Appl Math Optim (2012) 65:371–402
DOI 10.1007/s00245-011-9160-y

Optimality Conditions for Semilinear Hyperbolic Equations with Controls in Coefficients

Bo Li · Hongwei Lou

Published online: 11 January 2012

Abstract An optimal control problem for semilinear hyperbolic partial differential equations is considered. The control variable appears in coefficients. Necessary conditions for optimal controls are established by method of two-scale convergence and homogenized spike variation. Results for problems with state constraints are also stated.

Keywords Optimal control · Necessary conditions · Hyperbolic equation · Two-scale convergence · Homogenized spike variation

1 Introduction

The main purpose of this paper is to give necessary conditions of optimal controls for semilinear hyperbolic partial differential equation (PDE) with coefficients containing controls. Let us consider the following controlled hyperbolic PDE:

$$\begin{cases} \partial_{tt} z(t,x) - \nabla\cdot\big(A(t,x,u(t,x))\nabla z(t,x)\big) + B(t,x,u(t,x))\cdot\nabla z(t,x) \\ \quad -\nabla\cdot\big(D(t,x,u(t,x))z(t,x)\big) = f(t,x,z(t,x),u(t,x),v(t,x)), \quad \text{in } \Omega_T, \\ z(t,x) = 0, \quad \text{on } [0,T]\times\partial\Omega, \\ z(0,x) = z_0(x), \quad \text{in } \Omega, \\ \partial_t z(0,x) = z_1(x), \quad \text{in } \Omega, \end{cases} \tag{1.1}$$

Communicating Editor: Irena Lasiecka.

B. Li
School of Mathematical Sciences, Fudan University, Shanghai 200433, China
e-mail: 062018053@fudan.edu.cn

H. Lou (✉)
School of Mathematical Sciences, and LMNS, Fudan University, Shanghai 200433, China
e-mail: hwlou@fudan.edu.cn

where $\Omega_T = (0, T) \times \Omega$, $T > 0$ and $\Omega \subset \mathbb{R}^n$ is a bounded domain with a smooth boundary $\partial\Omega$, $A : \Omega_T \times U \to \mathcal{S}_+^n$, $B, D : \Omega_T \times U \to \mathbb{R}^n$, $f : \Omega_T \times \mathbb{R} \times U \times \times V \to \mathbb{R}$, with $\mathcal{S}_+^n$ being the set of all $n \times n$ (symmetric) positive definite matrices and $U \times V$ being supposed later by (S2).

The control function $(u(\cdot), v(\cdot))$ is taken from the set

$$\mathcal{U}_{ad} \equiv \{u : \Omega_T \to U \,|\, u(\cdot), \partial_t u(\cdot) \in L^\infty(\Omega_T; U)\} \times \mathcal{M}(\Omega_T; V),$$

where $\mathcal{M}(\Omega_T; V)$ denotes the set of all measurable functions from Ω_T to V.

Under some suitable conditions, for any $(u(\cdot), v(\cdot)) \in \mathcal{U}_{ad}$, (1.1) admits a unique weak solution $z(\cdot) \equiv z(\cdot; u(\cdot), v(\cdot))$, which is called the state function corresponding to $(u(\cdot), v(\cdot))$.

Consider the following cost functional:

$$J\big(u(\cdot), v(\cdot)\big) = \int_{\Omega_T} f^0(t, x, z(t, x), u(t, x), v(t, x))\, dt\, dx, \tag{1.2}$$

where $f^0 : \Omega_T \times \mathbb{R} \times U \times V \to \mathbb{R}$. Our optimal control problem can be stated as follows.

Problem (C). Find a $(\bar{u}(\cdot), \bar{v}(\cdot)) \in \mathcal{U}_{ad}$ such that

$$J(\bar{u}(\cdot), \bar{v}(\cdot)) = \inf_{(u(\cdot), v(\cdot)) \in \mathcal{U}_{ad}} J(u(\cdot), v(\cdot)). \tag{1.3}$$

We call $(\bar{u}(\cdot), \bar{v}(\cdot)) \in \mathcal{U}_{ad}$ an *optimal control* if it satisfies (1.3), and $(\bar{z}(\cdot), (\bar{u}(\cdot), \bar{v}(\cdot)))$ an *optimal pair* with $\bar{z}(\cdot) \equiv \bar{z}(\cdot; \bar{u}(\cdot), \bar{v}(\cdot))$ being the state corresponding to $(\bar{u}(\cdot), \bar{v}(\cdot))$.

If $A(t, x, u)$, $B(t, x, u)$ and $D(t, x, u)$ are independent of u, then Problem (C) becomes a classical one. Though it seems there is no research work covering it, similar problems were studied by many researchers before, see for examples, [8, 18] and the references there.

Many works are related to the elliptic cases with leading term containing controls, see [4, 5, 13] and [21] for examples. We would like to mention the special cases named "shape optimization" or "structural optimization". For such problems, control variables are contained only in the leading term, and take values in a finite set. Each element in control set stands for a kind of material and the optimization problem is to lay out several materials throughout a given domain to maximize an integral functional associated with the conductive state of an assembled medium. Many relevant works are devoted to those problems. We refer to the books [2, 6, 15] and references there. Works devoted to parabolic cases and hyperbolic cases are quite less than those for elliptic cases. We mention [12] and [19] for parabolic cases. While readers can find works for some special hyperbolic cases in [10] and [20]. See also [16].

For optimal control problems with coefficients containing controls, the spike variation technique does not work directly. We adopt the idea of homogenization for PDEs. By carefully selecting some special type spike variations of controls, we can obtain desired "differentiability" of the state with respect to the control. This method is useful for the case of elliptic and parabolic equations (see [12] and [13]). However, in this paper, the controls also appear in the first order coefficients. As far as we know, there is no homogenization result on such kind of equations yet. One can find a result for hyperbolic PDE without terms of order one in [7]. To research the corresponding homogenization equation, it is convenient to use the technique of two-scale convergence. On the other hand, unlike those for elliptic and parabolic cases, to guarantee the well-posedness of (1.1), one need additional assumptions on the smoothness of coefficients on t. This also brings our problem some difference from the elliptic and parabolic cases. For example, we need to suppose that $u(\cdot)$ has additional regularities on t.

In this paper, we make the following assumptions.

(S1) Let $T > 0$ and Ω be a bounded domain in $\mathbb{R}^n$ with a smooth boundary $\partial\Omega$.

(S2) Let U be a subset of U_0 with U_0 being a domain in $\mathbb{R}^m$ and V be a separable metric space. Moreover, for any $u_0, u_1 \in U$, there exists a C^1 map $\tau : [0,1] \to U$, such that $\tau(0) = u_0$, $\tau(1) = u_1$.

(S3) Functions $A(t,x,u) \in L^\infty(\Omega_T \times U_0; \mathcal{S}^n_+)$, $\partial_t A, \partial_u A \in L^\infty(\Omega_T \times U; \mathcal{S}^n)$ and there exist $\Lambda \ge \lambda > 0$ such that for almost all $(t,x) \in \Omega_T$,

$$\lambda|\xi|^2 \le A(t,x,u)\xi \cdot \xi \le \Lambda|\xi|^2, \quad \forall \xi \in \mathbb{R}^n, u \in U_0. \tag{1.4}$$

(S4) Functions $B, D \in L^\infty(\Omega_T \times U_0; \mathbb{R}^n)$. Moreover,

$$\partial_t B, \partial_u B, \partial_t D, \partial_u D \in L^\infty(\Omega_T \times U; \mathbb{R}^n).$$

(S5) Functions $f(t,x,z,u,v)$ and $f^0(t,x,z,u,v)$ are measurable in (t,x), continuous in $(z,u,v) \in \mathbb{R} \times U \times V$, continuously differentiable in $z \in \mathbb{R}$. Moreover, there exists a $K > 0$ such that

$$\begin{cases} |f(t,x,0,u,v)| + |f_z(t,x,z,u,v)| \le K, \\ |f^0(t,x,0,u,v)| + |f^0_z(t,x,z,u,v)| \le K, \end{cases}$$
$$\forall (t,x,z,u,v) \in \Omega_T \times \mathbb{R} \times U \times V. \tag{1.5}$$

Remark 1.1 It is easy to see that if (S2) holds, then for any $u_0, u_1 \in U$, there exists a C^1 map $\tau : [0,1] \to U$, such that $\tau(0) = u_0$, $\tau(1) = u_1$ and $\tau'(0) = \tau'(1) = 0$. Using this observation, one can easily prove that for any measurable function $u : \Omega_T \to U$, there exists a sequence of $u^k \in L^\infty(\Omega_T; U)$ such that $\partial_t u^k \in L^\infty(\Omega_T; U)$ and u^k converges to u almost everywhere in Ω_T.

Our main result is the following theorem.

Theorem 1.2 *Assume* (S1)–(S5) *hold*, $z_0 \in H^1_0(\Omega)$, $z_1 \in L^2(\Omega)$ *and* $(\bar z(\cdot), (\bar u(\cdot), \bar v(\cdot)))$ *be an optimal pair of problem* (C). *Let* $\bar\psi(\cdot)$ *be the weak solution of the following ad-*

joint equation

$$\begin{cases} \partial_{tt}\bar{\psi}(t,x) - \nabla\cdot\big(A(t,x,\bar{u}(t,x))\nabla\bar{\psi}(t,x)\big) - \nabla\cdot(B(t,x,\bar{u}(t,x))\bar{\psi}(t,x)) \\ \quad + D(t,x,\bar{u}(t,x))\cdot\nabla\bar{\psi}(t,x) - f_z(t,x,\bar{z}(t,x),\bar{u}(t,x),\bar{v}(t,x))\bar{\psi}(t,x) \\ \quad + f_z^0(t,x,\bar{z}(t,x),\bar{u}(t,x),\bar{v}(t,x)) = 0, \quad \text{in } \Omega_T, \\ \bar{\psi}(t,x) = 0, \quad \text{on } [0,T]\times\partial\Omega, \\ \bar{\psi}(T,x) = 0, \quad \text{in } \Omega, \\ \partial_t\bar{\psi}(T,x) = 0, \quad \text{in } \Omega. \end{cases} \tag{1.6}$$

Then when $n = 1$, *for almost all* $(t,x)\in\Omega_T$,

$$\begin{aligned} &H(t,x,\bar{z}(t,x),\bar{\psi}(t,x),\partial_x\bar{z}(t,x),\partial_x\bar{\psi}(t,x),\bar{u}(t,x),\bar{v}(t,x)) \\ &\quad - H(t,x,\bar{z}(t,x),\bar{\psi}(t,x),\partial_x\bar{z}(t,x),\partial_x\bar{\psi}(t,x),u,v) \\ &\quad \geq \frac{p(t,x,\bar{u}(t,x),u)q(t,x,\bar{u}(t,x),u)}{A(t,x,u)}, \quad \forall(u,v)\in U\times V; \end{aligned} \tag{1.7}$$

when $n\geq 2$, *for almost all* $(t,x)\in\Omega_T$,

$$\begin{aligned} &H(t,x,\bar{z}(t,x),\bar{\psi}(t,x),\nabla\bar{z}(t,x),\nabla\bar{\psi}(t,x),\bar{u}(t,x),\bar{v}(t,x)) \\ &\quad - H(t,x,\bar{z}(t,x),\bar{\psi}(t,x),\nabla\bar{z}(t,x),\nabla\bar{\psi}(t,x),u,v) \\ &\geq \frac{1}{2}\left|A(t,x,u)^{-\frac{1}{2}}p(t,x,\bar{u}(t,x),u)\right|\left|A(t,x,u)^{-\frac{1}{2}}q(t,x,\bar{u}(t,x),u)\right| \\ &\quad + \frac{1}{2}A(t,x,u)^{-1}p(t,x,\bar{u}(t,x),u)\cdot q(t,x,\bar{u}(t,x),u), \quad \forall(u,v)\in U\times V, \end{aligned} \tag{1.8}$$

where for $(t,x,z,\psi,\xi,\eta,u,v)\in\Omega_T\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times U\times V$,

$$\begin{aligned} H(t,x,z,\psi,\xi,\eta,u,v) &= \psi f(t,x,z,u,v) - f^0(t,x,z,u,v) - A(t,x,u)\xi\cdot\eta \\ &\quad - B(t,x,u)\psi\cdot\eta - D(t,x,u)z\cdot\xi, \end{aligned} \tag{1.9}$$

and for $(t,x,u_1,u_2)\in\Omega_T\times U\times U$,

$$\begin{cases} p(t,x,u_1,u_2) = \big(A(t,x,u_1) - A(t,x,u_2)\big)\nabla\bar{z}(t,x) \\ \qquad\qquad + \big(D(t,x,u_1) - D(t,x,u_2)\big)\bar{z}(t,x), \\ q(t,x,u_1,u_2) = \big(A(t,x,u_1) - A(t,x,u_2)\big)\nabla\bar{\psi}(t,x) \\ \qquad\qquad + \big(B(t,x,u_1) - B(t,x,u_2)\big)\bar{\psi}(t,x). \end{cases} \tag{1.10}$$

The rest of the paper is organized as follows. In Sect. 2, we present some preliminary results on two-scale convergence and give a homogenization theorem on a class of semilinear hyperbolic equations. Section 3 is devoted to a proof of our main result. Results for state constraints will be stated in Sect. 4. Finally, we will make a cursory discussion on the existence of optimal controls in Sect. 5.

2 Preliminaries

In this section, we will give some preliminary results needed in proving Theorem 1.2. For notation simplicity, in this section, functions A, B, etc. are similar to but different from those in other sections. The first result is concerned with the well-posedness and regularity of linear hyperbolic equation.

Proposition 2.1 *Assume* (S1) *hold. Let* $A \in L^\infty(\Omega_T; \mathcal{S}^n_+)$, $B \in L^\infty(\Omega_T; \mathbb{R}^n)$, $D \in L^\infty(\Omega_T; \mathbb{R}^n)$, $c \in L^\infty(\Omega_T; \mathbb{R})$, $f \in L^2(\Omega_T; \mathbb{R})$, $z_0 \in H_0^1(\Omega)$ *and* $z_1 \in L^2(\Omega)$. *Moreover, assume that* $\partial_t A \in L^\infty(\Omega_T; \mathcal{S}^n)$, $\partial_t D \in L^\infty(\Omega_T; \mathbb{R}^n)$ *and for some* $\Lambda \geq \lambda > 0$,

$$\lambda|\xi|^2 \leq A(t,x)\xi \cdot \xi \leq \Lambda|\xi|^2, \quad \forall \xi \in \mathbb{R}^n, (t,x) \in \Omega_T. \tag{2.1}$$

Then, the following linear hyperbolic equation

$$\begin{cases} \partial_{tt} z(t,x) - \nabla \cdot (A(t,x)\nabla z(t,x)) + B(t,x) \cdot \nabla z(t,x) \\ \quad - \nabla \cdot (D(t,x)z(t,x)) + c(t,x)z(t,x) = f(t,x), \quad \text{in } \Omega_T, \\ z(t,x) = 0, \quad \text{on } [0,T] \times \partial\Omega, \\ z(0,x) = z_0(x), \quad \text{in } \Omega, \\ \partial_t z(0,x) = z_1(x), \quad \text{in } \Omega \end{cases} \tag{2.2}$$

admits a unique weak solution z *such that*

$$\begin{cases} z \in L^\infty(0,T; H_0^1(\Omega)), \\ \partial_t z \in L^\infty(0,T; L^2(\Omega)), \\ \partial_{tt} z \in L^2(0,T; H^{-1}(\Omega)). \end{cases} \tag{2.3}$$

Furthermore, there exists a positive constant M *depending on* λ, Λ, Ω_T, $\|\partial_t A\|_{L^\infty(\Omega_T)}$, $\|B\|_{L^\infty(\Omega_T)}$, $\|D\|_{L^\infty(\Omega_T)}$, $\|\partial_t D\|_{L^\infty(\Omega_T)}$, $\|c\|_{L^\infty(\Omega_T)}$, $\|f\|_{L^2(\Omega_T)}$, $\|z_0\|_{H_0^1(\Omega)}$ *and* $\|z_1\|_{L^2(\Omega)}$, *such that*

$$\|z\|_{L^\infty(0,T;H^1(\Omega))} + \|\partial_t z\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_{tt} z\|_{L^2(0,T;H^{-1}(\Omega))} \leq M. \tag{2.4}$$

The existence of weak solution together with estimate (2.4) follows from standard Galerkin approximations and energy estimates. The uniqueness of weak solution is totally the same as that of Theorem 7.2.4 in [9].

In addition, (2.3) in fact implies

$$z \in C([0,T]; L^2(\Omega)), \qquad \partial_t z \in C([0,T]; H^{-1}(\Omega)).$$

The above proposition can be easily generalized to semilinear cases. We have

Proposition 2.2 *Assume that all assumptions of Proposition* 2.1 *hold. Moreover, suppose that function* $f(t,x,z,v)$ *is measurable in* (t,x), *continuously differentiable in* $z \in \mathbb{R}$ *and there exists a* $K > 0$ *such that*

$$|f(t,x,0)| + |f_z(t,x,z)| \leq K, \quad \forall (t,x,z) \in \Omega_T \times \mathbb{R}. \tag{2.5}$$

Then, the following semilinear hyperbolic equation

$$\begin{cases} \partial_{tt} z(t,x) - \nabla\cdot(A(t,x)\nabla z(t,x)) + B(t,x)\cdot\nabla z(t,x) \\ \quad - \nabla\cdot(D(t,x)z(t,x)) = f(t,x,z(t,x)), \quad \text{in } \Omega_T, \\ z(t,x) = 0, \quad \text{on } [0,T]\times\partial\Omega, \\ z(0,x) = z_0(x), \quad \text{in } \Omega, \\ \partial_t z(0,x) = z_1(x), \quad \text{in } \Omega \end{cases} \tag{2.6}$$

admits a unique weak solution z satisfying (2.3). *Furthermore, there exists a positive constant M depending on λ, Λ, K, Ω_T, $\|\partial_t A\|_{L^\infty(\Omega_T)}$, $\|B\|_{L^\infty(\Omega_T)}$, $\|D\|_{L^\infty(\Omega_T)}$, $\|\partial_t D\|_{L^\infty(\Omega_T)}$, $\|z_0\|_{H_0^1(\Omega)}$ and $\|z_1\|_{L^2(\Omega)}$, such that* (2.4) *holds.*

The following theorem is a homogenization result which plays a critical role in proving Theorem 1.2.

Theorem 2.3 *Assume* (S1) *hold. For $i = 1, 2$, $A_i(\cdot)$, $B_i(\cdot)$, $D_i(\cdot)$, $f_i(\cdot)$ satisfy conditions in Proposition* 2.2. *Let $\delta \in (0,1)$, $r : \mathbb{R} \to \mathbb{R}$ be a periodic function with period* 1 *and*

$$r(t) = \begin{cases} 0, & \text{if } \{t\} \in [0,\delta), \\ 1, & \text{if } \{t\} \in [\delta,1), \end{cases} \tag{2.7}$$

where $\{a\}$ denote the decimal part of a real number a. Define

$$\begin{cases} A(t,x,y) = (a_{ij}(t,x,y)) = A_1(t,x) + r(y_1)\big(A_2(t,x) - A_1(t,x)\big), \\ B(t,x,y) = (b_i(t,x,y)) = B_1(t,x) + r(y_1)\big(B_2(t,x) - B_1(t,x)\big), \\ D(t,x,y) = (l_i(t,x,y)) = D_1(t,x) + r(y_1)\big(D_2(t,x) - D_1(t,x)\big), \\ f(t,x,y,z) = f_1(t,x,z) + r(y_1)\big(f_2(t,x,z) - f_1(t,x,z)\big). \end{cases} \tag{2.8}$$

Let z_ε be the weak solution of

$$\begin{cases} \partial_{tt} z_\varepsilon(t,x) - \nabla\cdot\big[A\big(t,x,\frac{x}{\varepsilon}\big)\nabla z_\varepsilon(t,x)\big] + B\big(t,x,\frac{x}{\varepsilon}\big)\cdot\nabla z_\varepsilon(t,x) \\ \quad - \nabla\cdot\big[D\big(t,x,\frac{x}{\varepsilon}\big)z_\varepsilon(t,x)\big] = f\big(t,x,\frac{x}{\varepsilon},z_\varepsilon(t,x)\big), \quad \text{in } \Omega_T, \\ z_\varepsilon(t,x) = 0, \quad \text{on } [0,T]\times\partial\Omega, \\ z_\varepsilon(0,x) = z_0(x), \quad \text{in } \Omega, \\ \partial_t z_\varepsilon(0,x) = z_1(x), \quad \text{in } \Omega \end{cases} \tag{2.9}$$

with $z_0 \in H_0^1(\Omega)$, $z_1 \in L^2(\Omega)$. Then

$$z_\varepsilon \to z, \quad \text{weakly in } L^2(0,T;H_0^1(\Omega))$$

with $z(\cdot)$ being the weak solution of

$$\begin{cases} \partial_{tt} z(t,x) - \nabla\cdot(A^*(t,x)\nabla z(t,x)) + B^*(t,x)\cdot\nabla z(t,x) \\ \quad - \nabla\cdot(D^*(t,x)z(t,x)) = f^*(t,x,z(t,x)), \quad \textit{in } \Omega_T, \\ z(t,x) = 0, \quad \textit{on } [0,T]\times\partial\Omega, \\ z(0,x) = z_0(x), \quad \textit{in } \Omega, \\ \partial_t z(0,x) = z_1(x), \quad \textit{in } \Omega, \end{cases} \tag{2.10}$$

where

$$\begin{cases} A^*(t,x) = \delta A_1(t,x) + (1-\delta)A_2(t,x) \\ \qquad - \frac{\delta(1-\delta)[A_1(t,x)-A_2(t,x)]e_1 e_1^\top [A_1(t,x)-A_2(t,x)]}{\delta e_1^\top A_2(t,x)e_1 + (1-\delta)e_1^\top A_1(t,x)e_1}, \\ B^*(t,x) = \delta B_1(t,x) + (1-\delta)B_2(t,x) \\ \qquad - \frac{\delta(1-\delta)[A_1(t,x)-A_2(t,x)]e_1 e_1^\top [B_1(t,x)-B_2(t,x)]}{\delta e_1^\top A_2(t,x)e_1 + (1-\delta)e_1^\top A_1(t,x)e_1}, \\ D^*(t,x) = \delta D_1(t,x) + (1-\delta)D_2(t,x) \\ \qquad - \frac{\delta(1-\delta)[A_1(t,x)-A_2(t,x)]e_1 e_1^\top [D_1(t,x)-D_2(t,x)]}{\delta e_1^\top A_2(t,x)e_1 + (1-\delta)e_1^\top A_1(t,x)e_1}, \\ f^*(t,x,z) = \delta f_1(t,x,z) + (1-\delta)f_2(t,x,z) \\ \qquad + \frac{\delta(1-\delta)[B_1(t,x)-B_2(t,x)]^\top e_1 e_1^\top [D_1(t,x)-D_2(t,x)]}{\delta e_1^\top A_2(t,x)e_1 + (1-\delta)e_1^\top A_1(t,x)e_1} z. \end{cases} \tag{2.11}$$

In order to prove the above theorem, we need introduce briefly two-scale convergence first. The original idea of this kind of convergence was introduced by Nguetseng in 1989 (see [17]). Later in 1992, Allaire developed the theory further by studying some general properties of two-scale convergence (see [1]). Moreover he used two-scale convergence to analyze several homogenization problems. It seems that after that time, two-scale convergence method becomes the main tool to study homogenization problems. Let's recall the definition and some useful propositions of two-scale convergence. Let $Y = [0,1]^n$ and let $\{e_i\}_{1\le i\le n}$ be the canonical basis of $\mathbb{R}^n$. We call a function $f(y)$ Y-periodic if it is 1-periodic in each direction e_i. Denote by $\mathcal{M}_\#(Y)$ the space of all measurable Y-periodic functions in $\mathbb{R}^n$. Moreover, denote

$$\begin{aligned} C_\#(Y) &= C(\mathbb{R}^n)\cap\mathcal{M}_\#(Y), \\ C_\#^\infty(Y) &= C^\infty(\mathbb{R}^n)\cap\mathcal{M}_\#(Y), \\ L_\#^2(Y;C(\overline{\Omega})) &= \left\{ g:\overline{\Omega}\times\mathbb{R}^n\to\mathbb{R} \,\middle|\, g(x,\cdot)\in\mathcal{M}_\#(Y), \int_Y \|g(\cdot,y)\|_{C(\overline{\Omega})}^2\,dy<\infty \right\}, \\ D(\Omega;C_\#^\infty(Y)) &= \left\{ g\in C^\infty(\Omega\times\mathbb{R}^n) \,\middle|\, \bigcup_y \operatorname{supp} g(\cdot,y)\subset\subset\Omega, g(x,\cdot)\in\mathcal{M}_\#(Y) \right\}. \end{aligned}$$

Definition 2.4 We say that a sequence z_ε in $L^2(\Omega)$ two-scale converges to a limit $z \in L^2(\Omega \times Y)$ if

$$\lim_{\varepsilon\to 0} \int_\Omega z_\varepsilon(x)\varphi\Big(x, \frac{x}{\varepsilon}\Big)\,dx = \int_\Omega \int_Y z(x,y)\varphi(x,y)\,dy\,dx \tag{2.12}$$

for every $\varphi \in L^2(\Omega; C_\#(Y))$. We denote it by $z_\varepsilon \xrightarrow{T-S} z$ (in $\Omega \times Y$).

This notion of "two-scale convergence" makes sense because of the next compactness proposition (see [1]).

Proposition 2.5 *From each bounded sequence z_ε in $L^2(\Omega)$, one can extract a subsequence that two-scale converges to some $z \in L^2(\Omega \times Y)$.*

Now we will show some useful results about the two-scale convergence. They can be found in [14].

Proposition 2.6 *Let z_ε be a bounded sequence in $L^2(\Omega)$ such that*

$$\lim_{\varepsilon\to 0} \int_\Omega z_\varepsilon(x)\varphi\Big(x, \frac{x}{\varepsilon}\Big)\,dx = \int_\Omega \int_Y z(x,y)\varphi(x,y)\,dy\,dx$$

for every $\varphi \in D(\Omega; C^\infty_\#(Y))$. Then z_ε two-scale converges to z.

Proposition 2.7 *If $z_\varepsilon \in L^2(\Omega)$ two-scale converges to z, then (2.12) holds for every $\varphi \in L^2_\#(Y; C(\overline{\Omega}))$.*

Proposition 2.8 *Let z_ε be a sequence in $L^2(\Omega)$ which two-scale converges to $z \in L^2(\Omega \times Y)$. Then*

(i)

$$\liminf_{\varepsilon\to 0} \|z_\varepsilon\|_{L^2(\Omega)} \geq \|z\|_{L^2(\Omega\times Y)}.$$

(ii) *Assume that*

$$\lim_{\varepsilon\to 0} \|z_\varepsilon\|_{L^2(\Omega)} = \|z\|_{L^2(\Omega\times Y)}$$

and $w_\varepsilon \in L^2(\Omega)$ two-scale converges to $w \in L^2(\Omega \times Y)$. Then

$$\lim_{\varepsilon\to 0} \int_\Omega z_\varepsilon(x) w_\varepsilon(x)\varphi\Big(x, \frac{x}{\varepsilon}\Big)\,dx = \int_\Omega \int_Y z(x,y)w(x,y)\varphi(x,y)\,dy\,dx$$

for every $\varphi \in D(\Omega; C^\infty_\#(Y))$.

Proposition 2.9 *Let $z_\varepsilon \in L^2(\Omega)$ be a sequence strongly converges to z in $L^2(\Omega)$. Then z_ε two-scale converges to $\tilde{z} \in L^2(\Omega \times Y)$ with $\tilde{z}(x,y) = z(x)$. Particularly, for any $w \in L^2(\Omega)$ and $\varphi \in L^2(\Omega; C_\#(Y)) \cup L^2_\#(Y; C(\overline{\Omega}))$, we have*

$$\lim_{\varepsilon\to 0} \int_\Omega w(x)\varphi\Big(x, \frac{x}{\varepsilon}\Big)\,dx = \int_\Omega \int_Y w(x)\varphi(x,y)\,dy\,dx. \tag{2.13}$$

We see that the above proposition can be looked as a generalization of the well-known Riemann-Lebesgue's Lemma.

Proposition 2.10 *Let $z_\varepsilon \in H^1(\Omega)$ such that*

$$z_\varepsilon \to z, \quad \text{weakly in } H^1(\Omega).$$

Then z_ε two-scale converges to z and there exist a subsequence z_{ε_j} and a $z_1 \in L^2(\Omega; H^1_\#(Y))$ such that

$$\partial_i z_{\varepsilon_j}(x) \xrightarrow{T-S} \partial_i z(x) + \partial_{y_i} z_1(x,y), \quad i = 1, 2, \dots, n.$$

Denote $\widetilde{Y} = [0,1] \times Y = [0,1]^{n+1}$. We have

Lemma 2.11 *Let $h_1(\cdot), h_2(\cdot) \in L^2(\Omega_T)$ and $r(\cdot)$ be defined by (2.7). Define*

$$h(t,x,y) = h_1(t,x) + r(y_1)[h_2(t,x) - h_1(t,x)].$$

Then

$$h\Big(t, x, \frac{x}{\varepsilon}\Big) \xrightarrow{T-S} h(t,x,y), \quad \text{in } \Omega_T \times \widetilde{Y} \tag{2.14}$$

and

$$\Big\| h\Big(\cdot, \frac{\cdot}{\varepsilon}\Big)\Big\|_{L^2(\Omega_T)} = \|h\|_{L^2(\Omega_T \times \widetilde{Y})}. \tag{2.15}$$

Proof For any $\varphi \in D(\Omega_T; C^\infty_\#(\widetilde{Y}))$, we have

$$\begin{aligned}
&\int_{\Omega_T} h\Big(t,x,\frac{x}{\varepsilon}\Big)\varphi\Big(t,x,\frac{t}{\varepsilon},\frac{x}{\varepsilon}\Big)\,dx\,dt \\
&\quad = \int_{\Omega_T} \Big[h_1(t,x) + r\Big(\frac{x_1}{\varepsilon}\Big)(h_2(t,x) - h_1(t,x))\Big]\varphi\Big(t,x,\frac{t}{\varepsilon},\frac{x}{\varepsilon}\Big)\,dx\,dt \\
&\quad = \int_{\Omega_T} h_1(t,x)\varphi\Big(t,x,\frac{t}{\varepsilon},\frac{x}{\varepsilon}\Big)\,dx\,dt \\
&\qquad + \int_{\Omega_T} (h_2(t,x) - h_1(t,x)) r\Big(\frac{x_1}{\varepsilon}\Big)\varphi\Big(t,x,\frac{t}{\varepsilon},\frac{x}{\varepsilon}\Big)\,dx\,dt.
\end{aligned}$$

By Proposition 2.9 and noting that $r\varphi \in L^2_\#(\widetilde{Y}, C(\overline{\Omega}_T))$, we have

$$\begin{aligned}
&\lim_{\varepsilon \to 0} \int_{\Omega_T} h\Big(t,x,\frac{x}{\varepsilon}\Big)\varphi\Big(t,x,\frac{t}{\varepsilon},\frac{x}{\varepsilon}\Big)\,dx\,dt \\
&\quad = \int_{\Omega_T}\int_{\widetilde{Y}} h_1(t,x)\varphi(t,x,s,y)\,dy\,ds\,dx\,dt \\
&\qquad + \int_{\Omega_T}\int_{\widetilde{Y}} (h_2(t,x) - h_1(t,x)) r(y_1)\varphi(t,x,s,y)\,dy\,ds\,dx\,dt
\end{aligned}$$

$$= \int_{\Omega_T} \int_{\widetilde{Y}} h(t,x,y)\varphi(t,x,s,y)\, dy\, ds\, dx\, dt.$$

Moreover, $h(\cdot, \frac{\cdot}{\varepsilon})$ is bounded uniformly in $L^2(\Omega_T)$ since $h_1(\cdot)$ and $h_2(\cdot)$ are bounded in $L^2(\Omega_T)$. Then we obtain (2.14) by Proposition 2.6.

In order to prove (2.15), we calculate $\|h(\cdot, \frac{\cdot}{\varepsilon})\|_{L^2(\Omega_T)}$:

$$\begin{aligned}
\left\| h\left(\cdot, \frac{\cdot}{\varepsilon}\right)\right\|^2_{L^2(\Omega_T)} &= \int_{\Omega_T} \left[h_1(t,x) + r\left(\frac{x_1}{\varepsilon}\right)(h_2(t,x) - h_1(t,x))\right]^2 dx\, dt \\
&= \int_{\Omega_T} \left[h_1^2(t,x) + 2r\left(\frac{x_1}{\varepsilon}\right)(h_2(t,x) - h_1(t,x)) \right. \\
&\quad \left. + r\left(\frac{x_1}{\varepsilon}\right)(h_2(t,x) - h_1(t,x))^2\right] dx\, dt.
\end{aligned}$$

By Proposition 2.9, we have

$$\begin{aligned}
&\lim_{\varepsilon\to 0} \left\| h\left(\cdot, \frac{\cdot}{\varepsilon}\right)\right\|^2_{L^2(\Omega_T)} \\
&\quad = \int_{\Omega_T} \int_{\widetilde{Y}} \Big(h_1^2(t,x) + 2r(y_1)(h_2(t,x) - h_1(t,x)) \\
&\qquad + r(y_1)(h_2(t,x) - h_1(t,x))^2\Big)\, dy\, ds\, dx\, dt \\
&\quad = \int_{\Omega_T} \int_{\widetilde{Y}} h(t,x,y)^2\, dy\, ds\, dx\, dt = \|h\|^2_{L^2(\Omega_T\times\widetilde{Y})}.
\end{aligned}$$

□

Now we will prove Theorem 2.3.

Proof of Theorem 2.3. Since z_ε is the weak solution of (2.9),

$$\begin{cases} z_\varepsilon \in L^\infty(0,T; H_0^1(\Omega)), \\ \partial_t z_\varepsilon \in L^\infty(0,T; L^2(\Omega)), \\ \partial_{tt} z_\varepsilon \in L^2(0,T; H^{-1}(\Omega)) \end{cases} \tag{2.16}$$

and

$$\begin{aligned}
&\int_\Omega \Big[\partial_{tt} z_\varepsilon(t,x)\varphi(x) + A\left(t,x,\frac{x}{\varepsilon}\right)\nabla_x z_\varepsilon(t,x)\cdot\nabla\varphi(x) \\
&\qquad + \varphi(x) B\left(t,x,\frac{x}{\varepsilon}\right)\cdot\nabla_x z_\varepsilon(t,x) + z_\varepsilon(t,x) D\left(t,x,\frac{x}{\varepsilon}\right)\cdot\nabla\varphi(x)\Big]\, dx \\
&\quad = \int_\Omega f\left(t,x,\frac{x}{\varepsilon}, z_\varepsilon(t,x)\right)\varphi(x)\, dx
\end{aligned} \tag{2.17}$$

for all $\varphi \in C_c^\infty(\Omega)$, a.e. $0 \le t \le T$, where we denote $\partial_{tt} z_\varepsilon(t,\cdot)$ and $\varphi(\cdot)$ pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$ by $\int_\Omega \partial_{tt} z_\varepsilon(t,x)\varphi(x)\, dx$.

When (2.16) holds, the above equation is equivalent to

$$\int_{\Omega_T}\Big[\partial_{tt}z_\varepsilon(t,x)\varphi(x)+A\Big(t,x,\frac{x}{\varepsilon}\Big)\nabla_x z_\varepsilon(t,x)\cdot\nabla\varphi(x)$$
$$+\varphi(x)B\Big(t,x,\frac{x}{\varepsilon}\Big)\cdot\nabla_x z_\varepsilon(t,x)+z_\varepsilon(t,x)D\Big(t,x,\frac{x}{\varepsilon}\Big)\cdot\nabla\varphi(x)\Big]\rho(t)\,dx\,dt$$
$$=\int_{\Omega_T}f\Big(t,x,\frac{x}{\varepsilon},z_\varepsilon(t,x)\Big)\varphi(x)\rho(t)\,dx\,dt \tag{2.18}$$

for all $\rho\in C_c^\infty(0,T)$.

By Proposition 2.1, there exists a constant M which is independent of ε, such that

$$\|z_\varepsilon\|_{L^\infty(0,T;H^1(\Omega))}+\|\partial_t z_\varepsilon\|_{L^\infty(0,T;L^2(\Omega))}+\|\partial_{tt}z_\varepsilon\|_{L^2(0,T;H^{-1}(\Omega))}\le M. \tag{2.19}$$

Then, along a subsequence $\varepsilon\to 0$,

$$z_\varepsilon\longrightarrow z,\quad \text{weakly}^* \text{ in } L^\infty(0,T;H_0^1(\Omega)),\text{ strongly in } L^2(\Omega_T), \tag{2.20}$$
$$\partial_t z_\varepsilon\longrightarrow \partial_t z,\quad \text{weakly}^* \text{ in } L^\infty(0,T;L^2(\Omega)), \tag{2.21}$$
$$\partial_{tt} z_\varepsilon\longrightarrow \partial_{tt} z,\quad \text{weakly in } L^2(0,T;H^{-1}(\Omega)) \tag{2.22}$$

for some z. Further, it follows from (2.19)–(2.21) that along a subsequence $\varepsilon\to 0$, z_ε converges to z weakly in $H^1(\Omega_T)$. By Proposition 2.10, along a subsequence $\varepsilon\to 0$,

$$\partial_{x_i}z_\varepsilon(t,x)\xrightarrow{T-S}\partial_{x_i}z(t,x)+\partial_{y_i}z_1(t,x,s,y),\quad i=1,2,\dots,n \tag{2.23}$$

for some $z_1\in L^2(\Omega_T;H^1_\#(\widetilde{Y})\backslash\mathbb{R})$.

Now, choose $\varphi(x)=\varphi_0(x)+\varepsilon\varphi_1(x,\frac{x}{\varepsilon})$, $\varphi_0\in C_c^\infty(\Omega)$, $\varphi_1\in D(\Omega;C_\#^\infty(Y))$. Then

$$\int_{\Omega_T}\Big\{A\Big(t,x,\frac{x}{\varepsilon}\Big)\nabla_x z_\varepsilon(t,x)\cdot\nabla_x\Big[\varphi_0(x)+\varepsilon\varphi_1\Big(x,\frac{x}{\varepsilon}\Big)\Big]\Big\}\rho(t)\,dx\,dt$$
$$=\int_{\Omega_T}\Big\{A\Big(t,x,\frac{x}{\varepsilon}\Big)\nabla_x z_\varepsilon(t,x)\cdot\Big[\nabla_x\varphi_0(x)+\nabla_y\varphi_1\Big(x,\frac{x}{\varepsilon}\Big)$$
$$+\varepsilon\nabla_x\varphi_1\Big(x,\frac{x}{\varepsilon}\Big)\Big]\Big\}\rho(t)\,dx\,dt.$$

By Lemma 2.11, $a_{ij}(t,x,\frac{x}{\varepsilon})$ two-scale converges to $a_{ij}(t,x,y)$ and

$$\lim_{\varepsilon\to 0}\Big\|a_{ij}\Big(\cdot,\frac{\cdot}{\varepsilon}\Big)\Big\|_{L^2(\Omega_T)}=\|a_{ij}\|_{L^2(\Omega_T\times\widetilde{Y})}.$$

Thus, we deduce from Proposition 2.8 that

$$\lim_{\varepsilon\to 0}\int_{\Omega_T}\Big\{A\Big(t,x,\frac{x}{\varepsilon}\Big)\nabla_x z_\varepsilon(t,x)\cdot\nabla_x\Big[\varphi_0(x)+\varepsilon\varphi_1\Big(x,\frac{x}{\varepsilon}\Big)\Big]\Big\}\rho(t)\,dx\,dt$$
$$=\int_{\Omega_T}\int_{\widetilde{Y}}\Big[A(t,x,y)\Big(\nabla_x z(t,x)+\nabla_y z_1(t,x,s,y)\Big)$$

$$\times \Big(\nabla_x \phi_0(x) + \nabla_y \phi_1(x,y)\Big)\Big]\rho(t)\,dy\,ds\,dx\,dt. \tag{2.24}$$

Similarly,

$$\lim_{\varepsilon\to 0}\int_{\Omega_T}\Big[B\Big(t,x,\frac{x}{\varepsilon}\Big)\cdot\nabla_x z_\varepsilon(t,x)\Big]\Big[\varphi_0(x)+\varepsilon\varphi_1\Big(x,\frac{x}{\varepsilon}\Big)\Big]\rho(t)\,dx\,dt$$
$$= \int_{\Omega_T}\int_{\widetilde{Y}}\Big[B(t,x,y)\cdot\Big(\nabla_x z(t,x)+\nabla_y z_1(t,x,s,y)\Big)\Big]\varphi_0(x)\rho(t)\,dy\,ds\,dx\,dt \tag{2.25}$$

and

$$\lim_{\varepsilon\to 0}\int_{\Omega_T}\Big\{D\Big(t,x,\frac{x}{\varepsilon}\Big)\cdot\nabla_x\Big[\varphi_0(x)+\varepsilon\varphi_1\Big(x,\frac{x}{\varepsilon}\Big)\Big]\Big\}z_\varepsilon(t,x)\rho(t)\,dx\,dt$$
$$= \int_{\Omega_T}\int_{\widetilde{Y}}\Big[D(t,x,y)\cdot\Big(\nabla_x\varphi_0(x)+\nabla_y\varphi_1(x,y)\Big)\Big]z(t,x)\rho(t)\,dy\,ds\,dx\,dt. \tag{2.26}$$

Moreover, by assumptions on f_1 and f_2, one can see that z_ε converges strongly to z implies that $f_i(t,x,z_\varepsilon(t,x))$ converges strongly to $f_i(t,x,z(t,x))$. Thus,

$$\lim_{\varepsilon\to 0}\int_{\Omega_T} f\Big(t,x,\frac{x}{\varepsilon},z_\varepsilon(t,x)\Big)\Big[\varphi_0(x)+\varepsilon\varphi_1\Big(x,\frac{x}{\varepsilon}\Big)\Big]\rho(t)\,dx\,dt$$
$$= \lim_{\varepsilon\to 0}\int_{\Omega_T} f\Big(t,x,\frac{x}{\varepsilon},z_\varepsilon(t,x)\Big)\varphi_0(x)\,\rho(t)\,dx\,dt$$
$$= \lim_{\varepsilon\to 0}\int_{\Omega_T} f_1(t,x,z_\varepsilon(t,x))\varphi_0(x)\,\rho(t)\,dx\,dt$$
$$+ \lim_{\varepsilon\to 0}\int_{\Omega_T} r\Big(\frac{x_1}{\varepsilon}\Big)\Big(f_2(t,x,z_\varepsilon(t,x)) - f_1(t,x,z_\varepsilon(t,x))\Big)\varphi_0(x)\,\rho(t)\,dx\,dt$$
$$= \int_{\Omega_T}\Big(\delta f_1(t,x,z(t,x)) + (1-\delta)f_2(t,x,z(t,x))\Big)\varphi_0(x)\,\rho(t)\,dx\,dt. \tag{2.27}$$

In addition, from (2.22), one can deduce that

$$\lim_{\varepsilon\to 0}\int_{\Omega_T}\partial_{tt}z_\varepsilon(t,x)\Big[\varphi_0(x)+\varepsilon\varphi_1\Big(x,\frac{x}{\varepsilon}\Big)\Big]\rho(t)\,dx\,dt$$
$$= \int_{\Omega_T}\partial_{tt}z(t,x)\varphi_0(x)\rho(t)\ dx\,dt \tag{2.28}$$

for all $\varphi\in C_c^\infty(\Omega)$ and $\rho\in C_c^\infty(0,T)$.

Combining (2.18) with (2.24)–(2.28), we have

$$\int_{\Omega_T}\Big\{\partial_{tt}z(t,x)\varphi_0(x) + \int_{\widetilde{Y}}\Big[A(t,x,y)\big(\nabla_x z(t,x)+\nabla_y z_1(t,x,s,y)\big)$$

$$\times \left(\nabla_x \varphi_0(x) + \nabla_y \varphi_1(x,y)\right)\Big] dy\, ds$$

$$+ \int_{\widetilde{Y}} \Big[B(t,x,y) \cdot \left(\nabla_x z(t,x) + \nabla_y z_1(t,x,s,y)\right)\Big] \varphi_0(x)\, dy\, ds$$

$$+ \int_{\widetilde{Y}} \Big[D(t,x,y) \cdot \left(\nabla_x \varphi_0(x) + \nabla_y \varphi_1(x,y)\right)\Big] z(t,x)\, dy\, ds \Big\} \rho(t)\, dx\, dt$$

$$= \int_{\Omega_T} \Big(\delta f_1(t,x,z(t,x)) + (1-\delta) f_2(t,x,z(t,x))\Big) \varphi_0(x) \rho(t)\, dx\, dt. \tag{2.29}$$

Consequently, setting $\varphi_1 = 0$ in (2.29), we get that

$$\begin{aligned} &\partial_{tt} z(t,x) - \nabla \cdot \int_{\widetilde{Y}} A(t,x,y)(\nabla_x z(t,x) + \nabla_y z_1(t,s,x,y))\, dy\, ds \\ &\quad + \int_{\widetilde{Y}} B(t,x,y) \cdot (\nabla_x z(t,x) + \nabla_y z_1(t,s,x,y))\, dy\, ds \\ &\quad - \nabla \cdot \int_Y D(t,x,y) z(t,x)\, dy \\ &\quad = \delta f_1(t,x,z(t,x)) + (1-\delta) f_2(t,x,z(t,x)) \end{aligned} \tag{2.30}$$

holds in the weak sense. If we set $\varphi_0 = 0$ in (2.29), then we get that for almost all $(t,x) \in \Omega_T$,

$$\nabla_y \cdot \Big[A(t,x,y)(\nabla_x z(t,x) + \nabla_y z_1(t,x,s,y)) + D(t,x,y) z(t,x)\Big] = 0 \tag{2.31}$$

holds in the weak sense.

Now we will resolve (2.31). Noting that

$$\int_{\widetilde{Y}} \nabla_y \cdot (A(t,x,y)\nabla_x z(t,x) + B(t,x,y) z(t,x))\, dy\, ds = 0, \quad \text{a.e. } (t,x) \in \Omega_T,$$

we can get that up to an additive function of (t,x), $z_1(t,x,\cdot)$ is the unique periodic solution of (2.31) for fixed $z(t,x)$ (see Remark 2.1 in Chapter 1 of [3]). Since $A(t,x,y)$ and $B(t,x,y)$ do not depend on $y_2 \cdots y_n$ and s, we may let $z_1(t,x,s,y)$ be independent of $y_2 \cdots y_n$ and s, i.e. $z_1(t,x,s,y) = z_1(t,x,y_1)$. Then (2.31) can be rewritten as

$$\partial_{y_1} \Big[e_1^\top \Big(A(t,x,y)(\nabla_x z(t,x) + \partial_{y_1} z_1(t,x,y_1) e_1) + D(t,x,y) z(t,x)\Big)\Big] = 0,$$

which leads to

$$e_1^\top \Big[A(t,x,y)(\nabla_x z(t,x) + \partial_{y_1} z_1(t,x,y_1) e_1) + D(t,x,y) z(t,x)\Big] = \zeta(t,x).$$

Hence

$$\partial_{y_1} z_1(t,x,y_1) = \frac{\zeta(t,x) - e_1^\top [A(t,x,y)\nabla_x z(t,x) + D(t,x,y) z(t,x)]}{e_1^\top A(t,x,y) e_1}. \tag{2.32}$$

By the periodicity of z_1, we require

$$\int_0^1 \frac{\zeta(t,x) - e_1^\top[A(t,x,y)\nabla_x z(t,x) + D(t,x,y)z(t,x)]}{e_1^\top A(t,x,y)e_1}\,dy_1 = 0.$$

Thus,

$$\zeta(t,x) = \left[\int_0^1 \frac{1}{e_1^\top A(t,x,y)e_1}\,dy_1\right]^{-1} \times \int_0^1 \frac{e_1^\top[A(t,x,y)\nabla_x z(t,x) + D(t,x,y)z(t,x)]}{e_1^\top A(t,x,y)e_1}\,dy_1.$$

For notation simplicity, we denote $\delta_1 = \delta, \delta_2 = 1-\delta$ in the following. By (2.8), we have

$$\int_0^1 \frac{1}{e_1^\top A(t,x,y)e_1}\,dy_1 = \sum_{j=1}^2 \frac{\delta_j}{e_1^\top A_j(t,x)e_1}$$

and

$$\int_0^1 \frac{e_1^\top[A(t,x,y)\nabla_x z(t,x) + D(t,x,y)z(t,x)]}{e_1^\top A(t,x,y)e_1}\,dy_1 = \sum_{j=1}^2 \frac{\delta_j e_1^\top[A_j(t,x)\nabla_x z(t,x) + D_j(t,x)z(t,x)]}{e_1^\top A_j(t,x)e_1},$$

which means

$$\zeta(t,x) = \sum_{j=1}^2 \frac{\delta_j e_1^\top[A_j(t,x)\nabla_x z(t,x) + D_j(t,x)z(t,x)]}{e_1^\top A_j(t,x)e_1} \Big/ \sum_{j=1}^2 \frac{\delta_j}{e_1^\top A_j(t,x)e_1}. \tag{2.33}$$

Then

$$\int_{\widetilde{Y}} A(t,x,y)\big(\nabla_x z(t,x) + \nabla_y z_1(t,s,x,y)\big)\,dy\,ds = \delta A_1(t,x)\nabla_x z(t,x) + (1-\delta)A_2(t,x)\nabla_x z(t,x) + \int_Y A(t,x,y)e_1\partial_{y_1} z_1(t,x,y_1)\,dy.$$

By (2.32) and (2.33), we have

$$\int_Y A(t,x,y)e_1\partial_{y_1} z_1(t,x,y_1)\,dy = \int_Y A(t,x,y)e_1 \frac{\zeta(t,x) - e_1^\top(A(t,x,y)\nabla_x z(t,x) + D(t,x,y)z(t,x))}{e_1^\top A(t,x,y)e_1}\,dy$$

$$
\begin{aligned}
&= \sum_{j=1}^{2} \frac{\delta_j A_j(t,x)e_1[\zeta(t,x) - e_1^\top(A_j(t,x)\nabla_x z(t,x) + D_j(t,x)z(t,x))]}{e_1^\top A_j(t,x)e_1} \\
&= \left(\sum_{j=1}^{2} \frac{\delta_j A_j(t,x)e_1}{e_1^\top A_j(t,x)e_1}\right)\left(\sum_{j=1}^{2} \frac{\delta_j e_1^\top(A_j(t,x)\nabla_x z(t,x) + D_j z(t,x))}{e_1^\top A_j(t,x)e_1}\right) \\
&\quad \bigg/ \left(\sum_{j=1}^{2} \frac{\delta_j}{e_1^\top A_j(t,x)e_1}\right) \\
&\quad - \sum_{j=1}^{2} \frac{\delta_j A_j(t,x)e_1 e_1^\top(A_j(t,x)\nabla_x z(t,x) + D_j(t,x)z(t,x))}{e_1^\top A_j(t,x)e_1} \\
&= \sum_{k=1}^{2} \left\{\left[\left(\sum_{j=1}^{2} \frac{\delta_j A_j(t,x)e_1}{e_1^\top A_j(t,x)e_1}\right) \bigg/ \left(\sum_{j=1}^{2} \frac{\delta_j}{e_1^\top A_j(t,x)e_1}\right) - A_k(t,x)e_1\right]\right. \\
&\quad \left. \times \frac{\delta_k e_1^\top(A_k(t,x)\nabla_x z(t,x) + D_k(t,x)z(t,x))}{e_1^\top A_k(t,x)e_1}\right\} \\
&= -\frac{\delta(1-\delta)(A_1(t,x) - A_2(t,x))e_1 e_1^\top(A_1(t,x) - A_2(t,x))}{\delta e_1^\top A_2(t,x)e_1 + (1-\delta)e_1^\top A_1(t,x)e_1} \nabla_x z(t,x) \\
&\quad - \frac{\delta(1-\delta)(A_1(t,x) - A_2(t,x))e_1 e_1^\top(D_1(t,x) - D_2(t,x))}{\delta e_1^\top A_2(t,x)e_1 + (1-\delta)e_1^\top A_1(t,x)e_1} z(t,x).
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
&\int_{\widetilde{Y}} A(t,x,y)(\nabla_x z(t,x) + \nabla_y z_1(t,s,x,y))\,dy\,ds \\
&\quad = \delta A_1(t,x)\nabla_x z(t,x) + (1-\delta)A_2(t,x)\nabla_x z(t,x) \\
&\qquad - \frac{\delta(1-\delta)(A_1(t,x) - A_2(t,x))e_1 e_1^\top(A_1(t,x) - A_2(t,x))}{\delta e_1^\top A_2(t,x)e_1 + (1-\delta)e_1^\top A_1(t,x)e_1} \nabla_x z(t,x) \\
&\qquad - \frac{\delta(1-\delta)(A_1(t,x) - A_2(t,x))e_1 e_1^\top(D_1(t,x) - D_2(t,x))}{\delta e_1^\top A_2(t,x)e_1 + (1-\delta)e_1^\top A_1(t,x)e_1} z(t,x). \qquad (2.34)
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\int_{\widetilde{Y}} B(t,x,y)\cdot(\nabla_x z(t,x) + \nabla_y z_1(t,s,x,y))\,dy\,ds \\
&\quad = \delta B_1(t,x)\cdot\nabla_x z(t,x) + (1-\delta)B_2(t,x)\cdot\nabla_x z(t,x) \\
&\qquad - \frac{\delta(1-\delta)(B_1(t,x) - B_2(t,x))^\top e_1 e_1^\top(A_1(t,x) - A_2(t,x))}{\delta e_1^\top A_2(t,x)e_1 + (1-\delta)e_1^\top A_1(t,x)e_1} \nabla_x z(t,x)
\end{aligned}
$$

$$- \frac{\delta(1-\delta)(B_1(t,x)-B_2(t,x))^\top e_1 e_1^\top (D_1(t,x)-D_2(t,x))}{\delta e_1^\top A_2(t,x)e_1 + (1-\delta)e_1^\top A_1(t,x)e_1} z(t,x) \quad (2.35)$$

and

$$\int_Y D(t,x,y)z(t,x)\,dy\,ds = \delta D_1(t,x)z(t,x) + (1-\delta)D_2(t,x)z(t,x). \quad (2.36)$$

Combining (2.30) with (2.34)–(2.36), we obtain

$$\begin{aligned} &\partial_{tt} z(t,x) - \nabla\cdot\big(A^*(t,x)\nabla z(t,x)\big) + B^*(t,x)\cdot\nabla z(t,x) \\ &\quad - \nabla\cdot\big(D^*(t,x)z(t,x)\big) = f^*(t,x,z(t,x)), \end{aligned} \quad (2.37)$$

where $A^*(t,x)$, $B^*(t,x)$, $D^*(t,x)$ and $f^*(t,x,z)$ are defined by (2.11).

Now, we verify $z(0,x)=z_0(x)$ and $\partial_t z(0,x)=z_1(x)$. For any function $\rho\in C^\infty([0,T])$ with $\rho(T)=\rho'(T)=0$ and $\varphi\in C_c^\infty(\Omega)$, it follows from (2.37) that

$$\begin{aligned} &\int_{\Omega_T}\Big[z(t,x)\varphi(x)\rho''(t) + \Big(A^*(t,x)\nabla z(t,x)\cdot\nabla\varphi(x) + (B^*(t,x)\cdot\nabla z(t,x))\varphi(x) \\ &\qquad + (D^*(t,x)\cdot\nabla\varphi(x))z(t,x) + c^*(t,x)z(t,x)\varphi(x)\Big)\rho(t)\Big]\,dx\,dt \\ &\quad = \int_{\Omega_T} f(t,x)\varphi(x)\rho(t)\,dx\,dt - \int_\Omega z(0,x)\varphi(x)\rho'(0)\,dx \\ &\qquad + \int_\Omega \partial_t z(0,x)\varphi(x)\rho(0)\,dx. \end{aligned} \quad (2.38)$$

Similarly, we deduce from (2.9) and (2.17) that

$$\begin{aligned} &\int_{\Omega_T}\Big\{z_\varepsilon(t,x)\varphi(x)\rho''(t) + \Big[A\Big(t,x,\frac{x}{\varepsilon}\Big)\nabla z_\varepsilon(t,x)\cdot\nabla\varphi(x) \\ &\qquad + \varphi(x)B\Big(t,x,\frac{x}{\varepsilon}\Big)\cdot\nabla z_\varepsilon(t,x) \\ &\qquad + z_\varepsilon(t,x)D\Big(t,x,\frac{x}{\varepsilon}\Big)\cdot\nabla\varphi(x) + c\Big(t,x,\frac{x}{\varepsilon}\Big)z_\varepsilon(t,x)\varphi(x)\Big]\rho(t)\Big\}\,dx\,dt \\ &\quad = \int_{\Omega_T} f_\varepsilon(t,x)\varphi(x)\rho(t)\,dx\,dt - \int_\Omega z_\varepsilon(0,x)\varphi(x)\rho'(0)\,dx \\ &\qquad + \int_\Omega \partial_t z_\varepsilon(0,x)\varphi(x)\rho(0)\,dx \\ &\quad = \int_{\Omega_T} f_\varepsilon(t,x)\varphi(x)\rho(t)\,dx\,dt - \int_\Omega z_0(x)\varphi(x)\rho'(0)\,dx + \int_\Omega z_1(x)\varphi(x)\rho(0)\,dx. \end{aligned}$$

Recall (2.24)–(2.27) and (2.34)–(2.36), we have

$$\int_{\Omega_T}\Big[z(t,x)\varphi(x)\rho''(t) + \Big(A^*(t,x)\nabla z(t,x)\cdot\nabla\varphi(x) + \big(B^*(t,x)\cdot\nabla z(t,x)\big)\varphi(x)$$

$$+ \Big(D^*(t,x)\cdot\nabla\varphi(x)\big)z(t,x) + c^*(t,x)z(t,x)\varphi(x)\Big)\rho(t)\Big]\,dx\,dt$$
$$= \int_{\Omega_T} f(t,x)\varphi(x)\rho(t)\,dx\,dt - \int_\Omega z_0(x)\varphi(x)\rho'(0)\,dx + \int_\Omega z_1(x)\varphi(x)\rho(0)\,dx. \tag{2.39}$$

Comparing (2.38) with (2.39), we deduce $z(0,x) = z_0(x)$ and $\partial_t z(0,x) = z_1(x)$. This ends the proof of Theorem 2.3. □

In order to prove Theorem 1.2, we need the following lemma:

Lemma 2.12 *Let $n \geq 2$ and $\xi, \eta \in \mathbb{R}^n$. Then*

$$\sup_{|x|=1} x^\top \xi \eta^\top x = \frac{|\xi||\eta| + \xi^\top \eta}{2}.$$

The proof of the above lemma is easy. See [13], for example.

3 Proof of the Main Theorem

In this section, we present a proof of Theorem 1.2. Having established Theorem 2.3, we can prove Theorem 1.2 similarly as in [13]. The proof is divided into four steps. Let $(\bar{u}(\cdot), \bar{v}(\cdot)) \in \mathcal{U}_{ad}$ be an optimal control and $\bar{z}(\cdot)$ be the corresponding optimal state. Let $(u(\cdot), v(\cdot)) \in \mathcal{U}_{ad}$ be fixed in Steps I–III.

I. Homogenizing spike variation of the control. Let $\delta \in (0,1)$ and $\varepsilon > 0$. For any $(t,x) = (t, x_1, x_2, \dots, x_n) \in \Omega_T$, define

$$(u^{\varepsilon,\delta}(t,x), v^{\varepsilon,\delta}(t,x)) = \begin{cases} (u(t,x), v(t,x)), & \text{if } \{\frac{x_1}{\varepsilon}\} \in [0,\delta), \\ (\bar{u}(t,x), \bar{v}(t,x)), & \text{if } \{\frac{x_1}{\varepsilon}\} \in [\delta, 1). \end{cases} \tag{3.1}$$

Then $(u^{\varepsilon,\delta}, v^{\varepsilon,\delta}) \in \mathcal{U}_{ad}$. Let $z^{\varepsilon,\delta}$ be the state corresponding to $(u^{\varepsilon,\delta}, v^{\varepsilon,\delta})$, i.e.,

$$\begin{cases} \partial_{tt} z^{\varepsilon,\delta}(t,x) - \nabla\cdot(A(t,x,u^{\varepsilon,\delta}(t,x))\nabla z^{\varepsilon,\delta}(t,x)) \\ \quad + B(t,x,u^{\varepsilon,\delta}(t,x))\cdot\nabla z^{\varepsilon,\delta}(t,x) - \nabla\cdot(D(t,x,u^{\varepsilon,\delta}(t,x))z^{\varepsilon,\delta}(t,x)) \\ \quad = f(t,x,z^{\varepsilon,\delta}(t,x), u^{\varepsilon,\delta}(t,x), v^{\varepsilon,\delta}(t,x)), \quad \text{in } \Omega_T, \\ z^{\varepsilon,\delta}(t,x) = 0, \quad \text{on } [0,T]\times\partial\Omega, \\ z^{\varepsilon,\delta}(0,x) = z_0(x), \quad \text{in } \Omega, \\ \partial_t z^{\varepsilon,\delta}(0,x) = z_1(x), \quad \text{in } \Omega. \end{cases} \tag{3.2}$$

Noting that

$$\begin{aligned}
&f(t,x,z^{\varepsilon,\delta}(t,x),u^{\varepsilon,\delta}(t,x),v^{\varepsilon,\delta}(t,x))\\
&\quad= f(t,x,z^{\varepsilon,\delta}(t,x),u(t,x),v(t,x))+r\Big(\frac{x_1}{\varepsilon}\Big)\Big(f(t,x,z^{\varepsilon,\delta}(t,x),\bar u(t,x),\bar v(t,x))\\
&\qquad- f(t,x,z^{\varepsilon,\delta}(t,x),u(t,x),v(t,x))\Big)
\end{aligned}$$

and

$$A(t,x,u^{\varepsilon,\delta}(t,x))=A(t,x,u(t,x))+r\Big(\frac{x_1}{\varepsilon}\Big)\big(A(t,x,\bar u(t,x))-A(t,x,u(t,x))\big),$$

$$B(t,x,u^{\varepsilon,\delta}(t,x))=B(t,x,u(t,x))+r\Big(\frac{x_1}{\varepsilon}\Big)\big(B(t,x,\bar u(t,x))-B(t,x,u(t,x))\big),$$

$$D(t,x,u^{\varepsilon,\delta}(t,x))=D(t,x,u(t,x))+r\Big(\frac{x_1}{\varepsilon}\Big)\big(D(t,x,\bar u(t,x))-D(t,x,u(t,x))\big),$$

we can apply Theorem 2.3 to (3.2) and get that as $\varepsilon\to 0^+$,

$$z^{\varepsilon,\delta}(\cdot)\longrightarrow z^{\delta}(\cdot),\quad \text{weakly in } L^2(0,T;H_0^1(\Omega))$$

with $z^{\delta}(\cdot)$ being the weak solution of

$$\begin{cases}
\partial_{tt}z^{\delta}(t,x)-\nabla\cdot(A^{\delta}(t,x)\nabla z^{\delta}(t,x))+B^{\delta}(t,x)\cdot\nabla z^{\delta}(t,x)\\
\quad-\nabla\cdot(D^{\delta}(t,x)z^{\delta}(t,x))=f^{\delta}(t,x,z^{\delta}(t,x)),\quad \text{in } \Omega_T,\\
z^{\delta}(t,x)=0,\quad \text{on } [0,T]\times\partial\Omega,\\
z^{\delta}(0,x)=z_0(x),\quad \text{in } \Omega,\\
\partial_t z^{\delta}(0,x)=z_1(x),\quad \text{in } \Omega,
\end{cases}\tag{3.3}$$

where

$$\begin{aligned}
A^{\delta}(t,x)&=\delta A(t,x,u(t,x))+(1-\delta)A(t,x,\bar u(t,x))\\
&\quad-\frac{\delta(1-\delta)(A(t,x,u(t,x))-A(t,x,\bar u(t,x)))e_1e_1^{\top}(A(t,x,u(t,x))-A(t,x,\bar u(t,x)))}{\delta e_1^{\top}A(t,x,\bar u(t,x))e_1+(1-\delta)e_1^{\top}A(t,x,u(t,x))e_1},\\
B^{\delta}(t,x)&=\delta B(t,x,u(t,x))+(1-\delta)B(t,x,\bar u(t,x))\\
&\quad-\frac{\delta(1-\delta)(A(t,x,u(t,x))-A(t,x,\bar u(t,x)))e_1e_1^{\top}(B(t,x,u(t,x))-B(t,x,\bar u(t,x)))}{\delta e_1^{\top}A(t,x,\bar u(t,x))e_1+(1-\delta)e_1^{\top}A(t,x,u(t,x))e_1},\\
D^{\delta}(t,x)&=\delta D(t,x,u(t,x))+(1-\delta)D(t,x,\bar u(t,x))\\
&\quad-\frac{\delta(1-\delta)(A(t,x,u(t,x))-A(t,x,\bar u(t,x)))e_1e_1^{\top}(D(t,x,u(t,x))-D(t,x,\bar u(t,x)))}{\delta e_1^{\top}A(t,x,\bar u(t,x))e_1+(1-\delta)e_1^{\top}A(t,x,u(t,x))e_1},\\
f^{\delta}(t,x,z)&=\delta f(t,x,z,u(t,x))+(1-\delta)f(t,x,z,\bar u(t,x))\\
&\quad+\frac{\delta(1-\delta)(B(t,x,u(t,x))-B(t,x,\bar u(t,x)))^{\top}e_1e_1^{\top}(D(t,x,u(t,x))-D(t,x,\bar u(t,x)))}{\delta e_1^{\top}A(t,x,\bar u(t,x))e_1+(1-\delta)e_1^{\top}A(t,x,u(t,x))e_1}z.
\end{aligned}$$

By (1.5),

$$\begin{aligned}&\left|f^0(t,x,z^{\varepsilon,\delta}(t,x),u^{\varepsilon,\delta}(t,x),v^{\varepsilon,\delta}(t,x))-f^0(t,x,z^\delta(t,x),u^{\varepsilon,\delta}(t,x),v^{\varepsilon,\delta}(t,x))\right|\\&\quad\le K|z^{\varepsilon,\delta}(t,x)-z^\delta(t,x)|,\quad (t,x)\in\Omega_T\end{aligned}\tag{3.4}$$

and

$$\left|f^0(t,x,z^\delta(t,x),u^{\varepsilon,\delta}(t,x),v^{\varepsilon,\delta}(t,x))\right|\le K+K|z^\delta(t,x)|,\quad (t,x)\in\Omega_T.\tag{3.5}$$

On the other hand,

$$z^{\varepsilon,\delta}(\cdot)\longrightarrow z^\delta(\cdot),\quad \text{strongly in } L^2(Q_T)$$

since it follows from Proposition 2.1 that $z^{\varepsilon,\delta}(\cdot)$ is bounded uniformly in $H^1(Q_T)$. Thus, by (3.4),

$$\begin{aligned}\lim_{\varepsilon\to0^+}\int_{\Omega_T}\Big[&f^0(t,x,z^{\varepsilon,\delta}(t,x),u^{\varepsilon,\delta}(t,x),v^{\varepsilon,\delta}(t,x))\\&-f^0(t,x,z^\delta(t,x),u^{\varepsilon,\delta}(t,x),v^{\varepsilon,\delta}(t,x))\Big]\,dx\,dt=0.\end{aligned}$$

Therefore, by (3.5) and the optimality of $(\bar u(\cdot),\bar v(\cdot))$, we have

$$\begin{aligned}J(\bar u(\cdot),\bar v(\cdot))&\le J_\delta\equiv\lim_{\varepsilon\to0^+}J(u^{\varepsilon,\delta}(\cdot),v^{\varepsilon,\delta}(\cdot))\\&=\lim_{\varepsilon\to0^+}\int_{\Omega_T}f^0(t,x,z^{\varepsilon,\delta}(t,x),u^{\varepsilon,\delta}(t,x),v^{\varepsilon,\delta}(t,x))\,dx\,dt\\&=\lim_{\varepsilon\to0^+}\int_{\Omega_T}f^0(t,x,z^\delta(t,x),u^{\varepsilon,\delta}(t,x),v^{\varepsilon,\delta}(t,x))\,dx\,dt\\&=\int_{\Omega_T}\Big(\delta f^0(t,x,z^\delta(t,x),u(t,x),v(t,x))\\&\quad+(1-\delta)f^0(t,x,z^\delta(t,x),\bar u(t,x),\bar v(t,x))\Big)\,dx\,dt.\end{aligned}\tag{3.6}$$

II. Variations of state and cost functional. We try to calculate

$$\lim_{\delta\to0^+}\frac{J^\delta-J(\bar u(\cdot),\bar v(\cdot))}{\delta}.$$

To this aim, we need study the variation of state first. Denote

$$Z^\delta(t,x)=\frac{z^\delta(t,x)-\bar z(t,x)}{\delta},\quad (t,x)\in\Omega_T.$$

Then it follows from (3.3) that

$$\begin{cases} \partial_{tt} Z^\delta(t,x) - \nabla\cdot(A(t,x,\bar u(t,x))\nabla Z^\delta(t,x)) + B(t,x,\bar u(t,x))\cdot\nabla Z^\delta(t,x) \\ \qquad - \nabla\cdot(D(t,x,\bar u(t,x))Z^\delta(t,x)) - \nabla\cdot\big(\frac{A^\delta(t,x)-A(t,x,\bar u(t,x))}{\delta}\nabla z^\delta(t,x)\big) \\ \qquad + \big(\frac{B^\delta(t,x)-B(t,x,\bar u(t,x))}{\delta}\cdot\nabla z^\delta(t,x)\big) - \nabla\cdot\big(\frac{D^\delta(t,x)-D(t,x,\bar u(t,x))}{\delta} z^\delta(t,x)\big) \\ \quad = f(t,x,z^\delta(t,x),u(t,x),v(t,x)) - f(t,x,z^\delta(t,x),\bar u(t,x),\bar v(t,x)) \\ \qquad + \int_0^1 f_z\big(t,x,\bar z(t,x)+\beta(z^\delta(t,x)-\bar z(t,x)),\bar u(t,x),\bar v(t,x)\big)\,d\beta\, Z^\delta(t,x) \\ \qquad + \Gamma^\delta(t,x)z^\delta(t,x), \quad \text{in } \Omega_T, \\ Z^\delta(t,x)=0, \quad \text{on } [0,T]\times\partial\Omega, \\ Z^\delta(0,x)=0, \quad \text{in } \Omega, \\ \partial_t Z^\delta(0,x)=0, \quad \text{in } \Omega, \end{cases} \tag{3.7}$$

where

$$\Gamma^\delta(t,x) = \frac{(1-\delta)(B(t,x,u(t,x))-B(t,x,\bar u(t,x)))^\top e_1 e_1^\top (D(t,x,u(t,x))-D(t,x,\bar u(t,x)))}{\delta e_1^\top A(t,x,\bar u(t,x))e_1 + (1-\delta)e_1^\top A(t,x,u(t,x))e_1}.$$

One can verify directly that

$$\begin{cases} \frac{A^\delta(t,x)-A(t,x,\bar u(t,x))}{\delta} \longrightarrow \Theta(t,x), \\ \frac{B^\delta(t,x)-B(t,x,\bar u(t,x))}{\delta} \longrightarrow \Xi(t,x), \\ \frac{D^\delta(t,x)-D(t,x,\bar u(t,x))}{\delta} \longrightarrow \Psi(t,x), \\ \Gamma^\delta(t,x) \longrightarrow \Gamma(t,x), \end{cases} \quad \text{strongly in } L^\infty(\Omega_T), \tag{3.8}$$

where

$$\Theta(t,x) = A(t,x,u(t,x)) - A(t,x,\bar u(t,x)) - \frac{(A(t,x,u(t,x))-A(t,x,\bar u(t,x)))e_1 e_1^\top (A(t,x,u(t,x))-A(t,x,\bar u(t,x)))}{e_1^\top A(t,x,u(t,x))e_1},$$

$$\Xi(t,x) = B(t,x,u(t,x)) - B(t,x,\bar u(t,x)) - \frac{(A(t,x,u(t,x))-A(t,x,\bar u(t,x)))e_1 e_1^\top (B(t,x,u(t,x))-B(t,x,\bar u(t,x)))}{e_1^\top A(t,x,u(t,x))e_1},$$

$$\Psi(t,x) = D(t,x,u(t,x)) - D(t,x,\bar u(t,x)) - \frac{(A(t,x,u(t,x))-A(t,x,\bar u(t,x)))e_1 e_1^\top (D(t,x,u(t,x))-D(t,x,\bar u(t,x)))}{e_1^\top A(t,x,u(t,x))e_1},$$

$$\Gamma(t,x) = \frac{(B(t,x,u(t,x))-B(t,x,\bar u(t,x)))^\top e_1 e_1^\top (D(t,x,u(t,x))-D(t,x,\bar u(t,x)))}{e_1^\top A(t,x,u(t,x))e_1}.$$

On the other hand, $z^{\delta}(\cdot)$ is bounded uniformly in $L^2(0,T;H_0^1(\Omega))$. Thus, we can prove step-by-step that as $\delta\to 0^+$, Z^{δ} is bounded uniformly in $L^2(0,T;H_0^1(\Omega))$, $z^{\delta}(\cdot)$ converges strongly to $\bar z(\cdot)$ in $L^2(0,T;H_0^1(\Omega))$, and $Z^{\delta}(\cdot)$ converges weakly to $Z(\cdot)$ in $L^2(0,T;H_0^1(\Omega))$ with $Z(\cdot)$ being the weak solution of

$$\begin{cases} \partial_{tt}Z(t,x)-\nabla\cdot(A(t,x,\bar u(t,x))\nabla Z(t,x))+B(t,x,\bar u(t,x))\cdot\nabla Z(t,x)\\ \qquad -\nabla\cdot(D(t,x,\bar u(t,x))Z(t,x))-\nabla\cdot(\Theta(t,x)\nabla\bar z(t,x))\\ \qquad +(\Xi(t,x)\cdot\nabla\bar z(t,x))-\nabla\cdot(\Psi(t,x)\bar z(t,x))\\ \quad =f(t,x,\bar z(t,x),u(t,x),v(t,x))-f(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x))\\ \qquad +f_z(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x))\,Z(t,x)+\Gamma(t,x)\bar z(t,x),\quad \text{in } \Omega_T,\\ Z(t,x)=0,\quad \text{on } [0,T]\times\partial\Omega,\\ Z(0,x)=0,\quad \text{in } \Omega,\\ \partial_t Z(0,x)=0,\quad \text{in } \Omega. \end{cases} \tag{3.9}$$

Thus, it follows from (3.6) that

$$\begin{aligned} 0&\le \lim_{\delta\to 0^+}\lim_{\varepsilon\to 0^+}\frac{J(u^{\varepsilon,\delta}(\cdot),v^{\varepsilon,\delta}(\cdot))-J(\bar u(\cdot),\bar v(\cdot))}{\delta}=\lim_{\delta\to 0^+}\frac{J^{\delta}-J(\bar u(\cdot),\bar v(\cdot))}{\delta}\\ &=\lim_{\delta\to 0^+}\Big[\int_{\Omega_T}\Big(f^0(t,x,z^{\delta}(t,x),u(t,x),v(t,x))\\ &\qquad -f^0(t,x,z^{\delta}(t,x),\bar u(t,x),\bar v(t,x))\Big)\,dx\,dt\\ &\qquad +\int_{\Omega_T}dx\,dt\int_0^1 f_z^0(t,x,\bar z(t,x)+\beta(z^{\delta}(t,x)\\ &\qquad -\bar z(t,x)),\bar u(t,x),\bar v(t,x))Z^{\delta}(t,x)\,d\beta\Big]\\ &=\int_{\Omega_T}\Big(f^0(t,x,\bar z(t,x),u(t,x),v(t,x))-f^0(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x))\\ &\qquad +f_z^0(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x))Z(t,x)\Big)\,dx\,dt. \end{aligned} \tag{3.10}$$

III. Duality. We introduce the adjoint equation (1.6). By (S3), (S4) and Proposition 2.1, (1.6) has a unique weak solution $\bar\psi(\cdot)$. Then (3.10) can be rewritten as

$$\begin{aligned} 0\le \int_{\Omega_T}&\Big\{f^0(t,x,\bar z(t,x),u(t,x),v(t,x))-f^0(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x))\\ &-\Big[\partial_{tt}\bar\psi(t,x)-\nabla\cdot\big(A(t,x,\bar u(t,x))\nabla\bar\psi(t,x)\big)-\nabla\cdot\big(B(t,x,\bar u(t,x))\bar\psi(t,x)\big)\\ &+D(t,x,\bar u(t,x))\cdot\nabla\bar\psi(t,x))\\ &-f_z(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x))\bar\psi(t,x)\Big]Z(t,x)\Big\}\,dx\,dt \end{aligned}$$

$$
\begin{aligned}
&= \int_{\Omega_T} \Big\{ f^0(t,x,\bar z(t,x),u(t,x),v(t,x)) - f^0(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x)) \\
&\quad - \Big[\partial_{tt} Z(t,x) - \nabla \cdot \big(A(t,x,\bar u(t,x)) \nabla Z(t,x) \big) + B(t,x,\bar u(t,x)) \cdot \nabla Z(t,x) \\
&\quad - \nabla \cdot \big(D(t,x,\bar u(t,x)) Z(t,x) \big) - f_z(t,x,\bar z(t,x),\bar u(t,x)) Z(t,x) \Big] \Big\} \bar\psi(t,x)\, dx\, dt \\
&= \int_{\Omega_T} \Big\{ f^0(t,x,\bar z(t,x),u(t,x),v(t,x)) - f^0(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x)) \\
&\quad - \Big[\nabla \cdot \big(\Theta(t,x) \nabla \bar z(t,x) \big) - \big(\Xi(t,x) \cdot \nabla \bar z(t,x) \big) + \nabla \cdot \big(\Psi(t,x) \bar z(t,x) \big) \\
&\quad + \Gamma(t,x) \bar z(t,x) + f(t,x,\bar z(t,x),u(t,x),v(t,x)) \\
&\quad - f(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x)) \Big] \bar\psi(t,x) \Big\}\, dx\, dt \\
&= \int_{\Omega_T} \Big\{ \Big(f(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x)) \bar\psi(t,x) - f^0(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x)) \Big) \\
&\quad - \Big(f(t,x,\bar z(t,x),u(t,x),v(t,x)) \bar\psi(t,x) - f^0(t,x,\bar z(t,x),u(t,x),v(t,x)) \Big) \\
&\quad + \Big(\Theta(t,x) \nabla \bar z(t,x) \cdot \nabla \bar\psi(t,x) + \Xi(t,x) \bar\psi(t,x) \cdot \nabla \bar z(t,x) \\
&\quad + \Psi(t,x) \bar z(t,x) \cdot \nabla \bar\psi(t,x) - \Gamma(t,x) \bar z(t,x) \psi(t,x) \Big) \Big\}\, dx\, dt \\
&= \int_{\Omega_T} \Big\{ \Big[\Big(f(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x)) \bar\psi(t,x) - f^0(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x)) \\
&\quad - A(t,x,\bar u(t,x)) \nabla \bar z(t,x) \cdot \nabla \bar\psi(t,x) - B(t,x,\bar u(t,x)) \bar\psi(t,x) \cdot \nabla \bar z(t,x) \\
&\quad - D(t,x,\bar u(t,x)) \bar z(t,x) \cdot \nabla \bar\psi(t,x) \Big) \\
&\quad - \Big(f(t,x,\bar z(t,x),u(t,x),v(t,x)) \bar\psi(t,x) - f^0(t,x,\bar z(t,x),u(t,x),v(t,x)) \\
&\quad - A(t,x,u(t,x)) \nabla \bar z(t,x) \cdot \nabla \bar\psi(t,x) - B(t,x,u(t,x)) \bar\psi(t,x) \cdot \nabla \bar z(t,x) \\
&\quad - D(t,x,u(t,x)) \bar z(t,x) \cdot \nabla \bar\psi(t,x) \Big) \Big] \\
&\quad - \frac{e_1^\top p(t,x,\bar u(t,x),u(t,x)) q(t,x,\bar u(t,x),u(t,x))^\top e_1}{e_1^\top A(t,x,u(t,x)) e_1} \Big\}\, dx\, dt.
\end{aligned}
$$

That is,

$$
\begin{aligned}
0 \le \int_{\Omega_T} \Big[& H(t,x,\bar z(t,x),\bar\psi(t,x),\nabla \bar z(t,x),\nabla \bar\psi(t,x),\bar u(t,x),\bar v(t,x)) \\
& - H(t,x,\bar z(t,x),\bar\psi(t,x),\nabla \bar z(t,x),\nabla \bar\psi(t,x),u(t,x),v(t,x)) \\
& - \frac{e_1^\top p(t,x,\bar u(t,x),u(t,x)) q(t,x,\bar u(t,x),u(t,x))^\top e_1}{e_1^\top A(t,x,u(t,x)) e_1} \Big]\, dx\, dt, \qquad (3.11)
\end{aligned}
$$

where H, p and q are defined by (1.9)–(1.10).

IV. Maximum condition. When we try to yield (1.8) or (1.7) from (3.11), we need to mention the definition of $\mathcal{U}_{ad}$ is quite different from those in elliptic and parabolic cases. For a control $(u(\cdot), v(\cdot))$ in $\mathcal{U}_{ad}$, $u(\cdot)$ is demanded to be differentiable in t. This makes some difficulties. However, under assumption (S2), as Remark 1.1 indicated, we can see easily that for any

$$(u(\cdot), v(\cdot)) \in L^\infty(\Omega_T; U) \times \mathcal{M}(\Omega_T; V),$$

there exists a sequence $u^k(\cdot)$, such that $(u^k(\cdot), v(\cdot)) \in \mathcal{U}_{ad}$ and $u^k(\cdot)$ converges to $u(\cdot)$ almost everywhere in Ω_T. Consequently, we can deduce from (3.11) that

$$\begin{aligned} 0 \le \int_{\Omega_T} \Big[& H(t,x,\bar{z}(t,x), \bar{\psi}(t,x), \nabla\bar{z}(t,x), \nabla\bar{\psi}(t,x), \bar{u}(t,x), \bar{v}(t,x)) \\ & - H(t,x,\bar{z}(t,x), \bar{\psi}(t,x), \nabla\bar{z}(t,x), \nabla\bar{\psi}(t,x), u(t,x), v(t,x)) \\ & - \frac{e_1^\top p(t,x,\bar{u}(t,x),u(t,x)) q(t,x,\bar{u}(t,x),u(t,x))^\top e_1}{e_1^\top A(t,x,u(t,x)) e_1} \Big] dx\, dt, \\ & \forall (u(\cdot), v(\cdot)) \in L^\infty(\Omega_T; U) \times \mathcal{M}(\Omega_T; V). \end{aligned} \tag{3.12}$$

Then a standard discussion leads to

$$\begin{aligned} & H(t,x,\bar{z}(t,x), \bar{\psi}(t,x), \nabla\bar{z}(t,x), \nabla\bar{\psi}(t,x), \bar{u}(t,x), \bar{v}(t,x)) \\ & \quad - H(t,x,\bar{z}(t,x), \bar{\psi}(t,x), \nabla\bar{z}(t,x), \nabla\bar{\psi}(t,x), u, v) \\ & \ge \frac{e_1^\top p(t,x,\bar{u}(t,x),u) q(t,x,\bar{u}(t,x),u)^\top e_1}{e_1^\top A(t,x,u) e_1}, \quad \forall (u,v) \in U \times V, \text{ a.e. } (t,x) \in \Omega_T. \end{aligned} \tag{3.13}$$

If $n = 1$, the above gives (1.7).

If $n \ge 2$, noting that the unit sphere S^{n-1} of $\mathbb{R}^n$ is separable, we can generalize (3.13) to the following:

$$\begin{aligned} & H(t,x,\bar{z}(t,x), \bar{\psi}(t,x), \nabla\bar{z}(t,x), \nabla\bar{\psi}(t,x), \bar{u}(t,x), \bar{v}(t,x)) \\ & \quad - H(t,x,\bar{z}(t,x), \bar{\psi}(t,x), \nabla\bar{z}(t,x), \nabla\bar{\psi}(t,x), u, v) \\ & \ge \sup_{e \in S^{n-1}} \frac{e^\top p(t,x,\bar{u}(t,x),u) q(t,x,\bar{u}(t,x),u)^\top e}{e^\top A(t,x,u) e}, \\ & \forall (u,v) \in U \times V, \text{ a.e. } (t,x) \in \Omega_T. \end{aligned}$$

Then (1.8) follows from Lemma 2.12. This completes the proof of Theorem 1.2. □

Remark 3.1 If U is only a subset of U_0, we can still get (3.11). And then we can get

$$0 \le \int_0^T \Big[H(t,x,\bar{z}(t,x), \bar{\psi}(t,x), \nabla\bar{z}(t,x), \nabla\bar{\psi}(t,x), \bar{u}(t,x), \bar{v}(t,x))$$

$$
\begin{aligned}
&- \max_{v\in V} \Big(H(t,x,\bar z(t,x),\bar\psi(t,x),\nabla\bar z(t,x),\nabla\bar\psi(t,x),u(t),v) \Big) \\
&- \max_{e\in S^{n-1}} \frac{e_1^\top p(t,x,\bar u(t,x),u(t)) q(t,x,\bar u(t,x),u(t))^\top e_1}{e_1^\top A(t,x,u(t)) e_1} \Big) \Big] dt, \\
&\forall u(\cdot)\in W^{1,\infty}([0,T];U), \quad \text{a.e. } x\in\Omega.
\end{aligned}
$$

4 Problem with State Constraints

In this section, we will give a necessary optimality conditions for the cases of state constraint. We only state the result in this section since the proof is a combination of that for Theorem 1.2 and that for Theorem 5.1.2 in [11].

The state constraint is given by

$$
F(z(\cdot)) \in E, \tag{4.1}
$$

where $F : L^2(0,T;H_0^1(\Omega)) \to \mathcal{Z}$ with $\mathcal{Z}$ being a Banach space, and E is a subset of $\mathcal{Z}$. We need the following assumption about $\mathcal{Z}$, F and E:

(S6) Let $\mathcal{Z}$ be a Banach space with strictly convex dual $\mathcal{Z}^*$, $F : L^2(0,T;H_0^1(\Omega)) \to \mathcal{Z}$ be continuous Fréchet differentiable, and $E \subset \mathcal{Z}$ be closed and convex.

We call $(z(\cdot),(u(\cdot),v(\cdot))$ an admissible pair if it satisfies (1.1) and (4.1). Denote $\mathcal{P}_{ad}$ the set of all the admissible pair. Define the set of admissible controls by

$$
\mathcal{U}_{ad} \equiv \{(u(\cdot),v(\cdot)) \big| (z(\cdot),(u(\cdot),v(\cdot)) \in \mathcal{P}_{ad}\}.
$$

Problem (SC). Find a control $(\bar u(\cdot),\bar v(\cdot)) \in \mathcal{U}_{ad}$ such

$$
J(\bar u(\cdot),\bar v(\cdot)) = \inf_{(u(\cdot),v(\cdot))\in\mathcal{U}_{ad}} J(u(\cdot),v(\cdot)). \tag{4.2}
$$

Before stating necessary conditions for optimal control of Problem (SC), we need to recall the notion of finite co-dimensional (see Chapter 4 of [11], for example).

Definition 4.1 Let X be a Banach space and X_0 be a subspace of X. We say that X_0 is finite co-dimensional in X if there exist $x_1, x_2, \dots, x_m \in X$, such that

$$
\text{span}\,\{X_0, x_1, \dots, x_m\} \equiv \text{the space spanned by} \{X_0, x_1, \dots, x_n\} = X.
$$

A subset S of X is said to be finite co-dimensional in X if for some $x_0 \in S$, span $(S - \{x_0\}) \equiv$ the closed subspace spanned by $\{x - x_0 | x \in S\}$ is a finite co-dimensional subspace of X and $\overline{\text{co}}\,\text{S} \equiv$ the closed convex hull of $S - \{x_0\}$ has a nonempty interior in this subspace.

Let $(\bar z(\cdot),(\bar u(\cdot),\bar v(\cdot)))$ be an optimal pair of Problem (SC). Let $Z = Z(\cdot; u(\cdot), v(\cdot)) \in L^2(0,T;H_0^1(\Omega))$ be the unique weak solution of the variational equation (3.9). Define the reachable set of variational system (3.9) by

$$
\mathcal{R} = \{Z(\cdot; u(\cdot), v(\cdot)) \big| (u(\cdot),v(\cdot)) \in \mathcal{U}_{ad}\}.
$$

Now, we can state the analog result of Theorem 1.2 for Problem (SC) as follows:

Theorem 4.2 *Assume* (S1)–(S6) *hold. Let* $(\bar z(\cdot),(\bar u(\cdot),\bar v(\cdot)))$ *be an optimal pair of Problem* (SC) *and*

$$F'(\bar z(\cdot))\mathcal{R}-E\equiv\{\xi-\eta\,\big|\,\xi\in F'(\bar z(\cdot))\mathcal{R},\ \eta\in E\}$$

be finite co-dimensional in $\mathcal{Z}$. *Then there exists a triple* $(\bar\psi_0,\bar\psi(\cdot),\bar\varphi(\cdot))\in\mathbb{R}\times L^2(0,T;H_0^1(\Omega))\times\mathcal{Z}^*$ *satisfying*

$$\begin{cases}\bar\psi_0\le 0,\\ (\bar\psi_0,\bar\varphi(\cdot))\ne 0,\\ (\bar\psi_0,\bar\psi(\cdot))\ne 0, \quad \textit{if } F'(\bar z(\cdot))^* \textit{ is injective},\end{cases}$$

$$\langle\bar\varphi(\cdot),\eta-F(\bar z(\cdot))\rangle_{\mathcal{Z}^*,\mathcal{Z}}\le 0,\quad \forall\eta\in E,$$

$$\begin{cases}\partial_{tt}\bar\psi(t,x)-\nabla\cdot(A(t,x,\bar u(t,x))\nabla\bar\psi(t,x))-\nabla\cdot(B(t,x,\bar u(t,x))\bar\psi(t,x))\\ \quad +D(t,x,\bar u(t,x))\cdot\nabla\bar\psi(t,x)-f_z(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x))\bar\psi(t,x)\\ \quad -\bar\psi_0 f_z^0(t,x,\bar z(t,x),\bar u(t,x),\bar v(t,x))+F'(\bar z(\cdot))^*\bar\varphi=0,\quad \textit{in } \Omega_T,\\ \bar\psi(t,x)=0,\quad \textit{on } [0,T]\times\partial\Omega,\\ \bar\psi(T,x)=0,\quad \textit{in } \Omega,\\ \partial_t\bar\psi(T,x)=0,\quad \textit{in } \Omega,\end{cases}\tag{4.3}$$

such that when $n=1$,

$$\begin{aligned}&H(t,x,\bar z(t,x),\bar\psi(t,x),\partial_x\bar z(t,x),\partial_x\bar\psi(t,x),\bar u(t,x),\bar v(t,x))\\ &\quad -H(t,x,\bar z(t,x),\bar\psi(t,x),\partial_x\bar z(t,x),\partial_x\bar\psi(t,x),u,v)\\ &\ge\frac{p(t,x,\bar u(t,x),u)q(t,x,\bar u(t,x),u)}{A(t,x,u)},\quad \forall(u,v)\in U\times V, a.e.\ (t,x)\in\Omega_T,\end{aligned}\tag{4.4}$$

and when $n\ge 2$,

$$\begin{aligned}&H(t,x,\bar z(t,x),\bar\psi(t,x),\nabla\bar z(t,x),\nabla\bar\psi(t,x),\bar u(t,x),\bar v(t,x))\\ &\quad -H(t,x,\bar z(t,x),\bar\psi(t,x),\nabla\bar z(t,x),\nabla\bar\psi(t,x),u,v)\\ &\ge\frac12\left|A(t,x,u)^{-\frac12}p(t,x,,\bar u(t,x),u)\right|\left|A(t,x,u)^{-\frac12}q(t,x,\bar u(t,x),u)\right|\\ &\quad +\frac12 A^{-1}(t,x,u)p(t,x,\bar u(t,x),u)\cdot q(t,x,\bar u(t,x),u),\\ &\forall(u,v)\in U\times V,\ \text{a.e. } (t,x)\in\Omega_T,\end{aligned}\tag{4.5}$$

where for $(t,x,z,\psi,\xi,\eta,u,v)\in[0,T]\times\Omega\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times U\times V$,

$$
\begin{aligned}
H(t,x,z,\psi,\xi,\eta,u,v) = \psi f(t,x,z,u,v) + \bar{\psi}_0 f^0(t,x,z,u,v) - A(t,x,u)\xi\cdot\eta \\
- B(t,x,u)\psi\cdot\eta - D(t,x,u)z\cdot\xi
\end{aligned}
$$

and for $(t,x,u_1,u_2)\in[0,T]\times\Omega\times U\times U$,

$$
\begin{aligned}
p(t,x,u_1,u_2) &= [A(t,x,u_1) - A(t,x,u_2)]\nabla\bar{z}(t,x) \\
&\quad + [D(t,x,u_1) - D(t,x,u_2)]\bar{z}(t,x), \\
q(t,x,u_1,u_2) &= [A(t,x,u_1) - A(t,x,u_2)]\nabla\bar{\psi}(t,x) \\
&\quad + [B(t,x,u_1) - B(t,x,u_2)]\psi(t,x).
\end{aligned}
$$

5 A Cursory Discussion on the Existence of Optimal Controls

It is generally considered that usually optimal control does not exist for optimal control problem with leading term containing controls. However, we think that there are still many cases that admit optimal controls.

If f and f^0 are independent of (u,v), then an optimal control exists if

$$
\{(A(\cdot;u(\cdot)),B(\cdot;u(\cdot)),D(\cdot;u(\cdot)))|u(\cdot)\in\{w:\Omega_T\to U|w(\cdot),\partial_t w(\cdot)\in L^\infty(\Omega_T;U)\}\}
$$

is closed in some sense like G-closedness (see [4] for the elliptic cases).

If A, B, and D are independent of u, then a Cesari-type condition and some mild conditions would guarantee the existence of an optimal control.

For general cases, an effective method to study existence theory is the relaxed control theory. Usually, optimal relaxed control exists.

When A, B, D are independent of the control variable, a relaxed control can be described as a probability measure valued function. If we can prove that an optimal relaxed control takes Dirac measures almost everywhere, then we get an optimal control. In particular, if for any (t,x,z,ψ,ξ,η),

$$
H(t,x,z,\psi,\xi,\eta,\bar{u},\bar{v}) = \max_{(u,v)\in U\times V} H(t,x,z,\psi,\xi,\eta,u,v)
$$

always admits a unique solution $(\bar{u},\bar{v})$, then the support of an optimal relaxed control will be a singleton for any (t,x). And consequently, an optimal control exists. Particularly, this would happen when the control domain $U\times V$ is a convex subset of $\mathbb{R}^m\times\mathbb{R}^k$ and as a function of (u,v), $H(t,x,z,\psi,\xi,\eta,u,v)$ is strictly concave.

Unfortunately, when A, B and D depend on the control variable, relaxed control could not be described simply as a probability measure valued function. It should contain two parts, a probability measure valued function and the corresponding coefficients A^*, B^* and D^*. More precisely, the relaxed system of Problem (C) would be described in the following form

$$\begin{cases} \partial_{tt} z(t,x) - \nabla \cdot \big(A^*(t,x) \nabla z(t,x)\big) + B^*(t,x) \cdot \nabla z(t,x) - \nabla \cdot \big(D^*(t,x) z(t,x)\big) \\ \quad = \int_{U \times V} f(t,x,z(t,x),u,v) \sigma(t,x)(du)\, \mu(t,x)(dv), \quad \text{in } \Omega_T, \\ z(t,x) = 0, \quad \text{on } [0,T] \times \partial\Omega, \\ z(0,x) = z_0(x), \quad \text{in } \Omega, \\ \partial_t z(0,x) = z_1(x), \quad \text{in } \Omega \end{cases} \tag{5.1}$$

while the relaxed cost functional would be

$$J\big(\sigma, \mu, A^*, B^*, D^*\big) = \int_{\Omega_T} dt\, dx \int_{U \times V} f^0(t,x,z(t,x),u,v) \sigma(t,x)(du)\, \mu(t,x)(dv), \tag{5.2}$$

where $(\sigma(\cdot), \mu(\cdot))$ is a probability measure valued function on Ω_T and (A^*, B^*, D^*) depends on $(\sigma(\cdot), \mu(\cdot))$ but it is not determined uniquely by $(\sigma(\cdot), \mu(\cdot))$. In particular, usually,

$$(A^*(t,x), B^*(t,x), D^*(t,x)) \neq \int_U (A(t,x,u), B(t,x,u), D(t,x,u)) \sigma(t,x)(du),$$

which makes the relaxed control theory of Problem (C) much different from that of problems with no control in coefficients. Thus, suppose $n \geq 2$ for simplicity, even if

$$H(t,x,z,\psi,\xi,\eta,\bar{u},\bar{v}) = \max_{(u,v) \in U \times V} H(t,x,z,\psi,\xi,\eta,u,v)$$

always admits a unique solution, we can not get naturally that the support of $(\bar{\sigma}(t,x), \bar{\mu}(t,x))$ is a singleton for an optimal relaxed control $(\bar{\sigma}, \bar{\mu}, \overline{A}^*, \overline{B}^*, \overline{D}^*)$. Nevertheless, one can expect that

$$\begin{aligned} &\left\| A^*(t,x) - \int_U A(t,x,u) \sigma(t,x)(du) \right\| \\ &\quad \leq C \int_U \left\| A(t,x,w) - \int_U A(t,x,u) \sigma(t,x)(du) \right\|^2 \sigma(t,x)(dw) \end{aligned}$$

holds under some mild conditions (we can see from (2.11) that the above inequality holds for some simple cases). Then, if H is "concave enough" (for example, as a function of (u,v), $H + M\|u\|^2$ is strictly concave for some M large enough), we can get from optimal conditions of an optimal relaxed control $(\bar{\sigma}, \bar{\mu}, \overline{A}^*, \overline{B}^*, \overline{D}^*)$ that the corresponding $(\bar{\sigma}, \bar{\mu})$ takes Dirac measures. And then we will get the existence of optimal control.

Acknowledgements The authors thank the anonymous referees for their helpful suggestions.

This work was supported in part by NSFC (No. 61074047 and 10831007), and 973 Program (No. 2011CB808002).

Appendix

Proof of Proposition 2.1

Existence. Step 1: Galerkin approximations. Define the time-dependent bilinear form $\mathcal{B}[x, w; t]$ by

$$\mathcal{B}[z, w; t] = \int_\Omega \Big[A(t,x)\nabla z(t,x) \cdot \nabla w(t,x) + (B(t,x) \cdot \nabla z(t,x))w(t,x) + (D(t,x) \cdot \nabla w(t,x))z(t,x) + c(t,x)z(t,x)w(t,x) \Big]\, dx. \tag{A.1}$$

Write

$$z_m(t,x) = \sum_{k=1}^m d_{m,k}(t) w_k(x),$$

where $\{w_k\}_{k=1}^\infty$ is an orthogonal basis of $H_0^1(\Omega)$ and an orthonormal basis of $L^2(\Omega)$. It is well-known, such a base exists. Similar to the proof of Theorem 7.2.1 in [9], we can find unique $W^{2,2}(0,T)$ coefficients $d_{m,k}(t)$ $(0 \le t \le T,\ k = 1, 2, \ldots, m)$ satisfying

$$d_{m,k}(0) = \int_\Omega z_0(x) w_k(x)\, dx,$$
$$d'_{m,k}(0) = \int_\Omega z_1(x) w_k(x)\, dx$$

and

$$\int_\Omega \partial_{tt} z_m(t,x) w_k(x)\, dx + \mathcal{B}[z_m, w_k; t] = \int_\Omega f(t,x) w_k(x)\, dx. \tag{A.2}$$

Step 2: Energy estimates. By (A.2), we have

$$\sum_{k=1}^m d'_{m,k}(t) \Big(\int_\Omega \partial_{tt} z_m(t,x) w_k(x)\, dx + \mathcal{B}[z_m, w_k; t] \Big) = \sum_{k=1}^m d'_{m,k}(t) \int_\Omega f(t,x) w_k(x)\, dx, \quad \text{a.e. } t \in [0,T].$$

That is,

$$\int_\Omega \partial_{tt} z_m(t,x) \partial_t z_m(x)\, dx + \mathcal{B}[z_m, \partial_t z_m; t] = \int_\Omega f(t,x) \partial_t z_m(t,x)\, dx, \quad \text{a.e. } t \in [0,T]. \tag{A.3}$$

Since A is symmetric,

$$
\begin{aligned}
&\mathcal{B}[z_m, \partial_t z_m; t] \\
&\quad = \frac{d}{dt}\Big[\int_\Omega \Big(\frac{1}{2}A(t,x)\nabla z_m(t,x)\cdot\nabla z_m(t,x) + (D(t,x)\cdot\nabla z_m(t,x))z_m(t,x)\Big)\,dx\Big] \\
&\qquad - \int_\Omega \Big(\frac{1}{2}\partial_t A(t,x)\nabla z_m(t,x)\cdot\nabla z_m(t,x) + (\partial_t D(t,x)\cdot\nabla z_m(t,x))z_m(t,x) \\
&\qquad + (D(t,x)\cdot\nabla z_m(t,x))\partial_t z_m(t,x) - (B(t,x)\cdot\nabla z_m(t,x))\partial_t z_m(t,x) \\
&\qquad - c(t,x)z_m(t,x)\partial_t z_m(t,x)\Big)\,dx. \qquad \text{(A.4)}
\end{aligned}
$$

Noting that

$$
\int_\Omega \partial_{tt} z_m(t,x)\partial_t z_m(x)\,dx = \frac{d}{dt}\Big(\frac{1}{2}\|\partial_t z_m(t,\cdot)\|^2_{L^2(\Omega)}\Big) \qquad \text{(A.5)}
$$

and combining (A.3)–(A.5), we get

$$
\begin{aligned}
&\frac{d}{dt}\Big(\|\partial_t z_m(t,\cdot)\|^2_{L^2(\Omega)} + \int_\Omega \big(A(t,x)\nabla z_m(t,x)\cdot\nabla z_m(t,x) \\
&\qquad + 2(D(t,x)\cdot\nabla z_m(t,x))z_m(t,x)\big)\,dx\Big) \\
&\quad \le C(\|\partial_t z_m(t,\cdot)\|^2_{L^2(\Omega)} + \|z_m(t,\cdot)\|^2_{H_0^1(\Omega)} + \|f(t,\cdot)\|^2_{L^2(\Omega)}),
\end{aligned}
$$

where C depends on $\|\partial_t A\|_{L^\infty(\Omega_T)}$, $\|B\|_{L^\infty(\Omega_T)}$, $\|D\|_{L^\infty(\Omega_T)}$, $\|\partial_t D\|_{L^\infty(\Omega_T)}$ and $\|c\|_{L^\infty(\Omega_T)}$.

Then, by the ellipticity of A and Poincaré's inequality, we deduce

$$
\begin{aligned}
&\|\partial_t z_m(t,\cdot)\|^2_{L^2(\Omega)} + \|z_m(t,\cdot)\|^2_{H_0^1(\Omega)} \\
&\quad \le C\Big(\int_0^t (\|\partial_t z_m(s,\cdot)\|^2_{L^2(\Omega)} + \|z_m(s,\cdot)\|^2_{H_0^1(\Omega)} + \|f(s,\cdot)\|^2_{L^2(\Omega)})\,ds \\
&\qquad + \|z_m(t,\cdot)\|^2_{L^2(\Omega)} + \|z_0\|^2_{H_0^1(\Omega)} + \|z_1\|^2_{L^2(\Omega)}\Big).
\end{aligned}
$$

Since

$$
\begin{aligned}
&\|z_m(t,\cdot)\|^2_{L^2(\Omega)} \\
&\quad = \int_\Omega \Big(\int_0^t \frac{d}{ds} z_m^2(s,x)\,ds + z_m^2(0,x)\Big)\,dx \\
&\quad = 2\int_0^t \int_\Omega z_m(s,x)\partial_t z_m(s,x)\,dx\,ds + \|z_0\|^2_{L^2(\Omega)} \\
&\quad \le \Big(\int_0^t (\|\partial_t z_m(s,\cdot)\|^2_{L^2(\Omega)} + \|z_m(s,\cdot)\|^2_{L^2(\Omega)})\,ds + \|z_0\|^2_{L^2(\Omega)}\Big),
\end{aligned}
$$

we get

$$
\|\partial_t z_m(t,\cdot)\|^2_{L^2(\Omega)} + \|z_m(t,\cdot)\|^2_{H_0^1(\Omega)}
$$

$$\leq C\Big(\int_0^t (\|\partial_t z_m(s,\cdot)\|^2_{L^2(\Omega)} + \|z_m(s,\cdot)\|^2_{H_0^1(\Omega)})\,ds + \|f\|^2_{L^2(\Omega_T)}$$
$$+ \|z_0\|^2_{H_0^1(\Omega)} + \|z_1\|^2_{L^2(\Omega)}\Big).$$

Thus, by Gronwall's inequality,

$$\|\partial_t z_m(t,\cdot)\|^2_{L^2(\Omega)} + \|z_m(t,\cdot)\|^2_{H_0^1(\Omega)} \leq C\Big(\|f\|^2_{L^2(\Omega_T)} + \|z_0\|^2_{H_0^1(\Omega)} + \|z_1\|^2_{L^2(\Omega)}\Big). \tag{A.6}$$

Now fix $v \in H_0^1(\Omega)$, $\|v\|_{H_0^1(\Omega)} \leq 1$, and write $v = v_1 + v_2$, where $v_1 \in \operatorname{span}\{w_k\}_{k=1}^m$ and

$$\int_\Omega v_2(x) w_k(x) = 0, \quad k = 1, \dots, m.$$

Since $\{w_k\}_{k=1}^\infty$ is an orthogonal basis of $H_0^1(\Omega)$, we have $\|v_1\|_{H_0^1(\Omega)} \leq 1$. Moreover,

$$\langle \partial_{tt} z_m(t,\cdot), v(\cdot)\rangle = \int_\Omega \partial_{tt} z_m(t,x) v(x)\,dx$$
$$= \int_\Omega \partial_{tt} z_m(t,x) v_1(x)\,dx = \int_\Omega f(t,x) v_1(x)\,dx - \mathcal{B}[z_m, v_1; t],$$

where $\langle\,,\,\rangle$ denote the pair between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. Thus

$$|\langle \partial_{tt} z_m(t,\cdot), v(\cdot)\rangle| \leq C(\|f(t,\cdot)\|_{L^2(\Omega)} + \|z_m(t,\cdot)\|_{H_0^1(\Omega)}).$$

Consequently,

$$\int_0^T \|\partial_{tt} z_m(t,\cdot)\|^2_{H^{-1}(\Omega)}\,dt$$
$$\leq C \int_0^T \Big(\|f(t,\cdot)\|^2_{L^2(\Omega)} + \|z_m(t,\cdot)\|^2_{H_0^1(\Omega)}\Big)\,dt$$
$$\leq C\Big(\|f\|^2_{L^2(\Omega_T)} + \|z_0\|^2_{H_0^1(\Omega)} + \|z_1\|^2_{L^2(\Omega)}\Big). \tag{A.7}$$

Combining (A.6) with (A.7), we have

$$\|\partial_t z_m\|_{L^\infty(0,T;L^2(\Omega))} + \|z_m\|_{L^\infty(0,T;H_0^1(\Omega))} + \|\partial_{tt} z_m\|_{L^2(0,T;H^{-1}(\Omega))}$$
$$\leq C\Big(\|f\|_{L^2(\Omega_T)} + \|z_0\|_{H_0^1(\Omega)} + \|z_1\|_{L^2(\Omega)}\Big), \tag{A.8}$$

where the constant C depend on ł, Λ, Ω_T, $\|\partial_t A\|_{L^\infty(\Omega_T)}$, $\|B\|_{L^\infty(\Omega_T)}$, $\|D\|_{L^\infty(\Omega_T)}$, $\|\partial_t D\|_{L^\infty(\Omega_T)}$, $\|c\|_{L^\infty(\Omega_T)}$.

Step 3: According to the energy estimates (A.8), there exists a subsequence $\{z_{m_k}\}$ and $z \in L^2(0,T;H_0^1(\Omega))$, with $\partial_t z \in L^2(0,T;L^2(\Omega))$, $\partial_{tt} z \in L^2(0,T;H^{-1}(\Omega))$ such that

$$z_{m_k} \longrightarrow z, \quad \text{weakly in } L^2(0,T;H_0^1(\Omega)),$$

$$\partial_t z_{m_k} \longrightarrow \partial_t z, \quad \text{weakly in } L^2(0,T;L^2(\Omega)),$$
$$\partial_{tt} z_{m_k} \longrightarrow \partial_{tt} z, \quad \text{weakly in } L^2(0,T;H^{-1}(\Omega)).$$

Then we can prove that z is an weak solution of (2.2) in the same way of Theorem 7.2.3 in [9].

Uniqueness. We omit the proof since it is totally the same as that of Theorem 7.2.4 in [9].

By the energy estimate (A.8), we can deduce (2.4) by passing limits to m. □

Proof of Proposition 2.2

Existence. Let $\phi_0 = 0$. Define ϕ_{k+1} be the weak solution of

$$\begin{cases} \partial_{tt}\phi_{k+1}(t,x) - \nabla\cdot(A(t,x)\nabla\phi_{k+1}(t,x)) + B(t,x)\cdot\nabla\phi_{k+1}(t,x) \\ \quad -\nabla\cdot(D(t,x)\phi_{k+1}(t,x)) \\ \quad -\int_0^1 f_z(t,x,\beta\phi_k(t,x))\,d\beta\,\phi_{k+1}(t,x) = f(t,x,0), \quad \text{in} \\ O_T, \\ \phi_{k+1}(t,x) = 0, \quad \text{on } [0,T]\times\partial\Omega, \\ \phi_{k+1}(0,x) = z_0(x), \quad \text{in } \Omega, \\ \partial_t\phi_{k+1}(0,x) = z_1(x), \quad \text{in } \Omega. \end{cases} \tag{A.9}$$

We have

$$\|\phi_{k+1}\|_{L^\infty(0,T;H^1(\Omega))} + \|\partial_t\phi_{k+1}\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_{tt}\phi_{k+1}\|_{L^2(0,T;H^{-1}(\Omega))} \le M. \tag{A.10}$$

for some positive constant M depending on λ, Λ, K, Ω_T, $\|\partial_t A\|_{L^\infty(\Omega_T)}$, $\|B\|_{L^\infty(\Omega_T)}$, $\|D\|_{L^\infty(\Omega_T)}$, $\|\partial_t D\|_{L^\infty(\Omega_T)}$, $\|c\|_{L^\infty(\Omega_T)}$, $\|z_0\|_{H_0^1(\Omega)}$ and $\|z_1\|_{L^2(\Omega)}$.

Then along a subsequence,

$$\phi_k \longrightarrow \phi, \quad \text{weakly in } L^2(0,T;H_0^1(\Omega)), \text{ strongly in } L^2(\Omega_T),$$
$$\partial_t\phi_k \longrightarrow \partial_t\phi, \quad \text{weakly in } L^2(0,T;L^2(\Omega)),$$
$$\partial_{tt}\phi_k \longrightarrow \partial_{tt}\phi, \quad \text{weakly in } L^2(0,T;H^{-1}(\Omega)).$$

Then it follows easily from (A.9) that

$$\begin{cases} \partial_{tt}z(t,x) - \nabla\cdot(A(t,x)\nabla z(t,x)) + B(t,x)\cdot\nabla z(t,x) - \nabla\cdot(D(t,x)z(t,x)) \\ \quad -\int_0^1 f_z(t,x,\beta z(t,x))\,d\beta\, z(t,x) = f(t,x,0), \quad \text{in } \Omega_T, \\ z(t,x) = 0, \quad \text{on } [0,T]\times\partial\Omega, \\ z(0,x) = z_0(x), \quad \text{in } \Omega, \\ \partial_t z(0,x) = z_1(x), \quad \text{in } \Omega. \end{cases} \tag{A.11}$$

That is, z is a weak solution of (2.6).

Uniqueness. Let $\hat{z}$ be an another weak solution of (2.6). Then $\widehat{Z} \equiv \hat{z} - z$ satisfies

$$\begin{cases} \partial_{tt}\widehat{Z}(t,x) - \nabla\cdot(A(t,x)\nabla\widehat{Z}(t,x)) + B(t,x)\cdot\nabla\widehat{Z}(t,x) - \nabla\cdot(D(t,x)\widehat{Z}(t,x)) \\ \quad = \int_0^1 f_z(t,x,z(t,x)+\beta(\hat{z}(t,x)-z(t,x)))\,d\beta\,\widehat{Z}(t,x), \quad \text{in } \Omega_T, \\ \widehat{Z}(t,x) = 0, \quad \text{on } [0,T]\times\partial\Omega, \\ \widehat{Z}(0,x) = 0, \quad \text{in } \Omega, \\ \partial_t\widehat{Z}(0,x) = 0, \quad \text{in } \Omega. \end{cases} \tag{A.12}$$

Then Proposition 2.2 implies $\widehat{Z} = 0$ and we get the uniqueness. □

References

1. Allaire, G.: Homogenization and two-scale convergence. SIAM J. Math. Anal. **23**(6), 1482–1518 (1992)
2. Allaire, G.: Shape Optimization by the Homogenization Method. Springer, New York (2002)
3. Bensoussan, A., Lions, J.L., Papanicolaou, G.: Asymptotic Analysis for Periodic Structures. North-Holland, Amsterdam (1978)
4. Casado-Diaz, J., Couce-Calvo, J., Martin-Gomez, J.D.: Optimality conditions for nonconvex multi-state control problems in the coefficients. SIAM J. Control Optim. **43**(1), 216–239 (2004)
5. Casas, E.: Optimal control in coefficients of elliptic equations with state constraints. Appl. Math. Optim. **26**(1), 21–37 (1992)
6. Cherkaev, A.: Variational Methods for Structural Optimization. Springer, New York (2000)
7. Colombini, F., Spagnolo, S.: On the convergence of solutions of hyperbolic equations. Commun. Partial Differ. Equ. **3**, 77–103 (1978)
8. Debińska-Nagórska, A., Just, A., Stempień, Z.: Analysis and semidiscrete Galerkin approximation of a class of nonlinear hyperbolic optimal control problems. Optimization **48**, 177–190 (2000)
9. Evans, L.C.: Partial Differential Equations. AMS, Providence (1998)
10. Kuliev, G.F.: The problem of optimal control of coefficients for equations of hyperbolic type. Izv. Vysš. Učebn. Zaved., Mat. **3**, 39–44 (1985). See also page 87 (Russian)
11. Li, X., Yong, J.: Optimal Control Theory for Infinite Dimensional Systems. Birkhäuser, Boston (1995)
12. Lou, H.: Optimality conditions for semilinear parabolic equations with controls in leading term. ESAIM: Control Optim. Calc. Var. **17**(4), 975–994 (2011)
13. Lou, H., Yong, J.: Optimality conditions for semilinear elliptic equations with leading term containing controls. SIAM J. Control Optim. **48**(4), 2366–2387 (2009)
14. Lukkassen, D., Nguetseng, G., Wall, P.: Two-scale convergence. Int. J. Pure Appl. Math. **2**(1), 35–86 (2002)
15. Lurie, K.A.: Applied Optimal Control Theory of Distributed Systems. Plenum, New York (1993)
16. Lurie, K.A.: Control in the coefficients of linear hyperbolic equations via spacio-temporal components. In: Homogenization, Ser. Adv. Math. Appl. Sci., vol. 50, pp. 285–315. World Sci., River Edge (1999)
17. Nguetseng, G.: A general convergence result for a functional related to the theory of homogenization. SIAM J. Math. Anal. **20**(3), 608–623 (1989)
18. Petukhov, L.V.: Optimal control of processes described by equations of hyperbolic type. J. Appl. Math. Mech. **41**(3), 385–396 (1978); translated from Prikl. Mat. Meh. **41**(3), 387–398 (1977) (Russian)
19. Tagiyev, R.K.: Optimal control by the coefficients of a parabolic equation. Trans. NAS Azerb., Isc. Math. Mech. **24**(4), 247–256 (2004)
20. Tagiyev, R.K.: On the optimal control problem by coefficients of a hyperbolic equation. Trans. Acad. Sci. Azerb. Ser. Phys.-Tech. Math. Sci. **21**(4), 230–235 (2001)
21. Murat, F., Tartar, L.: Calculus of variations and homogenization. In: Cherkaev, L., Kohn, R.V. (eds.) Collection d'tudes d'electricit de France, rewritten in Topics in the Mathematical Modelling of Composite Materials, Progress in Nonlinear Differential Equations and their Applications, vol. 31, pp. 139–174. Birkaüser, Boston (1998)