

# 一种基于动态规划策略的离散动态 大系统递阶优化神经网络<sup>1)</sup>

侯增广 吴沧浦

(北京理工大学自动控制系 北京 100081)

**摘 要** 给出了一种具有集成化特征的、快速求解大规模系统动态规划问题的神经网络模型(LDPNN),该神经网络将大系统的各子系统的动态方程约束嵌入局部优化子网络,使得整个网络的结构简洁、紧凑,便于硬件实现,该神经网络计算模型克服了数值方法迭代计算的缺陷,求解效率高,适宜于大规模动态系统实时优化应用.

**关键词** 动态规划,大系统,动态系统,递阶优化,神经网络.

## A DYNAMIC PROGRAMMING NEURAL NETWORK FOR LARGE-SCALE OPTIMIZATION PROBLEMS

HOU Zengguang WU Cangpu

(Department of Automatic Control, Beijing Institute of Technology, Beijing 100081)

**Abstract** To overcome the difficulties existing in the conventional numerical methods, a novel neural network (LDPNN) for solving large-scale dynamic programming problems is developed in the paper. The dimension of the neural network is reduced greatly by imbedding the dynamic equations of the subsystems into the corresponding local optimization networks. Moreover, the coordination and local optimization are carried out simultaneously. Therefore, the LDPNN not only has concise and compact structure, but also has high efficiency for problem-solving. An demonstrative example shows that the LDPNN is suitable for real-time optimization of large-scale systems.

**Key words** Dynamic programming, large-scale systems, dynamic systems, hierarchical optimization, neural networks.

## 1 引言

现在,动态规划理论已经成为运筹学与现代控制理论的一个重要分支,并在经济、军

1)国家自然科学基金资助项目.

事等诸多领域广泛应用. 然而, 常规的动态规划数值计算方法在实现时, 往往存在令现有的及可预见的计算机难以承受的巨大的内存空间及计算时间要求, 即“维数灾”问题. 尽管已提出了许多改进算法, 如微分动态规划、增量动态规划等等, 然而, 这些算法大多以牺牲计算时间和计算精度为代价, 来换取较小的高速内存需求, 难以满足实际工程问题对计算速度和计算精度愈来愈高的要求.

本文用神经网络技术求解大系统动态规划问题, 以克服传递数值计算方法的困难, 所提出的神经网络模型由局部优化子网络和协调网络构成, 局部优化子网络和协调网络的求解计算同步进行, 不存在数值迭代中的交替等待过程, 求解效率高; 将子系统的动态方程及其初始条件嵌入相应的局部优化子网络中, 避免了因该项约束造成的求解困难; 该网络仅把控制变量作为其状态变量, 使得网络具有低的维数; 该网络具有全集成化特点, 结构更加简洁、紧凑, 便于硬件实现.

## 2 问题描述

考虑由  $N$  个子系统构成的离散时间大系统动态规划(LDFDP)问题:

1) 子系统的动态方程

$$x_i(k+1) = f_i[x_i(k), u_i(k), z_i(k), k], i = 1, \dots, N, k = 0, \dots, K-1, x_i(0) \text{ 给定}; \quad (1)$$

2) 子系统之间的关联方程

$$z_i(k) = \sum_{j=1}^N g_{ij}[x_j(k), u_j(k), k], i = 1, \dots, N, k = 0, \dots, K-1; \quad (2)$$

3) 子系统的目标函数

$$J_i[x_i(k), u_i(k), \dots, x_i(K), k] = \Gamma_i[x_i(k), u_i(k), J_i(k+1), k], k = 0, \dots, K-1, \quad (3)$$

$$J_i[x_i(K), K] = \Gamma_i[x_i(K), K], i = 1, \dots, N. \quad (4)$$

其中  $x_i(k) \in R^{n_i}$ ,  $u_i(k) \in R^{m_i}$ ,  $z_i(k) \in R^{l_i}$  分别为第  $i$  个子系统在  $k$  时刻的状态变量、控制变量及关联输入变量;  $f_i: R^{n_i} \times R^{m_i} \times R^{l_i} \times T \rightarrow R^{n_i}$ ,  $g_{ij}: R^{n_j} \times R^{m_j} \times T \rightarrow R^{l_i}$ , 其中  $T = \{k | k = 0, 1, \dots, K-1\}$ ,  $\Gamma_i(k): R^{n_i} \times R^{m_i} \times R \rightarrow R$ ,  $k \in T$ ,  $\Gamma_i(K): R^{n_i} \rightarrow R$ . 则一个离散时间大系统动态规划问题可以描述为: 求控制向量序列  $u_i(0), u_i(1), \dots, u_i(K-1)$ ,  $i = 1, \dots, N$ , 在满足1),

2) 约束下, 使系统目标函数  $J(k) = \sum_{i=1}^N J_i(k)$  ( $k = 0, \dots, K-1$ ) 极小化.

设问题(LDFDP)满足条件: 1) 对于任一子优化问题  $\min J_i[x_i(k), u_i(k), \dots, x_i(K), k]$ , ( $x_i(k)$  为一可达状态) 至少有一个最优解存在, 即指标函数下界存在; 2) 函数  $f_i, g_{ij}, \Gamma_i$  为二阶连续可微函数; 3) 函数  $\Gamma_i(k)$  为关于其自变量  $J_i(k+1)$  的单调非减函数.

## 3 大系统动态规划神经网络模型(LDPNN)

构造一个具有集成化特征的大系统动态规划神经网络如

$$C_i^k \frac{du_i(k)}{dt} = - \frac{\partial J(k)}{\partial u_i(k)}, \quad i = 1, \dots, N, k = 0, \dots, K-1, \quad (5)$$



其中  $C_i^k \in R^{m_i \times m_i}$  是正定对角矩阵, 为收敛速率控制因子.

$$\frac{\partial J(k)}{\partial u_i(k)} = \frac{\partial \Gamma_i(k)}{\partial u_i(k)} + \frac{\partial f_i(x_i, u_i, z_i, k)}{\partial u_i(k)} \cdot \frac{\partial J_i(k+1)}{\partial x_i(k+1)} \cdot \frac{\partial \Gamma_i(k)}{\partial J_i(k+1)} - \sum_{j=1}^N \frac{\partial g_{ji}(x_i, u_i, k)}{\partial u_i(k)} \cdot \frac{\partial f_j(x_j, u_j, z_j, k+1)}{\partial z_j(k+1)} \cdot \frac{\partial J_j(k+1)}{\partial x_j(k+1)} \cdot \frac{\partial \Gamma_j(k)}{\partial J_j(k+1)}, \quad (6)$$

注意到式(6)的特点, 令

$$p_i(k) = \frac{\partial J_i(k)}{\partial x_i(k)}, \quad i = 1, \dots, N, k = 0, \dots, K-1, \quad (7)$$

$$\lambda_i(k) = - \frac{\partial f_i(x_i, u_i, z_i, k)}{\partial z_j(k)} p_i(k+1), \quad i = 1, \dots, N, k = 0, \dots, K-1, \quad (8)$$

其中  $p_i(k) \in R^{n_i}$  和  $\lambda_i(k) \in R^{l_i}$  分别为共态向量和 Lagrange 乘子向量. 则式(6)简化为

$$\frac{\partial J(k)}{\partial u_i(k)} = \frac{\partial \Gamma_i(k)}{\partial u_i(k)} + \frac{\partial f_i(x_i, u_i, z_i, k)}{\partial u_i(k)} \cdot p_i(k+1) \cdot \frac{\partial \Gamma_i(k)}{\partial J_i(k+1)} - \sum_{j=1}^N \frac{\partial g_{ji}(x_i, u_i, k)}{\partial u_i(k)} \cdot \lambda_j(k) \cdot \frac{\partial \Gamma_j(k)}{\partial J_j(k+1)}. \quad (9)$$

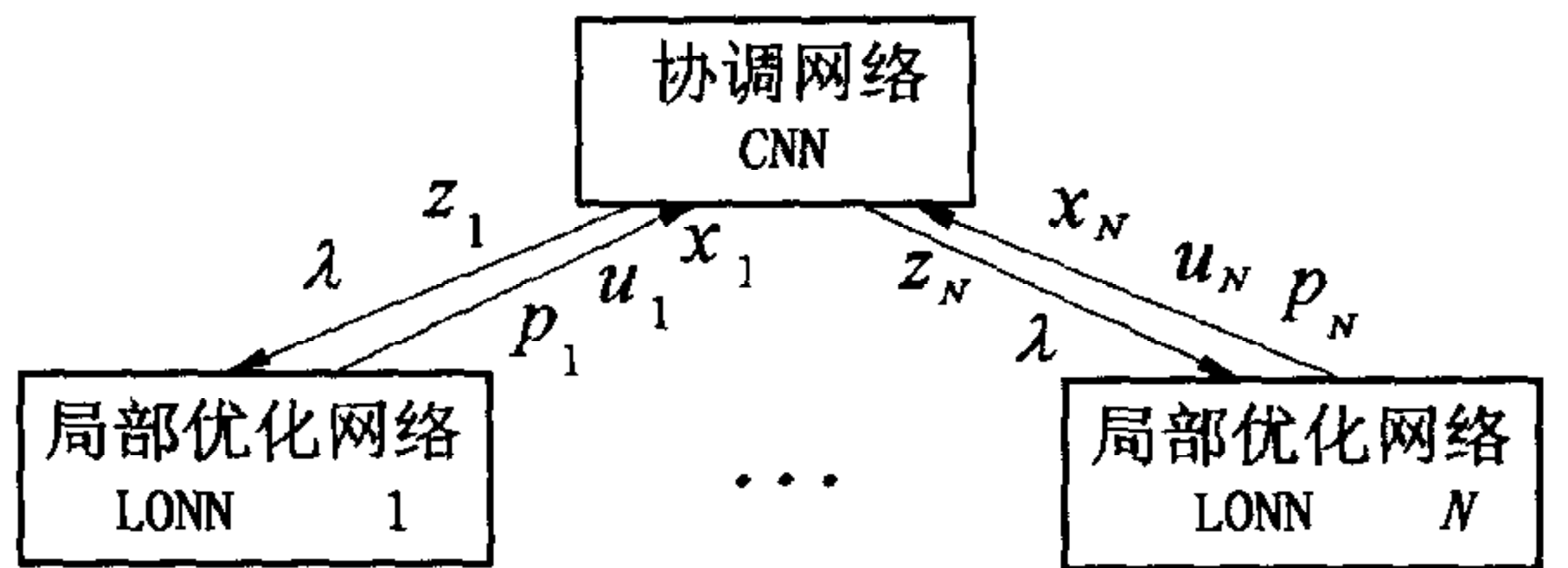


图1 神经网络(LDPNN)的基本结构

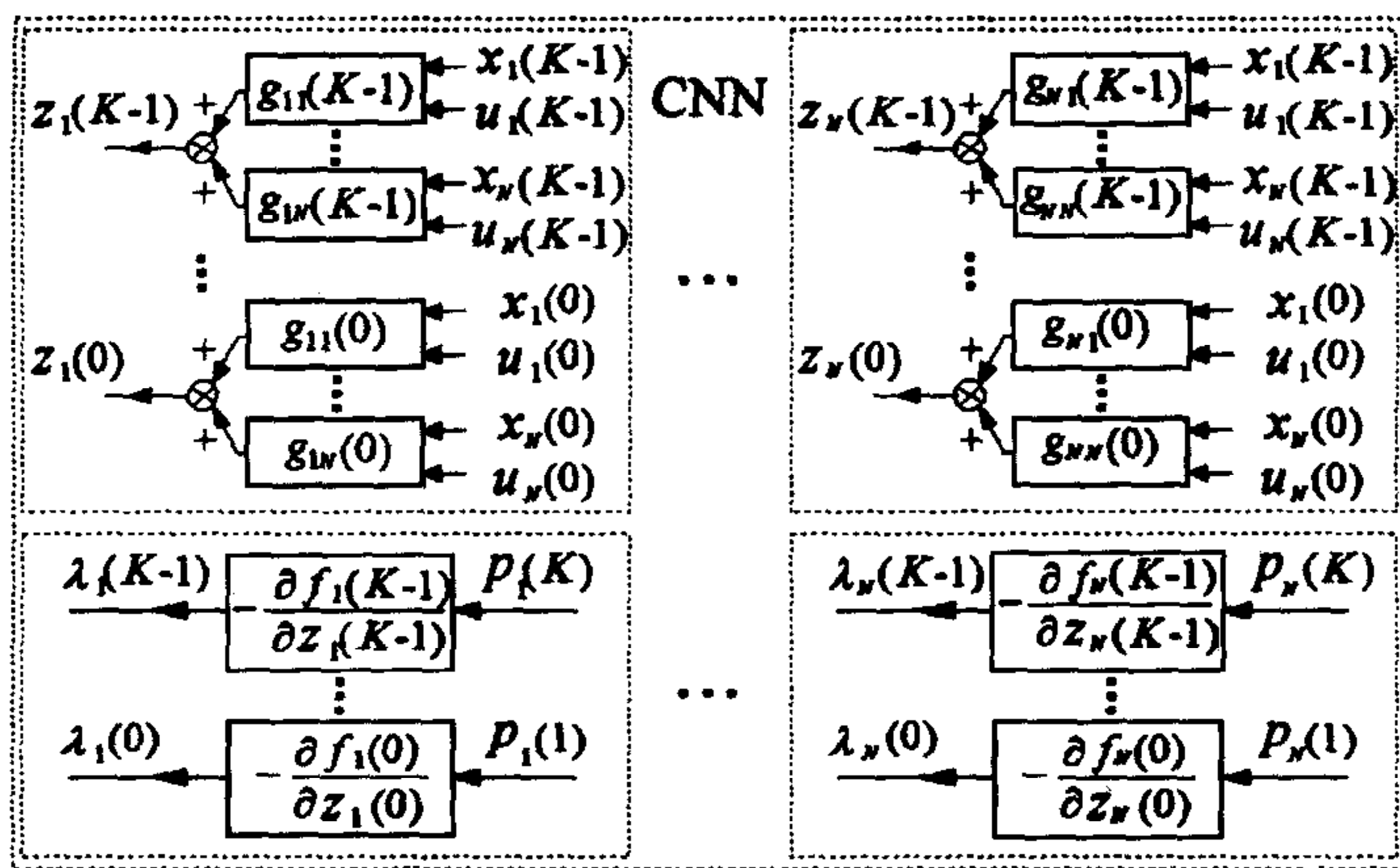


图2(a) 协调网络(CNN)

式(9)中

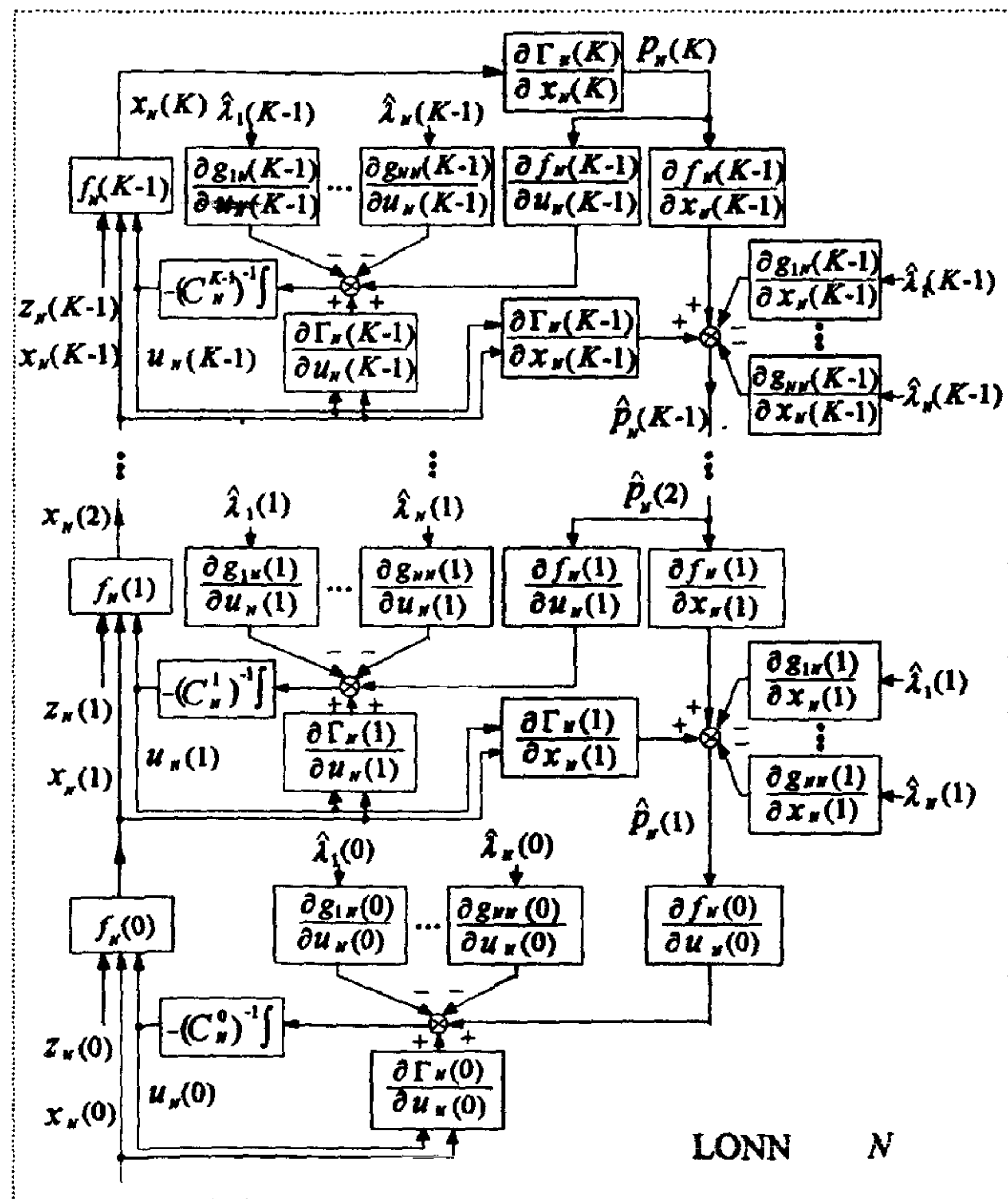
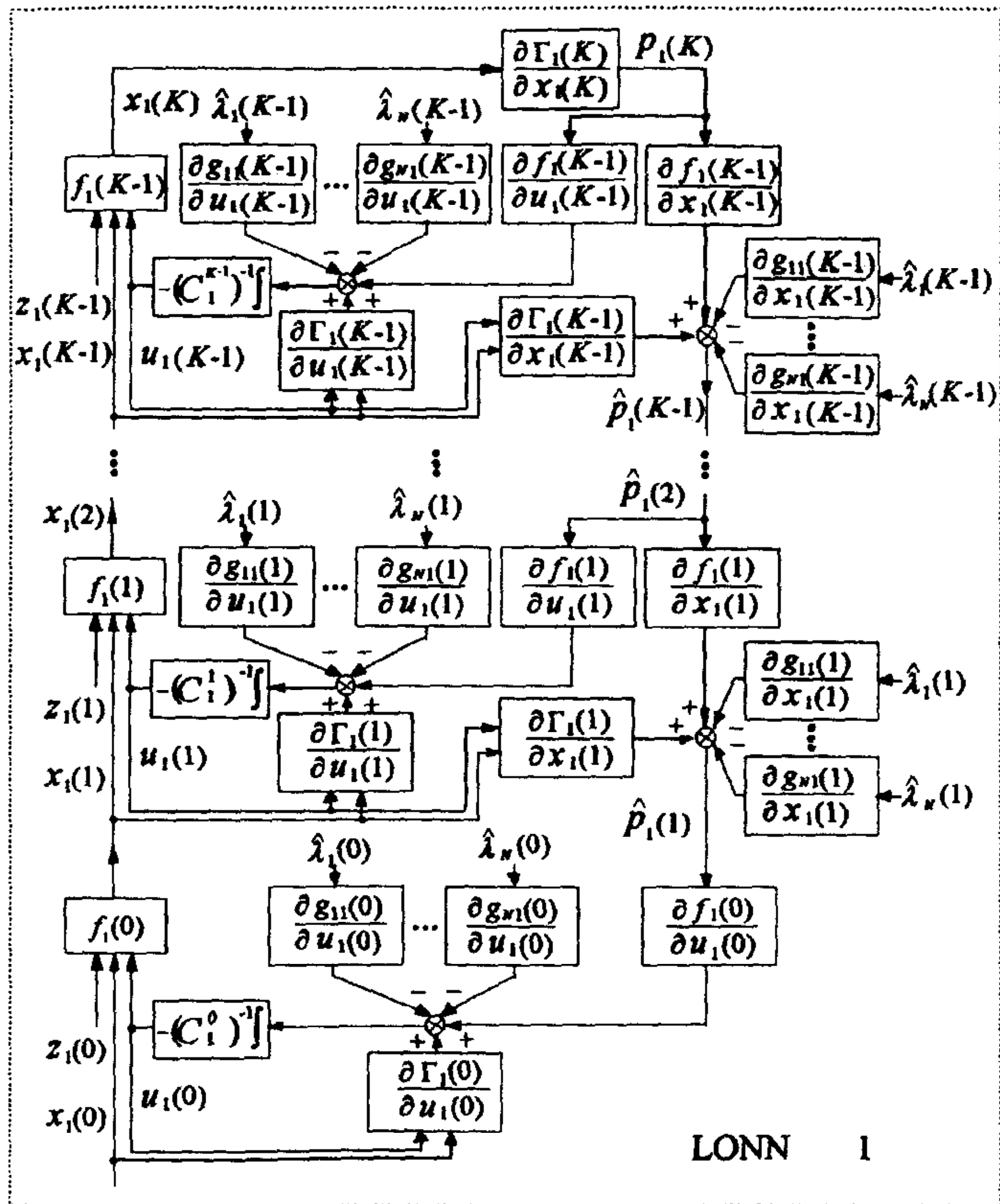
$$p_i(k) = \frac{\partial \Gamma_i(k)}{\partial x_i(k)} + \frac{\partial f_i(x_i, u_i, z_i, k)}{\partial x_i(k)} \cdot p_i(k+1) \cdot \frac{\partial \Gamma_i(k)}{\partial J_i(k+1)} - \sum_{j=1}^N \frac{\partial g_{ji}(x_i, u_i, k)}{\partial x_i(k)} \cdot \lambda_j(k) \cdot \frac{\partial \Gamma_j(k)}{\partial J_j(k+1)}, \quad (10)$$

$$i = 1, \dots, N, k = 1, \dots, K-1;$$

$$p_i(K) = \frac{\partial \Gamma_i[x_i(K)]}{\partial x_i(K)}, \quad i = 1, \dots, N. \quad (11)$$

图2 神经网络(LDPNN)实现原理图

(b)局部优化网络(LONN)



注意到,在递阶优化过程中,协调器的任务是:在每一时刻,根据子系统的状态信息,调整协调变量使系统整体最优;而局部优化单元的任务是:在每一时刻,按照协调器给定的协调变量进行各子系统的局部优化计算.考虑到本文中网络构成及信息传输的特点,可以看出,式(8)和(2)完成的是协调功能,协调向量为 $[\lambda^T, z^T]^T$ ,所以,式(8),(2)构成了协调子网络(CNN);而式(5)和式(9),(10),(11)则完成相应子系统的局部优化功能,它们构成局部优化子网络(LONN).神经网络(LDPNN)的基本结构如图1所示,图2的(a)和(b)为其实现原理图.注意,在图2(b)中, $\hat{p}_i(k+1) = p_i(k+1) \cdot \frac{\partial \Gamma_i(k)}{\partial J_i(k+1)}$ ,  $\hat{\lambda}_j(k) = \lambda_j(k) \cdot \frac{\partial \Gamma_j(k)}{\partial J_j(k+1)}$ .

#### 4 神经网络(LDPNN)的稳定性及其与原问题的等价性

**定理1.** 在满足本文的设计条件下,神经网络(LDPNN)是完全稳定的,即对于神经网络的任意初始状态,该网络将渐近收敛于一个与大系统动态规划问题的初始状态相对应的最优解.

证明,定义神经网络(LDPNN)的能量函数为

$$E(u) = J(0) = \sum_{i=1}^N J_i[x_i(0, u_i(0), \dots, x_i(K), 0)]. \quad (12)$$

由于

$$\begin{aligned} \frac{dE}{dt} = & \sum_{i=1}^N \sum_{k=0}^{K-1} \left( \frac{\partial E}{\partial u_i(k)} \right)^T \frac{du_i(k)}{dt} = \\ & \sum_{i=1}^N \left[ \left( \sum_{j=1}^N \frac{\partial J_j(0)}{\partial u_i(0)} \right)^T (-C_i^0)^{-1} \left( \sum_{j=1}^N \frac{\partial J_j(0)}{\partial u_i(0)} \right) + \right. \\ & \left. \sum_{k=1}^{K-1} \left( \sum_{j=1}^N \frac{\partial J_j(k)}{\partial u_i(k)} \cdot \frac{\partial \Gamma_j(0)}{\partial J_j(k)} \right)^T (-C_i^k)^{-1} \left( \sum_{j=1}^N \frac{\partial J_j(k)}{\partial u_i(k)} \right) \right], \end{aligned} \quad (13)$$

式中第一项  $\left( \sum_{j=1}^N \frac{\partial J_j(0)}{\partial u_i(0)} \right)^T (C_i^0)^{-1} \left( \sum_{j=1}^N \frac{\partial J_j(0)}{\partial u_i(0)} \right) \geq 0$  显然成立,

第二项  $\left( \sum_{j=1}^N \frac{\partial J_j(k)}{\partial u_i(k)} \cdot \frac{\partial \Gamma_j(0)}{\partial J_j(k)} \right)^T (-C_i^k)^{-1} \left( \sum_{j=1}^N \frac{\partial J_j(k)}{\partial u_i(k)} \right) =$

$$- \left[ \left( \frac{\partial J_1(k)}{\partial u_i(k)} \right)^T \dots \left( \frac{\partial J_N(k)}{\partial u_i(k)} \right)^T \right] \Lambda_i^k \left[ \left( \frac{\partial J_1(k)}{\partial u_i(k)} \right)^T \dots, \left( \frac{\partial J_N(k)}{\partial u_i(k)} \right)^T \right]^T, \quad (14)$$

其中  $\Lambda_i^k = \left[ \text{block diag} \left\{ \underbrace{\frac{\partial \Gamma_1(0)}{\partial J_1(k)}, \dots, \frac{\partial \Gamma_1(0)}{\partial J_1(k)}}_{m_i}, \dots, \underbrace{\frac{\partial \Gamma_N(0)}{\partial J_N(k)}, \dots, \frac{\partial \Gamma_N(0)}{\partial J_N(k)}}_{m_i} \right\} \right] \Pi_i^k,$

$\Pi_i^k = [I_{i1} \dots I_{iN}]^T (C_i^k)^{-1} [I_{i1} \dots I_{iN}]$ ,  $I_{i1}, \dots, I_{iN}$  为  $N$  个  $m_i \times m_i$  单位对角阵块.

注意到  $C_i^k$  为正定阵,则  $\Pi_i^k$  为半正定阵,又  $\frac{\partial \Gamma_j(0)}{\partial J_j(k)}$  ( $j=1, \dots, N$ ) 为标量,由条件3)可知  $\frac{\partial \Gamma_j(0)}{\partial J_j(k)} \geq 0$ ,则  $\Lambda_i^k$  为半正定,则式(14)为非负数,所以有

$$dE/dt \leq 0. \quad (15)$$

若式(15)中等式成立,当且仅当  $\frac{\partial J_j(k)}{\partial u_i(k)} = 0$  成立,即  $\frac{du_i(k)}{dt} = 0$  成立,其中  $i, j=1, \dots, N, k$



$=1, \dots, K-1$ . 由条件1)知,  $E(u)$  有下界,  $dE/dt \leq 0$ , 所以  $E(u)$  为神经网络(LDPNN)的一个 Lyapunov 函数. 因此, 对于任意初始状态, 神经网络(LDPNN)将渐近收敛于一个稳定的平衡状态, 并且该平衡状态与大系统动态规划问题的一个最优解相对应.

## 5 例子

考虑如下的经典算例<sup>[5]</sup>, 一个由两个子系统构成的非线性大系统优化问题:

子系统1

$$\begin{aligned} \min J_1 &= \sum_{k=0}^{50} [0.05x_1^2(k) + 0.1u_1^2(k)], \\ \text{s. t. } x_1(k+1) &= 0.9x_1(k) + 0.1z_1(k), \\ x_1(0) &= 10.0. \end{aligned}$$

子系统2

$$\begin{aligned} \min J_2 &= \sum_{k=0}^{50} [0.05x_2^2(k) + 0.05u_2^2(k)], \\ \text{s. t. } x_2(k+1) &= [0.1 - 0.1x_2(k)] \times x_2(k) + 0.2z_2(k), \\ x_2(0) &= 4.5. \end{aligned}$$

子系统间的关联作用

$$z_1(k) = x_2(k), \quad z_2(k) = x_1(k).$$

系统的总体目标函数为

$$\min J = J_1 + J_2.$$

按照前述办法建立一个神经网络(LDPNN), 其仿真结果如图3、图4、图5所示, 其中图3为最优状态轨迹, 图4为最优控制轨线, 图5为目标函数值的动态演化过程. 按照文献[5]方法, 要1.63s 才收敛, 最优值为  $J = 30.9765$ , 而按本文方法在0.06s 时最优值已为  $J = 30.9790$ , 仿真结果表明神经网络(LDPNN)不但求解速度快、动态特性好, 而且具有很高的求解精度, 适合于实际系统的实时优化.

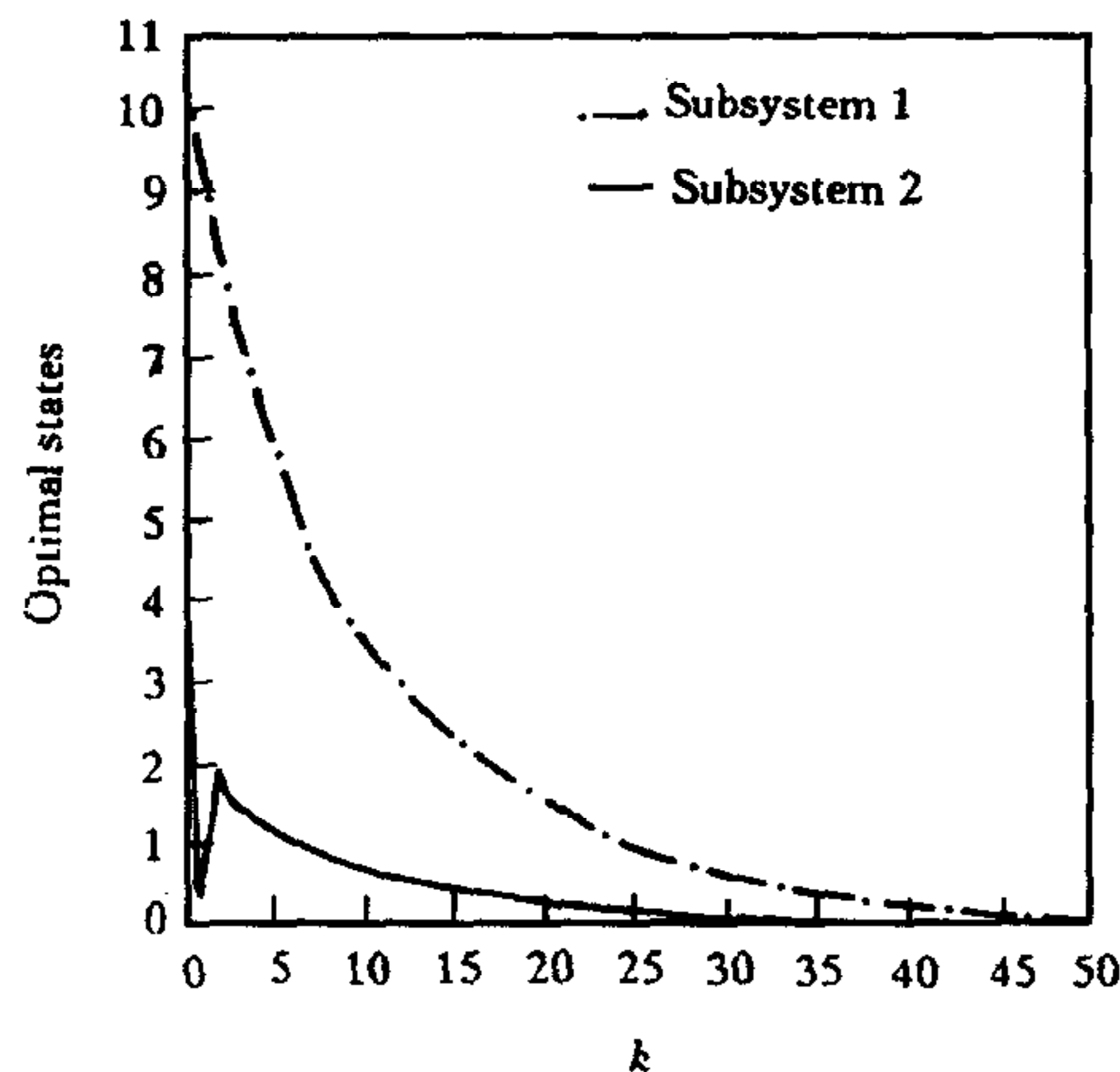


图3 最优状态轨迹

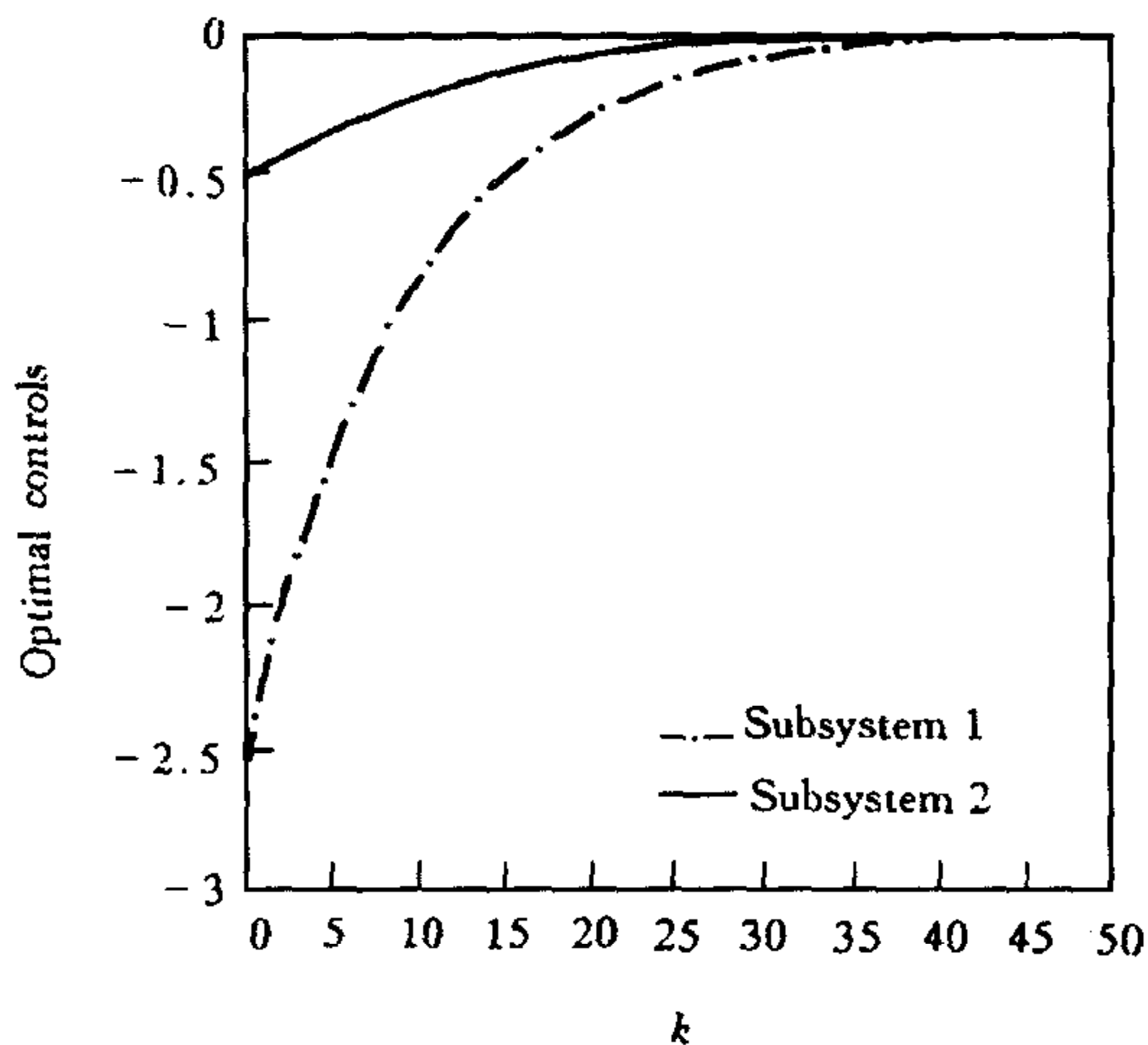


图4 最优控制轨迹

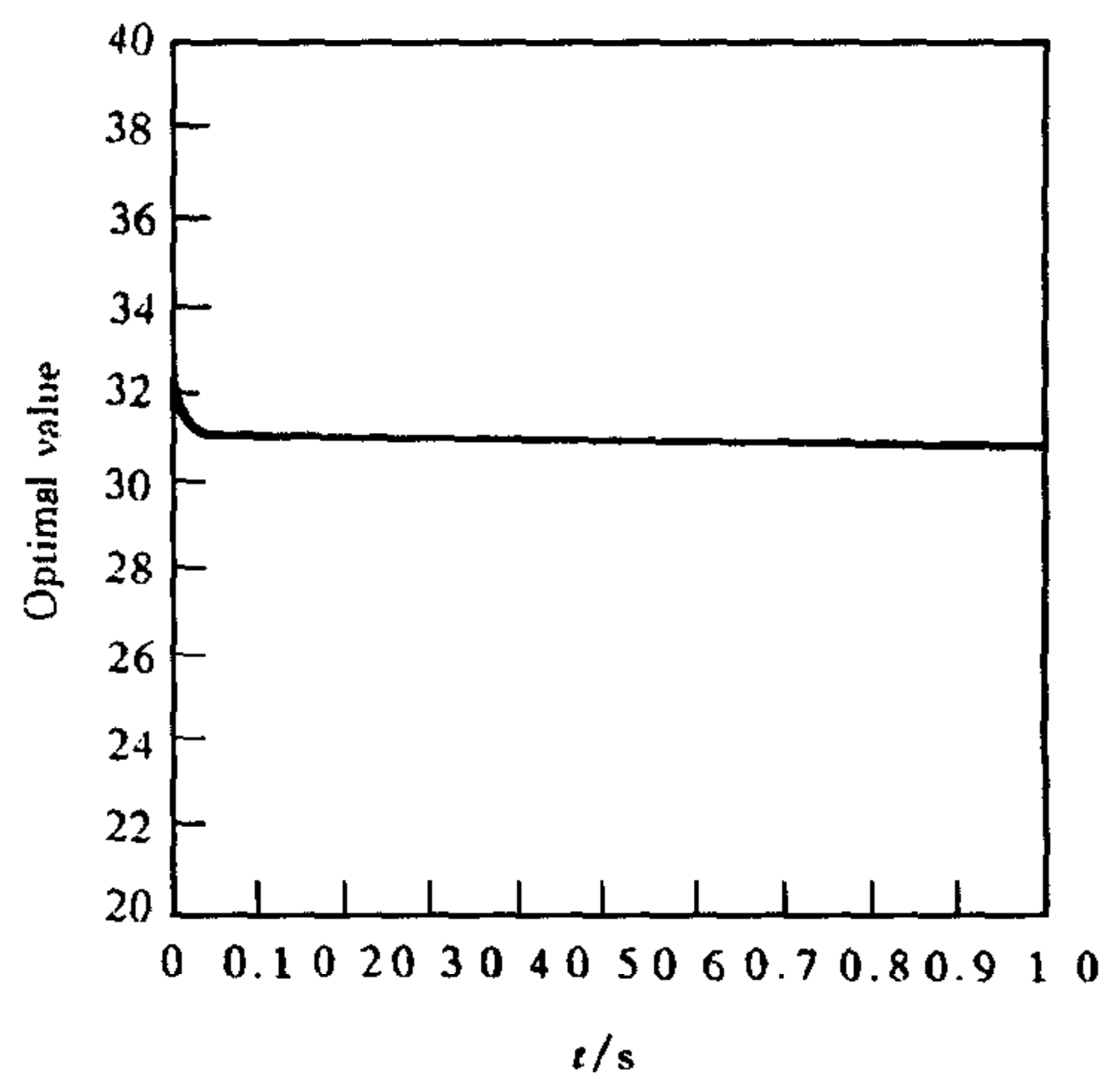


图5 目标函数值的动态演化过程

## 6 结束语

以神经网络求解大系统动态规划问题,可以有效地克服传统动态规划计算方法所存在的“维数灾”问题,由于神经网络的并行计算结构,使得其计算求解速度几乎与动态规划问题的维数无关,这对于大系统优化问题尤其重要,本文所给出的神经网络(LDPNN)具有全集成化的特征,结构简洁、紧凑、高效,适宜于硬件实现及实际应用。

## 参 考 文 献

- 1 Bellman R E. Dynamic programming. Princeton: Princeton University Press, 1957
- 2 Bellman R E, Dreyfus S. Applied dynamic programming. Princeton: Princeton University Press, 1962
- 3 Bertsekas D P. Dynamic programming and stochastic control. New York: Academic, 1976
- 4 Larson R E, Casti J L. Principles of dynamic programming, II. New York: Marcel Dekker, 1982
- 5 Singh M G. Dynamical hierarchical control. Amsterdam: North Holland, 1980
- 6 Hopfield J J, Tank D W. "Neural" computation of decisions in optimization problems. *Biol Cybern*, 1985, **52**: 141—152
- 7 Wu C P, Huang X. A neural network for dynamic programming problems. in: Proc 1st CWC ICIA. Beijing: Science Press, 1993, **1**: 512—517

**侯增广** 1969年生,1993年在燕山大学自动化系获硕士学位,1997年在北京理工大学自动控制系获博士学位,现在中科院系统所做博士后研究.研究兴趣是系统优化的理论与方法、神经网络技术、机器人与非线性控制.

**吴沧浦** 简介见本刊第20卷第6期.