数理解析研究所講究録 1216 巻 2001 年 189-200

Optimal Control Problems for Distributed Hopfield-type Neural Networks

神戸大学自然科学研究科　王 全芳 (Quan-Fang Wang)
神戸大学工学部　中桐 信一 (Shin-ichi Nakagiri)

1 Introduction

The stability properties of lumped Hopfield-type neural networks have been studied extensively and the results as well as the application to optimization problems are reported in many references (see the references in Vanualailai, Nakagiri and Soma [12]). In this paper we study a model which involves spatial distributions of neural networks described by the dynamics of n-numbers of neurons. The distributive model may be consider as an analogous model of the Hodgkin-Huxley equation and the Fitz-Hugh-Nagumo equation which describe the nerve impulse transmissions (cf. Hodgkin and Huxley [6], Fitz-Hugh [5], Nagumo [9]). The purpose of this paper is to study the optimal control problems for the systems governed by distributed models of Hopfield-type neural networks.

Let Ω be an open bounded domain of $\mathbf{R}^m$ and $\partial\Omega=\Gamma$ be the boundary of Ω. Let $T>0$ and let $Q=\Omega\times(0,T)$ and $\Sigma=\Gamma\times(0,T)$. Let y_i denotes the activation potential of the i-th neurons, which is a functions of time t and the place $x\in\Omega$, $i=1,2,...n$. The distributed Hopfield-type model of coupled n-numbers of neurons is described by

$$\begin{cases} \dfrac{\partial y_i}{\partial t}-d_i\Delta y_i=-a_iy_i+\displaystyle\sum_{j=1}^{n}c_{ij}F_j(y_j)+g_i \quad in\ Q, \\ \dfrac{\partial y_i}{\partial\eta}=k_i \quad on\ \ \Sigma, \\ y_i(0,x)=y_0^i \quad in\ \ \Omega, \qquad i=1,2,...,n. \end{cases} \tag{1.1}$$

Here in (1.1) the constants $a_i>0$, c_{ij} are the same as explained in [12], $d_i>0$ are diffusion constants, g_i are forcing input, k are the Neumann inputs, y_0^i are initial values and $F_j:\mathbf{R}=(-\infty,\infty)\to(-1,1)$ are nonlinear activation functions.

We consider the quadratic optimal control problem for (1.1). The control system under consideration is given by (1.1) in which g_i, k_i and y_0^i are replaced by the control variables $B_i^0u_i^0$, $B_i^1u_i^1$ and E_iw_i, respectively. Here B_i^0,B_i^1 and E_i are distributed, boundary and initial controllers and u_i^0,u_i^1 and w_i are respective control variables. Let $\mathcal{U}_i^0$, $\mathcal{U}_i^1$ and $\mathcal{W}_i$ be the Hilbert spaces of control variables u_i^0, u_i^1 and w_i. We denote

$$\mathcal{U}^0=\prod_{i=1}^{n}\mathcal{U}_i^0,\ \ \mathcal{U}^1=\prod_{i=1}^{n}\mathcal{U}_i^1,\ \ \mathcal{W}=\prod_{i=1}^{n}\mathcal{W}_i$$

and $\mathbf{u}_0=(u_1^0,u_2^0,...,u_n^0)$, $\mathbf{u}_1=(u_1^1,u_2^1,...,u_n^1)$, $\mathbf{w}=(w_1,w_2,...,w_n)$. We set the product Hilbert space $\mathcal{U}=\mathcal{U}^0\times\mathcal{U}^1\times\mathcal{W}$ of control variables $\mathbf{u}=(\mathbf{u}_0,\mathbf{u}_1,\mathbf{w})\in\mathcal{U}$.

Let $\mathbf{y} = \mathbf{y}(\mathbf{u}) = (y_1(\mathbf{u}), y_2(\mathbf{u}), ..., y_n(\mathbf{u}))^T$ be the solution state of control system for a given $\mathbf{u} \in \mathcal{U}$. The quadratic cost function attached to the system is given by

$$J(\mathbf{u}) = \sum_{i=1}^{n} \int_Q |y_i(\mathbf{u}) - z_{id}^0|^2 dxdt + \sum_{i=1}^{n} \int_\Omega |y_i(\mathbf{u}, T) - z_{id}^1|^2 dx + (\mathbf{N}\mathbf{u}, \mathbf{u})_{\mathcal{U}},$$
$$\forall \mathbf{u} = (\mathbf{u}_0, \mathbf{u}_1, \mathbf{w}) \in \mathcal{U}, \qquad (1.2)$$

where $z_{id}^0 \in L^2(Q)$ and $z_{id}^1 \in L^2(\Omega), i = 1, 2, ..., n$ are desired values, and $\mathbf{N} \in \mathcal{L}(\mathcal{U})$ is symmetric and positive. Let $\mathcal{U}_{ad}$ be an admissible subset of $\mathcal{U}$. The optimal control problem is to find and characterize an element $\mathbf{u}^* = (\mathbf{u}_0^*, \mathbf{u}_1^*, \mathbf{w}^*) \in \mathcal{U}_{ad}$, called the optimal control, such that

$$\inf_{\mathbf{u} \in \mathcal{U}_{ad}} J(\mathbf{u}) = J(\mathbf{u}^*). \qquad (1.3)$$

In this paper we shall solve this quadratic cost optimal control problem for the distributed model of Hopfield-type neural networks (1.1). For the purpose we state the basic results on the existence and uniqueness of weak solutions for the nonlinear system (1.1) in the framework of Dautray and Lions [4]. After that we prove the existence of an optimal control for (1.2). The main contribution is to construct the adjoint state systems and to establish the necessary conditions of optimality for the quadratic cost (1.2). For the related works on optimal control theory of nonlinear parabolic equations, we refer to [1], [2], [3], [13], [14].

2 Existence and Uniqueness of weak solutions

In this section we shall give the results on existence, uniqueness and regularity of solution for the uncontrolled (free) system (1.1) based on the variational formulation of systems due to Dautray and Lions [4].

For the nonlinear function $F_j(s)$, we suppose the uniform Lipschitz continuity:

$$\exists K > 0 : \quad |F_j(s) - F_j(r)| \leq K|s - r|, \quad j = 1, 2, ..., n. \qquad (2.1)$$

For the evolution equation setting of (1.1), we introduce two Hilbert spaces $H = L^2(\Omega)$ and $V = H^1(\Omega)$ according to the Neumann boundary condition in (1.1). We endow those space with the usual inner products and norms

$$(\psi, \phi) = \int_\Omega \psi(x)\phi(x)dx, \quad |\psi| = (\psi, \psi)^{\frac{1}{2}} \quad \text{for all } \psi, \phi \in H,$$
$$((\psi, \phi)) = \int_\Omega \psi(x)\phi(x)dx + \sum_{i=1}^{n} \int_\Omega \frac{\partial}{\partial x_i}\psi(x)\frac{\partial}{\partial x_i}\phi(x)dx \quad \text{for all } \psi, \phi \in H^1(\Omega),$$

respectively. Let us introduce the product Hilbert spaces $\mathcal{V} = (H^1(\Omega))^n$, $\mathcal{H} = (L^2(\Omega))^n$ with the inner products defined by

$$(\boldsymbol{\phi}, \boldsymbol{\psi})_{\mathcal{H}} = \sum_{i=1}^{n} (\phi_i, \psi_i), \quad \boldsymbol{\phi} = (\phi_1, \phi_2, ..., \phi_n)^T, \ \boldsymbol{\psi} = (\psi_1, \psi_2, ..., \psi_n)^T \in \mathcal{H}$$

$$(\phi,\psi)_{\mathcal{V}}=\sum_{i=1}^{n}(\!(\phi_i,\psi_i)\!)+\sum_{i=1}^{n}(\phi_i,\psi_i),\quad \phi=(\phi_1,\phi_2,...,\phi_n)^T,\ \psi=(\psi_1,\psi_2,...,\psi_n)^T\in\mathcal{V},$$

respectively. Then the dual space of $\mathcal{V}$ is given by $\mathcal{V}'=(V')^n$ and the dual pairing between $\mathcal{V}'$ and $\mathcal{V}$ is given by

$$\langle\phi,\psi\rangle_{\mathcal{V},\mathcal{V}'}=\sum_{i=1}^{n}\langle\phi_i,\psi_i\rangle,\quad \forall\phi=(\phi_1,\phi_2,...,\phi_n)^T\in\mathcal{V},\ \psi=(\psi_1,\psi_2,...,\psi_n)^T\in\mathcal{V}',$$

where $\langle\phi_i,\psi_i\rangle$ denotes the dual pairing between V and V' of $\phi_i\in V$ and $\psi_i\in V'$. The norms of $\mathcal{V}$ and $\mathcal{H}$ are denoted by $\|\psi\|_{\mathcal{V}}$ and $|\psi|_{\mathcal{H}}$, respectively.

For the sake of simplicity of notations, we introduce the following vector and matrix representations:

$$\mathbf{y}=\begin{bmatrix} y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix},\ \Delta\mathbf{y}=\begin{bmatrix}\Delta y_1\\ \Delta y_2\\ \vdots\\ \Delta y_n\end{bmatrix},\ \mathbf{C}=\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n}\\ c_{21} & c_{22} & \cdots & c_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ c_{n1} & c_{n2} & \cdots & c_{nn}\end{bmatrix},\ \mathbf{F}(\mathbf{y})=\begin{bmatrix} F_1(y_1)\\ F_2(y_2)\\ \vdots\\ F_n(y_n)\end{bmatrix},\qquad (2.2)$$

$\mathbf{D}=diag\ \{d_1,d_2,...,d_n\},\mathbf{A}=diag\ \{a_1,a_2,...,a_n\}$, $\mathbf{g}=(g_1,g_2,...,g_n)^T$, $\mathbf{k}=(k_1,k_2,...,k_n)^T$ and $\mathbf{y}_0=(y_0^1,y_0^2,...,y_0^n)^T$. Then the system (1.1) is rewritten simply by

$$\begin{cases}\dfrac{\partial\mathbf{y}}{\partial t}-\mathbf{D}\Delta\mathbf{y}=\mathbf{A}\mathbf{y}+\mathbf{C}\mathbf{F}(\mathbf{y})+\mathbf{g}\quad in\ Q,\\ \dfrac{\partial\mathbf{y}}{\partial\eta}=\mathbf{k}\quad on\ \Sigma,\\ \mathbf{y}(0,x)=\mathbf{y}_0\quad in\ \Omega.\end{cases}\qquad (2.3)$$

Now we give the definition of a weak solution of (2.3). First we introduce the Hilbert space $W(0,T;\mathcal{V},\mathcal{V}')$, which will be a solution space, by

$$W(0,T;\mathcal{V},\mathcal{V}')=\{\mathbf{g}|\ \mathbf{g}\in L^2(0,T;\mathcal{V}),\ \mathbf{g}'\in L^2(0,T;\mathcal{V}')\}.$$

The inner product and the induced norm in $W(0,T;\mathcal{V},\mathcal{V}')$ are defined respectively by

$$(\mathbf{g}_1,\mathbf{g}_2)_{W(0,T;\mathcal{V},\mathcal{V}')}=\int_0^T\{(\mathbf{g}_1(t),\mathbf{g}_2(t))_{\mathcal{V}}+(\mathbf{g}_1'(t),\mathbf{g}_2'(t))_{\mathcal{V}'}\}dt,$$

$$\|\mathbf{g}\|_{W(0,T;\mathcal{V},\mathcal{V}')}=\left(\|\mathbf{g}\|^2_{L^2(0,T;\mathcal{V})}+\|\mathbf{g}'\|^2_{L^2(0,T;\mathcal{V}')}\right)^{\frac{1}{2}}.$$

The space $W(0,T;\mathcal{V},\mathcal{V}')$ can be identified with $W(0,T;V,V')^n$, and for simplify, we denote it by $=\mathbf{W}(0,T)$. Also we define the Hilbert spaces $\mathcal{H}^{\frac{1}{2}}(\Gamma)$ and its dual $\mathcal{H}^{-\frac{1}{2}}(\Gamma)$ by $(H^{\frac{1}{2}}(\Gamma))^n$ and $(H^{-\frac{1}{2}}(\Gamma))^n$, respectively. The dual pairing between $\mathcal{H}^{\frac{1}{2}}(\Gamma)$ and $\mathcal{H}^{-\frac{1}{2}}(\Gamma)$ is defined by

$$\langle\phi,\psi\rangle_{\mathcal{H}^{\frac{1}{2}}(\Gamma),\mathcal{H}^{-\frac{1}{2}}(\Gamma)}=\sum_{i=1}^{n}\langle\phi_i,\psi_i\rangle_\Gamma,\forall\phi=(\phi_1,\phi_2,...,\phi_n)^T\in\mathcal{H}^{\frac{1}{2}}(\Gamma),\ \psi=(\psi_1,\psi_2,...,\psi_n)^T\in\mathcal{H}^{-\frac{1}{2}}(\Gamma$$

where $\langle \phi_i, \psi_i \rangle_\Gamma$ denotes the dual pairing between $H^{\frac{1}{2}}(\Gamma)$ and $H^{-\frac{1}{2}}(\Gamma)$ of ϕ_i and ψ_i. Since $F_j : R \to (-1,1), j = 1,2,...,n$, we see from (2.1) that $\mathbf{F} : \mathcal{H} \to \mathcal{H}$ and

$$|\mathbf{F}(\psi)|^2_{\mathcal{H}} = \int_\Omega |\mathbf{F}(\psi)|^2 dx \leq n|\Omega|, \quad \forall \psi \in \mathcal{H}, \tag{2.4}$$

and by (2.1), we have

$$|\mathbf{F}(\phi) - \mathbf{F}(\psi)|_{\mathcal{H}} \leq K|\phi - \psi|_{\mathcal{H}}, \ \forall \phi, \psi \in \mathcal{H}. \tag{2.5}$$

Definition 1 *A function* $\mathbf{y}$ *is said to be a weak solution of* (2.3) *if* $\mathbf{y} \in \mathbf{W}(0,T)$ *and* $\mathbf{y}$ *satisfies*

$$\begin{cases} \langle \mathbf{y}', \mathbf{v} \rangle_{\mathcal{V}',\mathcal{V}} + (\mathbf{D}\nabla \mathbf{y}, \nabla \mathbf{v})_{\mathcal{H}} = (\mathbf{A}\mathbf{y}, \mathbf{v})_{\mathcal{H}} + (\mathbf{C}\mathbf{F}(\mathbf{y}), \mathbf{v})_{\mathcal{H}} + \langle \mathbf{g}, \mathbf{v} \rangle_{\mathcal{V}',\mathcal{V}} + \langle \mathbf{k}, \mathbf{v}|_\Gamma \rangle_{\mathcal{H}^{-\frac{1}{2}}(\Gamma), \mathcal{H}^{\frac{1}{2}}(\Gamma)} \\ \qquad\qquad \text{for all } \mathbf{v} \in \mathcal{V} \text{ in the sense of } \mathcal{D}'(0,T), \\ \mathbf{y}(0) = \mathbf{y}_0 \in \mathcal{H}. \end{cases} \tag{2.6}$$

Here in (2.6), $\mathcal{D}'(0,T)$ denotes the space of distributions on $(0,T)$. Also we note that the trace $\mathbf{v}|_\Gamma$ of $\mathbf{v} \in \mathcal{V}$ on Γ belongs to $\mathcal{H}^{\frac{1}{2}}(\Gamma)$.

For the existence and uniqueness of weak solution for (1.1), we can give the following theorem.

Theorem 1 *Assume that* $\mathbf{y}_0 \in \mathcal{H}$, $\mathbf{g} \in L^2(0,T;\mathcal{V}')$, $\mathbf{k} \in L^2(0,T;\mathcal{H}^{-\frac{1}{2}}(\Gamma))$. *Then the problem* (2.3) *has a unique weak solution* $\mathbf{y}$ *in* $\mathbf{W}(0,T)$, *which belongs to* $C([0,T];\mathcal{H})$. *Further, we have the estimate*

$$\|\mathbf{y}\|^2_{L^\infty(0,T;\mathcal{H})}, \ \|\mathbf{y}\|^2_{\mathbf{W}(0,T)} \leq C(1 + |\mathbf{y}_0|^2_{\mathcal{H}} + \|\mathbf{g}\|^2_{L^2(0,T;\mathcal{V}')} + \|\mathbf{k}\|^2_{L^2(0,T;\mathcal{H}^{-\frac{1}{2}}(\Gamma))}), \tag{2.7}$$

where $C > 0$ *depends only on* a_i, c_{ij} *and* d_i.

3 Optimal control problems

In this section we study the quadratic optimal control problem for (1.1) by means of distributive, boundary and initial controls. Let $B_i^0 \in \mathcal{L}(\mathcal{U}_i^0, L^2(0,T;V'))$, $B_i^1 \in \mathcal{L}(\mathcal{U}_i^1, L^2(0,T;H^{\frac{1}{2}}(\Gamma))$ and $E_i \in \mathcal{L}(W_i, H)$ for $i = 1,2,...,n$. We denote

$$\mathbf{B}^0 = diag\ \{B_1^0, B_2^0, ..., B_n^0\}, \ \mathbf{B}^1 = diag\ \{B_1^1, B_2^1, ..., B_n^1\}, \ \mathbf{E} = diag\ \{E_1, E_2, ..., E_n\}, \tag{3.1}$$

respectively. Then $\mathbf{B}^0, \mathbf{B}^1$ and $\mathbf{E}$ are operators satisfying $\mathbf{B}^0 \in \mathcal{L}(\mathcal{U}^0, L^2(0,T;\mathcal{V}'))$, $\mathbf{B}^1 \in \mathcal{L}(\mathcal{U}^1, L^2(0,T;\mathcal{H}^{\frac{1}{2}}(\Gamma)))$ and $\mathbf{E} \in \mathcal{L}(\mathcal{W}, \mathcal{H})$, which are called the controllers. We consider the following control system described by

$$\begin{cases} \dfrac{\partial \mathbf{y}}{\partial t} - \mathbf{D}\Delta \mathbf{y} = \mathbf{A}\mathbf{y} + \mathbf{C}\mathbf{F}(\mathbf{y}) + \mathbf{B}^0\mathbf{u}_0 \quad in\ Q, \\ \dfrac{\partial \mathbf{y}}{\partial \eta} = \mathbf{B}^1\mathbf{u}_1 \quad on\ \Sigma, \\ \mathbf{y}(0,x) = \mathbf{E}\mathbf{w} \quad in\ \Omega. \end{cases} \tag{3.2}$$

For any $\mathbf{u} = (\mathbf{u}_0, \mathbf{u}_1, \mathbf{w}) \in \mathcal{U}$, by virtue of Theorem 1, we have a unique weak solution $\mathbf{y} = \mathbf{y}(\mathbf{u})$ of (3.2) in $\mathbf{W}(0,T)$. Hence we can define the solution map $\mathbf{u} \to \mathbf{y}(\mathbf{u})$ of $\mathcal{U}$ into $\mathbf{W}(0,T)$. We shall call $\mathbf{y}(\mathbf{u})$ the state of the control system (3.2). The quadratic cost function associated with the control system (3.2) is given by (1.2), and is written in compact form

$$J(\mathbf{u}) = \|\mathbf{y}(\mathbf{u}) - \mathbf{z}_d^0\|^2_{(L^2(Q))^n} + \|\mathbf{y}(\mathbf{u},T) - \mathbf{z}_d^1\|^2_{(L^2(\Omega))^n} + (\mathbf{N}\mathbf{u},\mathbf{u})_{\mathcal{U}},\ \forall \mathbf{u} = (\mathbf{u}_0, \mathbf{u}_1, \mathbf{w}) \in \mathcal{U}, \quad (3.3)$$

where $\mathbf{z}_d^0 = (z_{1d}^0, z_{2d}^0, ..., z_{nd}^0)^T \in (L^2(Q))^n$, $\mathbf{z}_d^1 = (z_{1d}^1, z_{2d}^1, ..., z_{nd}^1)^T \in (L^2(\Omega))^n$ are desired values and $\mathbf{N} = (\mathbf{N}^0, \mathbf{N}^1, \mathbf{N}^2) \in \mathcal{L}(\mathcal{U}) = \mathcal{L}(\mathcal{U}^0) \times \mathcal{L}(\mathcal{U}^1) \times \mathcal{L}(\mathcal{W})$. We suppose that each $\mathbf{N}^j$ $(j = 0,1,2)$ is symmetirc and positive. Since $\mathbf{y} \in \mathbf{W}(0,T) \subset L^2(0,T;\mathcal{H}) = (L^2(Q))^n$ and $\mathbf{y}(\mathbf{u},T) \in \mathcal{H} = (L^2(\Omega))^n$ by Theorem 1, the cost (3.3) is meaningful for any $\mathbf{u} \in \mathcal{U}$.

Let $\mathcal{U}_{ad}$ be a closed convex subset of $\mathcal{U}$, which is called the admissible set. We shall solve the following two fundamental problems for the control system (3.2) attached the quadratic cost (3.3):

(i) Existence problem of an element $\mathbf{u}^* = (\mathbf{u}_0^*, \mathbf{u}_1^*, \mathbf{w}^*) \in \mathcal{U}_{ad}$ such that

$$\inf_{\mathbf{u}\in\mathcal{U}_{ad}} J(\mathbf{u}) = J(\mathbf{u}^*);$$

(ii) Characterization problem of such $\mathbf{u}^*$.

Such a $\mathbf{u}^*$ in (i) is called the optimal control for the system (3.2) with the cost (3.3).

3.1 Existence of optimal control

First we solve the existence problem (i) in the following theorem.

Theorem 2 *Assume that $\mathcal{U}_{ad}$ is a non-empty bounded closed convex set of $\mathcal{U}$. Then there exists at least one optimal control $\mathbf{u}^*$ for the control problem* (3.2) *with the cost* (3.3).

Proof: Set $J = \inf_{\mathbf{u}\in\mathcal{U}_{ad}} J(\mathbf{u})$. Since $\mathcal{U}_{ad}$ is non-empty, there is a sequence $\{\mathbf{u}_n\}$ in $\mathcal{U}_{ad}$ such that

$$\inf_{\mathbf{u}\in\mathcal{U}_{ad}} J(\mathbf{u}) = \lim_{n\to\infty} J(\mathbf{u}_n) = J.$$

Obviously, $\{J(\mathbf{u}_n)\}$ is bounded in $\mathbf{R}^+$. Since $\mathcal{U}_{ad}$ is bounded closed and convex, we can choose a subsequence $\{\mathbf{u}_m\} = \{\mathbf{u}_m^0, \mathbf{u}_m^1, \mathbf{w}_m\}$ of $\{\mathbf{u}_n\}$ and find a $\mathbf{u}^* = (\mathbf{u}_0^*, \mathbf{u}_1^*, \mathbf{w}^*) \in \mathcal{U}_{ad}$ such that

$$\mathbf{u}_m \to \mathbf{u}^* \textit{ weakly in } \mathcal{U} \textit{ as } m \to \infty. \quad (3.4)$$

By the estimate (2.7) in Theorem 1, we have for $\mathbf{y} = \mathbf{y}(\mathbf{u})$ that

$$\|\mathbf{y}\|^2_{L^\infty(0,T;\mathcal{H})},\ \|\mathbf{y}\|^2_{\mathbf{W}(0,T)} \leq C(1 + |\mathbf{E}\mathbf{w}|^2_{\mathcal{H}} + \|\mathbf{B}^0\mathbf{u}_0\|^2_{L^2(0,T;\mathcal{V}')} + \|\mathbf{B}^1\mathbf{u}_1\|^2_{L^2(0,T;\mathcal{H}^{-\frac{1}{2}}(\Gamma))}). \quad (3.5)$$

Since $\mathcal{U}_{ad}$ is bounded, we see from (3.5) that $\{\mathbf{y}(\mathbf{u_m})\}$ is bounded in $\mathbf{W}(0,T)$. Hence we can choose a subsequence $\{\mathbf{y}(\mathbf{u_{mk}})\}$ of $\{\mathbf{y}(\mathbf{u_m})\}$ and find a $\mathbf{z} \in \mathbf{W}(0,T)$ such that

$$\mathbf{y}(\mathbf{u_{mk}}) \to \mathbf{z} \quad \textit{weakly in } \mathbf{W}(0,T). \quad (3.6)$$

For simplicity let us denote $\mathbf{y_{mk}} = \mathbf{y}(\mathbf{u_{mk}})$. By the compactness of embedding $V \hookrightarrow H$, the embedding $\mathcal{V} \hookrightarrow \mathcal{H}$ is compact too. Thus, by the compactness embedding theorem due to the Aubin-Lions-Temam (cf. Temam [11, p.274]). we can suppose

$$\mathbf{y_{mk}} \to \mathbf{z} \ \textit{strongly in } L^2(0,T;\mathcal{H}). \quad (3.7)$$

By the uniformly Lipschitz continuity (2.5), it follows from (3.7) that

$$\mathbf{F}(\mathbf{y_{mk}}) \to \mathbf{F}(\mathbf{z}) \ \ \textit{strongly in} \ \ L^2(0,T;\mathcal{H}). \tag{3.8}$$

By (3.6), we see that $\mathbf{y}'_{\mathbf{mk}} \to \mathbf{z}'$ weakly in $L^2(0,T;\mathcal{V}')$ and $\nabla \mathbf{y}_{mk} \to \nabla \mathbf{z}$ weakly in $L^2(0,T;\mathcal{H})$. Hence by the definition of weak solutions, we have

$$\begin{aligned} &\int_0^T \{\langle \mathbf{y_{mk}}', \phi\rangle_{\mathcal{V}',\mathcal{V}} + (\mathbf{D}\nabla \mathbf{y_{mk}}, \nabla\phi)_{\mathcal{H}}\}dt \\ = &\int_0^T \{(\mathbf{A}\mathbf{y_{mk}}, \phi)_{\mathcal{H}} + (\mathbf{CF}(\mathbf{y_{mk}}), \phi)_{\mathcal{H}} + \langle \mathbf{B^0 u^0_{mk}}, \phi\rangle_{\mathcal{V}',\mathcal{V}} + \langle \mathbf{B^1 u^1_{mk}}, \phi|_\Gamma\rangle_{\mathcal{H}^{-\frac{1}{2}}(\Gamma),\mathcal{H}^{\frac{1}{2}}(\Gamma)}\}dt, \\ &\qquad \forall \phi \in L^2(0,T;\mathcal{V}). \end{aligned} \tag{3.9}$$

Therefore, by taking $k \to \infty$ in (3.9) and using (3.4) and (3.8), we can deduce

$$\begin{aligned} &\int_0^T \{\langle \mathbf{z}', \phi\rangle_{\mathcal{V},\mathcal{V}'} + (\mathbf{D}\nabla \mathbf{z}, \nabla\phi)_{\mathcal{H}}\}dt \\ = &\int_0^T \{(\mathbf{A}\mathbf{z}, \phi)_{\mathcal{H}} + (\mathbf{CF}(\mathbf{z}), \phi)_{\mathcal{H}} + \langle \mathbf{B^0 u_0^*}, \phi\rangle_{\mathcal{V},\mathcal{V}'} + \langle \mathbf{B^1 u_1^*}, \phi|_\Gamma\rangle_{\mathcal{H}^{-\frac{1}{2}}(\Gamma),\mathcal{H}^{\frac{1}{2}}(\Gamma)}\}dt, \\ &\qquad \forall \phi \in L^2(0,T;\mathcal{V}). \end{aligned} \tag{3.10}$$

This implies, by the standard manupulation as in Dautray and Lions [4], that $\mathbf{z}$ satisfies

$$\begin{aligned} &\langle \mathbf{z}', \mathbf{v}\rangle_{\mathcal{V}',\mathcal{V}} + (\mathbf{D}\nabla \mathbf{z}, \nabla \mathbf{v})_{\mathcal{H}} \\ = &(\mathbf{A}\mathbf{z}, \mathbf{v})_{\mathcal{H}} + (\mathbf{CF}(\mathbf{z}), \mathbf{v})_{\mathcal{H}} + \langle \mathbf{B^0 u_0^*}, \mathbf{v}\rangle_{\mathcal{V}',\mathcal{V}} + \langle \mathbf{B^1 u_1^*}, \mathbf{v}|_\Gamma\rangle_{\mathcal{H}^{-\frac{1}{2}}(\Gamma),\mathcal{H}^{\frac{1}{2}}(\Gamma)}, \quad \forall \mathbf{v} \in \mathcal{V}, \end{aligned}$$

in $\mathcal{D}(0,T)$. It is not difficult to verified that $(\mathbf{y_{mk}}(0), \mathbf{v})_{\mathcal{H}} = (\mathbf{E}\mathbf{w}_{mk}, \mathbf{v})_{\mathcal{H}} \to (\mathbf{z}(0), \mathbf{v})_{\mathcal{H}} = (\mathbf{E}\mathbf{w}^*, \mathbf{v})_{\mathcal{H}}$ for any $\mathbf{v} \in \mathcal{V}$. So that $\mathbf{z}(0) = \mathbf{E}\mathbf{w}^*$.

Hence from the uniqueness of weak solution for the system (3.2), we have $\mathbf{z} = \mathbf{y}(\mathbf{u}^*)$. Then from (3.6) and (3.7) we see

$$\mathbf{y}(\mathbf{u_{mk}}) \longrightarrow \mathbf{y}(\mathbf{u}^*) \ \ \textit{strongly in} \ L^2(0,T;\mathcal{H}) = (L^2(Q))^n, \tag{3.11}$$

$$\mathbf{y}(\mathbf{u_{mk}}, T) \longrightarrow \mathbf{y}(\mathbf{u}^*, T) \ \ \textit{weakly in} \ \mathcal{H} = (L^2(\Omega))^n. \tag{3.12}$$

By (3.11) and (3.12), we have

$$\lim_{m\to\infty} \|\mathbf{y}(\mathbf{u_{mk}}) - \mathbf{z}_d^0\|_{(L^2(Q))^n} = \|\mathbf{y}(\mathbf{u}^*) - \mathbf{z}_d^0\|_{(L^2(Q))^n} \tag{3.13}$$

Since the norm $\|\cdot\|_{(L^2(\Omega))^n}$ is lower semi-continuous in the weak topology of $\mathcal{H} = (L^2(\Omega))^n$, we have

$$\liminf_{m\to\infty} \|\mathbf{y}(\mathbf{u_{mk}}, T) - \mathbf{z}_d^1\|_{(L^2(\Omega))^n} \geq \|\mathbf{y}(\mathbf{u}^*, T) - \mathbf{z}_d^1\|_{(L^2(\Omega))^n}. \tag{3.14}$$

On the other hand, the weak convergence (3.4) and boundedness of $\mathbf{N}$ imply

$$\liminf_{m\to\infty} (\mathbf{N}\mathbf{u_{mk}}, \mathbf{u_{mk}})_{\mathcal{U}} \geq (\mathbf{N}\mathbf{u}^*, \mathbf{u}^*). \tag{3.15}$$

Therefore $J = \liminf_{m\to\infty} J(\mathbf{u_{mk}}) \geq J(\mathbf{u}^*)$, and hence $J(\mathbf{u}^*) = \inf_{\mathbf{u}\in\mathcal{U}_{ad}} J(\mathbf{u})$. This proves that $\mathbf{u}^*$ is an optimal control for the cost (3.3). This completes the proof of Theorem 2.

3.2 Necessary optimality conditions

In this subsection we consider the problem (ii). It is well known (cf. Lions [7]) that the optimality condition for $\mathbf{u}^*$ is given by the variational inequality

$$J'(\mathbf{u}^*)(\mathbf{u}-\mathbf{u}^*) \geq 0 \quad \text{for all} \quad \mathbf{u} \in \mathcal{U}_{ad}, \tag{3.16}$$

where $J'(\mathbf{u}^*)$ denotes the Gâteaux derivative of $J(\mathbf{u})$ in (3.3) at $\mathbf{u}^*$. The objective of this subsection is to write down the optimality condition (3.16) in terms of proper adjoint state systems. So we need to calculate the Gâteaux derivative of $\mathbf{y}(\mathbf{u})$ at $\mathbf{u} = \mathbf{u}^*$. For the purpose we have to prepare the following propositions:

Proposition 1 *Let $\mathbf{y}_1$ and $\mathbf{y}_2$ be two weak solutions of* (3.2) *with control variables $\mathbf{u}_1 = (\mathbf{u}_0^1, \mathbf{u}_1^1, \mathbf{w}^1)$ and $\mathbf{u}_2 = (\mathbf{u}_0^2, \mathbf{u}_1^2, \mathbf{w}^2)$, respectively. Then $\mathbf{z} = \mathbf{y}_1 - \mathbf{y}_2$ satisfies the following equality for all $t \in [0,T]$:*

$$\begin{aligned}
& \frac{1}{2}|\mathbf{z}(t)|_{\mathcal{H}}^2 + \int_0^t |\sqrt{\mathbf{D}}\nabla \mathbf{z}|_{\mathcal{H}}^2 dt \\
= \; & \frac{1}{2}|\mathbf{E}\mathbf{w}^1 - \mathbf{E}\mathbf{w}^2|_{\mathcal{H}}^2 + \int_0^t (\mathbf{A}\mathbf{z},\mathbf{z})_{\mathcal{H}} dt + \int_0^t (\mathbf{C}\mathbf{F}(\mathbf{y}_1) - \mathbf{C}\mathbf{F}(\mathbf{y}_2), \mathbf{z})_{\mathcal{H}} dt \\
& + \int_0^t \langle \mathbf{B}^0\mathbf{u}_0^1 - \mathbf{B}^0\mathbf{u}_0^2, \mathbf{z}\rangle_{\mathcal{V}',\mathcal{V}} dt + \int_0^t \langle \mathbf{B}^1\mathbf{u}_1^1 - \mathbf{B}^1\mathbf{u}_0^2, \mathbf{z}|_\Gamma\rangle_{\mathcal{H}^{-\frac{1}{2}}(\Gamma),\mathcal{H}^{\frac{1}{2}}(\Gamma)} dt.
\end{aligned} \tag{3.17}$$

This proposition follows from the energy equalities for $\mathbf{y}_1$ and $\mathbf{y}_2$.

Proposition 2 *Let $\mathbf{v}_0 \in \mathcal{U}$ be fixed. Then*

$$\mathbf{y}(\mathbf{u}+\lambda\mathbf{v}_0) \to \mathbf{y}(\mathbf{u}) \quad \textit{strongly in} \quad C([0,T];\mathcal{H}) \textit{ and } L^2(0,T;\mathcal{V}) \quad \textit{as } \lambda \to 0. \tag{3.18}$$

Proof: For $\lambda \in (0,1]$, let $\mathbf{y}(\mathbf{u}+\lambda\mathbf{v}_0)$ and $\mathbf{y}(\mathbf{u})$ be the weak solutions of (3.2) for $\mathbf{u}$ and $\mathbf{v}_0 = (\mathbf{u}_0^0, \mathbf{u}_1^0, \mathbf{w}^0)$ in $\mathcal{U}$. Set $\mathbf{y}_\lambda = \mathbf{y}(\mathbf{u}+\lambda\mathbf{v}_0) - \mathbf{y}(\mathbf{u})$. Then $\mathbf{y}_\lambda$ is a weak solution of

$$\begin{cases}
\dfrac{\mathbf{y}_\lambda}{dt} - \mathbf{D}\Delta\mathbf{y}_\lambda = \mathbf{A}\mathbf{y}_\lambda + \mathbf{C}(\mathbf{F}(\mathbf{y}(\mathbf{u}+\lambda\mathbf{v}_0)) - \mathbf{C}\mathbf{F}(\mathbf{y}(\mathbf{u})) + \lambda\mathbf{B}^0\mathbf{u}_0^0 & \textit{in } Q, \\
\dfrac{\partial\mathbf{y}_\lambda}{\partial\eta} = \lambda\mathbf{B}^1\mathbf{u}_1^0 \quad \textit{on } \Sigma, \\
\mathbf{y}_\lambda(0,x) = \lambda\mathbf{E}\mathbf{w}^0 \quad \textit{in } \Omega.
\end{cases} \tag{3.19}$$

Since the nonlinear term satisfies the Lipschitz continuity (2.5), we have

$$\begin{aligned}
& |(\mathbf{C}\mathbf{F}(\mathbf{y}(\mathbf{u}+\lambda\mathbf{v}_0) - \mathbf{C}\mathbf{F}(\mathbf{y}(\mathbf{u})), \mathbf{y}_\lambda)_{\mathcal{H}}| \\
\leq \; & \|\mathbf{C}\| \; |\mathbf{F}(\mathbf{y}(\mathbf{u}+\lambda\mathbf{v}_0) - \mathbf{F}(\mathbf{y}(\mathbf{u}))|_{\mathcal{H}} \; |\mathbf{y}_\lambda|_{\mathcal{H}} \leq K\|\mathbf{C}\| \; |\mathbf{y}_\lambda|_{\mathcal{H}}^2
\end{aligned} \tag{3.20}$$

Using (3.20), we can deduce by Proposition 1 that

$$\|\mathbf{y}_\lambda\|_{L^2(0,T;\mathcal{V})}^2, \; \|\mathbf{y}_\lambda\|_{L^\infty(0,T;\mathcal{H})}^2 \leq C\lambda(|\mathbf{E}\mathbf{w}^0|_{\mathcal{H}}^2 + \|\mathbf{B}^0\mathbf{u}_0^0\|_{L^2(0,T;\mathcal{V})}^2 + \|\mathbf{B}^1\mathbf{u}_1^0\|_{L^2(0,T;\mathcal{H}^{-\frac{1}{2}}(\Gamma))}^2). \tag{3.21}$$

This means that

$$\mathbf{y}_\lambda \to 0 \quad \textit{strongly in} \quad L^\infty(0,T;\mathcal{H}) \cap L^2(0,T;\mathcal{V})$$

as $\lambda \to 0$. Now it is obviously by the inclusion $C([0,T];\mathcal{H}) \subset \mathbf{W}(0,T)$ that

$$\mathbf{y}_\lambda \to 0 \quad \textit{strongly in} \quad C([0,T];\mathcal{H})$$

as $\lambda \to 0$. This proves Proposition 2.

In this subsection we further assume that $F_j \in C^1(\mathbf{R})$ and

$$\exists K > 0 : \quad |F_j'(s)| \le K, \quad j = 1,2,...,n. \tag{3.22}$$

Then for any $\mathbf{y} \in \mathcal{H}$, $\mathbf{F}$ is Gâteaux differentiable and the derivative $\partial_\mathbf{y}\mathbf{F}(\mathbf{y}) \in \mathcal{L}(\mathcal{H})$ is given by

$$\partial_\mathbf{y}\mathbf{F}(\mathbf{y})\mathbf{z} = \begin{pmatrix} F_1'(y_1)z_1 \\ F_1'(y_2)z_2 \\ \vdots \\ F_n'(y_n)z_n \end{pmatrix}, \quad \forall \mathbf{z} = (z_1, z_2, ..., z_n)^T \in \mathcal{H}, \tag{3.23}$$

where $\mathbf{y} = (y_1, y_2, ..., y_n)^T \in \mathcal{H}$ and the multiplication operators $F_j'(y_j) : H \to H$ are bounded for all $j = 1,2,...,n$. We remark that (3.22) is stronger than (2.5) and $\|\partial_\mathbf{y}\mathbf{F}(\mathbf{y})\|_{\mathcal{L}(\mathcal{H})} \le K$.

Theorem 3 *Assume* (3.22). *Then the map* $\mathbf{u} \to \mathbf{y}(\mathbf{u})$ *of* $\mathcal{U}$ *into* $\mathbf{W}(0,T)$ *is weakly Gâteaux differentiable at* $\mathbf{u}^* = (\mathbf{u}_0^*, \mathbf{u}_1^*, \mathbf{w}^*)$ *and such the Gâteaux derivative of* $\mathbf{y}(\mathbf{u})$ *at* $\mathbf{u}^*$ *in the direction* $\mathbf{u} - \mathbf{u}^* \in \mathcal{U}$, *say* $\mathbf{z} = D\mathbf{y}(\mathbf{u}^*)(\mathbf{u} - \mathbf{u}^*) \in \mathbf{W}(0,T)$, *is a unique weak solution of the following equation*

$$\begin{cases} \dfrac{\partial \mathbf{z}}{\partial t} - \mathbf{D}\Delta\mathbf{z} = \mathbf{A}\mathbf{z} + \mathbf{C}\partial_\mathbf{y}\mathbf{F}(\mathbf{y}(\mathbf{u}))\mathbf{z} + \mathbf{B}^0(\mathbf{u}_0 - \mathbf{u}_0^*) \quad in \;\; Q, \\ \dfrac{\partial \mathbf{z}}{\partial \eta} = \mathbf{B}^1(\mathbf{u}_1 - \mathbf{u}_1^*) \quad on \;\; \Sigma, \\ \mathbf{z}(0,x) = \mathbf{E}(\mathbf{w} - \mathbf{w}^*) \quad in \;\; \Omega. \end{cases} \tag{3.24}$$

Proof: Let $\mathbf{u} = (\mathbf{u}_0, \mathbf{u}_1, \mathbf{w}), \mathbf{u}^* = (\mathbf{u}_0^*, \mathbf{u}_1^*, \mathbf{w}^*)$ and $\mathbf{v}_0 = \mathbf{u} - \mathbf{u}^* = (\mathbf{u}_0^0, \mathbf{u}_1^0, \mathbf{w}^0)$. Let $\mathbf{y}(\mathbf{u}^* + \lambda\mathbf{v}_0)$ and $\mathbf{y}(\mathbf{u}^*)$ be the solution of (3.1) corresponding to $\mathbf{u}^* + \lambda\mathbf{v}_0$ and $\mathbf{u}^*$ in $\mathcal{U}_{ad}$, respectively. Set $\mathbf{y}_\lambda = \mathbf{y}(\mathbf{u}^* + \lambda\mathbf{v}_0) - \mathbf{y}(\mathbf{u}^*)$ and $\mathbf{z}_\lambda = \dfrac{\mathbf{y}_\lambda}{\lambda}$, then $\mathbf{z}_\lambda$ satisfies

$$\begin{cases} \dfrac{\partial \mathbf{z}_\lambda}{\partial t} - \mathbf{D}\Delta\mathbf{z}_\lambda = \mathbf{A}\mathbf{z}_\lambda + \mathbf{C}\lambda^{-1}(\mathbf{F}(\mathbf{y}(\mathbf{u}^* + \lambda\mathbf{v}_0)) - \mathbf{F}(\mathbf{y}(\mathbf{u}^*))) + \mathbf{B}^0\mathbf{u}_0^0 \quad in \;\; Q, \\ \dfrac{\partial \mathbf{z}_\lambda}{\partial \eta} = \mathbf{B}^1\mathbf{u}_1^0 \quad on \;\; \Sigma, \\ \mathbf{z}_\lambda(0) = \mathbf{E}\mathbf{w}^0 \quad in \; \Omega \end{cases} \tag{3.25}$$

in the weak sense. By substituting $\mathbf{v} = \mathbf{z}_\lambda$ in the weak form of (3.25) and integrating it over [0,T], we have

$$\begin{aligned} &\int_0^T \langle \mathbf{z}_\lambda', \mathbf{z}_\lambda \rangle_{\mathcal{V},\mathcal{V}'} + (\mathbf{D}\nabla\mathbf{z}_\lambda, \nabla\mathbf{z}_\lambda)_\mathcal{H} dt \\ = &\int_0^T (\mathbf{A}\mathbf{z}_\lambda, \mathbf{z}_\lambda)_\mathcal{H} dt + \int_0^T (\mathbf{C}\lambda^{-1}(\mathbf{F}(\mathbf{y}(\mathbf{u}^* + \lambda\mathbf{u}_0)) - \mathbf{F}(\mathbf{y}(\mathbf{u}^*))), \mathbf{z}_\lambda)_\mathcal{H} dt \\ &+ \int_0^T \langle \mathbf{B}^0\mathbf{u}_0^0, \mathbf{z}_\lambda \rangle dt + \int_0^T \langle \mathbf{B}^1\mathbf{u}_1^0, \mathbf{z}_\lambda|_\Gamma \rangle_{\mathcal{H}^{-\frac{1}{2}}(\Gamma), \mathcal{H}^{\frac{1}{2}}(\Gamma)} dt. \end{aligned} \tag{3.26}$$

From (2.5), we have the estimate

$$|(\mathbf{C}\lambda^{-1}(\mathbf{F}(\mathbf{y}(\mathbf{u}^* + \lambda \mathbf{v_0})) - \mathbf{F}(\mathbf{y}(\mathbf{u}^*))), \mathbf{z}_\lambda)_\mathcal{H}| \le K\|\mathbf{C}\| \, |\mathbf{z}_\lambda|^2_\mathcal{H}. \tag{3.27}$$

By (3.26) and (3.27), we can verify the boundedness of $\{\mathbf{z}_\lambda\}$ in $\mathbf{W}(0,T)$. Hence by the compactness theorem there exists a $\mathbf{z} \in \mathbf{W}(0,T)$ and a subsequence $\{\mathbf{z}_{\lambda_k}\}$ of $\{\mathbf{z}_\lambda\}$ such that

$$\mathbf{z}_{\lambda_k} \to \mathbf{z} \textit{ weakly in } \mathbf{W}(0,T), \tag{3.28}$$

$$\mathbf{z}_{\lambda_k} \to \mathbf{z} \textit{ strongly in } L^2(0,T;\mathcal{H}) \tag{3.29}$$

as $\lambda_k \to 0$. Since $L^2(0,T;\mathcal{H}) = (L^2(Q))^n$, it follows from (3.29) that, if necessary by taking subsequence of $\{\lambda_k\}$,

$$\mathbf{z}_{\lambda_k} \to \mathbf{z} \quad \textit{a.e. in } Q \tag{3.30}$$

as $\lambda_k \to \infty$. For any $\phi \in L^2(0,T;\mathcal{V}) \subset L^2(0,T;\mathcal{H})$, $\mathbf{z}_{\lambda_k}$ satisfies

$$\begin{aligned}
&\int_0^T \langle \mathbf{z}'_{\lambda_k}(t) + \mathbf{D}\Delta \mathbf{z}_{\lambda_k}(t), \phi(t)\rangle_{\mathcal{V}',\mathcal{V}} dt \\
= \; &\int_0^T (\mathbf{A}\mathbf{z}_{\lambda_k}, \phi(t))_\mathcal{H} dt + \int_0^T (\mathbf{C}\frac{\mathbf{F}(\mathbf{y}(\mathbf{u}^* + \lambda_k \mathbf{v_0})) - F(\mathbf{y}(\mathbf{u}^*))}{\lambda_k}, \phi(t))_\mathcal{H} dt \\
&+ \int_0^T \langle \mathbf{B}^0\mathbf{u}_0^0, \phi(t)\rangle_{\mathcal{V}',\mathcal{V}} dt + \int_0^T \langle \mathbf{B}^1\mathbf{u}_1^0, \phi(t)|_\Gamma\rangle_{\mathcal{H}^{-\frac{1}{2}}(\Gamma),\mathcal{H}^{\frac{1}{2}}(\Gamma)} dt.
\end{aligned} \tag{3.31}$$

Now we shall prove that

$$\frac{1}{\lambda_k}\{\mathbf{F}(\mathbf{y}(\mathbf{u}^* + \lambda_k \mathbf{v_0})) - \mathbf{F}(\mathbf{y}(\mathbf{u}^*))\} \to \partial_\mathbf{y}\mathbf{F}(\mathbf{y}(\mathbf{u}))\mathbf{z} \textit{ strongly in } L^2(0,T;\mathcal{H}). \tag{3.32}$$

By the mean value theorem, there exist $\theta_1, \theta_2, ..., \theta_n \in [0,1]$ such that

$$\frac{1}{\lambda_k}\{\mathbf{F}(\mathbf{y}(\mathbf{u}^* + \lambda_k \mathbf{v_0}; t,x)) - \mathbf{F}(\mathbf{y}(\mathbf{u}^*; t,x))\} = \partial_\mathbf{y}\mathbf{F}(\mathbf{y}_{\theta_k}(t,x))\mathbf{z}_{\lambda_k}(t,x) \tag{3.33}$$

where $\mathbf{y}_{\theta_k}(t,x) = (\theta_1 y_1(\mathbf{u}^*;t,x) + (1-\theta_1)y_1(\mathbf{u}^* + \lambda_k \mathbf{v}_0;t,x), ..., \theta_n y_n(\mathbf{u}^*;t,x) + (1-\theta_n)y_n(\mathbf{u}^* + \lambda_k \mathbf{v}_0;t,x))^T$. By (3.30), we can see easily that

$$\mathbf{y}_{\theta_k}(t,x) \to \mathbf{y}(\mathbf{u}^*;t,x) \quad \textit{a.e. in } Q. \tag{3.34}$$

Since F_j are continuously differentiable for all $j = 1,2,...,n$, we have by (3.30), (3.33) and (3.34) that

$$\frac{1}{\lambda_k}\{\mathbf{F}(\mathbf{y}(\mathbf{u}^* + \lambda_k \mathbf{v_0}) - \mathbf{F}(\mathbf{y}(\mathbf{u}^*))\} \to \partial_\mathbf{y}\mathbf{F}(\mathbf{y}(\mathbf{u}^*)\mathbf{z} \quad \textit{a.e. in } Q. \tag{3.35}$$

It is verified by $\|\partial_\mathbf{y}\mathbf{F}(\mathbf{y}_{\theta_k})\|_{\mathcal{L}(\mathcal{H})}, \|\partial_\mathbf{y}\mathbf{F}(\mathbf{y}(\mathbf{u}^*)\|_{\mathcal{L}(\mathcal{H})} \le K$ that

$$|\partial_\mathbf{y}\mathbf{F}(\mathbf{y}_{\theta_k})\mathbf{z}_{\lambda_k} - \partial_\mathbf{y}\mathbf{F}(\mathbf{y}(\mathbf{u}^*)\mathbf{z}|^2 \le 2K^2(|\mathbf{z}_{\lambda_k}|^2_\mathcal{H} + |\mathbf{z}|^2_\mathcal{H}). \tag{3.36}$$

Hence by applying the Lebesgue dominated convergence theorem, we deduce from (3.35) and (3.36) that

$$\begin{aligned}
&\lim_{\lambda_k \to 0}\int_0^T\int_\Omega |\partial_\mathbf{y}\mathbf{F}(\mathbf{y}_{\theta_k})\mathbf{z}_\lambda - \partial_\mathbf{y}\mathbf{F}(\mathbf{y}(\mathbf{u}^*)\mathbf{z}|^2 dxdt \\
= \; &\int_Q \lim_{\lambda_k \to 0} |\partial_\mathbf{y}\mathbf{F}(\mathbf{y}_{\theta_k})\mathbf{z}_\lambda - \partial_\mathbf{y}\mathbf{F}(\mathbf{y}(\mathbf{u}^*)\mathbf{z}|^2 dxdt = 0.
\end{aligned} \tag{3.37}$$

This completes the proof of (3.32). Then we have the convergence

$$\int_0^T (\mathbf{C}\frac{\mathbf{F}(\mathbf{y}(\mathbf{u}^* + \lambda_k \mathbf{v_0}) - F(\mathbf{y}(\mathbf{u}^*))}{\lambda_k}, \phi(t))_{\mathcal{H}} dt \to \int_0^T (\mathbf{C}\partial_{\mathbf{y}}\mathbf{F}(\mathbf{y}(\mathbf{u}^*))\mathbf{z}, \phi(t))_{\mathcal{H}} dt \qquad (3.38)$$

as $\lambda_k \to 0$ for any $\phi \in L^2(0,T;\mathcal{H})$. Next we show that $\mathbf{z}$ is weak solution of (3.24). By taking $\lambda_k \to 0$ in (3.31), we deduce from (3.28) and (3.38) that

$$\begin{aligned}
&\int_0^T \langle \mathbf{z}'(t), \phi(t)\rangle_{\mathcal{V},\mathcal{V}'} dt + \int_0^T (\mathbf{D}\nabla\mathbf{z}, \nabla\phi(t))_{\mathcal{H}} dt \\
= &\int_0^T (\mathbf{Az}, \phi(t))_{\mathcal{H}} dt + \int_0^T (\mathbf{C}\partial_{\mathbf{y}}\mathbf{F}(\mathbf{y}(\mathbf{u}^*))\mathbf{z}, \phi(t))_{\mathcal{H}} dt \\
&+ \int_0^T \langle \mathbf{B}^0\mathbf{u}_0^0, \phi(t)\rangle_{\mathcal{V}',\mathcal{V}} dt + \int_0^T \langle \mathbf{B}^1\mathbf{u}_1^0, \phi(t)|_\Gamma\rangle_{\mathcal{H}^{-\frac{1}{2}}(\Gamma),\mathcal{H}^{\frac{1}{2}}(\Gamma)} dt. \qquad (3.39)
\end{aligned}$$

Hence we can conclude from (3.39) that z satisfies the equation

$$\begin{aligned}
&\langle \mathbf{z}'(t), \mathbf{v}\rangle_{\mathcal{V},\mathcal{V}'} + (\mathbf{D}\nabla\mathbf{z}, \nabla\mathbf{v}) \\
= \ &(\mathbf{Az}, \mathbf{v})_{\mathcal{H}} + (\mathbf{C}\partial_{\mathbf{y}}\mathbf{F}(\mathbf{y}(\mathbf{u}^*))\mathbf{z}, \mathbf{v})_{\mathcal{H}} + \langle \mathbf{B}^0\mathbf{u}_0^0, \mathbf{v}\rangle_{\mathcal{V}',\mathcal{V}} + \langle \mathbf{B}^1\mathbf{u}_1^0, \mathbf{v}|_\Gamma\rangle_{\mathcal{H}^{-\frac{1}{2}}(\Gamma),\mathcal{H}^{\frac{1}{2}}(\Gamma)}, \quad \forall \mathbf{v} \in \mathcal{V}
\end{aligned}$$

in the sense of $\mathcal{D}'(0,T)$, Using integration by parts in (3.39) for $\phi \in C^1([0,T];\mathcal{V})$ we can show $\mathbf{z}(0) = \mathbf{Ew}^0$. Therefore $\mathbf{z}$ is the weak solution of (3.24). This proves Theorem 3.

By Theorem 3, the cost $J(\mathbf{u})$ is weakly Gâteaux differentiable at $\mathbf{u}$ in the direction $\mathbf{u} - \mathbf{u}^*$ and the optimality condition (3.16) is rewritten by

$$\begin{aligned}
&J'(\mathbf{u}^*)(\mathbf{u} - \mathbf{u}^*) \\
= \ &(\mathbf{y}(\mathbf{u}^*) - \mathbf{z}_d^0, D\mathbf{y}(\mathbf{u}^*)(\mathbf{u} - \mathbf{u}^*))_{(L^2(Q))^n} + (\mathbf{y}(\mathbf{u}^*, T) - \mathbf{z}_d^1, D\mathbf{y}(\mathbf{u}^*, T)(\mathbf{u} - \mathbf{u}^*))_{(L^2(\Omega))^n} \\
&+(\mathbf{Nu}^*, \mathbf{u} - \mathbf{u}^*)_{\mathcal{U}}, \qquad \forall \mathbf{u} = (\mathbf{u_0}, \mathbf{u_1}, \mathbf{w}) \in \mathcal{U}_{ad}. \qquad (3.40)
\end{aligned}$$

Now we can give the necessary condition of optimality for the distributed Hopfield-type neural networks. The condition is represented in terms of the state and adjoint systems and the variational inequality.

Theorem 4 *Assume that all assumptions in Section 3 hold. Then the optimal control* $\mathbf{u}^* \in \mathcal{U}_{ad}$ *for* (3.3) *is characterized by the following system of equations and inequality:*

$$\begin{cases}
\dfrac{\partial \mathbf{y}}{\partial t} - \mathbf{D}\Delta\mathbf{y} = \mathbf{Ay} + \mathbf{CF}(\mathbf{y}) + \mathbf{B}^0\mathbf{u}_0^* \quad in \ Q, \\
\dfrac{\partial \mathbf{y}}{\partial \eta} = \mathbf{B}^1\mathbf{u}_1^* \quad on \ \Sigma, \\
\mathbf{y}(0,x) = \mathbf{Ew}^* \quad in \ \Omega.
\end{cases} \qquad (3.41)$$

$$\begin{cases}
-\dfrac{\partial \mathbf{p}}{\partial t} + \mathbf{D}\Delta\mathbf{p} = \mathbf{Ap} + \partial_{\mathbf{y}}\mathbf{F}(\mathbf{y}(\mathbf{u}^*))^*\mathbf{C}^*\mathbf{p} + \mathbf{y}(\mathbf{u}^*) - \mathbf{z}_d^0 \quad in \ Q, \\
\dfrac{\partial \mathbf{p}}{\partial \eta} = 0 \quad on \ \Sigma, \\
\mathbf{p}(\mathbf{u}^*, T, x) = \mathbf{y}(\mathbf{u}^*, T) - \mathbf{z}_d^1 \quad in \ \Omega.
\end{cases} \qquad (3.42)$$

$$\begin{aligned}
&(\mathbf{p}(\mathbf{u}^*, 0), \mathbf{E}(\mathbf{w} - \mathbf{w}^*))_{\mathcal{H}} + \int_0^T \langle \mathbf{p}(\mathbf{u}^*), \mathbf{B}^0(\mathbf{u_0} - \mathbf{u}_0^*)\rangle_{\mathcal{V}',\mathcal{V}} dt \\
&+ \int_0^T \langle \mathbf{p}(\mathbf{u}^*)|_\Gamma, \mathbf{B}^1(\mathbf{u_1} - \mathbf{u}_1^*)\rangle_{\mathcal{H}^{\frac{1}{2}}(\Gamma),\mathcal{H}^{-\frac{1}{2}}(\Gamma)} dt + (\mathbf{Nu}^*, \mathbf{u} - \mathbf{u}^*) \geq 0, \quad \forall \mathbf{u} \in \mathcal{U}_{ad}. \qquad (3.43)
\end{aligned}$$

3.3 Bang-Bang property

In this subsection, we consider the special case where $n = 1$, $\mathcal{U} = L^2(Q) \times L^2(\Sigma) \times L^2(\Omega)$ and

$$\begin{aligned}\mathcal{U}_{ad} &= \mathcal{U}_{ad}^0 \times \mathcal{U}_{ad}^1 \times \mathcal{W}_{ad} = \{u_0 \mid u_a^0 \le u_0 \le u_b^0, \ \ a.e. \ on \ Q\} \\ &\quad \times \{u_1 \mid u_a^1 \le u_1 \le u_b^1, \ \ a.e. \ on \ \Sigma\} \times \{w \mid w_a \le w \le w_b, \ \ a.e. \ on \ \Omega\}, \qquad (3.44)\end{aligned}$$

with $u_a^0,\ u_b^0 \in L^\infty(Q),\ u_a^1,\ u_b^1 \in L^\infty(\Sigma),\ w_a,\ w_b \in L^\infty(\Omega)$. Assume that $\mathbf{N} = 0, \mathbf{E} = \mathbf{B}^0 = \mathbf{B}^1 = I$. Since $\mathcal{U}_{ad}$ is closed and convex in $\mathcal{U}$, then from the necessary condition (3.43), we have

$$\begin{aligned}(p(\mathbf{u}^*,0), w - w^*)_{L^2(\Omega)} + \int_0^T (p(\mathbf{u}^*), u_0 - u_0^*)_{L^2(\Omega)} dt + \int_0^T (p(\mathbf{u}^*)|_\Gamma, u_1 - u_1^*)_{L^2(\Gamma)} dt \ge 0, \\ \forall \mathbf{u} = (u_0, u_1, w) \in \mathcal{U}_{ad}, \qquad (3.45)\end{aligned}$$

where $\mathbf{u}^* = (u_0^*, u_1^*, w^*) \in \mathcal{U}_{ad}$. By setting $(u_0^*, u_1^*, w) \in \mathcal{U}_{ad}$ in (3.45),we get

$$(p(\mathbf{u}^*,0), w - w^*)_{L^2(\Omega)} \ge 0, \qquad \forall w \in \mathcal{W}_{ad}. \qquad (3.46)$$

Similarly by the Lebesgue convergence theorem, we have from (3.45)

$$\begin{aligned}&(p(\mathbf{u}^*,t), u_0 - u_0^*)_{L^2(\Omega)} \ge 0, \quad a.e. \ \ t \in [0,T], \ \ \forall u_0 \in \mathcal{U}_{ad}^0, \\ &(p(\mathbf{u}^*,t)|_\Gamma, u_1 - u_1^*)_{L^2(\Gamma)} \ge 0 \quad a.e. \ \ t \in [0,T], \ \ \forall u_1 \in \mathcal{U}_{ad}^1.\end{aligned}$$

Then we can deduce the following property of $\mathbf{u}^*$:

$$\begin{array}{llll} i) & \text{if } p(\mathbf{u}^*,0,x) > 0, & x \in \Omega, & \text{then } w^*(x) = w_a(x); \\ & \text{if } p(\mathbf{u}^*,0,x) < 0, & x \in \Omega, & \text{then } w^*(x) = w_b(x). \\ ii) & \text{if } p(\mathbf{u}^*,t,x) > 0, & (t,x) \in Q, & \text{then } u_0^*(t,x) = u_a^0(t,x); \\ & \text{if } p(\mathbf{u}^*,t,x) < 0, & (t,x) \in Q, & \text{then } u_0^*(t,x) = u_b^0(t,x). \\ iii) & \text{if } p(\mathbf{u}^*,t,\xi) < 0, & (t,\xi) \in \Sigma, & \text{then } u_1^*(t,\xi) = u_a^1(t,\xi); \\ & \text{if } p(\mathbf{u}^*,t,\xi) < 0, & (t,\xi) \in \Sigma, & \text{then } u_1^*(t,\xi) = u_b^1(t,\xi). \end{array} \qquad (3.47)$$

This fact (3.47) is well known as the Bang-Bang property of optimal control $\mathbf{u}^*$. In the case of $n \ge 2$, we can obtain the similar Bang-Bang property for $\mathbf{u}^*$ with respect to each component of $\mathbf{u}^* = (\mathbf{u}_1^*, \mathbf{u}_2^*, \mathbf{w}^*)$.

References

[1] N. Arada and J.-P.Raymond, *Dirichlet Boundary Control of Semilinear Parabolic Equations.* Part 1: Problem with No State Constrains, Report 98-05,UMR-CNRS 5640,Toulouse, France.

[2] V. Barbu, *Analysis and Control of Nonlinear Infinite Dimensional Systems,* Academic Press, New York, 1993.

[3] E. Casas, *Pontryagin's principle for state-constrained boundary control problem of semilinear parabolic equations,* SIAM J. Control Optim., vol.35, pp.1297-1327, 1997.

[4] R. Dautray and J-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology, Vol. 5, Evolution Problems I*, Springer-Verlag, Berlin-Heidelberg-New York, 1992.

[5] R. Fitz-Hugh *Mathematical models of threshold phenomena in the nerve membrane*, Bull. Math. Biophys 17, 1955.

[6] A. L. Hodgkin and A. F. Huxley, *A qualitative decription of membrane current and its application to conduction and excitation in nerve*, J. Physiol., Vol.117,500-544, 1952.

[7] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin-Heidelberg-New York, 1971.

[8] J. L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications I, II*, Springer-Verlag, Berlin-Heidelberg-New York, 1972.

[9] J. Nagumo, S. Arimoto and S. Yoshizawa, *An active pulse transmission line simulating nerve axon*, Proc. IRE 50, 2061-2070; 1962. *Bistable transmission line*, IEEE Trans. Circuit Theory, CT-12,pp.400-412, 1965.

[10] R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physis*, Second Edition, Applied Mathematica Sciences. Springer-Verlag, Berlin-Heidelberg-New York, Vol. 68,1997.

[11] R. Temam, *Navier stokes equation*, North-Holland, 1984.

[12] J. Vanualailai, S. Nakagiri and T. Soma, *Practical stability of Hopfield-type neural networks*, this issue.

[13] Q. F. Wang, *Optimal control problems of systems governed by nonlinear evolution equations of parabolic type*, Master Thesis, Graduate School Science and Technology, Kobe University, 1999.

[14] Q. F. Wang and S. Nakagiri, *Optimal control problems for parabolic distributed parameter systems with uniform Lipschitz continuous nonlinearities*, Memo. Grad. School Sci. and Technol., Kobe Univ., 18-A, pp.113-129, 2000.