsilon_n}) \to (\theta^{**}, \chi_1^{**}, \chi_2^{**}) \text{ in } (C([0,T];H))^3 \text{ as } n \to \infty. \tag{4.15}$$

From (4.10)–(4.15) and the weak lower semicontinuity of $L^2$-norm, we infer that

$$\begin{aligned} J(f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**}) &\le \liminf_{n\to\infty} J_0^{\varepsilon_n}(f^{*,\varepsilon_n}, g^{*,\varepsilon_n}, \gamma_0^{*,\varepsilon_n}, \gamma_1^{*,\varepsilon_n}) \\ &\le \lim_{n\to\infty} J_0^{\varepsilon_n}(f, g, \gamma_0, \gamma_1) = J(f, g, \gamma_0, \gamma_1). \end{aligned}$$

Since $(f, g, \gamma_0, \gamma_1)$ is arbitrary function in $\mathcal{U}_{ad}^M$, we conclude from the above inequality that $(f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**})$ is the optimal control for (OP). Hence, Theorem 2.3(I) holds.

We now show (II). Assume $\delta > 0$. Let $(f^*, g^*, \gamma_0^*, \gamma_1^*)$ be the optimal control for (OP) obtained in Theorem 2.1. Let $(\theta^\varepsilon, \chi_1^\varepsilon, \chi_2^\varepsilon)$ be a unique solution to the approximating state system $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^*, g^*, \gamma_0^*, \gamma_1^*)^\varepsilon$ on $[0,T]$. In addition, let $(\theta^*, \chi_1^*, \chi_2^*)$ be a unique solution to the original state problem $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^*, g^*, \gamma_0^*, \gamma_1^*)$ on $[0,T]$. Then, we observe from Proposition 4.1 that

$$(\theta^\varepsilon, \chi_1^\varepsilon, \chi_2^\varepsilon) \to (\theta^*, \chi_1^*, \chi_2^*) \text{ in } (C([0,T];H))^3 \text{ as } \varepsilon \to 0. \tag{4.16}$$

On the other hand, since $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})$ is an optimal control for $(\mathrm{OP})_\delta^\varepsilon$, we observe that

$$\begin{aligned} & J_\delta^\varepsilon(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \\ \le & J_\delta^\varepsilon(f^*, g^*, \gamma_0^*, \gamma_1^*) \\ = & \frac{c_0}{2}\int_0^T |(\theta^\varepsilon - \theta_d)(t)|_H^2 dt + \frac{c_1}{2}\int_0^T |(\chi_1^\varepsilon - \chi_{1,d})(t)|_H^2 dt \\ & + \frac{c_2}{2}\int_0^T |(\chi_2^\varepsilon - \chi_{2,d})(t)|_H^2 dt \\ & + \frac{m_0}{2}\int_0^T a_0^2 |f^*(t)|_H^2 dt + \frac{m_1}{2}\int_0^T a_1^2 |g^*(t)|^2 dt \\ & + \frac{m_2}{2}\int_0^T a_2^2 |\gamma_0^*(t)|^2 dt + \frac{m_3}{2}\int_0^T a_3^2 |\gamma_1^*(t)|^2 dt. \end{aligned} \tag{4.17}$$

It follows from (1.8), (2.11), (4.16), and (4.17) that $\{(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})\}_{\varepsilon\in(0,1]}$ is bounded in $\mathcal{U}_{ad}^M$ with respect to $\varepsilon \in (0,1]$. Therefore, taking account of the compact embedding $H^1(0,T) \hookrightarrow C([0,T])$, there are a subsequence $\{\varepsilon_n\}_{n\in\mathbb{N}} \subset \{\varepsilon\}_{\varepsilon\in(0,1]}$ and the quadruplet of functions $(\widetilde{f}, \widetilde{g}, \widetilde{\gamma}_0, \widetilde{\gamma}_1) \in \mathcal{U}_{ad}^M$ such that $\varepsilon_n \to 0$,

$$f^{*,\varepsilon_n} \to \widetilde{f} \text{ weakly in } L^2(0,T;H), \tag{4.18}$$

$$g^{*,\varepsilon_n} \to \widetilde{g} \text{ weakly in } H^1(0,T), \text{ in } C([0,T]), \tag{4.19}$$

$$(\gamma_0^{*,\varepsilon_n}, \gamma_1^{*,\varepsilon_n}) \to (\widetilde{\gamma}_0, \widetilde{\gamma}_1) \text{ weakly in } (H^1(0,T))^2, \text{ in } (C([0,T]))^2, \tag{4.20}$$

as $n \to \infty$.

For any $n \in \mathbb{N}$, let $(\theta^{*,\varepsilon_n}, \chi_1^{*,\varepsilon_n}, \chi_2^{*,\varepsilon_n})$ be a unique solution to the approximating state system (SMA;$\theta_0, \chi_{1,0}, \chi_{2,0}, f^{*,\varepsilon_n}, g^{*,\varepsilon_n}, \gamma_0^{*,\varepsilon_n}, \gamma_1^{*,\varepsilon_n})^{\varepsilon_n}$ on $[0,T]$. Then, from (4.18)–(4.20) and Proposition 4.1, we observe that $(\theta^{*,\varepsilon_n}, \chi_1^{*,\varepsilon_n}, \chi_2^{*,\varepsilon_n})$ converges to the unique solution $(\widetilde{\theta}, \widetilde{\chi}_1, \widetilde{\chi}_2)$ to the original state system (SMA;$\theta_0, \chi_{1,0}, \chi_{2,0}, \widetilde{f}, \widetilde{g}, \widetilde{\gamma}_0, \widetilde{\gamma}_1)$ on $[0,T]$ in the sense that

$$(\theta^{*,\varepsilon_n}, \chi_1^{*,\varepsilon_n}, \chi_2^{*,\varepsilon_n}) \to (\widetilde{\theta}, \widetilde{\chi}_1, \widetilde{\chi}_2) \text{ in } (C([0,T];H))^3 \text{ as } n \to \infty. \tag{4.21}$$

From (2.11), (4.16)–(4.21), and the weak lower semicontinuity of $L^2$-norm, we infer that

$$\begin{aligned}
&\frac{\delta}{2} \limsup_{n\to\infty} \int_0^T |(f^{*,\varepsilon_n} - f^*)(t)|_H^2 dt \\
\le &\limsup_{n\to\infty} \Bigg( J_\delta^{\varepsilon_n}(f^*, g^*, \gamma_0^*, \gamma_1^*) - \frac{c_0}{2} \int_0^T |(\theta^{*,\varepsilon_n} - \theta_d)(t)|_H^2 dt \\
&\qquad - \frac{c_1}{2} \int_0^T |(\chi_1^{*,\varepsilon_n} - \chi_{1,d})(t)|_H^2 dt - \frac{c_2}{2} \int_0^T |(\chi_2^{*,\varepsilon_n} - \chi_{2,d})(t)|_H^2 dt \\
&\qquad - \frac{m_0}{2} \int_0^T a_0^2 |f^{*,\varepsilon_n}(t)|_H^2 dt - \frac{m_1}{2} \int_0^T a_1^2 |g^{*,\varepsilon_n}(t)|^2 dt \\
&\qquad - \frac{m_2}{2} \int_0^T a_2^2 |\gamma_0^{*,\varepsilon_n}(t)|^2 dt - \frac{m_3}{2} \int_0^T a_3^2 |\gamma_1^{*,\varepsilon_n}(t)|^2 dt \Bigg) \\
\le &J(f^*, g^*, \gamma_0^*, \gamma_1^*) - \frac{c_0}{2} \int_0^T |(\widetilde{\theta} - \theta_d)(t)|_{L^2(0,1)}^2 dt \\
&- \frac{c_1}{2} \int_0^T |(\widetilde{\chi}_1 - \chi_{1,d})(t)|_H^2 dt - \frac{c_2}{2} \int_0^T |(\widetilde{\chi}_2 - \chi_{2,d})(t)|_H^2 dt \\
&- \frac{m_0}{2} \int_0^T a_0^2 |\widetilde{f}(t)|_H^2 dt - \frac{m_1}{2} \int_0^T a_1^2 |\widetilde{g}(t)|^2 dt \\
&- \frac{m_2}{2} \int_0^T a_2^2 |\widetilde{\gamma}_0(t)|^2 dt - \frac{m_3}{2} \int_0^T a_3^2 |\widetilde{\gamma}_1(t)|^2 dt \\
= &J(f^*, g^*, \gamma_0^*, \gamma_1^*) - J(\widetilde{f}, \widetilde{g}, \widetilde{\gamma}_0, \widetilde{\gamma}_1).
\end{aligned}$$

Thus, we have

$$J(\widetilde{f}, \widetilde{g}, \widetilde{\gamma}_0, \widetilde{\gamma}_1) + \frac{\delta}{2} \limsup_{n\to\infty} \int_0^T |(f^{*,\varepsilon_n} - f^*)(t)|_H^2 dt \le J(f^*, g^*, \gamma_0^*, \gamma_1^*).$$

Since $(f^*, g^*, \gamma_0^*, \gamma_1^*)$ is the optimal control for (OP), we observe that

$$\frac{\delta}{2} \limsup_{n\to\infty} \int_0^T |(f^{*,\varepsilon_n} - f^*)(t)|_H^2 dt = 0. \tag{4.22}$$

Therefore, we conclude from (4.18) and (4.22) that $\widetilde{f} = f^*$ and the convergence (2.16) holds, i.e.,

$$f^{*,\varepsilon_n} \to f^* \text{ in } L^2(0,T;H) \text{ as } n \to \infty.$$

By the same arguments as above, we observe that $\widetilde{g} = g^*$, $\widetilde{\gamma}_0 = \gamma_0^*$, $\widetilde{\gamma}_1 = \gamma_1^*$, and the convergence (2.17) and (2.18) hold.

In addition, due to the uniqueness of solutions to $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^*, g^*, \gamma_0^*, \gamma_1^*)$ on $[0,T]$, we infer that $(\widetilde{\theta}, \widetilde{\chi}_1, \widetilde{\chi}_2) = (\theta^*, \chi_1^*, \chi_2^*)$ and the convergence (2.19) holds. Hence, Theorem 2.3(II) holds.

Thus, the proof of Theorem 2.3 is complete. □

# 5 Optimality condition for (OP)$_\delta^\varepsilon$

In previous Section 4, we proved the existence of an optimal control $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})$ for the approximating problem (OP)$_\delta^\varepsilon$ for each $\varepsilon \in (0,1]$ and $\delta \geq 0$. In this section, we show the main result (Theorem 2.4) concerning the necessary condition of the optimal control for (OP)$_\delta^\varepsilon$.

Throughout this section, we suppose that all the assumptions of Theorem 2.4 are made. In addition, we fix $\varepsilon \in (0,1]$ and $\delta \geq 0$.

For the space $\mathcal{U} := L^2(0,T;H) \times H^1(0,T) \times H^1(0,T) \times H^1(0,T)$ (cf. (1.8)), we define the control-to-state mapping as follows.

**Definition 5.1.** **(I)** *We denote by* $\Lambda^\varepsilon : \mathcal{U} \to (L^2(0,T;H))^3$ *the control-to-state mapping that assigns to any control* $(f,g,\gamma_0,\gamma_1) \in \mathcal{U}$ *the solution* $(\theta, \chi_1, \chi_2) := \Lambda^\varepsilon(f,g,\gamma_0,\gamma_1)$ *to the approximating state system* $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)^\varepsilon$ *on* $[0,T]$.
**(II)** *Let* $(f^*, g^*, \gamma_0^*, \gamma_1^*) \in \mathcal{U}_{ad}^M$ *be the optimal control for* $(\mathrm{OP})_\delta^\varepsilon$. *Then,*

$$(\theta^*, \chi_1^*, \chi_2^*, f^*, g^*, \gamma_0^*, \gamma_1^*) = (\Lambda^\varepsilon(f^*, g^*, \gamma_0^*, \gamma_1^*), f^*, g^*, \gamma_0^*, \gamma_1^*)$$

*is called the optimal pair for the optimal control problem* $(\mathrm{OP})_\delta^\varepsilon$.

For a moment, we often omit the subscript $\varepsilon \in (0,1]$.

We first show the Gâteaux differentiability of $\Lambda^\varepsilon$ and $J_\delta^\varepsilon$.

Note from Proposition 2.2 that for any control $(f,g,\gamma_0,\gamma_1) \in \mathcal{U}$, the approximating state system $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f, g, \gamma_0, \gamma_1)^\varepsilon$ has a unique solution $(\theta, \chi_1, \chi_2)$ on $[0,T]$. Therefore, for any $(f,g,\gamma_0,\gamma_1) \in \mathcal{U}$, any direction $(\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}$, and any number $\lambda \in [-1,1] \setminus \{0\}$, we here put $(\theta_\lambda, \chi_{1,\lambda}, \chi_{2,\lambda}) := \Lambda^\varepsilon(f + \lambda\check{f}, g + \lambda\check{g}, \gamma_0 + \lambda\check{\gamma}_0, \gamma_1 + \lambda\check{\gamma}_1)$, $(\theta, \chi_1, \chi_2) := \Lambda^\varepsilon(f, g, \gamma_0, \gamma_1)$, $u_\lambda := (\theta_\lambda - \theta)/\lambda$, $w_{1,\lambda} = (\chi_{1,\lambda} - \chi_1)/\lambda$, and $w_{2,\lambda} = (\chi_{2,\lambda} - \chi_2)/\lambda$.

Note that the triplet of functions $(u_\lambda, w_{1,\lambda}, w_{2,\lambda})$ satisfies the following system:

$$(L_0 u_\lambda - L_1 w_{1,\lambda})_t - h(u_\lambda)_{xx} = a_0 \check{f}(t,x) \text{ in } Q = (0,T) \times \Omega, \tag{5.1}$$

$$\begin{aligned} &\mu_0 \begin{pmatrix} w_{1,\lambda} \\ w_{2,\lambda} \end{pmatrix}_t - \mu_1 \begin{pmatrix} w_{1,\lambda} \\ w_{2,\lambda} \end{pmatrix}_{xx} + \begin{pmatrix} \overline{K}^\varepsilon_{11,\lambda}(t,x) w_{1,\lambda} + \overline{K}^\varepsilon_{12,\lambda}(t,x) w_{2,\lambda} \\ \overline{K}^\varepsilon_{21,\lambda}(t,x) w_{1,\lambda} + \overline{K}^\varepsilon_{22,\lambda}(t,x) w_{2,\lambda} \end{pmatrix} \\ &= \begin{pmatrix} -l u_\lambda \\ -\beta a_1 \check{g}(t) \alpha(\theta_\lambda) - \beta a_1 g(t) \overline{\alpha}_\lambda(t,x) u_\lambda + \beta \overline{\overline{\alpha}}_\lambda(t,x) u_\lambda \chi_{2,\lambda} + \beta \alpha(\theta)^2 w_{2,\lambda} \end{pmatrix} \text{ in } Q, \end{aligned} \tag{5.2}$$

$$\begin{aligned} -h(u_\lambda)_x(t,0) + k(u_\lambda(t,0) - a_2\check{\gamma}_0(t)) = h(u_\lambda)_x(t,1) + k(u_\lambda(t,1) - a_3\check{\gamma}_1(t)) = 0, \\ \text{for } t \in (0,T), \end{aligned} \tag{5.3}$$

$$(w_{i,\lambda})_x(t,0) = (w_{i,\lambda})_x(t,1) = 0, \quad t \in (0,T),\ i = 1,2, \tag{5.4}$$

$$u_\lambda(0,x) = 0, \quad w_{i,\lambda}(0,x) = 0, \quad x \in \Omega,\ i = 1,2, \tag{5.5}$$

where notations $\overline{K}^{\varepsilon}_{ij,\lambda}$ $(i,j=1,2)$, $\overline{\alpha}_\lambda$, and $\overline{\overline{\alpha}}_\lambda$ are functions on $Q$, given as:

$$\overline{K}^{\varepsilon}_{i1,\lambda}(t,x) = \int_0^1 \partial_1\partial_i\widehat{K}^{\varepsilon}(\chi_1(t,x) + s(\chi_{1,\lambda}(t,x) - \chi_1(t,x)), \chi_{2,\lambda}(t,x))ds, \quad (i=1,2);$$

$$\overline{K}^{\varepsilon}_{i2,\lambda}(t,x) = \int_0^1 \partial_2\partial_i\widehat{K}^{\varepsilon}(\chi_1(t,x), \chi_2(t,x) + s(\chi_{2,\lambda}(t,x) - \chi_2(t,x)))ds, \quad (i=1,2);$$

$$\overline{\alpha}_\lambda(t,x) = \int_0^1 \alpha'(\theta(t,x) + s(\theta_\lambda(t,x) - \theta(t,x)))ds;$$

$$\overline{\overline{\alpha}}_\lambda(t,x) = \int_0^1 2\alpha'(\theta(t,x) + s(\theta_\lambda(t,x) - \theta(t,x)))\alpha(\theta(t,x) + s(\theta_\lambda(t,x) - \theta(t,x)))ds;$$

for $(t,x) \in Q$, with use of the partial derivative $\partial_i\partial_j\widehat{K}^{\varepsilon}$ of the convex function $\widehat{K}^{\varepsilon}$ $(i,j=1,2)$ and the derivative $\alpha'$ of the single-valued function $\alpha$.

We now give a uniform estimate of solutions $(u_\lambda, w_{1,\lambda}, w_{2,\lambda})$ to (5.1)–(5.5) with respect to $\lambda \in [-1,1]\setminus\{0\}$.

**Lemma 5.1.** *Suppose that all the assumptions of Theorem 2.4 are satisfied. Then, there is a positive number $N_3 > 0$, independent of $\lambda \in [-1,1]\setminus\{0\}$, such that*

$$\begin{aligned}
&\sup_{t\in[0,T]} |u_\lambda(t)|_V^2 + \sum_{i=1}^2 \sup_{t\in[0,T]} |w_{i,\lambda}(t)|_V^2 + \int_0^T |(u_\lambda)_t(t)|_H^2 dt + \sum_{i=1}^2 \int_0^T |(w_{i,\lambda})_t(t)|_H^2 dt \\
\le& N_3\left(a_0^2|\check{f}|^2_{L^2(0,T;H)} + a_1^2|\check{g}|^2_{L^2(0,T)} + a_2^2|\check{\gamma}_0|^2_{W^{1,2}(0,T)} + a_3^2|\check{\gamma}_1|^2_{W^{1,2}(0,T)}\right)
\end{aligned} \tag{5.6}$$

*for any $(\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}$.*

*Proof.* Note from (A2) that

$$|\overline{\alpha}_\lambda(t,x)| \le \nu_1, \quad |\overline{\overline{\alpha}}_\lambda(t,x)| \le \nu_1, \quad \forall (t,x) \in Q \tag{5.7}$$

for some positive constant $\nu_1 > 0$ independent of $\lambda \in [-1,1]\setminus\{0\}$.

In addition, from the assumption (A5) for $\widehat{K}^{\varepsilon}$, we observe that (cf. (2.4)):

$$|\overline{K}^{\varepsilon}_{ij,\lambda}(t,x)| \le \frac{1}{\varepsilon}, \quad \text{a.a. } (t,x) \in Q,\ (i,j=1,2). \tag{5.8}$$

Here, from the boundedness (2.10) of solutions to $(\mathrm{SMA})^{\varepsilon}$, we note that

$$\begin{aligned}
&\sup_{t\in[0,T]} |\theta_\lambda(t)|_V^2 + \sum_{i=1}^2 \sup_{t\in[0,T]} |\chi_{i,\lambda}(t)|_V^2 + \int_0^T |(\theta_\lambda)_t(\tau)|_H^2 d\tau + \sum_{i=1}^2 \int_0^T |(\chi_{i,\lambda})_t(\tau)|_H^2 d\tau \\
\le& \nu_2\big(|\theta_0|_V^2 + |\chi_{1,0}|_V^2 + |\chi_{2,0}|_V^2 + a_2^2|\gamma_0(0)|^2 + a_2^2|\check{\gamma}_0(0)|^2 + a_3^2|\gamma_1(0)|^2 + a_3^2|\check{\gamma}_1(0)|^2 \\
&\quad + a_0^2|f|^2_{L^2(0,T;H)} + a_0^2|\check{f}|^2_{L^2(0,T;H)} + a_1^2|g|^2_{L^2(0,T)} + a_1^2|\check{g}|^2_{L^2(0,T)} \\
&\quad + a_2^2|\gamma_0|^2_{W^{1,2}(0,T)} + a_2^2|\check{\gamma}_0|^2_{W^{1,2}(0,T)} + a_3^2|\gamma_1|^2_{W^{1,2}(0,T)} + a_3^2|\check{\gamma}_1|^2_{W^{1,2}(0,T)} + 1\big)
\end{aligned} \tag{5.9}$$

where $\nu_2 > 0$ is a positive constant, independent of $(f, g, \gamma_0, \gamma_1) \in \mathcal{U}$, $(\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}$, and $\lambda \in [-1, 1] \setminus \{0\}$.

Since the embedding $V \hookrightarrow L^\infty(\Omega)$ is compact, we infer from (5.9) that

$$\begin{aligned}
&\sum_{i=1}^{2} \sup_{t\in[0,T]} |\chi_{i,\lambda}(t)|^2_{L^\infty(\Omega)} \\
\le&\nu_2' \big( |\theta_0|^2_V + |\chi_{1,0}|^2_V + |\chi_{2,0}|^2_V + a_2^2|\gamma_0(0)|^2 + a_2^2|\check{\gamma}_0(0)|^2 + a_3^2|\gamma_1(0)|^2 + a_3^2|\check{\gamma}_1(0)|^2 \\
&\quad + a_0^2|f|^2_{L^2(0,T;H)} + a_0^2|\check{f}|^2_{L^2(0,T;H)} + a_1^2|g|^2_{L^2(0,T)} + a_1^2|\check{g}|^2_{L^2(0,T)} \\
&\quad + a_2^2|\gamma_0|^2_{W^{1,2}(0,T)} + a_2^2|\check{\gamma}_0|^2_{W^{1,2}(0,T)} + a_3^2|\gamma_1|^2_{W^{1,2}(0,T)} + a_3^2|\check{\gamma}_1|^2_{W^{1,2}(0,T)} + 1 \big)
\end{aligned} \tag{5.10}$$

where $\nu_2' > 0$ is a positive constant, independent of $(f, g, \gamma_0, \gamma_1) \in \mathcal{U}$, $(\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}$, and $\lambda \in [-1, 1] \setminus \{0\}$.

On account of (5.7), (5.8), and (5.10), the uniform estimate (5.6) can be shown in a similar manner to the proof of (2.10) (cf. (2.3)). Therefore, we omit the detailed calculations. Thus, the proof of Lemma 5.1 is complete. □

We now state the result of the Gâteaux differentiability of $\Lambda^\varepsilon$.

**Proposition 5.1.** *Suppose that all the assumptions of Theorem 2.4 are satisfied. Then, the control-to-state mapping $\Lambda^\varepsilon$ admits the Gâteaux derivative at any $(f, g, \gamma_0, \gamma_1) \in \mathcal{U}$. More precisely, for arbitrary $(f, g, \gamma_0, \gamma_1) \in \mathcal{U}$, there exists a triplet of functions $(u, w_1, w_2) \in (W^{1,2}(0, T; H) \cap L^\infty(0, T; V))^3$ such that:*

$$\begin{aligned}
D_{(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1)}\Lambda^\varepsilon(f, g, \gamma_0, \gamma_1) &:= \lim_{\lambda\to 0} \frac{\Lambda^\varepsilon(f + \lambda\check{f}, g + \lambda\check{g}, \gamma_0 + \lambda\check{\gamma}_0, \gamma_1 + \lambda\check{\gamma}_1) - \Lambda^\varepsilon(f, g, \gamma_0, \gamma_1)}{\lambda} \\
&= (u, w_1, w_2) \quad \text{for any direction } (\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U},
\end{aligned} \tag{5.11}$$

*and $(u, w_1, w_2)$ solves the following linear system:*

$$(L_0 u - L_1 w_1)_t - h u_{xx} = a_0 \check{f}(t, x) \ \text{ in } Q, \tag{5.12}$$

$$\begin{aligned}
&\mu_0 \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}_t - \mu_1 \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}_{xx} + \begin{pmatrix} \partial_1^2 \widehat{K}^\varepsilon(\chi_1, \chi_2) w_1 + \partial_2\partial_1 \widehat{K}^\varepsilon(\chi_1, \chi_2) w_2 \\ \partial_1\partial_2 \widehat{K}^\varepsilon(\chi_1, \chi_2) w_1 + \partial_2^2 \widehat{K}^\varepsilon(\chi_1, \chi_2) w_2 \end{pmatrix} \\
=& \begin{pmatrix} -lu \\ -\beta a_1 \check{g}(t)\alpha(\theta) - \beta a_1 g(t)\alpha'(\theta)u + 2\beta\alpha'(\theta)\alpha(\theta)u\chi_2 + \beta\alpha(\theta)^2 w_2 \end{pmatrix} \ \text{ in } Q,
\end{aligned} \tag{5.13}$$

$$-hu_x(t, 0) + k(u(t, 0) - a_2\check{\gamma}_0(t)) = hu_x(t, 1) + k(u(t, 1) - a_3\check{\gamma}_1(t)) = 0 \ \text{ for } t \in (0, T), \tag{5.14}$$

$$(w_i)_x(t, 0) = (w_i)_x(t, 1) = 0, \quad t \in (0, T),\ i = 1, 2, \tag{5.15}$$

$$u(0, x) = 0, \quad w_i(0, x) = 0, \quad x \in \Omega,\ i = 1, 2. \tag{5.16}$$

*Proof.* Let $(f, g, \gamma_0, \gamma_1)$ be any function in $\mathcal{U}$. For all directions $(\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}$ and small numbers $\lambda \in [-1, 1] \setminus \{0\}$, we put $(\theta_\lambda, \chi_{1,\lambda}, \chi_{2,\lambda}) := \Lambda^\varepsilon(f + \lambda\check{f}, g + \lambda\check{g}, \gamma_0 + \lambda\check{\gamma}_0, \gamma_1 + \lambda\check{\gamma}_1)$, $(\theta, \chi_1, \chi_2) := \Lambda^\varepsilon(f, g, \gamma_0, \gamma_1)$, $u_\lambda := (\theta_\lambda - \theta)/\lambda$, $w_{1,\lambda} = (\chi_{1,\lambda} - \chi_1)/\lambda$, $w_{2,\lambda} = (\chi_{2,\lambda} - \chi_2)/\lambda$.

Then, by the uniform estimate (5.6) for $(u_\lambda, w_{1,\lambda}, w_{2,\lambda})$, it turns out that there are a subsequence $\{\lambda_n\}_{n\in\mathbb{N}} \subset \{\lambda\}_{\lambda\in[-1,1]\setminus\{0\}}$ and the triplet $(u, w_1, w_2) \in (W^{1,2}(0, T; H) \cap L^\infty(0, T; V))^3$ such that $\lambda_n \to 0$,

$$
(u_{\lambda_n}, w_{1,\lambda_n}, w_{2,\lambda_n}) \to (u, w_1, w_2) \quad \left.\begin{array}{l} \text{in } (C([0, T]; H))^3, \\ \text{weakly in } (W^{1,2}(0, T; H))^3, \\ \text{weakly-}* \text{ in } (L^\infty(0, T; V))^3, \end{array}\right\} \tag{5.17}
$$

as $n \to \infty$, and by the lower semicontinuity of the norm:

$$
\begin{aligned}
&\sup_{0\le t\le T} |u(t)|_V^2 + \sum_{i=1}^{2} \sup_{t\in[0,T]} |w_i(t)|_V^2 + \int_0^T |u_t(t)|_H^2 dt + \sum_{i=1}^{2} \int_0^T |(w_i)_t(t)|_H^2 dt \\
\le\ & N_3 \left( a_0^2 |\check{f}|^2_{L^2(0,T;H)} + a_1^2 |\check{g}|^2_{L^2(0,T)} + a_2^2 |\check{\gamma}_0|^2_{W^{1,2}(0,T)} + a_3^2 |\check{\gamma}_1|^2_{W^{1,2}(0,T)} \right),
\end{aligned} \tag{5.18}
$$

where $N_3$ is the same constant as in Lemma 5.1.

We now prove that the limit triplet $(u, w_1, w_2)$ of $(u_{\lambda_n}, w_{1,\lambda_n}, w_{2,\lambda_n})$ satisfies (5.12)–(5.16) in the variational sense. To this end, note from (5.6) that

$$
\begin{aligned}
&|\theta_\lambda - \theta|_{L^2(0,T;H)} + \sum_{i=1}^{2} |\chi_{i,\lambda} - \chi_i|_{L^2(0,T;H)} \\
&= \lambda |u_\lambda|_{L^2(0,T;H)} + \lambda \sum_{i=1}^{2} |w_{i,\lambda}|_{L^2(0,T;H)} \\
&\le \lambda\sqrt{T} |u_\lambda|_{L^\infty(0,T;H)} + \lambda\sqrt{T} \sum_{i=1}^{2} |w_{i,\lambda}|_{L^\infty(0,T;H)} \\
&\le 3\lambda\sqrt{TN_3} \left( a_0^2 |\check{f}|^2_{L^2(0,T;H)} + a_1^2 |\check{g}|^2_{L^2(0,T)} + a_2^2 |\check{\gamma}_0|^2_{W^{1,2}(0,T)} + a_3^2 |\check{\gamma}_1|^2_{W^{1,2}(0,T)} \right)^{\frac{1}{2}} \\
&\to 0 \ \text{ as } \lambda \to 0.
\end{aligned} \tag{5.19}
$$

Taking a subsequence if necessary, we observe from (5.19), (A2), (A5), and the continuity of functions $\partial_i\partial_j\widehat{K}^\varepsilon(\cdot, \cdot)$ $(i, j = 1, 2)$, $\alpha'(\cdot)$, and $\alpha(\cdot)$ that

$$
\begin{cases}
\theta_{\lambda_n}(t, x) \to \theta(t, x), \\
\chi_{i,\lambda_n}(t, x) \to \chi_i(t, x), \ (i = 1, 2), \\
\overline{K}^\varepsilon_{ij,\lambda_n}(t, x) \to \partial_j\partial_i\widehat{K}^\varepsilon(\chi_1(t, x), \chi_2(t, x)), \ (i, j = 1, 2), \\
\overline{\alpha}_{\lambda_n}(t, x) \to \alpha'(\theta(t, x)), \\
\overline{\overline{\alpha}}_{\lambda_n}(t, x) \to 2\alpha'(\theta(t, x))\alpha(\theta(t, x)),
\end{cases}
$$

for a.a. $(t, x) \in Q$ in the pointwise senses, as $n \to \infty$.

Here, let us fix arbitrary $0 \le t_0 < t_1 \le T$. Since functions $\overline{K}^\varepsilon_{ij,\lambda}$, $\overline{\alpha}_\lambda$, and $\overline{\overline{\alpha}}_\lambda$ $(\lambda \in [-1, 1] \setminus \{0\})$ are respectively bounded in senses of (5.8) and (5.7), we can apply Lebesgue's dominated convergence theorem to show that

$$
\begin{cases}
\overline{K}^\varepsilon_{ij,\lambda_n} \to \partial_j\partial_i\widehat{K}^\varepsilon(\chi_1, \chi_2), \ (i, j = 1, 2), \\
\overline{\alpha}_{\lambda_n} \to \alpha'(\theta), \\
\overline{\overline{\alpha}}_{\lambda_n} \to 2\alpha'(\theta)\alpha(\theta), \\
\alpha(\theta_{\lambda_n}) \to \alpha(\theta),
\end{cases}
\quad \text{in } L^2(t_0, t_1; H), \ \text{as } n \to \infty. \tag{5.20}
$$

Combining (5.6), (5.10), (5.17), (5.19), and (5.20), and by (A2), (A5), and the compact embeddings $V \hookrightarrow L^\infty(\Omega)$ and $H^1(0,T) \hookrightarrow C([0,T])$, it is deduced that:

$$u_{\lambda_n} \to u \text{ weakly in } L^2(t_0,t_1;V), \tag{5.21}$$

$$(u_{\lambda_n})_t \to u_t \text{ weakly in } L^2(t_0,t_1;H), \tag{5.22}$$

$$w_{i,\lambda_n} \to w_i \text{ weakly in } L^2(t_0,t_1;V), \quad (i=1,2), \tag{5.23}$$

$$(w_{i,\lambda_n})_t \to (w_i)_t \text{ weakly in } L^2(t_0,t_1;H), \quad (i=1,2), \tag{5.24}$$

$$\overline{K}^\varepsilon_{ij,\lambda_n} w_{j,\lambda_n} \to \partial_j\partial_i \widehat{K}^\varepsilon(\chi_1,\chi_2) w_j \text{ in } L^2(t_0,t_1;H), \quad (i,j=1,2), \tag{5.25}$$

$$\check{g}\alpha(\theta_{\lambda_n}) \to \check{g}\alpha(\theta) \text{ in } L^2(t_0,t_1;H), \tag{5.26}$$

$$g\overline{\alpha}_{\lambda_n} u_{\lambda_n} \to g\alpha'(\theta)u \text{ in } L^2(t_0,t_1;H), \tag{5.27}$$

$$\overline{\overline{\alpha}}_{\lambda_n} u_{\lambda_n} \chi_{2,\lambda_n} \to 2\alpha'(\theta)\alpha(\theta)u\chi_2 \text{ in } L^2(t_0,t_1;H), \tag{5.28}$$

and

$$\alpha(\theta)^2 w_{2,\lambda_n} \to \alpha(\theta)^2 w_2 \text{ in } L^2(t_0,t_1;H), \tag{5.29}$$

as $n \to \infty$, $(i,j=1,2)$.

Note from (5.1)–(5.5) that $(u_{\lambda_n}, w_{1,\lambda_n}, w_{2,\lambda_n})$ satisfies the following variational identities:

$$\begin{gathered}\int_{t_0}^{t_1} L_0((u_{\lambda_n})_t(t), z)_H dt - L_1 \int_{t_0}^{t_1} ((w_{1,\lambda_n})_t(t), z)_H dt + h \int_{t_0}^{t_1} (u_{\lambda_n}(t), z)_V dt \\ = \int_{t_0}^{t_1} (a_0 \check{f}(t), z)_H dt + k \int_{t_0}^{t_1} a_2 \check{\gamma}_0(t) z(0) dt + k \int_{t_0}^{t_1} a_3 \check{\gamma}_1(t) z(1) dt \\ \text{for all } z \in V \text{ and } n = 1,2,3,\cdots \end{gathered} \tag{5.30}$$

and

$$\begin{aligned} &\mu_0 \int_{t_0}^{t_1} \sum_{i=1}^2 ((w_{i,\lambda_n})_t(t), z_i)_H dt + \mu_1 \int_{t_0}^{t_1} \sum_{i=1}^2 ((w_{i,\lambda_n})_x(t), (z_i)_x)_H dt \\ &\quad + \int_{t_0}^{t_1} \sum_{i=1}^2 (\overline{K}^\varepsilon_{i1,\lambda}(t,x) w_{1,\lambda}(t) + \overline{K}^\varepsilon_{i2,\lambda}(t,x) w_{2,\lambda}(t), z_i)_H dt \\ &= \int_{t_0}^{t_1} (-l u_{\lambda_n}(t), z_1)_H dt \\ &\quad + \int_{t_0}^{t_1} (-\beta a_1 \check{g}(t) \alpha(\theta_{\lambda_n}(t)) - \beta a_1 g(t) \overline{\alpha}_{\lambda_n}(t,x) u_{\lambda_n}(t), z_2)_H dt \\ &\quad + \int_{t_0}^{t_1} (\beta \overline{\overline{\alpha}}_{\lambda_n}(t,x) u_{\lambda_n}(t) \chi_{2,\lambda_n}(t) + \beta \alpha(\theta(t))^2 w_{2,\lambda_n}(t), z_2)_H dt \\ &\qquad \text{for all } (z_1,z_2) \in V \times V \text{ and } n = 1,2,3,\cdots. \end{aligned} \tag{5.31}$$

On account of (5.17) and (5.21)–(5.29), and taking the limits in (5.30) and (5.31) as $n \to \infty$, we observe that the limit triplet $(u, w_1, w_2)$ of $(u_{\lambda_n}, w_{1,\lambda_n}, w_{2,\lambda_n})$ satisfies (5.12)–(5.15) in the variational sense.

In addition, it follows from (5.5) and (5.17) that the initial conditions (5.16) hold:

$$u(0,\cdot) = \lim_{n\to\infty} u_{\lambda_n}(0,\cdot) = 0 \text{ in } H,$$
$$w_i(0,\cdot) = \lim_{n\to\infty} w_{i,\lambda_n}(0,\cdot) = 0 \text{ in } H,\ (i = 1, 2).$$

Note that the Hessian matrix of $\widehat{K}^\varepsilon$ is positive semi-definite (cf. (6.11) below), since $\widehat{K}^\varepsilon$ is the convex function on $\mathbb{R}^2$ (cf. (A5)). Therefore, by the usual method with helps from the fact that $g \in C([0,T])$ (cf. (1.8)), $\alpha \in W^{2,\infty}(\mathbb{R})$ (cf. (A2)), and $\chi_2 \in L^\infty(Q)$ (cf. (5.10)), more precisely, by argument similar to [22, Theorem 2.1], we can prove that the solutions to the Cauchy problem (5.12)–(5.16) are uniquely determined. Hence, the uniqueness of solution to (5.12)–(5.16) guarantees that of cluster points of the sequence $(u_\lambda, w_{1,\lambda}, w_{2,\lambda})$ as $\lambda \to 0$:

$(*)$ $(u_\lambda, w_{1,\lambda}, w_{2,\lambda})$ originally converges to the unique solution $(u, w_1, w_2)$ to (5.12)–(5.16), in the variational sense (cf. (5.30) and (5.31)), as $\lambda \to 0$, and hence the operator $\mathcal{X}_{(f,g,\gamma_0,\gamma_1)} : \mathcal{U} \to (L^2(0,T;H))^3$, defined by $\mathcal{X}_{(f,g,\gamma_0,\gamma_1)}(\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) := D_{(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1)}\Lambda^\varepsilon(f, g, \gamma_0, \gamma_1)$ for all direction $(\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}$, is well-defined.

On account of the linearity inherent in (5.12)–(5.16), and the estimate (5.18), we observe that each operator $\mathcal{X}_{(f,g,\gamma_0,\gamma_1)}$ ($(f, g, \gamma_0, \gamma_1) \in \mathcal{U}$) is a bounded and linear operator from $\mathcal{U}$ into $(L^2(0,T;H))^3$, and hence, the control-to-state mapping $\Lambda^\varepsilon$ admits the Gâteaux derivative at any $(\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}$. Thus, the proof of Proposition 5.1 is complete. □

We now state the Gâteaux differentiability of the cost function $J_\delta^\varepsilon$, which is a direct consequence of Proposition 5.1.

**Corollary 5.1.** *Suppose that all the assumptions of Theorem 2.4 are satisfied. Then, the cost function $J_\delta^\varepsilon$ admits the Gâteaux derivative at any $(f, g, \gamma_0, \gamma_1) \in \mathcal{U}$. More precisely,*

$$\begin{aligned}
&D_{(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1)} J_\delta^\varepsilon(f, g, \gamma_0, \gamma_1) \\
:=& \lim_{\lambda\to 0} \frac{J_\delta^\varepsilon(f + \lambda\check{f}, g + \lambda\check{g}, \gamma_0 + \lambda\check{\gamma}_0, \gamma_1 + \lambda\check{\gamma}_1) - J_\delta^\varepsilon(f, g, \gamma_0, \gamma_1)}{\lambda} \\
=& c_0 \int_0^T ((\theta - \theta_d)(t), u(t))_H dt + c_1 \int_0^T ((\chi_1 - \chi_{1,d})(t), w_1(t))_H dt \\
&+ c_2 \int_0^T ((\chi_2 - \chi_{2,d})(t), w_2(t))_H dt \\
&+ m_0 a_0^2 \int_0^T (f(t), \check{f}(t))_H dt + m_1 a_1^2 \int_0^T g(t)\check{g}(t) dt \\
&+ m_2 a_2^2 \int_0^T \gamma_0(t)\check{\gamma}_0(t) dt + m_3 a_3^2 \int_0^T \gamma_1(t)\check{\gamma}_1(t) dt \qquad (5.32) \\
&+ \delta \int_0^T ((f - f^*)(t), \check{f}(t))_H dt + \delta \int_0^T (g - g^*)(t)\check{g}(t) dt \\
&+ \delta \int_0^T (\gamma_0 - \gamma_0^*)(t)\check{\gamma}_0(t) dt + \delta \int_0^T (\gamma_1 - \gamma_1^*)(t)\check{\gamma}_1(t) dt
\end{aligned}$$

*for any $(f,g,\gamma_0,\gamma_1) \in \mathcal{U}$ and any direction $(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1) \in \mathcal{U}$, where in the above formula $(\theta,\chi_1,\chi_2) = \Lambda^\varepsilon(f,g,\gamma_0,\gamma_1)$ denotes the unique solution to the approximating state system (SMA;$\theta_0,\chi_{1,0},\chi_{2,0},f,g,\gamma_0,\gamma_1)^\varepsilon$ on $[0,T]$, $\theta_d \in L^2(0,T;H)$, $\chi_{1,d} \in L^2(0,T)$, and $\chi_{2,d} \in L^2(0,T)$ are the given desired target profiles, and $(u,w_1,w_2)(= D_{(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1)}\Lambda^\varepsilon(f,g,\gamma_0,\gamma_1))$ is the triplet of functions obtained in Proposition 5.1.*

*Proof.* By virtue of Proposition 5.1 and (5.19), we can prove the Gâteaux differentiability of the cost function $J^\varepsilon_\delta$. Indeed, let $(u,w_1,w_2) := D_{(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1)}\Lambda^\varepsilon(f,g,\gamma_0,\gamma_1)$. Then, we have:

$$
\begin{aligned}
&D_{(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1)}J^\varepsilon_\delta(f,g,\gamma_0,\gamma_1)\\
&:= \lim_{\lambda\to 0}\frac{J^\varepsilon_\delta(f+\lambda\check{f},g+\lambda\check{g},\gamma_0+\lambda\check{\gamma}_0,\gamma_1+\lambda\check{\gamma}_1) - J^\varepsilon_\delta(f,g,\gamma_0,\gamma_1)}{\lambda}\\
&= \lim_{\lambda\to 0}\left\{\frac{c_0}{2}\int_0^T((\theta_\lambda+\theta-2\theta_d)(t),u_\lambda(t))_H dt\right.\\
&\qquad + \frac{c_1}{2}\int_0^T((\chi_{1,\lambda}+\chi_1-2\chi_{1,d})(t),w_{1,\lambda}(t))_H dt\\
&\qquad + \frac{c_2}{2}\int_0^T((\chi_{2,\lambda}+\chi_2-2\chi_{2,d})(t),w_{2,\lambda}(t))_H dt\\
&\qquad + \frac{m_0a_0^2}{2}\int_0^T((2f+\lambda\check{f})(t),\check{f}(t))_H dt + \frac{m_1a_1^2}{2}\int_0^T(2g+\lambda\check{g})(t)\check{g}(t)dt\\
&\quad + \frac{m_2a_2^2}{2}\int_0^T(2\gamma_0+\lambda\check{\gamma}_0)(t)\check{\gamma}_0(t)dt + \frac{m_3a_3^2}{2}\int_0^T(2\gamma_1+\lambda\check{\gamma}_1)(t)\check{\gamma}_1(t)dt\\
&\quad + \frac{\delta}{2}\int_0^T((2(f-f^*)+\lambda\check{f})(t),\check{f}(t))_H dt + \frac{\delta}{2}\int_0^T(2(g-g^*)+\lambda\check{g})(t)\check{g}(t)dt\\
&\quad \left. + \frac{\delta}{2}\int_0^T(2(\gamma_0-\gamma_0^*)+\lambda\check{\gamma}_0)(t)\check{\gamma}_0(t)dt + \frac{\delta}{2}\int_0^T(2(\gamma_1-\gamma_1^*)+\lambda\check{\gamma}_1)(t)\check{\gamma}_1(t)dt\right\}\\
&= c_0\int_0^T((\theta-\theta_d)(t),u(t))_H dt + c_1\int_0^T((\chi_1-\chi_{1,d})(t),w_1(t))_H dt\\
&\quad + c_2\int_0^T((\chi_2-\chi_{2,d})(t),w_2(t))_H dt\\
&\quad + m_0a_0^2\int_0^T(f(t),\check{f}(t))_H dt + m_1a_1^2\int_0^T g(t)\check{g}(t)dt\\
&\quad + m_2a_2^2\int_0^T\gamma_0(t)\check{\gamma}_0(t)dt + m_3a_3^2\int_0^T\gamma_1(t)\check{\gamma}_1(t)dt\\
&\quad + \delta\int_0^T((f-f^*)(t),\check{f}(t))_H dt + \delta\int_0^T(g-g^*)(t)\check{g}(t)dt\\
&\quad + \delta\int_0^T(\gamma_0-\gamma_0^*)(t)\check{\gamma}_0(t)dt + \delta\int_0^T(\gamma_1-\gamma_1^*)(t)\check{\gamma}_1(t)dt
\end{aligned}
$$

for any $(f,g,\gamma_0,\gamma_1) \in \mathcal{U}$ and any direction $(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1) \in \mathcal{U}$.

From Proposition 5.1 and (5.18), it follows that for any $(f,g,\gamma_0,\gamma_1) \in \mathcal{U}$, the functional:

$$(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1) \in \mathcal{U} \mapsto D_{(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1)}J^\varepsilon_\delta(f,g,\gamma_0,\gamma_1)$$

will form a bounded linear functional on $\mathcal{U}$. Hence, the cost functional $J_\delta^\varepsilon$ admits the Gâteaux derivative at any $(f, g, \gamma_0, \gamma_1) \in \mathcal{U}$ with the directional derivative as in (5.32). Thus, the proof of Corollary 5.1 is complete. □

On account of Proposition 5.1 and Corollary 5.1, we can prove Theorem 2.4 concerning the necessary condition of an optimal pair

$$(\theta^{*,\varepsilon}, \chi_1^{*,\varepsilon}, \chi_2^{*,\varepsilon}, f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) = (\Lambda^\varepsilon(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}), f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})$$

for the approximating problem (OP)$_\delta^\varepsilon$.

*Proof of Theorem 2.4.* Let

$$(\theta^{*,\varepsilon}, \chi_1^{*,\varepsilon}, \chi_2^{*,\varepsilon}, f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) = (\Lambda^\varepsilon(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}), f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})$$

be the optimal pair for (OP)$_\delta^\varepsilon$. Then, note from (2.10) and the compact embedding $V \hookrightarrow L^\infty(\Omega)$ that $\chi_2^{*,\varepsilon} \in L^\infty(Q)$ (cf. (5.10)).

On account of (2.4), $g^{*,\varepsilon} \in C([0,T])$ (cf. (1.8) and the compact embedding $H^1(0,T) \hookrightarrow C([0,T])$), $\alpha \in W^{2,\infty}(0,T)$ (cf. (A2)), and $\chi_2^{*,\varepsilon} \in L^\infty(Q)$, we can construct the unique solution to the adjoint equations (2.20)–(2.26) by applying Schauder's fixed point theorem and the theory of abstract nonlinear evolution equations (cf. [32, 34, 39, 44]). For such arguments, we refer to [34, Theorem 2.1], for instance. Thus, we omit the detailed proof of the existence-uniqueness of the solutions $(p^\varepsilon, q_1^\varepsilon, q_2^\varepsilon)$ to the adjoint equations (2.20)–(2.26).

We now show the necessary condition (2.27) for (OP)$_\delta^\varepsilon$. To this end, from the convexity of $\mathcal{U}_{ad}^M$, note that

$$\begin{aligned}
&(f, g, \gamma_0, \gamma_1) + \lambda(\check{f} - f, \check{g} - g, \check{\gamma}_0 - \gamma_0, \check{\gamma}_1 - \gamma_1) \\
&= (1-\lambda)(f, g, \gamma_0, \gamma_1) + \lambda(\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \\
&\in \mathcal{U}_{ad}^M, \quad \forall \lambda \in [0,1],\ \forall (f, g, \gamma_0, \gamma_1) \in \mathcal{U}_{ad}^M,\ \forall (\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}_{ad}^M.
\end{aligned} \tag{5.33}$$

In addition, from Corollary 5.1 it follows that $J_\delta^\varepsilon$ is Gâteaux differentiable at any quadruplet $(f, g, \gamma_0, \gamma_1) \in \mathcal{U}_{ad}^M$.

Furthermore, we observe from Proposition 5.1 that $\Lambda^\varepsilon$ is Gâteaux differentiable at $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \in \mathcal{U}_{ad}^M$. Therefore, we here put:

$$(u^{*,\varepsilon}, w_1^{*,\varepsilon}, w_2^{*,\varepsilon}) := D_{(\check{f} - f^{*,\varepsilon}, \check{g} - g^{*,\varepsilon}, \check{\gamma}_0 - \gamma_0^{*,\varepsilon}, \check{\gamma}_1 - \gamma_1^{*,\varepsilon})} \Lambda^\varepsilon(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}).$$

Since $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})$ is a minimizer for $J_\delta^\varepsilon(\cdot,\cdot,\cdot)$, we infer from (5.33) that

$$\begin{aligned}
&J_\delta^\varepsilon(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \\
\le & J_\delta^\varepsilon(f^{*,\varepsilon} + \lambda(\check{f} - f^{*,\varepsilon}), g^{*,\varepsilon} + \lambda(\check{g} - g^{*,\varepsilon}), \gamma_0^{*,\varepsilon} + \lambda(\check{\gamma}_0 - \gamma_0^{*,\varepsilon}), \gamma_1^{*,\varepsilon} + \lambda(\check{\gamma}_1 - \gamma_1^{*,\varepsilon})) \\
&\qquad \text{for all } \lambda \in [0,1] \text{ and all } (\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}_{ad}^M.
\end{aligned}$$

Therefore, we observe from Corollary 5.1, the adjoint system (2.21)–(2.26), and the linear system

(5.12)–(5.16) that

$$
\begin{aligned}
0 \leq & D_{(\check{f}-f^{*,\varepsilon},\check{g}-g^{*,\varepsilon},\check{\gamma}_0-\gamma_0^{*,\varepsilon},\check{\gamma}_1-\gamma_1^{*,\varepsilon})} J_\delta^\varepsilon(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \\
= & c_0 \int_0^T ((\theta^{*,\varepsilon} - \theta_d)(t), u^{*,\varepsilon}(t))_H dt + c_1 \int_0^T ((\chi_1^{*,\varepsilon} - \chi_{1,d})(t), w_1^{*,\varepsilon}(t))_H dt \\
& + c_2 \int_0^T ((\chi_2^{*,\varepsilon} - \chi_{2,d})(t), w_2^{*,\varepsilon}(t))_H dt \\
& + m_0 a_0^2 \int_0^T (f^{*,\varepsilon}(t), (\check{f} - f^{*,\varepsilon})(t))_H dt + m_1 a_1^2 \int_0^T g^{*,\varepsilon}(t)(\check{g} - g^{*,\varepsilon})(t) dt \\
& + m_2 a_2^2 \int_0^T \gamma_0^{*,\varepsilon}(t)(\check{\gamma}_0 - \gamma_0^{*,\varepsilon})(t) dt + m_3 a_3^2 \int_0^T \gamma_1^{*,\varepsilon}(t)(\check{\gamma}_1 - \gamma_1^{*,\varepsilon})(t) dt \\
& + \delta \int_0^T ((f^{*,\varepsilon} - f^*)(t), (\check{f} - f^{*,\varepsilon})(t))_H dt + \delta \int_0^T (g^{*,\varepsilon} - g^*)(t)(\check{g} - g^{*,\varepsilon})(t) dt \\
& + \delta \int_0^T (\gamma_0^{*,\varepsilon} - \gamma_0^*)(t)(\check{\gamma}_0 - \gamma_0^{*,\varepsilon})(t) dt + \delta \int_0^T (\gamma_1^{*,\varepsilon} - \gamma_1^*)(t)(\check{\gamma}_1 - \gamma_1^{*,\varepsilon})(t) dt \\
= & \int_0^T (-L_0 p_t^\varepsilon(t), u^{*,\varepsilon}(t))_H dt + \int_0^T h(p_x^\varepsilon(t), u_x^{*,\varepsilon}(t))_H dt \\
& + \int_0^T k p^\varepsilon(t,0) u^{*,\varepsilon}(t,0) dt + \int_0^T k p^\varepsilon(t,1) u^{*,\varepsilon}(t,1) dt \\
& + \int_0^T (l q_1^\varepsilon(t), u^{*,\varepsilon}(t))_H dt + \int_0^T (\beta a_1 g^{*,\varepsilon}(t) \alpha'(\theta^{*,\varepsilon}(t)) q_2^\varepsilon(t), u^{*,\varepsilon}(t))_H dt \\
& - \int_0^T (2\beta \alpha'(\theta^{*,\varepsilon}(t)) \alpha(\theta^{*,\varepsilon}(t)) \chi_2^{*,\varepsilon}(t) q_2^\varepsilon(t), u^{*,\varepsilon}(t))_H dt \\
& + \int_0^T (-\mu_0 (q_1^\varepsilon)_t(t), w_1^{*,\varepsilon}(t))_H dt + \int_0^T (L_1 p_t^\varepsilon(t), w_1^{*,\varepsilon}(t))_H dt \\
& + \int_0^T \mu_1((q_1^\varepsilon)_x(t), (w_1^{*,\varepsilon})_x(t))_H dt + \int_0^T (\partial_1^2 \widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t), \chi_2^{*,\varepsilon}(t)) q_1^\varepsilon(t), w_1^{*,\varepsilon}(t))_H dt \\
& + \int_0^T (\partial_1 \partial_2 \widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t), \chi_2^{*,\varepsilon}(t)) q_2^\varepsilon(t), w_1^{*,\varepsilon}(t))_H dt \\
& + \int_0^T (-\mu_0 (q_2^\varepsilon)_t(t), w_2^{*,\varepsilon}(t))_H dt + \int_0^T \mu_1((q_2^\varepsilon)_x(t), (w_2^{*,\varepsilon})_x(t))_H dt \\
& + \int_0^T (\partial_2 \partial_1 \widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t), \chi_2^{*,\varepsilon}(t)) q_1^\varepsilon(t), w_2^{*,\varepsilon}(t))_H dt \\
& + \int_0^T (\partial_2^2 \widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t), \chi_2^{*,\varepsilon}(t)) q_2^\varepsilon(t), w_2^{*,\varepsilon}(t))_H dt \\
& - \int_0^T (\beta \alpha(\theta^{*,\varepsilon}(t))^2 q_2^\varepsilon(t), w_2^{*,\varepsilon}(t))_H dt \\
& + m_0 a_0^2 \int_0^T (f^{*,\varepsilon}(t), (\check{f} - f^{*,\varepsilon})(t))_H dt + m_1 a_1^2 \int_0^T g^{*,\varepsilon}(t)(\check{g} - g^{*,\varepsilon})(t) dt \\
& + m_2 a_2^2 \int_0^T \gamma_0^{*,\varepsilon}(t)(\check{\gamma}_0 - \gamma_0^{*,\varepsilon})(t) dt + m_3 a_3^2 \int_0^T \gamma_1^{*,\varepsilon}(t)(\check{\gamma}_1 - \gamma_1^{*,\varepsilon})(t) dt \\
& + \delta \int_0^T ((f^{*,\varepsilon} - f^*)(t), (\check{f} - f^{*,\varepsilon})(t))_H dt + \delta \int_0^T (g^{*,\varepsilon} - g^*)(t)(\check{g} - g^{*,\varepsilon})(t) dt
\end{aligned}
$$

$$
\begin{aligned}
&+\delta\int_0^T(\gamma_0^{*,\varepsilon}-\gamma_0^*)(t)(\check\gamma_0-\gamma_0^{*,\varepsilon})(t)dt+\delta\int_0^T(\gamma_1^{*,\varepsilon}-\gamma_1^*)(t)(\check\gamma_1-\gamma_1^{*,\varepsilon})(t)dt\\
=&\int_0^T(L_0u_t^{*,\varepsilon}(t)-L_1(w_1^{*,\varepsilon})_t(t),p^\varepsilon(t))_Hdt+\int_0^T h(u_x^{*,\varepsilon}(t),p_x^\varepsilon(t))_Hdt\\
&+\int_0^T ku^{*,\varepsilon}(t,0)p^\varepsilon(t,0)dt+\int_0^T ku^{*,\varepsilon}(t,1)p^\varepsilon(t,1)dt\\
&+\int_0^T(\mu_0(w_1^{*,\varepsilon})_t(t),q_1^\varepsilon(t))_Hdt+\int_0^T\mu_1((w_1^{*,\varepsilon})_x(t),(q_1^\varepsilon)_x(t))_Hdt\\
&+\int_0^T(\partial_1^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t),\chi_2^{*,\varepsilon}(t))w_1^{*,\varepsilon}(t),q_1^\varepsilon(t))_Hdt\\
&+\int_0^T(\partial_2\partial_1\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t),\chi_2^{*,\varepsilon}(t))w_2^{*,\varepsilon}(t),q_1^\varepsilon(t))_Hdt\\
&+\int_0^T(lu^{*,\varepsilon}(t),q_1^\varepsilon(t))_Hdt\\
&+\int_0^T(\mu_0(w_2^{*,\varepsilon})_t(t),q_2^\varepsilon(t))_Hdt+\int_0^T\mu_1((w_2^{*,\varepsilon})_x(t),(q_2^\varepsilon)_x(t))_Hdt\\
&+\int_0^T(\partial_1\partial_2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t),\chi_2^{*,\varepsilon}(t))w_1^{*,\varepsilon}(t),q_2^\varepsilon(t))_Hdt\\
&+\int_0^T(\partial_2^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t),\chi_2^{*,\varepsilon}(t))w_2^{*,\varepsilon}(t),q_2^\varepsilon(t))_Hdt\\
&+\int_0^T(\beta a_1g^{*,\varepsilon}(t)\alpha'(\theta^{*,\varepsilon}(t))u^{*,\varepsilon}(t),q_2^\varepsilon(t))_Hdt\\
&-\int_0^T(2\beta\alpha'(\theta^{*,\varepsilon}(t))\alpha(\theta^{*,\varepsilon}(t))u^{*,\varepsilon}(t)\chi_2^{*,\varepsilon}(t),q_2^\varepsilon(t))_Hdt\\
&-\int_0^T(\beta\alpha(\theta^{*,\varepsilon}(t))^2w_2^{*,\varepsilon}(t),q_2^\varepsilon(t))_Hdt\\
&+m_0a_0^2\int_0^T(f^{*,\varepsilon}(t),(\check f-f^{*,\varepsilon})(t))_Hdt+m_1a_1^2\int_0^T g^{*,\varepsilon}(t)(\check g-g^{*,\varepsilon})(t)dt\\
&+m_2a_2^2\int_0^T\gamma_0^{*,\varepsilon}(t)(\check\gamma_0-\gamma_0^{*,\varepsilon})(t)dt+m_3a_3^2\int_0^T\gamma_1^{*,\varepsilon}(t)(\check\gamma_1-\gamma_1^{*,\varepsilon})(t)dt\\
&+\delta\int_0^T((f^{*,\varepsilon}-f^*)(t),(\check f-f^{*,\varepsilon})(t))_Hdt+\delta\int_0^T(g^{*,\varepsilon}-g^*)(t)(\check g-g^{*,\varepsilon})(t)dt\\
&+\delta\int_0^T(\gamma_0^{*,\varepsilon}-\gamma_0^*)(t)(\check\gamma_0-\gamma_0^{*,\varepsilon})(t)dt+\delta\int_0^T(\gamma_1^{*,\varepsilon}-\gamma_1^*)(t)(\check\gamma_1-\gamma_1^{*,\varepsilon})(t)dt\\
=&\int_0^T a_0(a_0m_0f^{*,\varepsilon}(t)+p^\varepsilon(t),(\check f-f^{*,\varepsilon})(t))_Hdt\\
&+\int_0^T a_1(a_1m_1g^{*,\varepsilon}(t)-(\beta\alpha(\theta^{*,\varepsilon}(t)),q_2^\varepsilon(t))_H)(\check g-g^{*,\varepsilon})(t)dt\\
&+\int_0^T a_2(a_2m_2\gamma_0^{*,\varepsilon}(t)+kp^\varepsilon(t,0))(\check\gamma_0-\gamma_0^{*,\varepsilon})(t)dt\\
&+\int_0^T a_3(a_3m_3\gamma_1^{*,\varepsilon}(t)+kp^\varepsilon(t,1))(\check\gamma_1-\gamma_1^{*,\varepsilon})(t)dt
\end{aligned}
$$

$$+\delta\int_0^T ((f^{*,\varepsilon}-f^*)(t),(\check{f}-f^{*,\varepsilon})(t))_H dt+\delta\int_0^T (g^{*,\varepsilon}-g^*)(t)(\check{g}-g^{*,\varepsilon})(t)dt$$
$$+\delta\int_0^T (\gamma_0^{*,\varepsilon}-\gamma_0^*)(t)(\check{\gamma}_0-\gamma_0^{*,\varepsilon})(t)dt+\delta\int_0^T (\gamma_1^{*,\varepsilon}-\gamma_1^*)(t)(\check{\gamma}_1-\gamma_1^{*,\varepsilon})(t)dt$$

for any $(\check{f},\check{g},\check{\gamma}_0,\check{\gamma}_1)\in\mathcal{U}_{ad}^M$. Thus, the proof of Theorem 2.4 is complete. □

# 6 Optimality condition to (OP)

In previous Section 5, we proved Theorem 2.4, which is concerned with the optimality condition for the approximating problem (OP)$_\delta^\varepsilon$. However, it is difficult to show the necessary condition of the optimal control for (OP) directly, because of the non-smooth constraint $\partial I_K(\cdot,\cdot)$ in (1.2). Therefore, through the limiting observation of approximating problems (OP)$_\delta^\varepsilon$, we derive the optimality condition for (OP).

We now prove the final main result (Theorem 2.5) of this paper, which is concerned with the necessary condition of the optimal control for (OP).

*Proof of Theorem 2.5.* On account of Theorem 2.3(II) and Theorem 2.4, we can prove Theorem 2.5.

Indeed, we assume $\delta>0$. Let $(f^*,g^*,\gamma_0^*,\gamma_1^*)\in\mathcal{U}_{ad}^M$ be any optimal control for (OP) obtained in Theorem 2.1. In addition, let $(f^{*,\varepsilon},g^{*,\varepsilon},\gamma_0^{*,\varepsilon},\gamma_1^{*,\varepsilon})$ be the optimal control for (OP)$_\delta^\varepsilon$ obtained in Theorem 2.2. Furthermore, let $(\theta^{*,\varepsilon},\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon})$ and $(\theta^*,\chi_1^*,\chi_2^*)$ be unique solutions to $(\mathrm{SMA};\theta_0,\chi_{1,0},\chi_{2,0},f^{*,\varepsilon},g^{*,\varepsilon},\gamma_0^{*,\varepsilon},\gamma_1^{*,\varepsilon})^\varepsilon$ and $(\mathrm{SMA};\theta_0,\chi_{1,0},\chi_{2,0},f^*,g^*,\gamma_0^*,\gamma_1^*)$ on $[0,T]$, respectively. Then, we observe from Theorem 2.3(II) that there is a subsequence of $\varepsilon$ (which we also denote $\varepsilon$ for simplicity) such that

$$f^{*,\varepsilon}\to f^* \text{ in } L^2(0,T;H), \tag{6.1}$$

$$g^{*,\varepsilon}\to g^* \text{ weakly in } H^1(0,T), \text{ and in } L^2(0,T), \tag{6.2}$$

$$\gamma_i^{*,\varepsilon}\to\gamma_i^* \text{ weakly in } H^1(0,T), \text{ and in } L^2(0,T),\ (i=0,1), \tag{6.3}$$

$$(\theta^{*,\varepsilon},\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon})\to(\theta^*,\chi_1^*,\chi_2^*) \text{ in } (C([0,T];H))^3 \tag{6.4}$$

as $\varepsilon\to 0$.

We now prove Theorem 2.5 by taking the limit with respect to $\varepsilon$. To this end, we give the a priori estimate of the solution $(p^\varepsilon,q_1^\varepsilon,q_2^\varepsilon)$ to the adjoint equations (2.21)–(2.26).

Note from (1.8), (2.10), and the embeddings $V\hookrightarrow L^\infty(\Omega)$ and $H^1(0,T)\hookrightarrow C([0,T])$ that

$$|\chi_2^{*,\varepsilon}|_{L^\infty(Q)}\le\nu_3,\quad |g^{*,\varepsilon}|_{C([0,T])}\le\nu_3,\quad \forall\varepsilon\in(0,1] \tag{6.5}$$

for some positive constant $\nu_3$ independent $\varepsilon\in(0,1]$.

Multiply (2.21) by $p^\varepsilon$. Then, by $\alpha\in W^{2,\infty}(\mathbb{R})$ (cf. (A2)), (6.5), and Young's inequality, we find positive constants $\nu_4$ and $\nu_5$, independent of $\varepsilon\in(0,1]$, such that

$$\begin{aligned}&-\frac{L_0}{2}\frac{d}{d\tau}|p^\varepsilon(\tau)|_H^2+h|p^\varepsilon(\tau)|_V^2\\ \le&\nu_4\left(|p^\varepsilon(\tau)|_H^2+|q_1^\varepsilon(\tau)|_H^2+|q_2^\varepsilon(\tau)|_H^2\right)+\nu_5|\theta^{*,\varepsilon}(\tau)-\theta_d(\tau)|_H^2\\ &\text{for a.a. } \tau\in(0,T).\end{aligned} \tag{6.6}$$

By integrating (6.6) in $\tau$ over $[T-t,T]$ $(t\in[0,T])$, we have

$$\begin{aligned}&\frac{L_0}{2}|p^\varepsilon(T-t)|_H^2+h\int_{T-t}^T|p^\varepsilon(\tau)|_V^2d\tau\\ \le&\nu_4\int_{T-t}^T\left(|p^\varepsilon(\tau)|_H^2+|q_1^\varepsilon(\tau)|_H^2+|q_2^\varepsilon(\tau)|_H^2\right)d\tau\\ &+\nu_5\int_{T-t}^T|\theta^{*,\varepsilon}(\tau)-\theta_d(\tau)|_H^2d\tau,\quad\forall t\in[0,T].\end{aligned}\tag{6.7}$$

Next, multiply (2.21) by $-p_t^\varepsilon$. Then, by $\alpha\in W^{2,\infty}(\mathbb{R})$ (cf. (A2)), (6.5), and Young's inequality, we find positive constants $\nu_6$ and $\nu_7$, independent of $\varepsilon\in(0,1]$, such that

$$\begin{aligned}&\frac{L_0}{2}|p_t^\varepsilon(\tau)|_H^2-\frac{h}{2}\frac{d}{d\tau}|p^\varepsilon(\tau)|_V^2\\ \le&\nu_6\left(|q_1^\varepsilon(\tau)|_H^2+|q_2^\varepsilon(\tau)|_H^2\right)+\nu_7|\theta^{*,\varepsilon}(\tau)-\theta_d(\tau)|_H^2\\ &\text{for a.a. }\tau\in(0,T).\end{aligned}\tag{6.8}$$

By integrating (6.8) in $\tau$ over $[T-t,T]$ $(t\in[0,T])$, we have

$$\begin{aligned}&\frac{L_0}{2}\int_{T-t}^T|p_t^\varepsilon(\tau)|_H^2d\tau+\frac{h}{2}|p^\varepsilon(T-t)|_V^2\\ \le&\nu_6\int_{T-t}^T\left(|q_1^\varepsilon(\tau)|_H^2+|q_2^\varepsilon(\tau)|_H^2\right)d\tau+\nu_7\int_{T-t}^T|\theta^{*,\varepsilon}(\tau)-\theta_d(\tau)|_H^2d\tau,\quad\forall t\in[0,T].\end{aligned}\tag{6.9}$$

Similarly, multiply (2.23) (resp. (2.24)) by $q_1^\varepsilon$ (resp. $q_2^\varepsilon$), and add the resultant to get:

$$\begin{aligned}&-\frac{\mu_0}{2}\sum_{i=1}^2\frac{d}{d\tau}|q_i^\varepsilon(\tau)|_H^2+\mu_1\sum_{i=1}^2|(q_i^\varepsilon)_x(\tau)|_H^2+(L_1p_t^\varepsilon(\tau),q_1^\varepsilon(\tau))_H\\ &+(\partial_1^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(\tau),\chi_2^{*,\varepsilon}(\tau))q_1^\varepsilon(\tau),q_1^\varepsilon(\tau))_H+(\partial_1\partial_2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(\tau),\chi_2^{*,\varepsilon}(\tau))q_2^\varepsilon(\tau),q_1^\varepsilon(\tau))_H\\ &+(\partial_2\partial_1\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(\tau),\chi_2^{*,\varepsilon}(\tau))q_1^\varepsilon(\tau),q_2^\varepsilon(\tau))_H+(\partial_2^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(\tau),\chi_2^{*,\varepsilon}(\tau))q_2^\varepsilon(\tau),q_2^\varepsilon(\tau))_H\\ &-(\beta\alpha(\theta^{*,\varepsilon}(\tau))^2q_2^\varepsilon(\tau),q_2^\varepsilon(\tau))_H\\ &=c_1(\chi_1^{*,\varepsilon}(\tau)-\chi_{1,d}(\tau),q_1^\varepsilon(\tau))_H+c_2(\chi_2^{*,\varepsilon}(\tau)-\chi_{2,d}(\tau),q_2^\varepsilon(\tau))_H\\ &\text{for a.a. }\tau\in(0,T).\end{aligned}\tag{6.10}$$

Since $\widehat{K}^\varepsilon$ is the convex function on $\mathbb{R}^2$ (cf. (A5)), note that the Hessian matrix of $\widehat{K}^\varepsilon$ is positive semi-definite, more precisely,

$$\begin{aligned}&(\partial_1^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(\tau),\chi_2^{*,\varepsilon}(\tau))q_1^\varepsilon(\tau),q_1^\varepsilon(\tau))_H+(\partial_1\partial_2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(\tau),\chi_2^{*,\varepsilon}(\tau))q_2^\varepsilon(\tau),q_1^\varepsilon(\tau))_H\\ &\quad+(\partial_2\partial_1\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(\tau),\chi_2^{*,\varepsilon}(\tau))q_1^\varepsilon(\tau),q_2^\varepsilon(\tau))_H+(\partial_2^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(\tau),\chi_2^{*,\varepsilon}(\tau))q_2^\varepsilon(\tau),q_2^\varepsilon(\tau))_H\\ =&\int_\Omega(q_1^\varepsilon(\tau,x),q_2^\varepsilon(\tau,x))\nabla^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(\tau,x),\chi_2^{*,\varepsilon}(\tau,x))\begin{pmatrix}q_1^\varepsilon(\tau,x)\\ q_2^\varepsilon(\tau,x)\end{pmatrix}dx\\ \ge&0,\end{aligned}\tag{6.11}$$

where $\nabla^2\widehat{K}^\varepsilon(\cdot,\cdot)$ is the Hessian matrix of $\widehat{K}^\varepsilon$ defined by

$$\nabla^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon}):=\begin{pmatrix}\partial_1^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon}) & \partial_1\partial_2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon})\\ \partial_2\partial_1\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon}) & \partial_2^2\widehat{K}^\varepsilon(\chi_1^{*,\varepsilon},\chi_2^{*,\varepsilon})\end{pmatrix}.$$

Therefore, by (6.10), (6.11), and Young's inequality, we find positive constants $\nu_8$ and $\nu_9$, independent of $\varepsilon \in (0,1]$, such that

$$\begin{aligned} & -\frac{\mu_0}{2}\sum_{i=1}^{2}\frac{d}{d\tau}|q_i^\varepsilon(\tau)|_H^2 + \mu_1\sum_{i=1}^{2}|(q_i^\varepsilon)_x(\tau)|_H^2 \\ \le & \frac{L_0}{4}|p_t^\varepsilon(\tau)|_H^2 + \nu_8\left(|q_1^\varepsilon(\tau)|_H^2 + |q_2^\varepsilon(\tau)|_H^2\right) \\ & + \nu_9\left(|\chi_1^{*,\varepsilon}(\tau) - \chi_{1,d}(\tau)|_H^2 + |\chi_2^{*,\varepsilon}(\tau) - \chi_{2,d}(\tau)|_H^2\right) \\ & \text{for a.a. } \tau \in (0,T). \end{aligned} \tag{6.12}$$

By integrating (6.12) in $\tau$ over $[T-t,T]$ $(t \in [0,T])$, we have

$$\begin{aligned} & \frac{\mu_0}{2}\sum_{i=1}^{2}|q_i^\varepsilon(T-t)|_H^2 + \mu_1\sum_{i=1}^{2}\int_{T-t}^{T}|(q_i^\varepsilon)_x(\tau)|_H^2 d\tau \\ \le & \frac{L_0}{4}\int_{T-t}^{T}|p_t^\varepsilon(\tau)|_H^2 d\tau + \nu_8\int_{T-t}^{T}\left(|q_1^\varepsilon(\tau)|_H^2 + |q_2^\varepsilon(\tau)|_H^2\right)d\tau \\ & + \nu_9\int_{T-t}^{T}\left(|\chi_1^{*,\varepsilon}(\tau) - \chi_{1,d}(\tau)|_H^2 + |\chi_2^{*,\varepsilon}(\tau) - \chi_{2,d}(\tau)|_H^2\right)d\tau, \quad \forall t \in [0,T]. \end{aligned} \tag{6.13}$$

Adding (6.7), (6.9), (6.13), we find positive constants $\nu_{10}$ and $\nu_{11}$, independent of $\varepsilon \in (0,1]$, such that

$$\begin{aligned} & \frac{L_0}{2}|p^\varepsilon(T-t)|_H^2 + h\int_{T-t}^{T}|p^\varepsilon(\tau)|_V^2 d\tau + \frac{L_0}{4}\int_{T-t}^{T}|p_t^\varepsilon(\tau)|_H^2 d\tau + \frac{h}{2}|p^\varepsilon(T-t)|_V^2 \\ & + \frac{\mu_0}{2}\sum_{i=1}^{2}|q_i^\varepsilon(T-t)|_H^2 + \mu_1\sum_{i=1}^{2}\int_{T-t}^{T}|(q_i^\varepsilon)_x(\tau)|_H^2 d\tau \\ \le & \nu_{10}\int_{T-t}^{T}\left(|p^\varepsilon(\tau)|_H^2 + |q_1^\varepsilon(\tau)|_H^2 + |q_2^\varepsilon(\tau)|_H^2\right)d\tau \\ & + \nu_{11}\int_{T-t}^{T}\left(|\theta^{*,\varepsilon}(\tau) - \theta_d(\tau)|_H^2 + |\chi_1^{*,\varepsilon}(\tau) - \chi_{1,d}(\tau)|_H^2 + |\chi_2^{*,\varepsilon}(\tau) - \chi_{2,d}(\tau)|_H^2\right)d\tau, \\ & \forall t \in [0,T]. \end{aligned} \tag{6.14}$$

Applying Gronwall-type inequality (e.g., [33, Proposition 0.4.1]) to (6.14), we infer from (6.4) that

$$\begin{aligned} & \int_0^T\left(|p^\varepsilon(\tau)|_H^2 + |q_1^\varepsilon(\tau)|_H^2 + |q_2^\varepsilon(\tau)|_H^2\right)d\tau \\ \le & \nu_{12}\left(|\theta^* - \theta_d|_{L^2(0,T;H)}^2 + |\chi_1^* - \chi_{1,d}|_{L^2(0,T;H)}^2 + |\chi_2^* - \chi_{2,d}|_{L^2(0,T;H)}^2 + 1\right) \end{aligned} \tag{6.15}$$

for some constant $\nu_{12} > 0$, independent of $\varepsilon \in (0,1]$, and dependent on $T$. Hence, we conclude from (6.14) and (6.15) that

$$\begin{aligned} & \sup_{t\in[0,T]}\left\{|p^\varepsilon(t)|_H^2 + |p^\varepsilon(t)|_V^2 + \sum_{i=1}^{2}|q_i^\varepsilon(t)|_H^2\right\} \\ & + \int_0^T|p^\varepsilon(t)|_V^2 dt + \int_0^T|p_t^\varepsilon(t)|_H^2 dt + \sum_{i=1}^{2}\int_0^T|(q_i^\varepsilon)_x(t)|_H^2 dt \\ \le & \nu_{13}\left(|\theta^* - \theta_d|_{L^2(0,T;H)}^2 + |\chi_1^* - \chi_{1,d}|_{L^2(0,T;H)}^2 + |\chi_2^* - \chi_{2,d}|_{L^2(0,T;H)}^2 + 1\right) \end{aligned} \tag{6.16}$$

for some constant $\nu_{13} > 0$, independent of $\varepsilon \in (0,1]$ and dependent on $T$.

For any $\varepsilon \in (0,1]$, let us now define a bounded and linear functional $\varsigma_1^\varepsilon \in W'$ on $W$, by putting:

$$\begin{aligned}\langle \varsigma_1^\varepsilon, \zeta\rangle_{W',W} := \int_0^T \Big\{&(\partial_1^2 \widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t), \chi_2^{*,\varepsilon}(t)) q_1^\varepsilon(t), \zeta(t))_H \\ &+(\partial_1\partial_2 \widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t), \chi_2^{*,\varepsilon}(t)) q_2^\varepsilon(t), \zeta(t))_H\Big\}\, dt, \quad \forall \zeta \in W.\end{aligned} \tag{6.17}$$

Similarly, we define a bounded and linear functional $\varsigma_2^\varepsilon \in W'$ on $W$ by:

$$\begin{aligned}\langle \varsigma_2^\varepsilon, \zeta\rangle_{W',W} := \int_0^T \Big\{&(\partial_2\partial_1 \widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t), \chi_2^{*,\varepsilon}(t)) q_1^\varepsilon(t), \zeta(t))_H \\ &+(\partial_2^2 \widehat{K}^\varepsilon(\chi_1^{*,\varepsilon}(t), \chi_2^{*,\varepsilon}(t)) q_2^\varepsilon(t), \zeta(t))_H\Big\}\, dt, \quad \forall \zeta \in W.\end{aligned} \tag{6.18}$$

On account of (2.23), (2.25), (6.4), (6.16), and (6.17), there exists a positive constant $\nu_{14}$, independent of $\varepsilon \in (0,1]$, such that:

$$\begin{aligned}|\langle \varsigma_1^\varepsilon, \zeta\rangle_{W',W}| \le\ & \left|\int_0^T (\mu_0 (q_1^\varepsilon)_t(t), \zeta(t))_H dt\right| + \left|\int_0^T (L_1 p_t^\varepsilon(t), \zeta(t))_H dt\right| \\ &+ \left|\int_0^T (\mu_1 (q_1^\varepsilon)_x(t), \zeta_x(t))_H\, dt\right| + \left|\int_0^T (c_1(\chi_1^{*,\varepsilon} - \chi_{1,d})(t), \zeta(t))_H\, dt\right| \\ =\ & \left|\int_0^T (-\mu_0 q_1^\varepsilon(t), \zeta_t(t))_H dt\right| + \left|\int_0^T (L_1 p_t^\varepsilon(t), \zeta(t))_H dt\right| \\ &+ \left|\int_0^T (\mu_1 (q_1^\varepsilon)_x(t), \zeta_x(t))_H\, dt\right| + \left|\int_0^T (c_1(\chi_1^{*,\varepsilon} - \chi_{1,d})(t), \zeta(t))_H\, dt\right| \\ \le\ & \nu_{14}\big(|\theta^* - \theta_d|_{L^2(0,T;H)} + |\chi_1^* - \chi_{1,d}|_{L^2(0,T;H)} + |\chi_2^* - \chi_{2,d}|_{L^2(0,T;H)} + 1\big)|\zeta|_W \\ & \text{for any } \zeta \in W = \{z \in H^1(Q)\,;\ z(0,x) = 0, \text{ a.e. } x \in \Omega\}.\end{aligned}$$

Therefore, we get

$$|\varsigma_1^\varepsilon|_{W'} \le \nu_{14}\left(|\theta^* - \theta_d|_{L^2(0,T;H)} + |\chi_1^* - \chi_{1,d}|_{L^2(0,T;H)} + |\chi_2^* - \chi_{2,d}|_{L^2(0,T;H)} + 1\right) \tag{6.19}$$

for all $\varepsilon \in (0,1]$.

In addition, from (2.24), (2.25), (6.4), (6.16), (6.18), and $\alpha \in W^{2,\infty}(\mathbb{R})$ (cf. (A2)), there exists a positive constant $\nu_{15}$, independent of $\varepsilon \in (0,1]$, such that:

$$\begin{aligned}|\langle \varsigma_2^\varepsilon, \zeta\rangle_{W',W}| \le & \left|\int_0^T (\mu_0 (q_2^\varepsilon)_t(t), \zeta(t))_H dt\right| + \left|\int_0^T (\mu_1 (q_2^\varepsilon)_x(t), \zeta_x(t))_H\, dt\right| \\ &+ \left|\int_0^T \big(\beta\alpha(\theta^{*,\varepsilon}(t))^2 q_2^\varepsilon(t), \zeta(t)\big)_H\, dt\right| + \left|\int_0^T (c_2(\chi_2^{*,\varepsilon} - \chi_{2,d})(t), \zeta(t))_H\, dt\right| \\ = & \left|\int_0^T (-\mu_0 q_2^\varepsilon(t), \zeta_t(t))_H dt\right| + \left|\int_0^T (\mu_1 (q_2^\varepsilon)_x(t), \zeta_x(t))_H\, dt\right| \\ &+ \left|\int_0^T \big(\beta\alpha(\theta^{*,\varepsilon}(t))^2 q_2^\varepsilon(t), \zeta(t)\big)_H\, dt\right| + \left|\int_0^T (c_2(\chi_2^{*,\varepsilon} - \chi_{2,d})(t), \zeta(t))_H\, dt\right| \\ \le & \nu_{15}\left(|\theta^* - \theta_d|_{L^2(0,T;H)} + |\chi_1^* - \chi_{1,d}|_{L^2(0,T;H)} + |\chi_2^* - \chi_{2,d}|_{L^2(0,T;H)} + 1\right)|\zeta|_W \\ & \text{for any } \zeta \in W.\end{aligned}$$

Therefore, we get

$$|\varsigma_2^\varepsilon|_{W'} \le \nu_{15}\left(|\theta^* - \theta_d|_{L^2(0,T;H)} + |\chi_1^* - \chi_{1,d}|_{L^2(0,T;H)} + |\chi_2^* - \chi_{2,d}|_{L^2(0,T;H)} + 1\right) \tag{6.20}$$

for all $\varepsilon \in (0,1]$.

By boundedness estimates (6.16), (6.19), and (6.20), there are a subsequence of $\varepsilon$ (which we also denote $\varepsilon$ for simplicity), the functions $p \in W^{1,2}(0,T;H) \cap L^\infty(0,T;V)$, $q_i \in L^2(0,T;V) \cap L^\infty(0,T;H)$ $(i=1,2)$, and elements $\varsigma_i \in W'$ $(i=1,2)$ such that

$$\left.\begin{array}{ll} p^\varepsilon \to p & \text{in } C([0,T];H), \\ & \text{weakly in } W^{1,2}(0,T;H), \\ & \text{weakly-}* \text{ in } L^\infty(0,T;V), \end{array}\right\} \tag{6.21}$$

$$p^\varepsilon(\cdot,0) \to p(\cdot,0) \ \text{ weakly in } L^2(0,T), \tag{6.22}$$

$$p^\varepsilon(\cdot,1) \to p(\cdot,1) \ \text{ weakly in } L^2(0,T), \tag{6.23}$$

$$\left.\begin{array}{ll} q_i^\varepsilon \to q_i & \text{weakly in } L^2(0,T;V), \\ & \text{weakly-}* \text{ in } L^\infty(0,T;H), \end{array}\right\} (i=1,2), \tag{6.24}$$

$$\varsigma_i^\varepsilon \to \varsigma_i \ \text{ weakly-}* \text{ in } W' \ (i=1,2) \tag{6.25}$$

as $\varepsilon \to 0$.

Taking account of the convergence (6.1)–(6.4) and (6.21)–(6.25), we can prove that the equations (2.28)–(2.32) hold. Indeed, in a similar manner to the proof of (5.20)–(5.29), we infer from (A2), (6.2), (6.4), (6.5), and Lebesgue's dominated convergence theorem that

$$\begin{aligned} g^{*,\varepsilon}(\cdot)\alpha'(\theta^{*,\varepsilon}(\cdot)) &\to g^*(\cdot)\alpha'(\theta^*(\cdot)) && \text{in } L^2(t_0,t_1;H), \\ \alpha'(\theta^{*,\varepsilon}(\cdot))\alpha(\theta^{*,\varepsilon}(\cdot))\chi_2^{*,\varepsilon} &\to \alpha'(\theta^*(\cdot))\alpha(\theta^*(\cdot))\chi_2^* && \text{in } L^2(t_0,t_1;H), \\ \alpha(\theta^{*,\varepsilon}(\cdot))^2 &\to \alpha(\theta^*(\cdot))^2 && \text{in } L^2(t_0,t_1;H), \end{aligned}$$

for arbitrary $0 \le t_0 < t_1 \le T$, as $\varepsilon \to 0$. Therefore, it follows from (A2), (6.5), and (6.24) that

$$g^{*,\varepsilon}(\cdot)\alpha'(\theta^{*,\varepsilon}(\cdot))q_2^\varepsilon \to g^*(\cdot)\alpha'(\theta^*(\cdot))q_2 \quad \text{weakly in } L^2(t_0,t_1;H), \tag{6.26}$$

$$\alpha'(\theta^{*,\varepsilon}(\cdot))\alpha(\theta^{*,\varepsilon}(\cdot))\chi_2^{*,\varepsilon}q_2^\varepsilon \to \alpha'(\theta^*(\cdot))\alpha(\theta^*(\cdot))\chi_2^* q_2 \quad \text{weakly in } L^2(t_0,t_1;H), \tag{6.27}$$

$$\alpha(\theta^{*,\varepsilon}(\cdot))^2 q_2^\varepsilon \to \alpha(\theta^*(\cdot))^2 q_2 \quad \text{weakly in } L^2(t_0,t_1;H), \tag{6.28}$$

for arbitrary $0 \le t_0 < t_1 \le T$, as $\varepsilon \to 0$.

In addition, the approximating adjoint system (2.21)–(2.25) is equivalent to the following variational identities:

$$\begin{aligned} &\int_0^T (-L_0 p_t^\varepsilon(t), \varpi(t))_H dt + \int_0^T h(p^\varepsilon(t), \varpi(t))_V dt + \int_0^T (l q_1^\varepsilon(t), \varpi(t))_H dt \\ &\quad + \int_0^T (\beta a_1 g^{*,\varepsilon}(t)\alpha'(\theta^{*,\varepsilon}(t))q_2^\varepsilon(t), \varpi(t))_H dt \\ &\quad - \int_0^T (2\beta\alpha'(\theta^{*,\varepsilon}(t))\alpha(\theta^{*,\varepsilon}(t))\chi_2^{*,\varepsilon}(t)q_2^\varepsilon(t), \varpi(t))_H dt \\ &= \int_0^T c_0(\theta^{*,\varepsilon}(t) - \theta_d(t), \varpi(t))_H dt \qquad \text{for all } \varpi \in L^2(0,T;V), \end{aligned} \tag{6.29}$$

$$\begin{aligned}&\int_0^T (\mu_0 q_1^\varepsilon(t), \zeta_t(t))_H dt + \int_0^T (L_1 p_t^\varepsilon(t), \zeta(t))_H dt + \int_0^T \mu_1((q_1^\varepsilon)_x(t), \zeta_x(t))_H dt \\ &\quad + \langle \varsigma_1^\varepsilon, \zeta \rangle_{W',W} \\ =& c_1 \int_0^T \left(\chi_1^{*,\varepsilon}(t) - \chi_{1,d}(t), \zeta(t)\right)_H dt \qquad \text{for all } \zeta \in W, \end{aligned} \tag{6.30}$$

and

$$\begin{aligned}&\int_0^T (\mu_0 q_2^\varepsilon(t), \zeta_t(t))_H dt + \int_0^T \mu_1((q_2^\varepsilon)_x(t), \zeta_x(t))_H dt + \langle \varsigma_2^\varepsilon, \zeta \rangle_{W',W} \\ &\quad - \int_0^T (\beta \alpha(\theta^{*,\varepsilon}(t))^2 q_2^\varepsilon(t), \zeta(t))_H dt \\ =& c_2 \int_0^T \left(\chi_2^{*,\varepsilon}(t) - \chi_{2,d}(t), \zeta(t)\right)_H dt \qquad \text{for all } \zeta \in W. \end{aligned} \tag{6.31}$$

Therefore, taking the limit of (6.29)–(6.31) as $\varepsilon \to 0$, we observe from (6.4) and (6.21)–(6.28) that the adjoint system (2.28)–(2.31) hold. In addition, we conclude from (2.26) and (6.21) that (2.32) holds.

Finally, we show (2.33). To this end, note from (A2), (6.2), (6.4), (6.5), and Lebesgue's dominated convergence theorem that

$$(\check{g} - g^{*,\varepsilon})(\cdot)\alpha(\theta^{*,\varepsilon}(\cdot)) \to (\check{g} - g^*)(\cdot)\alpha(\theta^*(\cdot)) \text{ in } L^2(0,T;H) \text{ as } \varepsilon \to 0.$$

Hence, we infer from (6.24) that

$$\begin{aligned}&\lim_{\varepsilon \to 0} \int_0^T (\beta a_1 \alpha(\theta^{*,\varepsilon}(t)), q_2^\varepsilon(t))_H (\check{g} - g^{*,\varepsilon})(t) dt \\ =& \lim_{\varepsilon \to 0} \int_0^T (\beta a_1 (\check{g} - g^{*,\varepsilon})(t) \alpha(\theta^{*,\varepsilon}(t)), q_2^\varepsilon(t))_H dt \\ =& \int_0^T (\beta a_1 (\check{g} - g^*)(t) \alpha(\theta^*(t)), q_2(t))_H dt \\ =& \int_0^T (\beta a_1 \alpha(\theta^*(t)), q_2(t))_H (\check{g} - g^*)(t) dt, \quad \forall \check{g} \in H^1(0,T). \end{aligned} \tag{6.32}$$

Therefore, taking the limit in (2.27) as $\varepsilon \to 0$, we conclude from (6.1)–(6.4), (6.21)–(6.24), and (6.32) that (2.33) hold. Hence, we see that Theorem 2.5 holds in the case when $\delta > 0$.

Similarly, we can consider the case when $\delta = 0$. Indeed, let $(f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**}) \in \mathcal{U}_{ad}^M$ be any optimal control for (OP) obtained in Theorem 2.3(I). Namely, there exists a subsequence of $\varepsilon$ (denoted by $\varepsilon$ for simplicity) such that $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \in \mathcal{U}_{ad}^M$ is the optimal control for the approximating problem $(\mathrm{OP})_0^\varepsilon$, $(\theta^{*,\varepsilon}, \chi_1^{*,\varepsilon}, \chi_2^{*,\varepsilon})$ is the unique solution to $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon})^\varepsilon$ on $[0,T]$, $(\theta^{**}, \chi_1^{**}, \chi_2^{**})$ is the unique solution to the original state system $(\mathrm{SMA};\theta_0, \chi_{1,0}, \chi_{2,0}, f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**})$ on $[0,T]$, and

$$f^{*,\varepsilon} \to f^{**} \text{ weakly in } L^2(0,T;H), \tag{6.33}$$

$$g^{*,\varepsilon} \to g^{**} \text{ weakly in } H^1(0,T), \text{ and in } L^2(0,T), \tag{6.34}$$

$$\gamma_i^{*,\varepsilon} \to \gamma_i^{**} \text{ weakly in } H^1(0,T), \text{ and in } L^2(0,T), \ (i = 0,1), \tag{6.35}$$

$$(\theta^{*,\varepsilon}, \chi_1^{*,\varepsilon}, \chi_2^{*,\varepsilon}) \to (\theta^{**}, \chi_1^{**}, \chi_2^{**}) \text{ in } (C([0,T];H))^3 \tag{6.36}$$

as $\varepsilon \to 0$.

On account of (6.21)–(6.28) and the arguments similar to the case $\delta > 0$, we can show that the adjoint system (2.28)–(2.33) works in the case when $\delta = 0$.

We now show that $(f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**}) \in \mathcal{U}_{ad}^M$ satisfies the optimality condition (2.33). To this end, we note from (2.27) and $\delta = 0$ that $(f^{*,\varepsilon}, g^{*,\varepsilon}, \gamma_0^{*,\varepsilon}, \gamma_1^{*,\varepsilon}) \in \mathcal{U}_{ad}^M$ satisfies the following:

$$\begin{aligned} &\int_0^T a_0((a_0 m_0 f^{*,\varepsilon} + p^\varepsilon)(t), \check{f}(t))_H dt - \int_0^T a_0(p^\varepsilon(t), f^{*,\varepsilon}(t))_H dt \\ &\quad + \int_0^T a_1(m_1 a_1 g^{*,\varepsilon}(t) - (\beta a_1 \alpha(\theta^{*,\varepsilon}(t)), q_2^\varepsilon(t))_H)(\check{g} - g^{*,\varepsilon})(t) dt \\ &\quad + \int_0^T a_2(m_2 a_2 \gamma_0^{*,\varepsilon}(t) + kp^\varepsilon(t,0))(\check{\gamma}_0 - \gamma_0^{*,\varepsilon})(t) dt \\ &\quad + \int_0^T a_3(m_3 a_3 \gamma_1^{*,\varepsilon}(t) + kp^\varepsilon(t,1))(\check{\gamma}_1 - \gamma_1^{*,\varepsilon})(t) dt \\ &\geq \int_0^T a_0(a_0 m_0 f^{*,\varepsilon}(t), f^{*,\varepsilon}(t))_H dt \quad \text{for any } (\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}_{ad}^M. \end{aligned} \tag{6.37}$$

Then, by (6.21)–(6.28), (6.32)–(6.36), and the weak lower semicontinuity of $L^2$-norm, we observe from (6.37) that

$$\begin{aligned} &\int_0^T a_0((a_0 m_0 f^{**} + p)(t), \check{f}(t))_H dt - \int_0^T a_0(p(t), f^{**}(t))_H dt \\ &\quad + \int_0^T a_1(m_1 a_1 g^{**}(t) - (\beta a_1 \alpha(\theta^{**}(t)), q_2(t))_H)(\check{g} - g^{**})(t) dt \\ &\quad + \int_0^T a_2(m_2 a_2 \gamma_0^{**}(t) + kp(t,0))(\check{\gamma}_0 - \gamma_0^{**})(t) dt \\ &\quad + \int_0^T a_3(m_3 a_3 \gamma_1^{**}(t) + kp(t,1))(\check{\gamma}_1 - \gamma_1^{**})(t) dt \\ &\geq \int_0^T a_0(a_0 m_0 f^{**}(t), f^{**}(t))_H dt \quad \text{for any } (\check{f}, \check{g}, \check{\gamma}_0, \check{\gamma}_1) \in \mathcal{U}_{ad}^M, \end{aligned}$$

which implies that $(f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**}) \in \mathcal{U}_{ad}^M$ satisfies (2.33). Hence, we conclude that Theorem 2.5 holds for the the optimal control $(f^{**}, g^{**}, \gamma_0^{**}, \gamma_1^{**}) \in \mathcal{U}_{ad}^M$ to (OP) obtained in Theorem 2.3(I).

Thus, the proof of Theorem 2.5 is complete. □

**Remark 6.1.** *In Theorems 2.3 and 2.5, we consider two cases: $\delta = 0$ and $\delta > 0$. If $\delta > 0$, then, for each optimal control $(f^*, g^*, \gamma_0^*, \gamma_1^*)$ to (OP), we can find the sequence of optimal controls for $(\mathrm{OP})_\delta^\varepsilon$ that converges to $(f^*, g^*, \gamma_0^*, \gamma_1^*)$ strongly in $L^2(0,T;H) \times (L^2(0,T))^3$. However, it is very difficult to give the numerical experiments for $(\mathrm{OP})_\delta^\varepsilon$, since the cost function $J_\delta^\varepsilon$ depends on the unknown optimal control $(f^*, g^*, \gamma_0^*, \gamma_1^*)$ for (OP). If $\delta = 0$, then, the cost function $J_0^\varepsilon$ is independent of the optimal control for (OP). Therefore, in the numerical analysis, we are forced to adopt $(\mathrm{OP})_0^\varepsilon$ as the approximating problem for (OP). Thus, it is worthy considering the case when $\delta = 0$ in Theorems 2.3 and 2.5.*

# Appendix

In this Appendix, we provide a typical example of $K^\varepsilon$ in Example 2.1, which is a smooth closed convex set in $\mathbb{R}^2$ such that $K^\varepsilon$ includes the convex set $K$ defined in (1.6). In addition, we give an example of construction of the convex function $\widehat{K}^\varepsilon$ satisfying assumption (A5).

We first define $K^\varepsilon$ in Example 2.1 for any $\varepsilon \in (0,1]$. To this end, we specify the boundary of $K^\varepsilon$. Indeed, we consider six boundary parts of $K^\varepsilon$ illustrated in Figure 3: first, we put

$$\mathrm{A}\left(-\frac{\varepsilon}{\sqrt[3]{2}}, -\frac{\varepsilon}{\sqrt[3]{2}}\right), \quad \mathrm{B}\left(1-\frac{\varepsilon}{\sqrt[3]{2}}, -1-\frac{\varepsilon}{\sqrt[3]{2}}\right), \quad \mathrm{C}(1+\varepsilon, -1),$$

$$\mathrm{D}(1+\varepsilon, 1), \quad \mathrm{E}\left(1-\frac{\varepsilon}{\sqrt[3]{2}}, 1+\frac{\varepsilon}{\sqrt[3]{2}}\right), \quad \mathrm{F}\left(-\frac{\varepsilon}{\sqrt[3]{2}}, \frac{\varepsilon}{\sqrt[3]{2}}\right);$$

then, we define the boundary $\partial K^\varepsilon$ of $K^\varepsilon$ as follows:

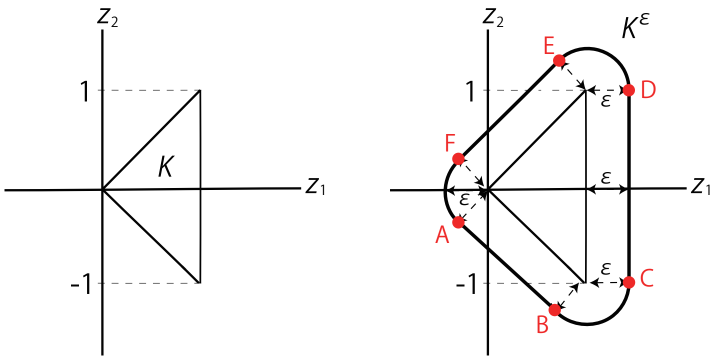


Figure 3: Convex sets $K$ and $K^\varepsilon$

$$\partial K^\varepsilon : \begin{cases} z_2 = -z_1 - \dfrac{2\varepsilon}{\sqrt[3]{2}} & \text{on } \partial K^\varepsilon|_{\mathrm{AB}}, \\ |z_1 - 1|^3 + |z_2 + 1|^3 = \varepsilon^3 & \text{on } \partial K^\varepsilon|_{\mathrm{BC}}, \\ z_1 = 1 + \varepsilon & \text{on } \partial K^\varepsilon|_{\mathrm{CD}}, \\ |z_1 - 1|^3 + |z_2 - 1|^3 = \varepsilon^3 & \text{on } \partial K^\varepsilon|_{\mathrm{DE}}, \\ z_2 = z_1 + \dfrac{2\varepsilon}{\sqrt[3]{2}} & \text{on } \partial K^\varepsilon|_{\mathrm{EF}}, \\ |z_1|^3 + |z_2|^3 = \varepsilon^3 & \text{on } \partial K^\varepsilon|_{\mathrm{FA}}, \end{cases} \tag{ap.1}$$

where $\partial K^\varepsilon|_{ij}$ $(i,j = \mathrm{A},\mathrm{B},\mathrm{C},\mathrm{D},\mathrm{E},\mathrm{F})$ indicates the boundary part of $K^\varepsilon$ from the points $i$ to $j$.

By (ap.1), we can define the smooth closed convex set $K^\varepsilon$ with the boundary $\partial K^\varepsilon$ illustrated in Figure 3 for any $\varepsilon \in (0,1]$.

**Remark Ap.1.** The convex set $K$ is line-symmetric relative to the $z_1$-axis, and hence, the set $K^\varepsilon$ defined as above is also line-symmetric (cf. Figure 3). In addition, $K^\varepsilon$ converges to $K$ in the sense of Hausdorff distance, thus, in the sense of Mosco [37] as $\varepsilon \to 0$.

**Remark Ap2.** By similar arguments as above, more precisely, by replacing $\varepsilon$ with $2\varepsilon$, we can define the smooth closed convex set $K^{2\varepsilon}$ in $\mathbb{R}^2$ such that $K^{2\varepsilon}$ includes $K^{\varepsilon}$ for any $\varepsilon \in (0,1]$ (cf. Figure 4 below).

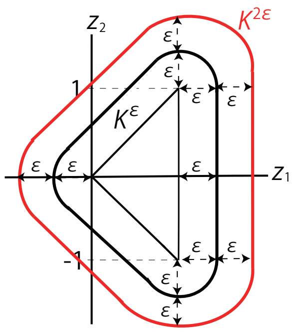


Figure 4: Convex set $K^{2\varepsilon}$

On account of Remarks Ap.1–2, we can define the smooth convex function $\widehat{K}^{\varepsilon}$ satisfying assumption (A5). Indeed, we consider seven regions of $\mathbb{R}^2$ illustrated in Figure 5:

$$\mathbb{R}^2 := K \cup D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5 \cup D_6.$$

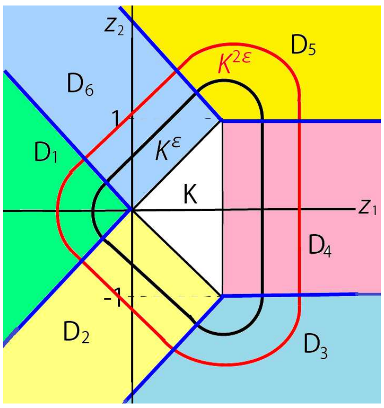


Figure 5: Decomposition of $\mathbb{R}^2$

At first, we define

$$\widehat{K}^{\varepsilon}(z_1, z_2) = 0 \quad \text{on } K.$$

Next, we define $\widehat{K}^\varepsilon(z_1,z_2)$ on $D_1$ as follows:

$$\widehat{K}^\varepsilon(z_1,z_2) := \begin{cases} 0 & \text{on } D_1 \cap K^\varepsilon, \\ \dfrac{1}{15\sqrt{2}}\left(\dfrac{|z_1|^3+|z_2|^3}{6\varepsilon^2} - \dfrac{\varepsilon}{6}\right) \\ \qquad - \dfrac{1}{30\sqrt{2}}\left(\sqrt[3]{|z_1|^3+|z_2|^3} - \varepsilon\right) & \text{on } D_1 \cap (K^{2\varepsilon} \setminus K^\varepsilon), \\ \dfrac{2}{45\sqrt{\varepsilon}}\sqrt{|z_1|^3+|z_2|^3} - \dfrac{4\sqrt{2}}{45}\varepsilon + \dfrac{7\varepsilon}{90\sqrt{2}} \\ \qquad - \dfrac{1}{30\sqrt{2}}\left(\sqrt[3]{|z_1|^3+|z_2|^3} - \varepsilon\right) & \text{on } D_1 \setminus K^{2\varepsilon}. \end{cases}$$

Similarly, we can define $\widehat{K}^\varepsilon(z_1,z_2)$ on $D_i$, $(i = 2,3,4,5,6)$ by the movement of the above function along the boundary $\partial K^\varepsilon$. Indeed, for instance, we have that

$$\widehat{K}^\varepsilon(z_1,z_2) := \begin{cases} 0 & \text{on } D_4 \cap K^\varepsilon, \\ \dfrac{1}{15\sqrt{2}}\left(\dfrac{|z_1-1|^3}{6\varepsilon^2} - \dfrac{\varepsilon}{6}\right) - \dfrac{1}{30\sqrt{2}}\left(|z_1-1| - \varepsilon\right) & \text{on } D_4 \cap (K^{2\varepsilon} \setminus K^\varepsilon), \\ \dfrac{2}{45\sqrt{\varepsilon}}\sqrt{|z_1-1|^3} - \dfrac{4\sqrt{2}}{45}\varepsilon + \dfrac{7\varepsilon}{90\sqrt{2}} \\ \qquad - \dfrac{1}{30\sqrt{2}}\left(|z_1-1| - \varepsilon\right) & \text{on } D_4 \setminus K^{2\varepsilon}, \end{cases}$$

$$\widehat{K}^\varepsilon(z_1,z_2) := \begin{cases} 0 & \text{on } D_5 \cap K^\varepsilon, \\ \dfrac{1}{15\sqrt{2}}\left(\dfrac{|z_1-1|^3+|z_2-1|^3}{6\varepsilon^2} - \dfrac{\varepsilon}{6}\right) \\ \qquad - \dfrac{1}{30\sqrt{2}}\left(\sqrt[3]{|z_1-1|^3+|z_2-1|^3} - \varepsilon\right) & \text{on } D_5 \cap (K^{2\varepsilon} \setminus K^\varepsilon), \\ \dfrac{2}{45\sqrt{\varepsilon}}\sqrt{|z_1-1|^3+|z_2-1|^3} - \dfrac{4\sqrt{2}}{45}\varepsilon + \dfrac{7\varepsilon}{90\sqrt{2}} \\ \qquad - \dfrac{1}{30\sqrt{2}}\left(\sqrt[3]{|z_1-1|^3+|z_2-1|^3} - \varepsilon\right) & \text{on } D_5 \setminus K^{2\varepsilon}, \end{cases}$$

and so on.

On account of the construction of $\widehat{K}^\varepsilon$ as above, assumption (A5) can be easily verified by standard calculations. Here, we omit the details.

## References


[1] T. Aiki, A. Kadoya and N. Sato, Optimal control problem for phase-field equations with nonlinear dynamic boundary conditions, Proceedings of the Third World Congress of Nonlinear Analysts, Part 5 (Catania, 2000), Nonlinear Anal., **47** (2001), 3183–3194.

[2] H. Attouch, *Variational Convergence for Functions and Operators*, Pitman Advanced Publishing Program, Boston-London-Melbourne, 1984.

[3] V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach spaces*, Noordhoff, Leyden, 1976.

[4] V. Barbu, *Optimal control of variational inequalities*, Research Notes in Mathematics, **100**, Pitman, London, 1984.

[5] V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*, Springer Monographs in Mathematics, 2010.

[6] V. Barbu, M. L. Bernardi, P. Colli and G. Gilardi, Optimal control problems of phase relaxation models, J. Optim. Theory Appl., **109** (2001), 557–585.

[7] E. Bonetti, Global solution to a Frémond model for shape memory alloys with thermal memory, Nonlinear Anal., **46** (2001), 535–565.

[8] E. Bonetti, Global solvability of a dissipative Frémond model for shape memory alloys. Part I: mathematical formulation and uniqueness, Quart. Appl. Math., **61** (2003), 759–781.

[9] E. Bonetti, Global solvability of a dissipative Frémond model for shape memory alloys. Part II: existence, Quart. Appl. Math., **62** (2004), 53–76.

[10] H. Brézis, *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, North-Holland, Amsterdam, 1973.

[11] Z. Chen and K.-H. Hoffmann, Numerical solutions of the optimal control problem governed by a phase field model, *Estimation and control of distributed parameter systems (Vorau, 1990)*, pp. 79–97, Internat. Ser. Numer. Math., Vol. 100, Birkhäuser, Basel, 1991.

[12] P. Colli, Mathematical study of an evolution problem describing the thermo-mechanical process in shape memory alloys, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl., **2** (1991), 55–64.

[13] P. Colli, Global existence for the three-dimensional Frémond model of shape memory alloys, Nonlinear Anal., **24** (1995), 1565–1579.

[14] P. Colli, M.H. Farshbaf-Shaker, G. Gilardi and J. Sprekels, Optimal boundary con- trol of a viscous Cahn–Hilliard system with dynamic boundary condition and dou- ble obstacle potentials, SIAM J. Control Optim., **53** (2015), 2696–2721.

[15] P. Colli, M.H. Farshbaf-Shaker, G. Gilardi and J. Sprekels, Second-order analysis of a boundary control problem for the viscous Cahn–Hilliard equation with dynamic boundary condition, Ann. Acad. Rom. Sci. Ser. Math. Appl., **7** (2015), 41–66.

[16] P. Colli, M.H. Farshbaf-Shaker and J. Sprekels, A deep quench approach to the optimal control of an Allen–Cahn equation with dynamic boundary conditions and double obstacles, Appl. Math. Optim., **71** (2015), 1–24.

[17] P. Colli, M. Frémond, E. Rocca and K. Shirakawa, Attractors for a three-dimensional thermo-mechanical model of shape memory alloys, Chinese Ann. Math. Ser. B, **27** (2006), 683–700.

[18] P. Colli, G. Gilardi and J. Sprekels, A boundary control problem for the pure Cahn–Hilliard equation with dynamic boundary conditions, Adv. Nonlinear Anal., **4** (2015), 311–325.

[19] P. Colli, G. Gilardi and J. Sprekels, A boundary control problem for the viscous Cahn–Hilliard equation with dynamic boundary conditions, Appl. Math. Optim., **73** (2016), 195–225.

[20] P. Colli and K.-H. Hoffmann, A nonlinear evolution problem describing multi-component phase changes with dissipation, Numer. Funct. Anal. Optim., **14**(1993), 275–297.

[21] P. Colli, Ph. Laurençot and U. Stefanelli, Long-time behavior for the full one-dimensional Frémond model for shape memory alloys, Contin. Mech. Thermodyn., **12** (2000), 423–433.

[22] P. Colli and K. Shirakawa, Attractors for the one-dimensional Frémond model of shape memory alloys, Asymptot. Anal., **40** (2004), 109–135.

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

[24] M.H. Farshbaf-Shaker, A relaxation approach to vector-valued Allen-Cahn MPEC problems, Appl. Math. Optim., **72** (2015), 325–351.

[25] M.H. Farshbaf-Shaker and C. Hecht, Optimal control of elastic vector-valued Allen-Cahn variational inequalities, SIAM J. Control Optim., **54** (2016), 129–152.

[26] M. Frémond, *Non-Smooth Thermomechanics*, Springer-Verlag, Berlin, 2002.

[27] M. Frémond and S. Miyazaki, *Shape Memory Alloys*, CISM Courses and Lectures, Vol. 351, Springer, Vienna, 1996.

[28] T. Fukao and N. Yamazaki, A boundary control problem for the equation and dynamic boundary condition of Cahn–Hilliard type, *Solvability, regularity, and optimal control of boundary value problems for PDEs*, pp. 255–280, Springer INdAM Ser., **22**, Springer, Cham, 2017.

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

[30] K.-H. Hoffmann and L. Jiang, Optimal control of a phase field model for solidification, Numer. Funct. Anal. Optim., **13** (1992), 11–27.

[31] K.-H. Hoffmann, M. Kubo and N. Yamazaki, Optimal control problems for elliptic-parabolic variational inequalities with time-dependent constraints, Numer. Funct. Anal. Optim., **27** (2006), 329–356.

[32] A. Ito, N. Yamazaki and N. Kenmochi, Attractors of nonlinear evolution systems generated by time-dependent subdifferentials in Hilbert spaces, *Dynamical systems and differential equations, Vol. I (Springfield, MO, 1996)*, pp. 327–349, Discrete Contin. Dynam. Systems, Added Volume I, 1998.

[33] N. Kenmochi, Solvability of nonlinear evolution equations with time-dependent constraints and applications, Bull. Fac. Education, Chiba Univ., **30** (1981), 1–87.

[34] N. Kenmochi and M. Niezgódka, Evolution systems of nonlinear variational inequalities arising from phase change problems, Nonlinear Anal., **22**(1994), 1163–1180.

[35] J.-L. Lions, *Optimal control of systems governed by partial differential equations*, Springer-Verlag, New York-Berlin, 1971.

[36] J.-L. Lions, *Contrôle des systèmes distribués singuliers,* Méthodes Mathématiques de l'Informatique, No.13, Gauthier-Villars, Montrouge, 1983.

[37] U. Mosco, Convergence of convex sets and of solutions of variational inequalities, Advances Math., **3** (1969), 510–585.

[38] P. Neittaanmäki, J. Sprekels and D. Tiba, *Optimization of Elliptic Systems: Theory and Applications*, Springer Monographs in Mathematics, Springer, New York, 2006.

[39] M. Ôtani, Nonmonotone perturbations for nonlinear parabolic equations associated with subdifferential operators, Cauchy problems, J. Differential Equations, **46** (1982), 268–299.

[40] T. Ohtsuka, K. Shirakawa and N. Yamazaki, Optimal control problems of singular diffusion equation with constraint, Adv. Math. Sci. Appl., **18** (2008), 1–28.

[41] T. Ohtsuka, K. Shirakawa and N. Yamazaki, Convergence of numerical algorithm for optimal control problem of Allen-Cahn type equation with constraint, *Proceedings of International Conference on: Nonlinear Phenomena with Energy Dissipation–Mathematical Analysis, Modelling and Simulation*, pp. 441–462, GAKUTO Intern. Ser. Math. Appl., vol **29**, Gakkotosho, Tokyo, 2008.

[42] I. Pawłow, *Analysis and Control of Evolution Multi-Phase Problems with Free Boundaries*, Prace habilitacyjne, Polska Akademia Nauk, Instytut Badań Systemowych, 1987.

[43] S.-U. Ryu and A. Yagi, Optimal control for an adsorbate-induced phase transition model, Appl. Math. Comput., **171** (2005), 420–432.

[44] K. Shirakawa, A, Ito, N. Yamazaki, and N. Kenmochi, Asymptotic stability for evolution equations governed by subdifferentials, *Recent Development in Domain Decomposition Methods and Flow Problems*, pp. 287–310, GAKUTO Internat. Ser. Math. Sci. Appl., **11**, Gakkōtosho, Tokyo, 1998.

[45] K. Shirakawa and N. Yamazaki, Optimal control problems of phase field system with total variation functional as the interfacial energy, Adv. Differential Equations, **18** (2013), 309–350.

[46] J. Sprekels and S. Zheng, Optimal control problems for a thermodynamically consistent model of phase-field type for phase transitions, Adv. Math. Sci. Appl., **1** (1992), 113–125.