Optimal control of elliptic variational-hemivariational inequalities

Zijia Peng[a,*], Karl Kunisch[b]

[a] *Guangxi Key Laboratory of Universities Optimization Control and Engineering Calculation, and College of Sciences, Guangxi University for Nationalities, Nanning, Guangxi 530006, P.R. China*

[b] *Institut für Mathematik, Karl-Franzens-Universität Graz, A-8010 Graz, Austria*

Abstract

We study the optimality systems of optimal control problems governed by nonlinear elliptic state systems whose variational formulations are variational-hemivariational inequalities. We first prove the existence of solutions to the inequality problems and show the solution mappings are weakly upper semi-continuous. Then we establish the existence theorem of optimal pairs to nonsmooth cost functionals. Under appropriate conditions, the approximation results and abstract necessary optimality conditions of first order are derived. Moreover, we take the obstacle problem with nonmonotone perturbation as an example and derive the optimality system precisely.

Keywords:
Hemivariational inequality, optimality system, necessary optimality condition

AMS Subject Classifications (2010) 47J20, 49J20, 49J40, 49K20

---

1. Introduction

Optimal control problems for equations and variational inequalities have been formulated and studied in numerous publications in the past several decades; see, e.g., [1–8] and the references therein. In this paper, we consider optimal control problems in which the variational formulations of the state

*Corresponding author
E-mail: pengzijia@126.com (Z. Peng), Karl.Kunisch@uni-graz.at (K. Kunisch)

equations are variational-hemivariational inequalities. The main emphasis is put on the existence of optimal pairs, approximations and necessary optimality conditions of first order. More precisely, we study the optimal control problems, (OP) for short, as follows.
Minimize the functional

$$G(y,u) = g(y) + h(u) \tag{1.1}$$

on all $(y,u) \in V \times U$, subject to

$$\text{(VI-HVI)}\begin{cases} Ay + \partial\varphi(y) + \overline{\partial} j(\cdot, y) \ni f + Bu \ \text{ in } \Omega \\ y = 0 \text{ on } \partial\Omega, \end{cases} \tag{1.2}$$

where y is the state, u is the control, and V and U are referred to as the state and control spaces, respectively. The cost functional $G(y,u)$ may be nonsmooth, and A is supposed to be a second order elliptic differential operators of the form

$$Ay(x) = \sum_{i=1}^{N} \frac{\partial}{\partial x_i} a_i(x, \nabla y) + a_0(x,y). \tag{1.3}$$

Further $\partial\varphi(y)$ stands for the convex subdifferential of $\varphi(y)$, while $\overline{\partial} j(\cdot, y)$ is understood to be $\overline{\partial} j(\cdot, y)(x) = \overline{\partial} j(x, y(x))$ a.e. $x \in \Omega$, and the latter denotes the Clarke generalized gradient of a locally Lipschitz continuous function j with respect to $y(x)$. Assume $V = H_0^1(\Omega)$ and $f + Bu \in V^* = H^{-1}(\Omega)$. Then we see that the variational formulation of the state system (1.2) is the following variational-hemivariational inequality

$$\langle Ay - f - Bu, v - y\rangle + \int_\Omega j^0(x, y; v - y)dx + \varphi(v) - \varphi(y) \geq 0, \ \forall\, v \in V, \tag{1.4}$$

where $j^0(x, y; v - y)$ is the Clarke generalized directional derivative of j. In particular, if $\varphi(u)$ is the indication function with respect to a nonempty, closed and convex subset $K \subset V$, i.e.,

$$\varphi(u) = I_K(u) = \begin{cases} \ 0, & \text{if } u \in K \\ +\infty, & \text{otherwise,} \end{cases} \tag{1.5}$$

then inequality (1.4) is equivalent to

$$\langle Ay - Bu - f, v - y\rangle + \int_\Omega j^0(x, y; v - y)dx \geq 0, \ \forall v \in K. \tag{1.6}$$

Further, if we define

$$K := \{v \in V : v(x) \geq \psi(x) \text{ a.e. } x \in \Omega\} \text{ for some } \psi \in V, \tag{1.7}$$

then (1.6) is well-known as the obstacle problem described by hemivariational inequality. We recall that the notion of hemivariational inequality was introduced by P.D. Panagiotopoulos in the 1980s for mechanical problems with nonconvex and nonsmooth energy functionals (so-called superpotentials); see [9, 10]. Based on the generalized gradient of F.H. Clarke, hemivariational inequalities generalize the classical variational ones and have been proved to be useful tools to deal with the problems involving nonmonotone and possibly multivalued relations. We refer the readers to [11–15] and the references therein for more details on the theory and applications in this field.

Now we give a brief remark on some related literature. Based on finite element approximation, the initial attempts towards the existence of optimal pairs of hemivariational inequalities can be found in [16–18], and the existence and necessary optimality conditions in [19, 20]. Thereafter, the optimal shape design of hemivariational inequalities was considered in [21, 22]. Recent studies in this field include [23–28] but they focus on existence results. In this work, however, we put the emphasis on existence and approximation results, and necessary optimality conditions of a class of variational-hemivariational inequalities.

The rest of this paper is organized as follow. In Section 2, we present some preliminary material on monotone operators, convex and nonsmooth analysis. Section 3 is devoted to the existence of solutions to the state system and optimal pairs to the optimal control problems (Theorems 3.2, 3.6). The approximation results and the abstract necessary optimality conditions are derived in Section 4 (Theorems 4.5, 4.6). In the last section, we focus on a class of obstacle problems with nonlinear and nonmonotone perturbation. The complete optimality system is presented (Theorem 5.4).

2. Preliminaries

In this section, we summarize some preliminary material on monotone operators, convex and nonsmooth analysis. We refer the reader to monographs and textbooks; e.g., [11, 12, 29, 30] for the notations and related proofs.

Let X be a real reflexive Banach space, X^* its dual and $\langle \cdot, \cdot \rangle$ the duality pairing between X and X^*. Let $\phi : X \to \mathbb{R} \cup \{+\infty\}$ be a convex functional with effective domain $D(\phi) := \{u \in X : \phi(u) < +\infty\}$. We say ϕ is proper if it is not identically equal to positive infinity. The subdifferential of ϕ at $u \in X$ is defined by

$$\partial\phi(u) := \{w \in X^* : \phi(v) - \phi(u) \geq \langle w, v - u\rangle, \text{ for all } v \in X\}. \tag{2.1}$$

We say a single-valued operator $T \colon X \to X^*$ is hemicontinuous if the real-valued function $\lambda \mapsto \langle T(u + \lambda v), w\rangle$ is continuous on $[0, 1]$ for all u, v, $w \in X$. The operator T is said to be demicontinuous if $u_n \to u$ in X implies $Tu_n \to Tu$ weakly in X^*, and it is said to be weakly-strongly continuous, if $u_n \to u$ weakly in X implies $Tu_n \to Tu$ in X^*. Moreover, T is called pseudomonotone, if it is bounded and if from $u_n \to u$ weakly in X and $\limsup\langle Tu_n, u_n - u\rangle \leq 0$, it follows that $\liminf\langle Tu_n, u_n - v\rangle_X \geq \langle Tu, u - v\rangle$ for all $v \in X$.

Remark 2.1. *An equivalent definition of pseudomonotonicity of $T \colon X \to X^*$ is given by: T is bounded and if $u_n \to u$ weakly in X and $\limsup\langle Tu_n, u_n - u\rangle_X \leq 0$, then we have*

$$\lim_{n\to\infty}\langle Tu_n, u_n - u\rangle = 0 \text{ and } Tu_n \to Tu \text{ weakly in } X^*.$$

Moreover, every monotone and hemicontinuous operator is pseudomonotone (see, e.g. [30, Proposition 27.6]).

Definition 2.2. *A multivalued operator $T : D(T) \subset X \to 2^{X^*}$ is said to be monotone, if*

$$\langle y_1 - y_2, x_1 - x_2\rangle \geq 0, \quad \forall\, x_1, x_2 \in D(T),\ \forall\, y_1 \in T(y_1),\ y_2 \in T(y_2).$$

T is said to be maximal monotone if T is monotone and there is no monotone extension in $X \times X^*$, i.e., for $x \in X, y \in X^*$ if

$$\langle y_1 - y, x_1 - x\rangle \geq 0, \quad \forall\, x_1 \in D(T),\ \forall\, y_1 \in T(x),$$

then we have $x \in D(T)$ and $y \in T(x)$.

It is well known that the subdifferential operator $\partial\phi$ defined by (2.1) is maximal monotone.

Let X, Y be Banach Spaces and $T \colon X \to 2^Y$ a multivalued operator. The inverse image of $E \subset Y$ under T is the set

$$T^{-1}(E) := \{x \in X : \ T(x) \cap E \neq \emptyset\},$$

and by $\mathrm{Gr}(T)$ we denote the graph of T, defined by

$$\mathrm{Gr}(T) := \{(x, y) \in X \times Y : \ y \in T(x)\}.$$

Then $T\colon X \to 2^Y$ is called to be upper semicontinuous (resp. weakly upper semicontinuous) if for all $D \subset Y$ closed (resp. weakly closed), we have that $F^{-1}(D)$ is closed (resp. weakly closed) in X.

Definition 2.3. *A multivalued operator $T : X \to 2^{X^*}$ is said to be pseudomonotone if the following conditions are satisfied:*

(i) for each $u \in X$, $T(u) \subset X^*$ is nonempty, bounded, closed and convex;

(ii) the restriction of T to each finite-dimensional subspace F of X is weakly upper semicontinuous as an operator from F to X^* ;

(iii) let $u_n \in X$ and $u_n^* \in X^*$ with $u_n^* \in T(u_n)$, from $u_n \to u$ weakly in X and $\limsup \langle u_n^*, u_n - u\rangle_X \leq 0$, it follows: to each element $v \in X$, there exists a $u^*(v) \in T(u)$ such that

$$\liminf_{n\to\infty} \langle u_n^*, u_n - v\rangle \geq \langle u^*(v), u - v\rangle.$$

Note that pseudomonotone mappings is invariant under addition of operators; see e.g., [11, Proposition 2.4].

Lemma 2.4. *(cf. [11, Theorem 2.12]) Let $T : X \to 2^{X^*}$ be a bounded and pseudomonotone operator, $T_1 : X \to 2^{X^*}$ be a maximal monotone operator with $v_0 \in D(T_1)$. If there exists a function $c : \mathbb{R}^+ \to R$ with $c(r) \to +\infty$ as $r \to +\infty$, such that for all $v \in X$ and $v^* \in T(v)$, $\langle v^*, v - v_0\rangle \geq c(\|v\|)\|v\|$. Then $T + T_1$ is surjective.*

Definition 2.5. *Let X be a Banach space and let $\phi\colon X \to \mathbb{R}$ be a locally Lipschitz function. The Clarke generalized directional derivative of ϕ at $x \in X$ in the direction $v \in X$, denoted by $\phi^0(x; v)$, is defined by*

$$\phi^0(x; v) := \limsup_{y\to x,\ \lambda\downarrow 0} \frac{\phi(y + \lambda v) - \phi(y)}{\lambda}$$

and the generalized gradient (subdifferential) of ϕ at x, denoted by $\overline{\partial}\phi(x)$, is a subset of a dual space X^* given by

$$\overline{\partial}\phi(x) := \{\, x^* \in X^* \mid \langle x^*, v\rangle \leq \phi^0(x; v) \ \ \textit{for all} \ \ v \in X \,\}.$$

It is known that for every $x \in X$, $\overline{\partial}\phi(x)$ is a nonempty, convex and w^*-compact subset in X^*; the mapping $\phi^0 : X \times X \to \mathbb{R}$ is upper semicontinuous; the graph of $\overline{\partial}\phi$ is sequentially closed in $X \times X^*$ with X^* equipped with weak star topology; see e.g., [12, Proposition 2.171]. In particular, if $\phi\colon X \to \mathbb{R}$ is a convex and continuous functional, then the generalized subdifferential of ϕ coincides with the subdifferential defined by (2.1).

3. Existence of solutions and optimal pairs

Let Ω be a bounded domain in $\mathbb{R}^N$ with sufficient smooth boundary $\partial\Omega$. Suppose that U is a reflexive Banach space. Assume that $V = H_0^1(\Omega)$ and $H = L^2(\Omega)$. Identifying H with its dual, then $V \subset H \subset V^*$ forms an evolution triple with all the embeddings being dense and compact. Throughout this paper, the symbol w-X is always used to indicate the space X equipped with weak topology. The symbol $\langle \cdot, \cdot \rangle$ stands for the duality pairing between V and V^*, and $(\cdot, \cdot)$ denotes the duality product between $L^p(\Omega)$ and its dual with $1 \le p < \infty$.

We now present the definitions and assumptions on the data of system (1.2). First of all, we suppose that the function $j : \Omega \times \mathbb{R} \to \mathbb{R}$ is given by

$$j(x, s) = \int_0^s \beta(x, \tau) d\tau, \tag{3.1}$$

and the following hypotheses on β are considered.
H(β) The function $\beta : \Omega \times \mathbb{R} \to \mathbb{R}$ is supposed to satisfy

(i) the mapping $x \mapsto \beta(x, s)$ is continuous for a.e. $s \in \mathbb{R}$ and the mapping $(x, s) \mapsto \beta(x, s)$ is measurable in $\Omega \times \mathbb{R}$,

(ii) there exist a function $\gamma_1 \in L^\infty(\Omega)$ and a positive constant $\bar{b}_1$ such that

$$|\beta(x, \tau)| \le \gamma_1(x) + \bar{b}_1 |\tau|, \ \forall\, (x, \tau) \in \Omega \times \mathbb{R},$$

(iii) there exist a function $\gamma_2 \in L^1(\Omega)$, and positive constants $1 \le \sigma < 2$ and $\bar{b}_2$ such that

$$\beta(x, \tau)\tau \ge -\gamma_2(x) - \bar{b}_2 |\tau|^\sigma, \ \forall\, (x, \tau) \in \Omega \times \mathbb{R},$$

(iv) there exists a positive constant m such that

$$\operatorname*{ess\,inf}_{s_1\neq s_2} \frac{\beta(x,s_1)-\beta(x,s_2)}{s_1-s_2} \geq -m, \quad \forall\, x\in\Omega,\ s_1,s_2\in\mathbb{R}.$$

Under H(β)(i)-(ii) the function j is well defined, $j(\cdot,s)$ is measurable for all $s\in\mathbb{R}$ and $j(x,\cdot)$ is locally Lipschitz continuous for a.e. $x\in\Omega$. Moreover, if $\lim\limits_{\tau\to s\pm}\beta(x,\tau)$ exists, then $\overline{\partial} j(x,s)=[\underline{\beta}(x,s),\,\overline{\beta}(x,s)]$ where

$$\underline{\beta}(x,s)=\min\big\{\beta(x,s-),\,\beta(x,s+)\big\},\quad \overline{\beta}(x,s)=\max\big\{\beta(x,s-),\,\beta(x,s+)\big\}.$$

Otherwise, $\overline{\partial} j(x,s)\subseteq[\beta_1(x,s),\,\beta_2(x,s)]$ where

$$\beta_1(x,s)=\lim_{\delta\to 0^+}\operatorname*{ess\,inf}_{|\tau-s|\leq\delta}\beta(x,\tau),\quad \beta_2(x,s)=\lim_{\delta\to 0^+}\operatorname*{ess\,sup}_{|\tau-s|\leq\delta}\beta(x,\tau).$$

In particular, if $s\mapsto\beta(x,s)$ is continuous, then $\overline{\partial} j(x,s)=\beta(x,s)$ is single-valued but in general nonlinear and nonmonotone. By means of j, we define a function $J:H\to\mathbb{R}$ given by

$$J(v)=\int_\Omega j(x,v(x))dx=\int_\Omega\int_0^{v(x)}\beta(x,s)dsdx,\ \forall\, v\in H. \tag{3.2}$$

Due to the hypotheses H(β)(i)-(iii) and Lebourg's mean value theorem, J is well defined and Lipschitz continuous on each bounded subset of H. Therefore, the Clarkes generalized gradient $\overline{\partial} J:H\to 2^H$ is well defined. Moreover, the Aubin-Clarke theorem (cf. [12, Theorem 2.181]) ensures that for each $v\in H$, we have

$$\xi\in\overline{\partial} J(v)\ \text{ implies }\ \xi(x)\in\overline{\partial} j(x,v(x))\ \text{ for a.e. } x\in\Omega, \tag{3.3}$$

and therefore

$$(\xi,w)\leq\int_\Omega j^0(x,v(x);w(x))\ \forall\, v,w\in H \text{ and } \xi\in\overline{\partial} J(v). \tag{3.4}$$

Using H(β)(ii)-(iv), we conclude the following properties.
From H(β)(ii) it follows that for some positive constant b_1

$$\|\xi\|_H\leq b_1(1+\|v\|_H),\ \forall\, v\in H,\,\xi\in\overline{\partial} J(v). \tag{3.5}$$

From H(β)(iii) there exists a positive constant b_2 such that

$$J^0(v;-v) \le b_2(1+\|v\|_H^\sigma),\ \forall\, v \in H. \tag{3.6}$$

From H(β)(iv), for all $\xi_i \in \overline{\partial} J(v_i)$ with $v_i \in H$, $i = 1, 2$, it follows

$$(\xi_1 - \xi_2, v_1 - v_2) \ge -m\|v_1 - v_2\|_H^2. \tag{3.7}$$

We remark that (3.7) or H(β)(iv) have been considered in numerous publications for the existence and uniqueness of solutions to hemivariational inequalities, see [20, (G) p. 69)], [12, (B1)(ii) p. 182], for example.

Next, let the nonlinear operator A be given by (1.3). We assume that A satisfies the following standard conditions H(A):

(A1) A is monotone, hemicontinuous and there exists a positive constant c_1 such that

$$\|Av\|_{V^*} \le c_1(1 + \|v\|_V),\ \forall\, v \in V.$$

(A2) A is coercive, i.e., there exist positive constants c_2 and $c_3 > 0$ such that

$$\langle Av, v\rangle \ge c_2\|v\|_V^2 - c_3,\ \forall\, v \in V.$$

Finally, let $f \in V^*$, and φ and B be given as follows.
H(φ) the function $\varphi : V \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex and lower semicontinuous functional with some element $y_0 \in D(\partial\varphi)$.

H(B) B is a linear and compact operator from U to V^*.

To solve the state system (1.2) we consider the following multivalued operator

$$\overline{\partial}(J|_V)(u) = (i^* \circ \overline{\partial} J \circ i)(u),\ \forall\, u \in V, \tag{3.8}$$

where i is the embedding operator from V to H and $i^* : H \to V^*$ denotes its adjoint, i.e.,

$$\langle i^* z, u\rangle = (z, iu) = \int_\Omega z(x)u(x)dx,\ \forall\, u \in V,\ z \in H.$$

It is well known that i and i^* are both linear, continuous and compact. In what follows, for simplicity we always omit the symbols i and i^* when no confusion arises. This means that the operator $\overline{\partial} J : H \to H$ stands for $\overline{\partial}(J|_V) : V \to V^*$ in the sense of (3.8) as well.

Then we consider the following nonlinear elliptic inclusion:

$$Ay + \overline{\partial} J(y) + \partial\varphi(y) \ni f + Bu \quad \text{in } V^*. \tag{3.9}$$

By means of (3.3)-(3.4), each solution of (3.9) is a solution a solution to (1.2) and the hemivariational inequality (1.4). Moreover, if j is regular in the sense of Clarke (cf. [29]), then they are equivalent.

Lemma 3.1. *Under the hypotheses H(A) and H(β)(i)–(iii), the sum* $A+\overline{\partial} J : V \to 2^{V^*}$ *is a bounded and pseudomonotone operator. Moreover, it is coercive in the sense that*

$$\liminf_{\|v\| \to +\infty,\ \eta \in A(v) + \overline{\partial} J(v)} \frac{\langle \eta, v - y_0 \rangle}{\|v\|} = +\infty. \tag{3.10}$$

Proof. From (A1) and (3.5) we see that $A + \overline{\partial} J$ is bounded, i.e., it maps each bounded subset of V into a bounded subset of V^*. Next, we first of all claim that $\overline{\partial} J : V \to 2^{V^*}$ is pseudomonotone. In fact, from the properties of Clarke generalized gradient, $\overline{\partial} J(u)$ is a nonempty, convex and a bounded subset of H for every $u \in H$ and the graph of $\overline{\partial} J$ is sequentially closed in $H \times w$-H. Due to $H \subset V^*$ compactly, the conditions in Definition 2.3 can be easily verified with T and X replaced by $\overline{\partial} J$ and V, respectively. Therefore, $\overline{\partial} J$ is pseudomonotone, as claimed. On the other hand, A is pseudomonotone because it is monotone and hemicontinuous. As the pseudomonotonicity is invariant under addition we thereby have that $A + \overline{\partial} J$ is pseudomonotone. Returning now to the coerciveness, we take $v \in V$, $\eta = Av + \xi$ with $\xi \in \overline{\partial} J(v)$ and find

$$\langle \eta, v - y_0 \rangle = \langle Av, v \rangle + (\xi, v) - \langle Av + \xi, y_0 \rangle. \tag{3.11}$$

Observe that $(\xi, -v) \leq J^0(v; -v)$ by definition. It follows $(\xi, v) \geq -J^0(v; -v)$. As y_0 is fixed, we deduce (3.10) from (A1), (A2), (3.5) and (3.6). □

Theorem 3.2. *Assume $u \in U$ and the hypotheses H(A), H(β)(i)–(iii) and H(φ) hold. Then the inclusion (3.9) admits at least one solution.*

Proof. As $\partial\varphi$ is a maximal monotone operator, this theorem follows directly from Lemmas 3.1 and 2.4. □

Remark 3.3. *The solution to ([3.9]) is in general not unique due to the non-monotonicity of* $\overline{\partial} J$.

Following Theorem 3.2 we denote by $S(u)$ the solution set of (3.9) for $u \in U$.

Lemma 3.4. *Under the assumptions of Theorem 3.2, we have*

$$\|y\|_V \le C(1 + \|u\|_U), \ \forall\, y \in S(u), \tag{3.12}$$

where C *is a constant depending on* c_1, c_2, c_3, b_1, b_2, f *and* y_0.

Proof. Let $y \in S(u)$. Then there exist $\xi \in \overline{\partial} J(u)$ and $\eta \in \partial\varphi(u)$ such that

$$Ay + \xi + \eta = f + Bu \ \text{ in } V^*. \tag{3.13}$$

Multiplying by y in the above equation, we find

$$\langle Ay + \xi + \eta, y\rangle = \langle f + Bu, y\rangle. \tag{3.14}$$

Then taking any $\eta_0 \in \partial\varphi(y_0)$ we can compute

$$\begin{aligned} \langle \eta, y\rangle &= \langle \eta - \eta_0, y - y_0\rangle + \langle \eta_0, y - y_0\rangle + \langle \eta, y_0\rangle \\ &\ge \langle \eta_0, y - y_0\rangle + \langle f + Bu - Ay - \xi, y_0\rangle \end{aligned} \tag{3.15}$$

since $\partial\varphi$ is monotone. We finally combine H(A), (3.5), (3.6), (3.14) and (3.15) to discover

$$\begin{aligned} c_2\|y\|^2 - b_2 k^\sigma \|y\|^\sigma \le{}& \|y\|\big(\|f\| + \|Bu\| + \|\eta_0\| + b_1 k\|y_0\| + c_1\|y_0\|\big) \\ &+ c_3 + b_2 + \|y_0\|\big(\|\eta_0\| + \|f\| + \|Bu\| + b_1 k + c_1\big) \end{aligned}$$

where k is the operator norm of $i : V \to H$. This implies (3.12) since B is bounded and $1 \le \sigma < 2$. ☐

In what follows, for simplicity, C always denotes a constant but may change from line to line.

Theorem 3.5. *Under H(B) and the assumptions of Theorem 3.2, the mapping* $u \mapsto S(u)$ *is weakly upper semicontinuous from* U *to* V.

Proof. Let D be arbitrary weakly closed set of V. We aim to show that

$$S^{-1}(D) := \{u \in U : S(u) \cap D \neq \emptyset\}$$

is weakly closed in U. To this end, we take any sequence $u_n \in S^{-1}(D)$ such that $u_n \to u$ weakly in U and aim at proving that $u \in S^{-1}(D)$. First of all, as $u_n \in S^{-1}(D)$ we can select $y_n \in V$ with $y_n \in S(u_n)$. Then there exist $\xi_n \in \overline{\partial} J(y_n)$ and $\eta_n \in \partial\varphi(y_n)$ such that

$$Ay_n + \xi_n + \eta_n = f + Bu_n, \ n = 1, 2, \cdots . \tag{3.16}$$

Using (A1), (3.5) and Lemma 3.4 we conclude that y_n, Ay_n and ξ_n are bounded in V, V^* and H, respectively. Then we further see from (3.16) that η_n is bounded in V^* since B is continuous. Therefore, by passing to a subsequence again denoted by n, there exist $y \in V$, $\zeta \in V^*$, $\xi \in H$ and $\eta \in V^*$ such that

$$y_n \to y \qquad \text{weakly in } V \tag{3.17}$$
$$Ay_n \to \zeta \qquad \text{weakly in } V^* \tag{3.18}$$
$$\xi_n \to \xi \qquad \text{weakly in } H \tag{3.19}$$
$$\eta_n \to \eta \qquad \text{weakly in } V^*. \tag{3.20}$$

From (3.17) we further have $y_n \to y$ in H. Since the graph of $\overline{\partial} J$ is sequentially closed in $H \times w\text{-}H$, we obtain $\xi \in \overline{\partial} J(y)$. Multiplying Eq. (3.16) by $y_n - y$ yields

$$\langle Ay_n + \xi_n + \eta_n, y_n - y\rangle = \langle f + Bu_n, y_n - y\rangle. \tag{3.21}$$

Note that the term on the right-hand side tends to zero because B is compact and $y_n \to y$ weakly in V. From (3.19) and $y_n \to y$ in H, it follows that

$$\lim_{n\to\infty} \langle \xi_n, y_n - y\rangle = \lim(\xi_n, y_n - y) = 0.$$

Moreover, as A is monotone, we have

$$\limsup_{n\to\infty} \langle Ay_n, y_n - y\rangle = \limsup_{n\to\infty} \langle Ay_n - Ay, y_n - y\rangle + \lim_{n\to\infty} \langle Ay, y_n - y\rangle \geq 0.$$

Thus we obtain $\limsup\langle \eta_n, y_n - y\rangle \leq 0$ from (3.21). Recalling (3.17) and (3.20) we therefore have $y \in D(\partial\varphi)$ and $\eta \in \partial\varphi(y)$ because $\partial\varphi$ is maximal monotone. Consequently, we now see

$$\limsup_{n\to\infty} \langle \eta_n, y_n - y\rangle = \limsup_{n\to\infty} \langle \eta_n - \eta, y_n - y\rangle + \lim_{n\to\infty} \langle \eta, y_n - y\rangle \geq 0.$$

Then we conclude from (3.21) that

$$\limsup_{n\to\infty}\langle Ay_n, y_n - y\rangle \le 0.$$

In view of (3.17), (3.18) and Remark 2.1, we therefore get $\zeta = Ay$. Passing to the limit in (3.16) we finally obtain

$$Ay + \xi + \eta = f + Bu,$$

with $\xi \in \overline{\partial} J(y)$ and $\eta \in \partial\varphi(y)$. This gives $y \in S(u)$, i.e., $u \in S^{-1}(D)$. □

Next, we turn to the existence of optimal pairs of (OP). We begin with the assumptions on the cost function.

H(h) $h : U \to \mathbb{R} \cup \{+\infty\}$ is a proper convex and lower semicontinuous functional satisfying

$$\lim_{\|u\|_U\to+\infty} h(u)/\|u\|_U = +\infty. \tag{3.22}$$

H(g) $g : H \to \mathbb{R}$ is a locally Lipschitz continuous functional (Lipschitz continuous on each bounded subset of H) and bounded from below by an affine function, i.e.,

$$g(y) \ge \langle w, y\rangle + C,\ \forall\, y \in H, \tag{3.23}$$

where $w \in H$ and C is a constant.

Theorem 3.6. *Suppose assumptions H(A), H(β)(i)–(iii), H(φ), H(B), H(h) and H(g) hold. Then there exists an optimal pair to problem (1.1)–(1.2).*

Proof. We see from H(h), H(g) and (3.12) that the cost function $G(y,u)$ is bounded from below. Then we set

$$d = \inf\left\{g(y^u) + h(u) : u \in U, y^u \in S(u)\right\} > -\infty. \tag{3.24}$$

Let $(u_n, y_n) \in U \times V$ be a minimization sequence of (3.24) such that

$$d \le g(y_n) + h(u_n) \le d + 1/n, \tag{3.25}$$

where $y_n := y^{u_n} \in S(u_n)$. Using H($h$), H($g$) and (3.12) once more, we deduce u_n and y_n are bounded in U and V, respectively. By passing to a subsequence, we may assume $u_n \to u$, $y_n \to y^u$ weakly in U and V, respectively. From Theorem 3.5, it is readily seen that the mapping $u \mapsto S(u)$ has a sequentially closed graph in w-$U \times w$-V. We therefore obtain $y^u \in S(u)$. Returning now to (3.25), we find

$$d \geq \liminf_{n\to\infty}\{g(y_n) + h(u_n)\} \geq \lim_{n\to\infty} g(y_n) + \liminf_{n\to\infty} h(u_n) \geq g(y^u) + h(u),$$

where H(h), H(g), $y_n \to y^u$ in H and $u_n \to u$ weakly in U are used. This shows that (y^u, u) is an optimal pair of (1.1)–(1.2). □

4. Approximation results and necessary optimality conditions

In this section, we study the optimality system of optimal control problem (1.1)–(1.2). This will be achieved by smooth approximations. It is then crucial to analyze the relations between approximate systems and the original ones. Note that in contrast to the existence of optimal pairs, the derivation of the necessary optimality conditions is in general much more complicated. Only a few studies have addressed this topic for hemivariational inequalities.

Lemma 4.1. *Assume $X = L^p(\Omega)$ with $1 < p < +\infty$, and $q(x,s) : \Omega \times \mathbb{R} \to \mathbb{R}$ is a measurable function such that for all $s, t \in \mathbb{R}$ and a.e. $x \in \Omega$*

$$|q(x,s) - q(x,t)| \leq c_q(x)(1 + |t|^{p-1} + |s|^{p-1})|t - s| \quad \text{with } c_q \in L^\infty_+(\Omega). \tag{4.1}$$

Let $\rho_\varepsilon(\tau) = \varepsilon^{-1}\rho(\tau/\varepsilon)$, where $\rho \in C_0^\infty(\mathbb{R})$ is the standard mollifier in $\mathbb{R}$. Define functionals Q and Q_ϵ by

$$Q(v) = \int_\Omega q(x, v(x))dx \quad \text{and} \quad Q_\varepsilon(v) = \int_\Omega q_\epsilon(x, v(x))dx \ \forall\, v \in X, \tag{4.2}$$

respectively, where $q_\varepsilon(x,s)$ is given by

$$q_\varepsilon(x,s) = \int_{-\infty}^{\infty} q(x, s-\tau)\rho_\varepsilon(\tau)d\tau. \tag{4.3}$$

Then Q_ε is Fréchet differentiable and

$$\lim_{\varepsilon\to 0} Q_\varepsilon(v) = Q(v) \ \forall\, v \in X. \tag{4.4}$$

Moreover, if $v_\varepsilon \to v$ in X and

$$\nabla Q_\varepsilon(v_\varepsilon) \to \xi \text{ weakly in } X^*, \tag{4.5}$$

then we have $\xi \in \overline{\partial} Q(v)$.

Proof. We rewrite Q_ε as

$$Q_\varepsilon(v) = \int_\Omega \int_{-\infty}^{\infty} q(x,\tau)\rho_\varepsilon(v(x)-\tau)d\tau dx, \tag{4.6}$$

from which it follows that Q_ε is Fréchet differentiable and ∇Q_ε is continuous. In view of (4.1), it follows that (4.4) holds and Q is Lipschitz continuous on each bounded subset of X. Therefore, the Clarke generalized derivative and gradient of Q are well defined. On the other hand, using Fubini's theorem we see that

$$\frac{(Q_\varepsilon(v_\varepsilon+\lambda z)-Q_\varepsilon(v_\varepsilon))}{\lambda} = \int_{-\infty}^{\infty} \frac{(Q(v_\varepsilon-\varepsilon\tau+\lambda z)-Q(v_\varepsilon-\varepsilon\tau))\rho(\tau)d\tau}{\lambda},$$

where $\varepsilon\tau$ is seen as a constant-valued function in X. Hence letting λ tend to zero in the last equation and using Fatou's lemma we have

$$\langle \nabla Q_\varepsilon(v_\varepsilon), z\rangle \le \int_{-\infty}^{\infty} Q^0(v_\varepsilon-\varepsilon\tau; z)\rho(\tau)d\tau. \tag{4.7}$$

Since the Clarke generalized gradient $Q^0 : X\times X \to \mathbb{R}$ is upper semicontinuous and $v_\varepsilon - \varepsilon\tau$ converges to v in X, we conclude from (4.5) that

$$\langle \xi, z\rangle \le Q^0(v; z), \tag{4.8}$$

which implies $\xi \in \overline{\partial} Q(v)$, as claimed. □

We now define function $J_\varepsilon : H \to \mathbb{R}$ by

$$J_\varepsilon(v) = \int_\Omega \int_0^{v(x)} \beta_\varepsilon(x,s)dsdx,\ \forall\, v \in H, \tag{4.9}$$

where $\beta_\varepsilon(x,s)$ is defined in the same way as in (4.3). Then a straightforward calculation gives

$$J_\varepsilon(v) = J_{1\varepsilon}(v) + M_1(\varepsilon,\beta) \tag{4.10}$$

where $J_{1\varepsilon}(v) = \int_\Omega j_\varepsilon(x, v(x))dx$ and $M_1(\varepsilon,\beta)$ is a constant defined by

$$M_1(\varepsilon,\beta) = \int_\Omega \int_{-\varepsilon}^{\varepsilon} \int_0^{\tau} \beta(x, s-\tau)\rho_\varepsilon(\tau)dsd\tau dx. \tag{4.11}$$

Clearly, $M_1(\varepsilon,\beta) \to 0$ as $\varepsilon \to 0$. According to H(β)(i)-(ii), Lemma 4.1 is applicable to $J_{1\varepsilon}$, and thus from (4.10) we have the following result.

Lemma 4.2. *Let the assumptions $H(\beta)$(i)-(ii) be satisfied and $J_\varepsilon : H \to \mathbb{R}$ be defined as above. Then J_ε is Fréchet differentiable and*

$$\lim_{\varepsilon\to 0} J_\varepsilon(v) = J(v) \ \forall\, v \in H. \tag{4.12}$$

Moreover, if $v_\varepsilon \to v$ in H and

$$\nabla J_\varepsilon(v_\varepsilon) \to \xi \text{ weakly in } H, \tag{4.13}$$

then we have $\xi \in \overline{\partial} J(v)$.

Lemma 4.3. *Suppose the hypotheses $H(\beta)$(i)-(iii) hold. Assume furthermore there exist a sufficiently small ε_0 and a sufficient large c_0 such that*

$$|\beta(x,s) - \beta(x,\tau)| \le c_0 \quad \text{a.e. } x \in \Omega, \forall\, s,\tau \in \mathbb{R} \text{ with } |s-\tau| \le \varepsilon_0. \tag{4.14}$$

Then J_ε has the same properties as J in (3.5) and (3.6) for $0 < \varepsilon \le \varepsilon_0$. Besides, from $H(\beta)$(iv) it follows $\dot{\beta}_\varepsilon(x,s) \ge -m$ where the dot above β_ε means the derivative with respect to s, and we have

$$(\nabla J_\varepsilon(v_1) - \nabla J_\varepsilon(v_2), v_1 - v_2) \ge -m\|v_1 - v_2\|_H^2, \ \forall\, v_1, v_2 \in H. \tag{4.15}$$

Proof. Since J_ε is smooth we have $\overline{\partial} J_\varepsilon(v) = \nabla J_\varepsilon(v)$. Then a straightforward calculation gives

$$|\beta_\varepsilon(x,s) - \beta(x,s)| \le \sup_{|\tau - s|\le\varepsilon} |\beta(x,\tau) - \beta(x,s)|, \tag{4.16}$$

and thus the property (3.5) with J replaced by J_ε follows from H(β)(ii), and (3.6) follows from (4.14) and H(β)(ii) (the constants may be changed but independent of ε with $\varepsilon \le \varepsilon_0$). On the other hand, we see from H(β)(iv) that for every $h \in \mathbb{R}$

$$\frac{\beta_\varepsilon(x,s+h) - \beta_\varepsilon(x,s)}{h} = \int_{-\varepsilon}^{\varepsilon} \frac{\beta(x,s+h-\tau) - \beta(x,s-\tau)}{h} \rho_\varepsilon(\tau) d\tau \ge -m.$$

This implies $\dot{\beta}_\varepsilon(x,s) \ge -m$ as h goes to zero. Moreover, from the above inequality, we also have

$$\begin{aligned}(\nabla J_\varepsilon(v_1) - \nabla J_\varepsilon(v_2), v_1 - v_2) &= \int_\Omega (\beta_\varepsilon(x,v_1) - \beta_\varepsilon(x,v_2))(v_1(x) - v_2(x))dx \\ &\ge -m \int_\Omega (v_1(x) - v_2(x))^2 dx \\ &= -m\|v_1 - v_2\|_H^2.\end{aligned}$$

This completes the proof. $\square$

In what follows, we always assume $0 < \varepsilon \leq \varepsilon_0$. We are now introduce the smooth approximation of g, see, e.g., [3]. Suppose that $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis of H and $H_n = \mathrm{Span}\{e_1, e_2, \cdots, e_n\}$. Let $P_n : H \to H_n$ and $\Lambda_n : \mathbb{R}^n \to H_n$ be given by

$$P_n v = \sum_{i=1}^{n} (v, e_i) e_i, \ \forall\, v \in H, \text{ and } \Lambda_n(\tau) = \sum_{i=1}^{n} \tau_i e_i, \ \tau = (\tau_1, \cdots, \tau_n), \tag{4.17}$$

respectively. Then we define the smooth functional $g_\varepsilon : H \to \mathbb{R}$ by

$$g_\varepsilon(v) = \int_{\mathbb{R}^n} g(P_n v - \varepsilon \Lambda_n \tau) \rho_n(\tau) d\tau, \tag{4.18}$$

where $n = [1/\varepsilon]$ and $\rho_n \in C_0^\infty(\mathbb{R}^n)$ is the standard mollifier in $\mathbb{R}^n$. From H(g), it follows that for all $y \in H$

$$g_\varepsilon(y) \geq \int_{|\tau| \leq 1} \langle w, P_n y - \varepsilon \Lambda_n \tau \rangle \rho_n(\tau) d\tau + C \geq -\|w\| \|y\| + C, \tag{4.19}$$

where w is given in (3.23). On the other hand, as g is locally Lipschitz continuous, it is readily seen from (4.18) that g_ε is still locally Lipschitz continuous with the same Lipschitz bound as g, uniformly for ε. Moreover, g_ε has the same properties as J_ε in Lemma 4.2.

Next, using H(φ) we can choose a family of convex functions $\varphi_\varepsilon : V \to \mathbb{R}$ of class C^2 such that

$$\varphi_\varepsilon(v) \geq -C(1 + \|v\|_V), \quad \lim_{\varepsilon \to 0} \varphi_\varepsilon(v) = \varphi(v), \tag{4.20}$$

and for any sequence $v_\varepsilon \to v$ weakly in V, one has

$$\liminf_{\varepsilon \to 0} \varphi_\varepsilon(v_\varepsilon) \geq \varphi(v). \tag{4.21}$$

Suppose that $h_\varepsilon : U \to \mathbb{R}$ approximates h and satisfies the same properties as φ_ε above.

We are now in a position to consider the approximate system $(\mathrm{OP})_\varepsilon$: Minimize the functional

$$G_\varepsilon(y, u) = g_\varepsilon(y) + h_\varepsilon(u) + \frac{1}{2}\|u - u^*\|_U^2 \tag{4.22}$$

on all $(y,u)\in V\times U$, subject to

$$Ay+\nabla J_\varepsilon(y)+\nabla\varphi_\varepsilon(y)=f+Bu, \tag{4.23}$$

where u^* is chosen from an optimal pair (y^*,u^*) of (1.1)–(1.2). In order to obtain the necessary optimality we substitute (A2)$'$ for (A2).

(A2)$'$ A is linear and there exists a constant $c_2>mk^2$ such that

$$\langle Av,v\rangle\ge c_2\|v\|_V^2,\ \forall\, v\in V. \tag{4.24}$$

Note that $H(\beta)$(iv) and (A2)$'$ are necessary to obtain the existence and boundedness of the solutions to the adjoint equation in the sequel.

Theorem 4.4. *Suppose $H(\beta)$(i)-(iv), $H(\varphi)$, $H(B)$, (A1) and (A2)$'$ are satisfied. Then problem (3.9) admits a unique solution and the mapping $u\mapsto S(u)$ is weakly-strongly continuous. On the other hand, assume that J_ε, φ_ε are defined as above and (4.14) holds, then the previous assertion holds for problem (4.23). Moreover, the solution y_ε^u of (4.23) has the bound*

$$\|y_\varepsilon^u\|_V\le C(1+\|u\|_U), \tag{4.25}$$

where C depends on $\hat b_1,\hat b_2,c_1,c_2,c_3,f,y_0$ but is independent of ε.

Proof. From Theorem 3.2, we see that $S(u)$ is nonempty for every $u\in U$. Assume y_1^u, $y_2^u\in S(u)$. Then there exist $\xi_i\in\overline{\partial}J(y_i^u)$ and $\eta_i\in\partial\varphi(y_i^u)$ such that

$$Ay_i^u+\xi_i+\eta_i=f+Bu,\ \ i=1,2. \tag{4.26}$$

Multiplying this equation by $y_1^u-y_2^u$ for $i=1,2$, we deduce that

$$\langle Ay_1^u-Ay_2^u,y_1^u-y_2^u\rangle+(\xi_1-\xi_2,y_1^u-y_2^u)+\langle\eta_1-\eta_2,y_1^u-y_2^u\rangle=0. \tag{4.27}$$

Using (A2)$'$, (3.7) and the monotonicity of $\partial\varphi$, we have

$$(c_2-mk^2)\|y_1^u-y_2^u\|_V^2\le 0.$$

This implies the uniqueness of the solution to (3.9). Next, assume that $u_n\to u$ weakly in U and $y_n=S(u_n)$. We aim to prove $y_n\to y$ and $y=S(u)$. In fact, we can use the same procedures as in Theorem 3.5 to obtain (3.18)–(3.20) and $y=S(u)$. Then, for $y_n=S(u_n)$ and $y_m=S(u_m)$ we have

$$\langle Ay_n-Ay_m+\xi_n-\xi_m+\eta_n-\eta_m,u_n-u_m\rangle=\langle Bu_n-Bu_m,u_n-u_m\rangle.$$

Using (A2)$'$ and (3.7) again, we therefore see that y_n is a Cauchy sequence in V as B is compact and $\partial\varphi$ is monotone. It follows that y_n converges to some y in V and $y = S(u)$ from Theorem 3.5. The conclusions for (4.23) can now be proved in the same way as above due to Lemma 4.3. Moreover, the bound (4.25) can be obtained as in the proof of Lemma 3.12. $\square$

Theorem 4.5. *Let J_ε, g_ε and φ_ε be defined as above. Assume that H(h), H(g), H(β)(i)-(iv), H(φ), H(B), (4.14), (A1) and (A2)$'$ hold. Then, (OP)$_\varepsilon$ admits at least one optimal pair $(y_\varepsilon, u_\varepsilon)$. Moreover, $u_\varepsilon \to u^*$ in U and $y_\varepsilon \to y^*$ weakly in V, where (y^*, u^*) is some optimal pair of (OP).*

Proof. Due to (4.19) and Theorem 4.4, the existence of $(y_\varepsilon, u_\varepsilon)$ can be proved in the same way as in the proof of Theorem 3.6. Now we aim to show that $u_\varepsilon \to u^*$ in U and $y_\varepsilon \to y^*$ weakly in V. Let $y_\varepsilon^{u^*}$ be the solution to (4.23) with u replaced by u^*. From (4.25), up to a subsequence, we have $y_\varepsilon^{u^*} \to y^{u^*}$ weakly in V and strongly in H. Since $(y_\varepsilon, u_\varepsilon)$ is an optimal pair, it follows that

$$g_\varepsilon(y_\varepsilon) + h(u_\varepsilon) + \frac{1}{2}\|u_\varepsilon - u^*\|_U^2 \leq g_\varepsilon(y_\varepsilon^{u^*}) + h(u^*). \tag{4.28}$$

Observe that

$$\begin{aligned}|g_\varepsilon(y_\varepsilon^{u^*}) - g(y^{u^*})| &\leq |g_\varepsilon(y_\varepsilon^{u^*}) - g_\varepsilon(y^{u^*})| + |g_\varepsilon(y^{u^*}) - g(y^{u^*})| \\ &\leq L\|y_\varepsilon^{u^*} - y^{u^*}\|_H + |g_\varepsilon(y^{u^*}) - g(y^{u^*})|\end{aligned}$$

for ε small enough, where L is the Lipschitz constant of g on the bounded subset $y_\varepsilon^{u^*}$. It follows that

$$\lim_{\varepsilon\to 0} g_\varepsilon(y_\varepsilon^{u^*}) = g(y^{u^*}). \tag{4.29}$$

Taking H(h), (4.19), (4.25), (4.28) and (4.29) into account, we have that u_ε and y_ε are bounded independently of ε in U and V, respectively. Recall that

$$Ay_\varepsilon + \nabla J_\varepsilon(y_\varepsilon) + \nabla\varphi_\varepsilon(y_\varepsilon) = f + Bu_\varepsilon. \tag{4.30}$$

Using hypothesis (A1), H(B) and Lemma 4.3, we have that Ay_ε, $\nabla J_\varepsilon(y_\varepsilon)$, and thus $\nabla\varphi_\varepsilon(y_\varepsilon)$ are bounded in V^*, H, and V^*, respectively. Consequently, up to another subsequence, there exist $\bar{u}, \bar{y}$, ξ and η such that

$$y_\varepsilon \to \bar{y} \qquad \text{weakly in } V \tag{4.31}$$

$$
\begin{aligned}
&u_\varepsilon \to \overline{u} && \text{weakly in } U && (4.32)\\
&Ay_\varepsilon \to A\overline{y} && \text{weakly in } V^* && (4.33)\\
&\nabla J_\varepsilon(y_\varepsilon) \to \xi && \text{weakly in } H && (4.34)\\
&\nabla \varphi_\varepsilon(y_\varepsilon) \to \eta && \text{weakly in } V^*. && (4.35)
\end{aligned}
$$

Passing to the limit in Eq. (4.30) we have

$$
A\overline{y} + \xi + \eta = f + B\overline{u} \quad \text{in } V^*. \tag{4.36}
$$

From (4.31), $y_\varepsilon \to \overline{y}$ in H. We thus get $\xi \in \overline{\partial} J(\overline{y})$ by using (4.34) and Lemma 4.2. Note that

$$
\begin{aligned}
&\limsup_{\varepsilon\to 0}\langle \nabla\varphi_\varepsilon(y_\varepsilon), y_\varepsilon - \bar{y}\rangle\\
=&\limsup_{\varepsilon\to 0}\langle f + Bu_\varepsilon - Ay_\varepsilon - \nabla J_\varepsilon(y_\varepsilon), y_\varepsilon - \bar{y}\rangle\\
=&\limsup_{\varepsilon\to 0}\langle A(y - y_\varepsilon) - Ay, y_\varepsilon - \bar{y}\rangle\\
\leq& 0.
\end{aligned} \tag{4.37}
$$

Here (4.31), (4.32), (4.34), (A2)$'$ and H(B) are used. From (4.35), it follows that

$$
\liminf_{\varepsilon\to 0}\langle \nabla\varphi_\varepsilon(y_\varepsilon), -y_\varepsilon\rangle \geq -\langle \eta, \bar{y}\rangle, \tag{4.38}
$$

and therefore we have

$$
\liminf_{\varepsilon\to 0}\langle \nabla\varphi_\varepsilon(y_\varepsilon), w - y_\varepsilon\rangle \geq \langle \eta, w - \bar{y}\rangle,\ \forall\, w \in V. \tag{4.39}
$$

On the other hand,

$$
\varphi_\varepsilon(w) - \varphi_\varepsilon(y_\varepsilon) \geq \langle \nabla\varphi_\varepsilon(y_\varepsilon), w - y_\varepsilon\rangle,\ \forall\, w \in V. \tag{4.40}
$$

Taking the upper and lower limits on the left and right hand sides of this inequality, respectively, and using (4.20), (4.21), (4.31), and (4.39), we have

$$
\varphi(w) - \varphi(\bar{y}) \geq \langle \eta, w - \bar{y}\rangle,\ \forall\, w \in V. \tag{4.41}
$$

This implies that $\eta \in \partial\varphi(\bar{y})$, and thus we now have $\bar{y} = S(\bar{u})$. To finish the proof, it suffice to prove that $y^* = \bar{y}, u^* = \bar{u}$. To this end, passing to the lower limit and limit on the left and right hand sides of (4.28), respectively, we have

$$
g(\bar{y}) + h(\bar{u}) + \frac{1}{2}\|\bar{u} - u^*\|_U^2 \leq g(y^{u^*}) + h(u^*). \tag{4.42}
$$

We claim that $y^{u^*}{=}y^*$. In fact, this can be proved in the same way as that we used just to show $\bar{y} = S(\bar{u})$, and the proof is much simpler since u^*, instead of u_ε, is fixed. Therefore, we further have $y^* = \bar{y}$ and $u^* = \bar{u}$ since (y^*, u^*) is an optimal pair. Returning now to (4.28), we see that

$$h(u^*) \leq \liminf_{\varepsilon \to 0} \big(h(u_\varepsilon) + \frac{1}{2} \|u_\varepsilon - u^*\|_U^2 \big) \leq h(u^*). \tag{4.43}$$

This gives

$$\lim_{\varepsilon \to 0} \|u_\varepsilon - u^*\|_U^2 = 0,$$

i.e., $u_\varepsilon \to u^*$ in U. Finally, by a standard argument, we conclude that $y_\varepsilon \to y^*$ weakly in V and $u_\varepsilon \to u^*$ in U hold for the whole sequence. This completes the proof. $\square$

Let A^* and B^* be the adjoint linear operators of A and B, respectively.

Theorem 4.6. *Assume that the hypotheses of Theorem 4.5 hold, and $(y_\varepsilon, u_\varepsilon)$, (y^*, u^*) are given as in that Theorem. Then there exists $p_\varepsilon \in V$ such that*

$$\begin{cases} A^* p_\varepsilon + \nabla^2 J_\varepsilon(y_\varepsilon) p_\varepsilon + \nabla^2 \varphi_\varepsilon(y_\varepsilon) p_\varepsilon = \nabla g_\varepsilon(y_\varepsilon) \\ B^* p_\varepsilon + \nabla h_\varepsilon(u_\varepsilon) + R(u_\varepsilon - u^*) = 0 \\ A y_\varepsilon + \nabla J_\varepsilon(y_\varepsilon) + \nabla \varphi_\varepsilon(y_\varepsilon) = f + B u_\varepsilon \end{cases} \tag{4.44}$$

where R is the duality mapping in U. Moreover, p_ε weakly converges to some p^* in V and we have

$$\begin{cases} A^* p^* + \zeta \in \overline{\partial} g(y^*) \\ B^* p^* + \partial h(u^*) \ni 0 \\ A y^* + \overline{\partial} J(y^*) + \partial \varphi(y^*) \ni f + B u^* \end{cases} \tag{4.45}$$

where ζ is the weak limit of $\nabla^2 J_\varepsilon(y_\varepsilon) p_\varepsilon + \nabla^2 \varphi_\varepsilon(y_\varepsilon) p_\varepsilon$ in V^*.

Note that (4.44) and (4.45) are referred to as the optimality systems of $(\mathrm{OP})_\varepsilon$ and (OP), respectively. The first and third equations and inclusions are understood to be satisfied in V^*, and the middle ones in U^*.

Proof. From Theorem 4.5 we know the existence of an optimal pair $(y_\varepsilon, u_\varepsilon)$ for every $\varepsilon > 0$. To verify (4.44) we use a Lagrangian approach. For this purpose we define the function $F_\varepsilon : V \times U \to V^*$,

$$F_\varepsilon(y, u) := Ay + \nabla J_\varepsilon(y) + \nabla \varphi_\varepsilon(y) - f - Bu,$$

and the Lagrange functional $L_\varepsilon : V \times U \times V \to \mathbb{R}$,

$$L_\varepsilon(y,u,p) := g_\varepsilon(y) + h_\varepsilon(u) + \frac{1}{2}\|u-u^*\|_U^2 - \langle F_\varepsilon(y,u), p\rangle.$$

Let us now argue that $F_\varepsilon'(y_\varepsilon, u_\varepsilon)$ is a bounded and linear operator from $V \times U$ onto V^*. This will follow from the surjectivity of $\partial F_\varepsilon(y_\varepsilon, u_\varepsilon)/\partial y$. Clearly,

$$\partial F_\varepsilon(y_\varepsilon, u_\varepsilon)/\partial y = A^* + \nabla^2 J_\varepsilon(y_\varepsilon) + \nabla^2 \varphi_\varepsilon(y_\varepsilon).$$

We now deduce that

$$\langle A^* p + \nabla^2 J_\varepsilon(y_\varepsilon)p + \nabla^2\varphi_\varepsilon(y_\varepsilon)p, p\rangle \geq (c_2 - k^2 m)\|p\|_V^2 \quad \forall\, p \in V, \tag{4.46}$$

where (A2)$'$, Lemma 4.3 as well as the positivity of $\nabla^2\varphi_\varepsilon(y_\varepsilon)$ are used. Then we conclude that, for arbitrary $v \in V^*$, the equation

$$A^* p + \nabla^2 J_\varepsilon(y_\varepsilon)p + \nabla^2\varphi_\varepsilon(y_\varepsilon)p = v$$

admits a unique solution by Lax-Milgram theorem. Therefore, there exists a $p_\varepsilon \in V$ and the necessary optimality conditions of $(\mathrm{OP})_\varepsilon$ are given by

$$\partial L_\varepsilon(y_\varepsilon, u_\varepsilon, p_\varepsilon)/\partial y = 0, \quad \partial L_\varepsilon(y_\varepsilon, u_\varepsilon, p_\varepsilon)/\partial u = 0, \;\; \partial L_\varepsilon(y_\varepsilon, u_\varepsilon, p_\varepsilon)/\partial p = 0,$$

from which it follows that $(y_\varepsilon, u_\varepsilon, p_\varepsilon)$ satisfies (4.44). Next, taking the duality product in $V \times V^*$ with p_ε in the first equation of (4.44) we find

$$\big\langle A p_\varepsilon + \nabla^2 J_\varepsilon(y_\varepsilon)p_\varepsilon + \nabla^2\varphi_\varepsilon(y_\varepsilon)p_\varepsilon, p_\varepsilon\big\rangle = \big(\nabla g_\varepsilon(y_\varepsilon), p_\varepsilon\big). \tag{4.47}$$

Then replacing p by p_ε in (4.46) we further deduce that

$$(c_2 - k^2 m)\|p_\varepsilon\|_V \leq \|\nabla g_\varepsilon(y_\varepsilon)\|_H. \tag{4.48}$$

Because $y_\varepsilon \to y^*$ in H, and g_ε is Lipschitz continuous on each bounded subset of H uniformly for ε, we have that $\|\nabla g_\varepsilon(y_\varepsilon)\|_H$ is bounded. Therefore p_ε is bounded independently of ε in V . By taking subsequences, we have

$$\begin{array}{ll} p_\varepsilon \to p^* & \text{weakly in } V \\ A^* p_\varepsilon \to A^* p^* & \text{weakly in } V^* \\ \nabla g_\varepsilon(y_\varepsilon) \to w & \text{weakly in } H \\ \nabla h_\varepsilon(y_\varepsilon) \to \eta & \text{weakly in } V^*. \end{array} \tag{4.49}$$

From Theorem 4.5, we see that $y_\varepsilon \to y^*$ weakly in V and $u_\varepsilon \to u^*$ in U. Thus $w \in \overline{\partial} g(y^*)$. On the other hand,

$$h_\varepsilon(v) - h_\varepsilon(u_\varepsilon) \geq \langle \nabla h_\varepsilon(u_\varepsilon), v - u_\varepsilon \rangle = -\langle B^* p_\varepsilon + F(u_\varepsilon - u^*), v - u_\varepsilon \rangle. \quad (4.50)$$

As $\lim_{\varepsilon \to 0} h_\varepsilon(v) = h(v)$ and $\liminf_{\varepsilon \to 0} h_\varepsilon(u_\varepsilon) \geq h(u)$, taking upper limit in (4.50) we have

$$h(v) - h(u^*) \geq \langle -B^* p^*, v - u^* \rangle, \forall\, v \in V, \quad (4.51)$$

where $u_\varepsilon \to u^*$ in U and $p_\varepsilon \to p$ weakly in V are used. Therefore, we get $-B^* p^* \in \partial h(u^*)$. Since $\nabla g_\varepsilon(y_\varepsilon)$ and $A^* p_\varepsilon$ are bounded in V^*, from (4.44) and up to another subsequence, we have

$$\nabla^2 J_\varepsilon(y_\varepsilon) p_\varepsilon + \nabla^2 \varphi_\varepsilon(y_\varepsilon) p_\varepsilon \to \zeta \quad \text{weakly in } V^*. \quad (4.52)$$

Thus we finally obtain (4.45). This completes the proof. □

Remark 4.7. *An important open question is whether one can further improve the weak limit ζ in the abstract necessary optimality system (4.45) under some appropriate conditions.*

5. Optimal control of obstacle problem with nonmonotone perturbation

In this section, we take the obstacle problem as an example to further explain the study of the abstract results in the last section. We still consider the optimal control problem (OP) but the state system is specifically given by

$$\text{(VI-HVI-1)} \begin{cases} Ay + \alpha(y - \psi) + \beta(\cdot, y) \ni f + Bu \text{ in } \Omega \\ y = 0 \quad \text{on } \partial\Omega. \end{cases} \quad (5.1)$$

Here $\alpha : \mathbb{R} \to 2^{\mathbb{R}}$ is a maximal monotone operator defined by

$$\alpha(s) = \begin{cases} 0, & s > 0 \\ \mathbb{R}^-, & s = 0 \\ \emptyset, & s < 0. \end{cases} \quad (5.2)$$

The function β, as before, is nonlinear and nonmonotone with respect to the second variable, but in this section we further assume that it is locally Lipschitz continuous. The operator A is given by

$$Ay = -\text{div}(M(x)\nabla y) + a_0 y \quad (5.3)$$

where a_0 is a positive constant and $M(x) \in \mathbb{R}^{n\times n}$ is a positive definite C^1-valued matrix and

$$\xi^T M(x)\xi \geq c_2, \quad \text{for all } \xi \in \mathbb{R}^n \text{ and a.e. } x \in \Omega. \tag{5.4}$$

Clearly, α is the convex subdifferential of the function ℓ given by

$$\ell(s) = \begin{cases} 0, & s \geq 0 \\ +\infty, & s < 0. \end{cases}$$

We define φ as in (1.5) with the convex set K given by (1.7) and then see that (5.1) is a special case of (1.2).

The following new assumptions are considered in this section.

(h1) The mapping $x \mapsto \beta(x, s)$ is continuous for a.e. $s \in \mathbb{R}$ and the mapping $s \mapsto \beta(x, s)$ is locally Lipschitz continuous for all $x \in \Omega$.

(h2) There exists a positive constant c_5 such that

$$-m \leq \xi \leq c_5(1 + |s|) \text{ for all } \xi \in \overline{\partial}\beta(x, s), s \in \mathbb{R} \text{ and a.e. } x \in \Omega.$$

(h3) The operator B is a linear and continuous operator from U to H; $f \in H$ and $\psi \in H^2(\Omega)$ with $\psi(x) \leq 0$ on $\partial\Omega$.

Using the function $-\varepsilon^{-1}s^-$, which is the Yosida approximation of $\alpha(s)$, we consider the smooth function α_ε defined by

$$\alpha_\varepsilon(s) = -\frac{1}{\varepsilon}\int_{-\infty}^{+\infty} \big((s - \varepsilon^2\theta)^- - \varepsilon^2\theta^-\big)\rho(\theta)d\theta. \tag{5.5}$$

Then φ_ε in the last section can be chosen as $\varphi_\varepsilon(v) = \int_\Omega \alpha_\varepsilon(v(x) - \psi(x))dx$.

Corollary 5.1. *Suppose the hypotheses $H(\beta)$(ii)-(iii), (4.14), (h1)-(h3) are satisfied. Assume that α and A are given by (5.2) and (5.3), respectively. Then the results in Theorems 4.4–4.6 hold for the optimal control problem (1.1) subject to (5.1).*

Proof. In fact, the operator A given by (5.3) satisfies (A1) and (A2)$'$, and (h3) implies H(B) since H embeds into V^* compactly. Therefore, it is readily seen that all the hypotheses of Theorems 4.4–4.6 are satisfied for the optimal control problem (1.1) subject to (5.1). □

Lemma 5.2. *Let α_ε be defined as above. Then $\dot{\alpha}_\varepsilon(s)$ and $\alpha_\varepsilon(s)$ are decreasing and increasing functions, respectively. Moreover, we have*

$$\dot{\alpha}_\varepsilon(s) = \begin{cases} 0, & s \geq \varepsilon^2 \\ 1/2\varepsilon, & s = 0 \\ 1/\varepsilon, & s \leq -\varepsilon^2 \end{cases} \quad \textit{and} \quad \alpha_\varepsilon(s) = \begin{cases} \varepsilon\rho_0, & s \geq \varepsilon^2 \\ 0, & s = 0 \\ s/\varepsilon + \varepsilon\rho_0, & s \leq -\varepsilon^2 \end{cases} \tag{5.6}$$

where $\rho_0 = \int_0^1 \tau\rho(\tau)d\tau < 1$.

The proof is straightforward and we omit it for simplicity.

Lemma 5.3. *Let $(y_\varepsilon, u_\varepsilon)$ be given as in Theorem 4.5. Then if (h3) holds, we have $\alpha_\varepsilon^2(y_\varepsilon - \psi)$ and y_ε are bounded in $L^2(\Omega)$ and $H^2(\Omega)$, respectively.*

Proof. As is seen $\nabla J_\varepsilon(y_\varepsilon)(x) = \beta_\varepsilon(x, y_\varepsilon(x))$ for a.e. $x \in \Omega$ and $\nabla\varphi_\varepsilon(y_\varepsilon)(x) = \alpha_\varepsilon(y_\varepsilon(x) - \psi(x))$ for a.e. $x \in \Omega$. Multiplying (4.30) by $\alpha_\varepsilon(y_\varepsilon - \psi) \in V$ we find

$$\int_\Omega \nabla y_\varepsilon^T M(x)\nabla(y_\varepsilon - \psi)\dot{\alpha}_\varepsilon(y_\varepsilon - \psi)dx + \int_\Omega \alpha_\varepsilon^2(y_\varepsilon - \psi)dx$$
$$= \int_\Omega (f + Bu_\varepsilon)\alpha_\varepsilon(y_\varepsilon - \psi)dx - \int_\Omega \beta_\varepsilon(x, y_\varepsilon)\alpha_\varepsilon(y_\varepsilon - \psi)dx.$$

Since $M(x)$ is a positive definite matrix and $\dot{\alpha}_\varepsilon(y_\varepsilon - \psi) \geq 0$, it follows that

$$\nabla(y_\varepsilon - \psi)^T M(x)\nabla(y_\varepsilon - \psi)\dot{\alpha}_\varepsilon(y_\varepsilon - \psi) \geq 0.$$

As $\psi \in H^2(\Omega)$, adding $-\int_\Omega A\psi\alpha_\varepsilon(y_\varepsilon - \psi)dx$ on each side of the last equation and integrating by parts we have

$$\int_\Omega \alpha_\varepsilon^2(y_\varepsilon - \psi)dx \leq \int_\Omega (f + Bu_\varepsilon - \beta_\varepsilon(x, y_\varepsilon) - A\psi)\alpha_\varepsilon(y_\varepsilon - \psi)dx.$$

Then Cauchy's inequality yields the bound

$$\int_\Omega \alpha_\varepsilon^2(y_\varepsilon - \psi)dx \leq \|f + Bu_\varepsilon + \beta_\varepsilon(x, y_\varepsilon) - A\psi)\|_H^2 \leq C. \tag{5.7}$$

Here the hypotheses H(β)(ii), (h3) and the properties of u_ε and y_ε from Theorem 4.5 are used. Consequently, $\alpha_\varepsilon^2(y_\varepsilon - \psi)$ belongs to H and is uniformly bounded with respect to ε. It also follows that

$$Ay_\varepsilon + \alpha_\varepsilon(y_\varepsilon - \psi) + \beta_\varepsilon(x, y_\varepsilon) = f + Bu_\varepsilon \quad \text{in } H. \tag{5.8}$$

This equation and (5.7) imply that y_ε is uniformly bounded in $H^2(\Omega)$ by standard regularity theory of elliptic partial differential equations. □

Theorem 5.4. *Suppose the hypotheses of Corollary 5.1 hold, and* $(y_\varepsilon, u_\varepsilon, p_\varepsilon)$ *and* (y^*, u^*, p^*) *are given as in Theorem 4.5. Then there exist* $\xi \in H$, $z \in (L^\infty(\Omega))^*$, *and* $w \in \overline{\partial} g(y^*)$ *such that the triple* (y^*, u^*, p^*) *satisfies*

$$\begin{cases} Ap^* + \xi p^* + z = w & \text{in } V^* \cap (L^\infty(\Omega))^* \\ p^*(Ay^* + \beta(x, y^*) - f - Bu^*) = 0 & \text{a.e. in } \Omega \\ B^* p^* + \partial h(u^*) \ni 0 & \text{in } U^* \\ Ay^* + \beta(x, y^*) = f + Bu^* & \text{a.e. in } \Omega(y^*) \\ (y^* - \psi)(Ay^* + \beta(x, y^*) - f - Bu^*) = 0 & \text{a.e. in } \Omega \\ \xi(x) \in \overline{\partial}\beta(x, y^*(x)) & \text{a.e. in } \Omega, \end{cases} \tag{5.9}$$

where $\Omega(y^*) = \{x \in \Omega : y^*(x) > \psi(x)\}$ *and* z *is the weak star limit of* $\dot{\alpha}_\varepsilon(y_\varepsilon - \psi)p_\varepsilon$ *in* $(L^\infty(\Omega))^*$. *If, in addition,* $1 \le N \le 3$, *then the first equation in (5.9) reads*

$$(Ap^* + \xi p^* - w)(y^* - \psi) = 0 \quad \text{a.e. in } \Omega. \tag{5.10}$$

Moreover, we have

$$\langle Ap^*, p^* \rangle + (\xi p^*, p^*) - (w, p^*) \le 0, \tag{5.11}$$

and for all $\phi \in C^1(\overline{\Omega})$

$$\langle Ap^*, (y^* - \psi)\phi \rangle + (\xi p^*, (y^* - \psi)\phi) - (w, (y^* - \psi)\phi) = 0. \tag{5.12}$$

Proof. From (5.3) it follows that $A^* = A$. As $\nabla^2 J_\varepsilon(y_\varepsilon)(x) = \dot{\beta}_\varepsilon(x, y_\varepsilon(x))$, $\nabla^2 \varphi_\varepsilon(y_\varepsilon)(x) = \dot{\alpha}_\varepsilon(y_\varepsilon(x) - \psi(x))$ for a.e. $x \in \Omega$, we see from (4.44) that

$$\begin{cases} Ap_\varepsilon + \dot{\alpha}_\varepsilon(y_\varepsilon - \psi)p_\varepsilon + \dot{\beta}_\varepsilon(\cdot, y_\varepsilon)p_\varepsilon = \nabla g_\varepsilon(y_\varepsilon) \\ B^* p_\varepsilon + \nabla h_\varepsilon(u_\varepsilon) + R(u_\varepsilon - u^*) = 0 \\ Ay_\varepsilon + \alpha_\varepsilon(y_\varepsilon - \psi) + \beta_\varepsilon(\cdot, y_\varepsilon) = f + Bu_\varepsilon. \end{cases} \tag{5.13}$$

This theorem will be proved by first verifying three claims.

Claim 1: The sequence $\dot{\alpha}_\varepsilon(y_\varepsilon - \psi)p_\varepsilon$ converges weakly star to z in the Banach space $(L^\infty(\Omega))^*$, by passing to a subsequence, if necessary.

To see this, for each $\lambda > 0$ we define the function $\delta_\lambda \in W^{1,\infty}(\mathbb{R})$ by

$$\delta_\lambda(s) = \begin{cases} 1, & s > \lambda \\ s/\lambda, & |s| \le \lambda \\ -1, & s < -\lambda. \end{cases} \tag{5.14}$$

As is seen δ_λ is an approximation of the signum function with $|\delta_\lambda(s)| \le 1$, and $|\dot{\delta}_\lambda(s)| \ge 0$ a.e. $s \in \mathbb{R}$. Multiplying by $\delta_\lambda(p_\varepsilon) \in V$ in the first equation of (5.13) and integrating by parts we have

$$\int_\Omega \nabla p_\varepsilon^T M(x) \nabla p_\varepsilon \dot{\delta}_\lambda(p_\varepsilon) dx + \int_\Omega \dot{\alpha}_\varepsilon(y_\varepsilon - \psi) p_\varepsilon \delta_\lambda(p_\varepsilon) dx \tag{5.15}$$

$$= \int_\Omega \nabla g_\varepsilon(y_\varepsilon) \delta_\lambda(p_\varepsilon) dx - \int_\Omega \dot{\beta}_\varepsilon(y_\varepsilon) p_\varepsilon \delta_\lambda(p_\varepsilon) dx. \tag{5.16}$$

From (h2), it follows that

$$\int_\Omega \dot{\alpha}_\varepsilon(y_\varepsilon - \psi) p_\varepsilon \delta_\lambda(p_\varepsilon) dx \le C\big(\|\nabla g_\varepsilon(y_\varepsilon)\|_H + \|y_\varepsilon\|_V \|p_\varepsilon\|_V + \|p_\varepsilon\|_V\big). \tag{5.17}$$

Letting λ tend to zero and using the Lebesgue dominated convergence theorem, we have

$$\int_\Omega |\dot{\alpha}_\varepsilon(y_\varepsilon - \psi) p_\varepsilon| dx \le \lim_{\lambda \to 0} \int_\Omega \dot{\alpha}_\varepsilon(y_\varepsilon - \psi) p_\varepsilon \delta_\lambda(p_\varepsilon) dx \le C, \tag{5.18}$$

where $\dot{\alpha}_\varepsilon(y_\varepsilon - \psi) \ge 0$ and $\lim_{\lambda \to 0} \delta_\lambda(p_\varepsilon) p_\varepsilon = |p_\varepsilon|$ are used. This estimate yields Claim 1.

Claim 2: The following two convergence results hold in $L^1(\Omega)$:

$$p_\varepsilon \alpha_\varepsilon(y_\varepsilon - \psi) \to 0 \text{ in } L^1(\Omega), \tag{5.19}$$

and

$$(y_\varepsilon - \psi) \dot{\alpha}_\varepsilon(y_\varepsilon - \psi) p_\varepsilon \to 0 \text{ in } L^1(\Omega). \tag{5.20}$$

In order to prove this assertion, for each $\varepsilon > 0$ we define

$$\Omega_{1\varepsilon} = \{x \in \Omega : y_\varepsilon - \psi \le -\varepsilon^2\} \text{ and } \Omega_{2\varepsilon} = \{x \in \Omega : y_\varepsilon - \psi \ge -\varepsilon^2\}.$$

Then using Lemma 5.2 we see that

$$\begin{aligned} \int_{\Omega_{1\varepsilon}} |p_\varepsilon \alpha_\varepsilon(y_\varepsilon - \psi)| dx &= \int_{\Omega_{1\varepsilon}} \varepsilon^{1/2} |\dot{\alpha}_\varepsilon^{1/2}(y_\varepsilon - \psi) p_\varepsilon \alpha_\varepsilon(y_\varepsilon - \psi)| dx \\ &\le \varepsilon^{1/2} \left(\int_\Omega \dot{\alpha}_\varepsilon(y_\varepsilon - \psi) p_\varepsilon^2 dx \right)^{1/2} \left(\int_\Omega \alpha_\varepsilon^2(y_\varepsilon - \psi) dx \right)^{1/2}. \end{aligned} \tag{5.21}$$

Note that from (4.47), (A2)$'$, (h2), and using that $\dot{\alpha}_\varepsilon(y_\varepsilon - \psi) = \nabla^2\varphi_\varepsilon(y_\varepsilon)$ we find

$$\int_\Omega \dot{\alpha}_\varepsilon(y_\varepsilon - \psi)p_\varepsilon^2 dx \leq \|\nabla g_\varepsilon(y_\varepsilon)\|_H \|p_\varepsilon\|_H \leq C.$$

We combine the last two estimates and (5.7) to discover

$$\int_{\Omega_{1\varepsilon}} |p_\varepsilon \alpha_\varepsilon(y_\varepsilon - \psi)| dx \to 0 \quad \text{for } \varepsilon \to 0.$$

On the other hand, we see from Lemma 5.2 that $|p_\varepsilon \alpha_\varepsilon(y_\varepsilon - \psi)| \leq \varepsilon |p_\varepsilon|$ on $\Omega_{2\varepsilon}$. Therefore, from the boundedness of p_ε in H it follows that

$$\int_\Omega |p_\varepsilon \alpha_\varepsilon(y_\varepsilon - \psi)| dx = \int_{\Omega_{1\varepsilon}} |p_\varepsilon \alpha_\varepsilon(y_\varepsilon - \psi)| dx + \int_{\Omega_{2\varepsilon}} |p_\varepsilon \alpha_\varepsilon(y_\varepsilon - \psi)| dx \to 0.$$

Moreover, from Lemma 5.2, we also have $|\alpha(s) - s\dot{\alpha}(s)| \leq \varepsilon$, and therefore,

$$|p_\varepsilon \alpha_\varepsilon(y_\varepsilon - \psi) - (y_\varepsilon - \psi)p_\varepsilon \dot{\alpha}_\varepsilon(y_\varepsilon - \psi)| \leq \varepsilon |p_\varepsilon| \ \text{ a.e. } x \in \Omega.$$

This gives (5.19), as claimed.

Claim 3: There exists $\xi \in H$ such that $\xi(x) \in \overline{\partial}\beta(x, y^*(x))$ a.e. $x \in \Omega$ and

$$\dot{\beta}_\varepsilon(\cdot, y_\varepsilon)p_\varepsilon \to \xi p^* \text{ weakly in } L^1(\Omega). \tag{5.22}$$

To prove it, for each $v \in H$ we now define

$$\Phi(v) = \int_\Omega \beta(x, v(x))dx \ \text{ and } \ \Phi_\varepsilon(v) = \int_\Omega \beta_\varepsilon(x, v(x))dx.$$

As is seen $\nabla\Phi_\varepsilon(y_\varepsilon) = \dot{\beta}_\varepsilon(\cdot, y_\varepsilon)$. From (h2) we have

$$\|\dot{\beta}_\varepsilon(\cdot, y_\varepsilon)\|_H \leq C(1 + \|y_\varepsilon\|_H) \leq C.$$

Then taking a subsequence, if necessary, we find

$$\dot{\beta}_\varepsilon(\cdot, y_\varepsilon) \to \xi \text{ weakly in } H. \tag{5.23}$$

We conclude from Lemma 4.1 with $Q = \Phi$ and $Q_\varepsilon = \Phi_\varepsilon$ that $\xi \in \overline{\partial}\Phi(y^*)$, i.e., $\xi(x) \in \overline{\partial}\beta(x, y^*(x))$ a.e. $x \in \Omega$. On the other hand, recalling from (4.49) that $p_\varepsilon \to p^*$ in H, it follows that $\dot{\beta}_\varepsilon(\cdot, y_\varepsilon)p_\varepsilon \to \xi p^*$ weakly in $L^1(\Omega)$, as claimed.

We proceed now with the proof. In view of (4.49), Claims 1, and 3, by passing to the limit in the first equation of (5.13), we see that the first equation in (5.9) holds. Returning now to Lemma 5.3, we find $y_\varepsilon \to y^*$ weakly in $H^2(\Omega)$ and

$$\alpha_\varepsilon(y_\varepsilon - \psi) \to f + Bu^* - \beta(\cdot, y^*) - Ay^* \text{ weakly in } H. \tag{5.24}$$

Recalling from (4.49) that $p_\varepsilon \to p^*$ in H, we further have

$$p_\varepsilon\alpha_\varepsilon(y_\varepsilon - \psi) \to p^*(f + Bu^* - \beta(\cdot, y^*) - Ay^*) \text{ weakly in } L^1(\Omega). \tag{5.25}$$

Thus the second equation in (5.9) follows from this and Claim 2. The remaining equations and inclusions in (5.9) can be proved by (4.45) and the properties of the functions α and β.

Now in case of $1 \le N \le 3$, it follows that $H^2(\Omega)$ and V are continuous and compactly embedded into $C(\overline{\Omega})$ and $L^4(\Omega)$, respectively. Then we have $y_\varepsilon - \psi \to y^* - \psi$ in $C(\overline{\Omega})$. We combine this fact, Claim 1, and (5.20) to find $z(y^* - \psi) = 0$. This implies (5.10) using the first equation in (5.9). Next, for each $\phi \in C^1(\overline{\Omega})$, taking the inner product in $L^2(\Omega)$ with $(y_\varepsilon - \psi)\phi$ in the first equation of (5.13) and passing to the limit, we obtain (5.12) because of Claims 1 and 2. Finally, as $p_\varepsilon \to p^*$ weakly in V, we see that $p_\varepsilon^2 \to p^{*2}$ in H. Therefore, multiplying by p_ε in the same equation and passing to the lower limit we have (5.11). □

Acknowledgements

This work was carried out while Z. Peng was a visiting associate professor at the Institute for Mathematics and Scientific Computing, University of Graz, Austria, supported by Guangxi Education Department. Z. Peng was also supported by NNSF of China under grant 11561007 and Special Funds of Guangxi Distinguished Experts Construction Engineering. K. Kunisch was supported by the ERC advanced grant 668998 (OCLOC) under the EUs H2020 research program.

References

[1] D. Tiba, Optimal Control of Nonsmooth Distributed Parameter Systems, Lecture Notes in Mathematics, 1459, Springer-Verlag, Berlin, 1990.

[2] F. Tröltzsch, Optimal Control of Partial Differential Equations, Theory, Methods and Applications, Graduate Studies in Mathematics, 112, AMS Providence, Rhode Island, 2010.

[3] V. Barbu, Analysis and Control of Nonlinear Infinite Dimensional Systems, Academic Press, San Diego, CA, 1993.

[4] K. Ito, K. Kunisch, Optimal control of elliptic variational inequalities, Appl. Math. Optim., 41 (2000) 343-364.

[5] M. Bergounioux, Optimal control of problems governed by abstract elliptic variational inequalities with state constraints, SIAM J. Control Optim., 36 (1998) 273-289.

[6] K. Ito, K. Kunisch, On the Lagrange multiplier approach to variational problems and applications, Monographs and Studies in Mathematics 24, SIAM, Philadelphia, 2008.

[7] K. Ito, K. Kunisch, Optimal control of parabolic variational inequalities, J. Math. Pures Appl., 93 (2010) 329-360.

[8] F. Rindler, Optimal control for nonconvex rate-independent evolution processes, SIAM J. Control Optim., 47 (2008) 2773-2794.

[9] P.D. Panagiotopoulos, Inequality Problems in Mechanics and Applications, Convex and Nonconvex Energy Functions, Birkhäuser, Basel, 1985.

[10] P.D. Panagiotopoulos, Hemivariational Inequalities, Applications in Mechanics and Engineering, Springer-Verlag, Berlin, 1993.

[11] Z. Naniewicz, P.D. Panagiotopoulos, Mathematical Theory of Hemivariational Inequalities and Applications, Marcel Dekker, New York, 1995.

[12] S. Carl, V. K. Le, D. Motreanu, Nonsmooth Variational Problems and Their Inequalities: Comparison Principles and Applications, Springer, 2007.

[13] S. Migorski, A. Ochal, M. Sofonea, Nonlinear Inclusions and Hemivariational Inequalities. Models and Analysis of Contact Problems, in: Advances in Mechanics and Mathematics, vol. 26, Springer, New York, 2013.

[14] Z. Peng, Existence and regularity results for doubly nonlinear inclusions with nonmonotone perturbation, Nonlinear Anal., 115 (2015) 71-88.

[15] Z. Peng, Z. Liu, X. Liu, Boundary hemivariational inequality problems with doubly nonlinear operators, Math. Ann., 356 (2013) 1339-1358.

[16] M. Miettinen, J. Haslinger, Approximation of optimal control problems of hemivariational inequalities, Numer. funct. Anal. Optimiz., 13 (1992) 43-68.

[17] J. Haslinger, P.D. Panagiotopoulous, Optimal control of hemivariational inequalities. Approximation Resultls., International Series of Numerical Mathematics, Vol. 99, Birkhäuser, Verlag, Basel, 1991, 165-173.

[18] P.D. Panagiotopoulous, Optimal control of systems governed by variational-hemivariational inequaties, International Series of Numerical Mathematics, Vol. 101, Birkhäuser, Verlag, Basel, 1991, 161-181.

[19] P.D. Panagiotopoulous, Optimal control of systems governed by hemivariational inequaties. Necessary conditions, International Series of Numerical Mathematics, Vol. 95, Birkhäuser, Verlag, Basel, 1990, 207-228.

[20] M. Miettinen, Approximation of Hemivariational Inequalities and Optimal Control Problems, University of Jyväskylä Department of Mathematics Report 59, 1993.

[21] L. Gasiński, Optimal shape design problems for a class of systems described by parabolic hemivariational inequality, J. Global Optim., 12 (1998) 299-317.

[22] Z. Denkowski, S. Migórski, Optimal shape design problems for a class of systems described by hemivariational inequalities. J. Global Optim., 12 (1998) 37-59.

[23] S. Migórski, A. Ochal, Optimal control of parabolic hemivariational inequalities, J. Global Optim., 17 (2000) 285-300.

[24] A.D. Nagórska, A. Just, Z. Stempień, the optimal control problem and approximation of some parabolic hemivariational inclusion, Z. Angew. Math. Mech., 88 (2008) 218-226.

[25] Z. Liu, B. Zeng, Optimal Control of generalized quasi-variational hemivariational inequalities and its applications, Appl. Math. Optim., 72 (2015) 305-323.

[26] J.Y. Park, J.U. Jeong, Optimal control of hemivariational inequalities with delays, Taiwanese J. Math., 15 (2011) 433-447.

[27] Z. Liu, X. Li, D. Motreanu, Approximate controllability for nonlinear evolution hemivariational inequalities in Hilbert spaces, SIAM J. Control Optim. 53 (2015) 3228-3244.

[28] Z. Denkowski, S. Migórski, A. Ochal, A class of optimal control problems for piezoelectric frictional contact models, Nonlinear Anal., Real World Appl., 12 (2011) 1883-1895.

[29] F. H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[30] E. Zeidler, Nonlinear functional analysis and applications, II A/B, Springer, New York, 1990.