# Multidimensional Taylor Network-Based MIMO Nonlinear Time-Delay System Identification: Postprint

**Authors:** Li Chenlong, Yan Hongsen

**Date:** 2022-04-07T00:00:00+00:00

## Abstract

To address the accuracy and real-time performance issues in the identification of Multiple Input Multiple Output (MIMO) nonlinear time-delay systems, an identification scheme based on Multi-dimensional Taylor Network (MTN) is proposed. MTN serves as the identification model, and the Weight-Elimination (WE) algorithm and Conjugate Gradient (CG) algorithm are comprehensively utilized, i.e., the WE-CG algorithm serves as the learning algorithm for the MTN identification model; the WE algorithm can effectively streamline the structure of the MTN identification model, thereby reducing computational complexity and improving the real-time performance of the model. Finally, a numerical simulation example and an engineering case are introduced to verify the effectiveness of the proposed identification scheme, and a comparison with the traditional MTN identification scheme is conducted, providing an analysis of the accuracy and complexity of both approaches, highlighting the accuracy and real-time performance of the proposed identification scheme. Experimental results demonstrate that the proposed scheme can accurately identify MIMO nonlinear time-delay systems. Meanwhile, compared with the traditional MTN identification scheme, the proposed identification scheme features a more streamlined structure and lower algorithmic complexity.

## Full Text

## Preamble

**Vol. 39 No. 8**
*Application Research of Computers* (ChinaXiv Cooperative Journal)
Accepted Paper

**Identification of MIMO Nonlinear Time-Delay Systems Using Multi-Dimensional Taylor Network**

Li Chenlong[1,2,3,4], Yan Hongsen[1,2]†
(1. School of Automation, Southeast University, Nanjing 210096, China;
2. Key Laboratory of Measurement & Control of Complex Systems of Engineering, Ministry of Education, Southeast University, Nanjing 210096, China;
3. Hangzhou Innovation Institute, Beihang University, Hangzhou 310051, China;
4. Beihang University, Beijing 100191, China)

**Abstract:** To address the accuracy and real-time performance challenges in identifying multiple-input multiple-output (MIMO) nonlinear time-delay systems, this paper proposes an identification scheme based on the Multi-dimensional Taylor Network (MTN). The MTN serves as the identification model, and a hybrid learning algorithm combining Weight-Elimination (WE) with Conjugate Gradient (CG)—denoted as WE-CG—is employed for training the MTN identification model. The WE algorithm effectively simplifies the MTN structure, thereby reducing computational complexity and improving real-time capability. Finally, both a numerical simulation example and an engineering case study are presented to validate the proposed scheme, with comparative analysis against traditional MTN identification methods in terms of accuracy and computational complexity, highlighting the superior accuracy and real-time performance of the proposed approach. Experimental results demonstrate that the proposed scheme can accurately identify MIMO nonlinear time-delay systems while achieving a more compact structure and lower algorithmic complexity compared to conventional MTN identification methods.

**Keywords:** multi-dimensional Taylor network; identification model; MIMO system; nonlinear time-delay system; weight-elimination algorithm
**Classification:** TP271
**doi:** 10.19734/j.issn.1001-3695.2022.01.0031

---

## 0 Introduction

Industrial systems typically exhibit nonlinear and time-delay characteristics, which pose significant challenges for modeling and control [1,2]. In traditional linear systems, precise mathematical models can be obtained, enabling conventional identification methods such as least squares and step response techniques to meet requirements [3,4]. However, for nonlinear time-delay systems, accurate mathematical models are difficult to derive, making conventional linear system identification methods inadequate and thus creating substantial identification difficulties [5]. The emergence of neural networks, with their powerful nonlinear approximation capabilities, has provided an effective solution pathway for nonlinear system identification [6–10]. For instance, reference [6] proposed a time-delay adaptive identification scheme for nonlinear systems based on artificial neural networks and Chebyshev functions; reference [7] developed a nonlinear system identification approach for engineering applications by exploiting the internal relationship between time-delay neural networks and kernel functions; reference [8] presented a time-delay neural network-based identification scheme for nonlinear continuous systems with unknown delays; reference [9] introduced a novel neural network approach for nonlinear system identification; and reference [10] proposed a data-knowledge-based fuzzy neural network identification scheme for nonlinear dynamic systems. Nevertheless, neural networks suffer from long training times, susceptibility to local minima, and high algorithmic complexity, which prevent them from meeting real-time predictive model construction requirements [11]. Moreover, multiple-input multiple-output (MIMO) systems are significantly more complex and difficult to model than single-input single-output (SISO) systems [12–14].

To address these challenges, Professor Yan Hongsen proposed the modeling concept of Multi-dimensional Taylor Network (MTN) in October 2010 [15,16], which was subsequently applied to nonlinear time series modeling and fault monitoring [17–21]. In December 2010, Professor Yan further proposed MTN optimal control [15,16], extending its application to nonlinear system control [22–24], with comprehensive studies on nonlinear time-varying systems [22], nonlinear time-delay systems [23], and nonlinear stochastic systems [24]. MTN and MTN optimal controllers exhibit several key characteristics [15–24]: (a) MTN possesses the capability to approximate any nonlinear function with arbitrary accuracy; (b) MTN models and controllers feature simple structures comprising only addition and multiplication operations, without transcendental or irrational functions, enabling convenient parameter adjustment, easy engineering implementation, and real-time modeling and control without complex mathematical computations; (c) MTN controllers inherently possess "learning" capabilities, allowing continuous self-improvement during control processes and enhancing robustness; and (d) MTN controllers demonstrate excellent scalability, enabling integration with other control methods and concepts while maintaining fault tolerance and evolutionary capabilities.

Building upon this foundation, this paper proposes an MTN-based identification scheme for MIMO nonlinear time-delay systems. First, MTN serves as the identification model. Second, a hybrid learning algorithm combining Weight-Elimination (WE) and Conjugate Gradient (CG)—denoted as WE-CG—is employed to train the MTN identification model. The WE algorithm simplifies the MTN structure to reduce computational complexity and satisfy real-time requirements. Finally, both a numerical simulation example and a continuous stirred tank reactor (CSTR) engineering case study are introduced to validate the proposed scheme, with comparative analysis against traditional MTN identification methods [17] regarding accuracy and complexity.

The main contributions of this paper are as follows:
(a) For MIMO nonlinear time-delay systems, leveraging MTN's powerful approximation capability, we propose an MTN-based identification scheme that is easy to implement;
(b) The WE-CG algorithm serves as the learning algorithm for the MTN identifi-

cation model, where the WE algorithm effectively simplifies the MTN structure, thereby reducing algorithmic complexity and improving real-time performance; (c) The proposed scheme exhibits low computational complexity and favorable real-time performance;
(d) The proposed identification scheme demonstrates generality and practical application value.

---

## 1.1 System Description

The MIMO discrete nonlinear time-delay system is represented as follows:

$$\begin{aligned} y(k+1) =& g(y(k), y(k-1), \dots, y(k-n_y), u(k-d), u(k-d-1), \dots, \\ & u(k-d-n_u)) \end{aligned}$$

where $g(\cdot)$ is an unknown nonlinear vector function; $y(k) \in \mathbb{R}^q$ and $u(k) \in \mathbb{R}^p$ represent the $q$-dimensional output and $p$-dimensional input of the system, respectively; $n_{yi}$ $(i = 1, 2, \dots, q)$ and $n_{uj}$ $(j = 1, 2, \dots, p)$ denote the orders of the $i$-th output component and $j$-th input component, respectively, with $n_y = \max\{n_{y1}, n_{y2}, \dots, n_{yq}\}$ and $n_u = \max\{n_{u1}, n_{u2}, \dots, n_{up}\}$; and $d$ represents the system time-delay.

**Lemma 1 [25]:** Any continuous function defined on a closed interval can be approximated arbitrarily accurately by a polynomial function.

**Lemma 2 [18]:** For a continuous function $f(x_1, x_2, \dots, x_n)$ defined on a closed interval, the approximation can be expressed as $f(x_1, x_2, \dots, x_n) = \sum_{t=1}^{N} w_t \prod_{i=1}^{n} x_i^{\lambda_{ti}}$, where $N$ is the total number of product terms in the approximation expansion, $w_t$ is the weight value preceding the $t$-th product term, and $\lambda_{ti}$ represents the power of variable $x_i$ in the $t$-th product term.

As shown in [Figure 1: see original paper], MTN adopts a forward single-hidden-layer structure comprising an input layer, hidden layer, and output layer. Based on Lemmas 1 and 2, with a sufficiently large $N$, MTN can approximate any model with adequate precision [18,25].

---

### 1.2.3 Multi-Dimensional Taylor Network Complexity Analysis

Since real-time analysis using identical computers and time frames is impractical, this paper employs floating-point operations to estimate computational load, i.e., real-time performance [26,27]. Both addition and multiplication operations are counted as one floating-point operation each, with the total number of additions and multiplications representing the overall floating-point count, which indicates computational complexity. When the expansion order is 2, Table 1 presents the computational complexity of MTN in a single iteration [17], highlighting MTN's excellent real-time performance due to its simple structure.

**Table 1** Computation complexity of MTN model

| Intermediate nodes | 0-order terms | 1-order terms | 2-order terms | Total |
|---|---|---|---|---|
| $n$ | 1 | $n$ | $n(n+1)/2$ | $(n+1)(n+2)/2$ |

---

### 1.2.1 MIMO Multi-Dimensional Taylor Network

For MIMO nonlinear systems, based on multi-dimensional Taylor network, the $j$-th output component can be expressed as [15]:

$$\hat{y}_j(k+1) = \sum_{t=1}^{\hat{N}} \hat{w}_{j,t}(k) \prod_{r=1}^{\hat{t}} \hat{z}_r^{\hat{\lambda}_{t,r}}(k)$$

where $\hat{N} = \sum_{t=0}^{\hat{m}} (n+1)^t$ represents the total number of terms, $\hat{w}_{j,t}(k)$ denotes the weight coefficient for the $t$-th product term in equation (3), $\hat{z}_r(k)$ represents the $r$-th input variable, and $\hat{\lambda}_{t,r}$ indicates the power of variable $\hat{z}_r(k)$ in the $t$-th product term, with $\sum_{r=1}^{\hat{t}} \hat{\lambda}_{t,r} \leq \hat{m}$.

MTN includes only addition and multiplication operations, and its computational complexity is equivalent to the Taylor expansion of a single neuron in a neural network.

---

## 2 MIMO Nonlinear Time-Delay System MTN Identification Scheme

The total number of terms is given by $\hat{N} = \sum_{t=0}^{\hat{m}} (n+1)^t$, with the weight vector $\hat{w}_j(k) = [\hat{w}_{j,1}(k), \hat{w}_{j,2}(k), \ldots, \hat{w}_{j,\hat{N}}(k)]^T$. The MIMO nonlinear system MTN structure is illustrated in [Figure 1: see original paper].

---

### 1.2.2 Multi-Dimensional Taylor Network Approximation Performance

This paper employs MTN as the identification model and the WE-CG algorithm as the learning algorithm, where the WE algorithm effectively simplifies the MTN network structure.

## 2.1 MTN Identification Model

For system (1), which can be approximated with arbitrary precision by MTN, the mapping relationship is denoted as $\hat{g}(\cdot)$. The MTN identification model can be expressed as:

$$\hat{y}_j(k+1) = \sum_{t=1}^{\hat{N}} \hat{w}_{j,t}(k) \prod_{r=1}^{\hat{t}} \hat{z}_r^{\hat{\lambda}_{t,r}}(k)$$

where $\hat{y}(k) \in \mathbb{R}^q$ is the MTN identification model output, with $\hat{y}(k) = [\hat{y}_1(k), \hat{y}_2(k), \ldots, \hat{y}_q(k)]^T$; $\hat{z}(k) \in \mathbb{R}^{q+p}$ represents the input vector; and $n_y$, $n_u$, and $d$ share the same meanings as in equation (1).

For convenience and without loss of generality, the weight vector for the $j$-th output component of MTN is denoted as $\hat{w}_j(k)$. However, the gradient method approaches the minimum in a zigzag manner, causing the search direction to remain orthogonal. Fortunately, the conjugate gradient method effectively resolves this issue, with weight updates performed as:

$$\hat{w}_j^{\tau+1} = \hat{w}_j^{\tau} + \mu^{\tau} p_j^{\tau}$$

where $p_j^{\tau} = -\hat{g}_j^{\tau} + \beta^{\tau} p_j^{\tau-1}$, with initial values $\hat{w}_j^0$ and $p_j^0 = -\hat{g}_j^0$.

## 2.2.2 WE-CG Algorithm

As the number of MTN input nodes increases, the "curse of dimensionality" occurs, necessitating network structure simplification. The weight elimination algorithm can be used to prune intermediate nodes, thereby reducing network complexity. This section combines the CG method with the WE method—designated as WE-CG—to serve as the learning algorithm for the MTN identification model.

**Lemma 3 [28]:** Assume the target function $h(x)$ can be represented by a minimal structure network. When training samples are sufficiently abundant, the decay and merging of intermediate nodes is inevitable. Therefore, pruning algorithms are necessary to simplify the network, as the network can still represent a target function after redundant node removal.

This paper implements MTN network structure simplification through regularization that eliminates redundant nodes. Since this regularization term possesses pruning characteristics [29,30], some redundant output weights gradually decay to zero, enabling the removal of connected intermediate nodes. Based on this, a weight elimination algorithm with Gaussian regularization is introduced into the CG algorithm as the MTN identification model learning algorithm.

The objective function is defined as:

$$J(\hat{w}) = E(\hat{w}) + \gamma C(\hat{w})$$

where $E(\hat{w})$ typically represents the sum of squared network errors; the second term is the regularization term; $\gamma$ is the regularization coefficient with $\gamma \geq 0$; and $w_0$ is the baseline weight value. Intermediate nodes are removed when their output weights satisfy:

$$|w_i| < w_{\min}$$

where $w_{\min}$ is the critical weight value. Since the regularization coefficient $\gamma$ significantly influences learning outcomes, a dynamic $\gamma$ adjustment scheme is proposed during the learning process, modifying $\gamma$ after each data center correction. The MTN identification model adjustment method employs:

$$\gamma(k+1) = \gamma(k) + \Delta\gamma(k)$$

where $\Delta\gamma(k)$ represents the increment, with the same meaning as in equation (11). Following Weigend et al. [29,30], the relationship between error quantities is monitored during learning: $E(k)$ is the error during current weight adjustment; $E(k-1)$ is the error during previous weight adjustment; $A(k)$ is the current weighted average error defined as $A(k) = \kappa A(k-1) + (1-\kappa)E(k)$, where $\kappa$ is a coefficient close to 1; and $D$ is the desired error value, which may result in longer computation time without prior knowledge.

The dynamic adjustment of $\gamma$ follows these rules:

1. If $E(k) \geq E(k-1)$ and $A(k) < D$, then $\gamma(k+1) = \gamma(k) - \Delta\gamma$
2. If $E(k) \geq E(k-1)$ and $A(k) \geq D$ and $E(k) \geq D$, then $\gamma(k+1) = \gamma(k) + \Delta\gamma$
3. If $E(k) < E(k-1)$ and $E(k) < D$, then $\gamma(k+1) = \gamma(k) - \sigma\gamma(k)$, where $\sigma$ is a coefficient close to 1.

---

### 2.2.1 Conjugate Gradient Method

MTN parameters can be obtained through sampling and appropriate learning algorithms. Let $u(k)$ be the system input and $y(k)$ the system output. The error variation equation is:

$$e(k) = y(k) - \hat{y}(k)$$

The error vector is defined as $e(k) = [e_1(k), e_2(k), \dots, e_q(k)]^T$. The gradient of the error with respect to weights is computed as:

$$\frac{\partial E}{\partial \hat{w}_j} = -\sum_{k=1}^{N} e_j(k) \frac{\partial \hat{y}_j(k)}{\partial \hat{w}_j}$$

The weight update direction is determined by the conjugate gradient method to ensure efficient convergence.

## 2.3 MTN Identification Scheme

This paper proposes an MTN-based identification scheme for MIMO nonlinear time-delay systems. First, MTN serves as the identification model. Second, the WE-CG algorithm acts as the learning algorithm, where the WE algorithm simplifies the MTN structure to reduce computational complexity and meet real-time requirements. The detailed procedure is presented in Algorithm 1.

**Algorithm 1: MTN-Based Identification for MIMO Nonlinear Time-Delay Systems**

**Step 1:** Determine the MTN identification model structure and initial weights.
**Step 2:** Train the MTN identification model using the WE-CG algorithm according to the error criterion function (7).
**Step 3:** Update the regularization coefficient $\gamma$ using equation (14).
**Step 4:** Repeat Steps 2–4 until the error meets the specified accuracy and training stops.

# 3 Experimental Results and Analysis

This section validates the accuracy and real-time performance of the proposed scheme through both a numerical simulation example and a continuous stirred tank reactor (CSTR) process case study.

## 3.1 Numerical Example

Consider the following MIMO nonlinear time-delay system:

$$\begin{aligned}
y_1(k+1) =& 0.3y_1(k) + 0.7y_1(k-1) + u_1(k-d) + 0.2u_1(k-d-1) \\
& + \sin(0.5y_2(k-1)) \\
y_2(k+1) =& 0.5y_2(k) + 0.5y_2(k-1) + u_2(k-d) + 0.2u_2(k-d-1) \\
& + \sin(0.5y_1(k-2))
\end{aligned}$$

The input signals are unit step functions. The MTN identification model adopts a 6-28-2 structure (6 input nodes, expanded to 2nd order). The WE-CG algorithm parameters are set as: iteration number = 200, reference weight = 0.1, initial regularization coefficient = 0, filter coefficient $\kappa = 0.995$, regularization coefficient increment $\Delta\gamma = 4 \times 10^{-5}$, and redundant weight elimination threshold = 0.05. The model is trained using 200 sampling instants and tested using 100 sampling instants.

[Figure 2: see original paper] and [Figure 3: see original paper] illustrate the identification process, where (a) shows identification results, (b) presents learning error curves, (c) displays regularization coefficient variation, and (d) depicts the network pruning process. Although the regularization coefficient exhibits non-smooth variation, the smooth learning curves of the WE algorithm indicate that MTN weight changes do not cause severe training error fluctuations, with minimal impact on final pruning results.

The simulation results show that for the first subsystem, the training MSEs for WE-CG-MTN and CG-MTN schemes are 0.0318 and 0.0285, respectively, while test MSEs are 0.0365 and 0.0364. For the second subsystem, training MSEs are 0.0172 for both schemes, and test MSEs are 0.0521 and 0.0520, respectively. Although WE algorithm introduction slightly affects learning error, the pruned networks retain only 11 and 19 intermediate nodes for the two subsystems, substantially reducing computation and enhancing real-time performance. The complexity analysis is presented in .

**Table 2** Computation complexity analysis

| Prediction Model Construction Scheme | Structure | Addition | Multiplication | Total Floating-Point Operations |
|---|---|---|---|---|
| WE-CG-MTN | 6-11-2 | 23 | 34 | 57 |
| CG-MTN | 6-28-2 | 59 | 86 | 145 |

The results demonstrate that the pruned MTN model requires significantly fewer floating-point operations than the classical MTN scheme, indicating superior real-time performance. Notably, MTN is equivalent to a single neuron in neural networks, with substantially lower computational requirements. Previous studies have shown that traditional MTN identification models outperform BP neural network identification models in both accuracy and real-time performance [19]. This paper's comparison reveals that the proposed scheme achieves comparable accuracy (training and test MSEs in the same order of magnitude) while requiring less computation than traditional MTN methods.

## 3.2 CSTR Process Case Study

The CSTR is an anaerobic treatment process involving complete mixing of crude oil and microorganisms [31–34], as shown in [Figure 4: see original paper]. The CSTR unit is a sealed tank with stirring equipment, representing a typical nonlinear chemical process with time-delay characteristics.

The CSTR model is given by [35]:

$$\begin{aligned}\dot{x}_1(t) =& \frac{1}{D}[x_{1f} - x_1(t)] - a x_1(t) \exp\left(-\frac{\phi}{x_2(t)}\right)\\ \dot{x}_2(t) =& \frac{1}{D}[x_{2f} - x_2(t)] + B a x_1(t) \exp\left(-\frac{\phi}{x_2(t)}\right) - C[x_2(t) - u(t)]\end{aligned}$$

Discretizing with sampling time $T_s = 0.1$s yields:

$$\begin{aligned}x_1(k+1) =& x_1(k) + \frac{T_s}{D}[x_{1f} - x_1(k)] - aT_s x_1(k) \exp\left(-\frac{\phi}{x_2(k)}\right)\\ x_2(k+1) =& x_2(k) + \frac{T_s}{D}[x_{2f} - x_2(k)] + BaT_s x_1(k) \exp\left(-\frac{\phi}{x_2(k)}\right)\\ &- CT_s[x_2(k) - u(k)]\end{aligned}$$

where $a = 0.008$, $D = 10$, $\phi = 6$, $B = 0.5$, $C = 6$, and $d = T_s$. The outputs are $y_1(k) = x_1(k)$ and $y_2(k) = x_2(k)$.

The input is a unit step signal. The MTN identification model uses a 6-28-2 structure expanded to 2nd order, trained with the WE-CG algorithm (iteration number = 1000, reference weight = 0.1, initial $\gamma = 0$, $\kappa = 0.995$, $\Delta\gamma = 4 \times 10^{-5}$, threshold = 0.05). Training uses 1000 sampling instants, with 200 instants for testing.

[Figure 5: see original paper] and [Figure 6: see original paper] show the identification process. For the first subsystem, WE-CG-MTN achieves training MSE = 0.0024 and test MSE = 0.0052. For the second subsystem, training MSE = 0.0065 and test MSE = 0.0108. After pruning, the networks retain only 2 and 7 intermediate nodes, respectively, dramatically reducing computational load and improving real-time performance.

---

## 4 Conclusion

This paper proposes an MTN-based identification scheme for MIMO nonlinear time-delay systems. The key contributions are: (a) utilizing MTN's strong approximation capability as the identification model; (b) employing WE-CG as the learning algorithm, where WE effectively simplifies the MTN structure to reduce complexity and enhance real-time performance.

Experimental results demonstrate that the proposed scheme accurately identifies MIMO nonlinear time-delay systems while achieving a more compact structure and lower computational complexity compared to traditional MTN methods, thereby offering better real-time performance and facilitating engineering implementation.

Future research directions include: (a) investigating other simplified identification schemes; (b) studying control problems for MIMO nonlinear time-delay systems; and (c) applying the proposed scheme to other engineering domains.

## References

[1] Liu X, Yang X Q, Zhu P B, et al. Robust identification of nonlinear time-delay system in state-space form [J]. *Journal of the Franklin Institute*, 2019, 356(16): 9953-9971.

[2] Hou Mingdong, Wang Yinsong. An adaptive integral sliding mode control for a class of nonlinear large-lag systems [J]. *Control Theory & Applications*, 2019, 36(7): 1182-1188.

[3] Zhu Kun, Yu Chengpu, Wan Yiming. Recursive least squares identification with variable-direction forgetting via oblique projection decomposition [J]. *IEEE/CAA Journal of Automatica Sinica*, 2022, 9(3): 805-814.

[4] Ahmed S. Step response-based identification of fractional order time delay models [J]. *Circuits Syst Signal Process*, 2020, 39: 3858-3874.

[5] Chen Ruwen, Zhan Shishi. Identification based on GNAR model for vibration system [J]. *Application Research of Computers*, 2016, 33(10): 3021-3025.

[6] Lu Lu, Yu Yi, Yang Xiaomin, et al. Time delay Chebyshev functional link artificial neural network [J]. *Neurocomputing*, 2019, 329: 153-164.

[7] de Paula N C G, Marques F D. Multi-variable Volterra kernels identification using time-delay neural networks: application to unsteady aerodynamic loading [J]. *Nonlinear Dynamics*, 2019, 97: 767-780.

[8] Ren Xuemei, Rad A B. Identification of nonlinear systems with unknown time delay based on time-delay neural networks [J]. *IEEE Transactions on Neural Networks*, 2007, 18(5): 1536-1541.

[9] Xu Jun, Tao Qinghua, Li Zhen, et al. Efficient hinging hyperplanes neural network and its application in nonlinear system identification [J]. *Automatica*, 2020, 116: 108906.

[10] Wu Xiaolong, Han Honggui, Liu Zheng, et al. Data-knowledge-based fuzzy neural network for nonlinear system identification [J]. *IEEE Transactions on Fuzzy Systems*, 2020, 28(9): 2209-2221.

[11] Rummelhart D E, Hinton G E, Williams R. J. Learning representations by back propagating errors [J]. *Nature*, 1986, 323: 533-536.

[12] Batselier K, Chen Zhongming, Wong N. A Tensor Network Kalman filter with an application in recursive MIMO Volterra system identification [J]. *Automatica*, 2017, 84: 17-25.

[13] Chee E, Wang Xiaonan. Generalized system identification for nonlinear MPC of highly nonlinear MIMO systems [J]. *IFAC-PapersOnLine*, 2021, 54(3): 366-371.

[14] Gray W S, Venkatesh G S, Espinosa L A D. Nonlinear system identification for multivariable control via discrete-time Chen-Fliess series [J]. *Automatica*, 2020, 119: 109085.

[15] Yan Hongsen. MTN (multi-dimensional Taylor network) optimal control [EB/OL]. https://automation.seu.edu.cn/yhs/list.htm, 2021-07-06.

[16] Yan Hongsen. Equivology and multi-dimensional Taylor network [EB/OL]. https://automation.seu.edu.cn/yhs/list.htm, 2021-07-06.

[17] Zhou Bo, Yan Hongsen. Financial time series forecasting based on wavelet and multi-dimensional Taylor network dynamics model [J]. *Systems Engineering-Theory & Practice*, 2013, 33(10): 2654-2662.

[18] Lin Yi, Yan Hongsen, Zhou Bo. Nonlinear time series prediction method based on multi-dimensional Taylor network and its applications [J]. *Control and Decision*, 2014, 29(5): 795-801.

[19] Li Chenlong, Yan Hongsen. Identification of nonlinear time-delay system using multi-dimensional Taylor network model [A]. In: *The 8th International Conference on Manipulation, Manufacturing and Measurement on the Nanoscale (IEEE 3M-NANO 2018)* [C]. Hangzhou, China, 2018: 87-90.

[20] Li Chenlong, Yan Hongsen. d-step-ahead predictive model based on multi-dimensional Taylor network [J]. *Control and Decision*, 2021, 36(2): 345-354.

[21] Li Chenlong, Yan Hongsen, Zhang Jiaojun. Multi-dimensional Taylor network adaptive predictive control for SISO nonlinear systems with input time-delay [J]. *Transactions of the Institute of Measurement and Control*, 2021, DOI: 10.1177/01423312211040294.

[22] Zhang Jiaojun, Yan Hongsen. MTN optimal control of MIMO non-affine nonlinear time-varying discrete systems for tracking only by output feedback [J]. *Journal of the Franklin Institute*, 2019, 356(8): 4304-4334.

[23] Li Chenlong, Yuan Changshun, Ma Xiaoshaung, et al. Integrated fault detection for nonlinear process monitoring based on Multi-dimensional Taylor Network [J]. *Assembly Automation*, 2022, https://doi.org/10.1108/AA-06-2021-0076.

[24] Han Yuqun, Yan Hongsen. Adaptive multi-dimensional Taylor network tracking control for SISO uncertain stochastic nonlinear systems [J]. *IET Control Theory and Applications*, 2018, 12(8): 1107-1115.

[25] Klambauer G. *Mathematical analysis* [M]. New York: Marcel Dekker INC, 1975: 236-237.

[26] Ding Feng. Complexity, convergence and computational efficiency for system identification algorithms [J]. *Control and Decision*, 2016, 31(10): 1729-1741.

[27] Golub G H, Van Loan C F. *Matrix computations* (Vol. 3) [M]. Baltimore: JHU Press, 2012.

[28] Wei Haikun. *Theory and method of neural network structure design* [M]. Beijing: National Defence Industry Press, 2005.

[29] Wei Haikun, Ding Weiming, Song Wenzhong, et al. Dynamic method for designing RBF neural networks [J]. *Control Theory and Applications*, 2002, 19(5): 673-680.

[30] Weigend A S, Rumelhart D E, Huberman B A. Generalization by weight-elimination with application to forecasting. In: *Advances in Neural Information Processing Systems* [C]. San Mateo, 1991, 3: 857-882.

[31] Zerari N, Chemachema, M. Robust adaptive neural network prescribed performance control for uncertain CSTR system with input nonlinearities and external disturbance [J]. *Neural Computing Applications*, 2020, 32: 11127-11138.

[32] Wu Qun, Du Wenli, Nagy Z. Steady-state target calculation integrating economic optimization for constrained model predictive control [J]. *Computers & Chemical Engineering*, 2021, 145: 107145.

[33] Yang Xiong, Wei Qinglai. Adaptive critic designs for optimal event-driven control of a CSTR system [J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(1): 484-493.

[34] Du Jing Jing, Johansen T A. A gap metric based weighting method for multimodel predictive control of MIMO nonlinear systems [J]. *Journal of Process Control*, 2014, 24: 1346-1357.

[35] Tan Yonghong, Cauwenberghe A V. Neural-network-based d-step-ahead predictors for nonlinear systems with time delay [J]. *Engineering Applications of Artificial Intelligence*, 1999, 12: 21-35.

*Note: Figure translations are in progress. See original paper for figures.*