



中國人民大學  
RENMIN UNIVERSITY OF CHINA

# 复杂批次系统智能学习控制

## 5. Constructive Wavelet Neural Networks



沈 楠

分布式人工智能实验室  
中国人民大学数学学院  
中国人民大学应用数学研究中心  
金融计算与数字工程教育部工程研究中心

# OUTLINE



**研究背景**

Background



**自构建小波**

Self-Constructive WNN



**自适应构建**

Adaptive Construction



**展望与拓展**

Perspective & Extension

# 人工智能的基石：神经网络

AI的能力边界正在不断扩展

AlphaGo   大语言模型   自动驾驶   蛋白质预测   .....

理论基石：神经网络架构

空间局部性

**CNN**

时间依赖性

**RNN**

全局交互性

**Transformer**

残差叠加性

**ResNet**

.....

神经网络的架构（深度、宽度、连接方式、基函数）及其规模，从根本上限制或赋予了网络逼近复杂函数空间的能力。

# »» 万能逼近定理

## 万能逼近定理

### Universal Approximation Theorem

一个单隐层前馈网络，若隐层使用连续、有界、非常数的激活函数，则只要隐层神经元数量足够多，就能以任意精度逼近定义在紧集上的任意给定连续函数。

## 实际应用

非构造性：逼近网络需要预先给定

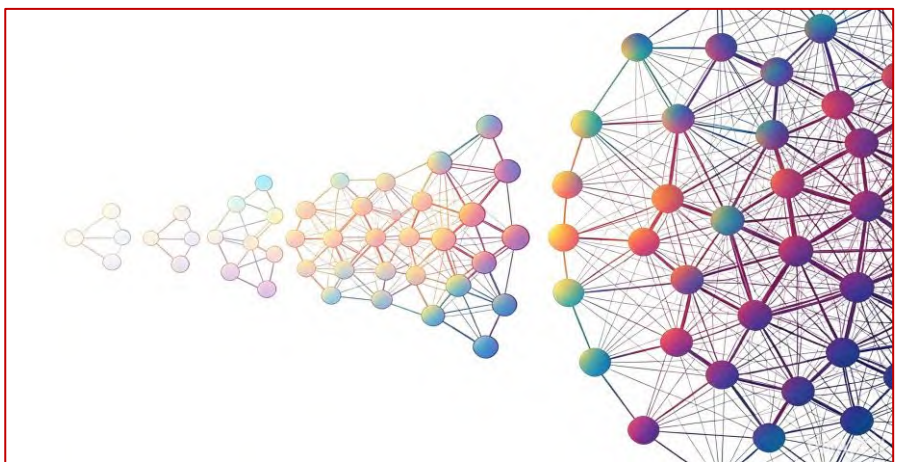
规模未知性：何为足够多？

存在  
最优网络

万能逼近定理是一张难以兑现的‘空白支票’——它承诺了任意精度的可能，却未附带任何关于网络如何构造、规模如何确定的‘操作手册’。

# 构造性学习

固定架构困境 从小规模开始，动态地、增量式地扩展 → 构建性学习



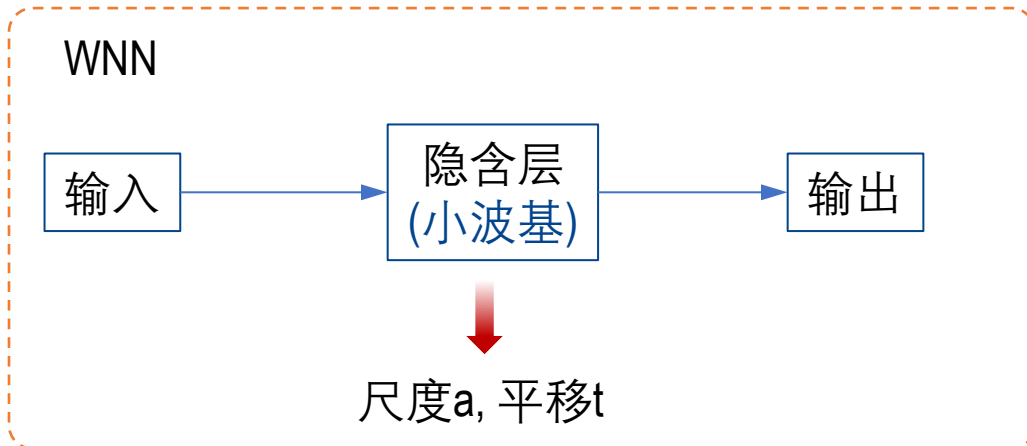
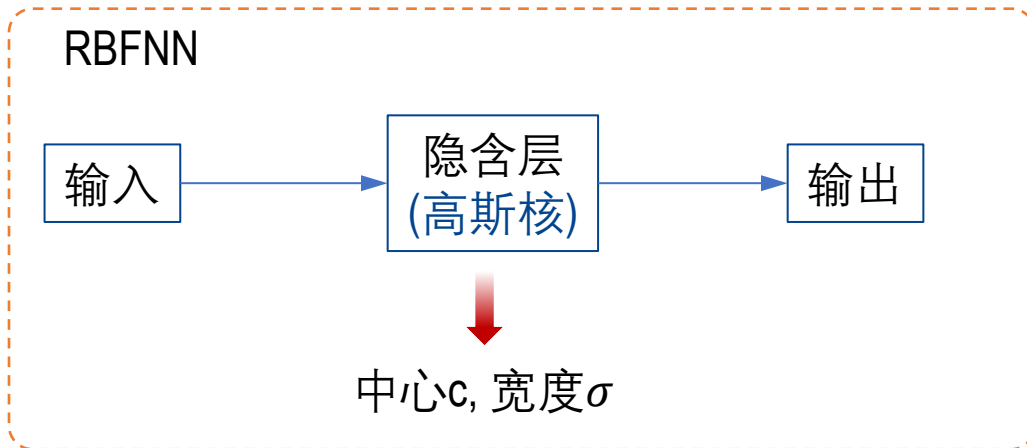
## ☑ 目标需求

- 增加基函数时，已训练好的参数不应被破坏
- 新加入的基函数应与旧基函数解耦，互不影响
- 每次增加后不要重新训练整个网络

小波神经网络

新增基函数时，不能破坏已学到的知识

# 为什么是小波神经网络



结构相同，但基函数的改变带来了逼近能力的质变

## 小波神经网络

### 多分辨率分析(MRA)

- ✓ 由粗到精的层级逼近
- ✓ 优先捕捉全局宏观趋势

### 时频域局部化分析

- ✓ 捕捉局部突变和扰动
- ✓ 结构稀疏性与计算高效性

### 正交小波的正交性

- ✓ 独立解耦  
与其他小波互不影响
- ✓ 网络结构变化不影响已训练好的参数

### 处理非平稳信号和复杂非线性问题

- 时间序列分析与预测
- 图像与信号处理
- 工业监控与故障诊断
- 复杂系统建模与控制

## Nonlinear Adaptive Wavelet Control Using Constructive Wavelet Networks

Jian-Xin Xu, Senior Member, IEEE, and Ying Tan, Member, IEEE

**Abstract**—In this paper, an adaptive wavelet-network-based control approach is proposed for highly nonlinear uncertain dynamical systems. Wavelet network, as a kind of universal approximator, has two novel properties—orthonormality and multiresolution. The orthonormal property ensures that adding a new resolution (new wavelets) does not affect the existing wavelet network that may have been well tuned. In the sequel, the online adjustment of the structure of the nonlinear adaptive wavelet controller (AWC) can be done in a constructive manner by gradually increasing the network resolution. The multiresolution property, on the other hand, assures a guaranteed improvement of the approximation precision when a new resolution is added. In real life problems we are unable to know the adequate size of a network, either a neural network (NN) or a wavelet network, to produce the required approximation precision. By virtue of the novel wavelet network properties, a coarse or very simple structure can be selected first. If the system fails to converge after the elapse of a dwell time, a new wavelet resolution is considered to be necessary and added directly. In this manner, the AWC can be easily constructed and tuned from the coarse to finer levels until the performance requirement is satisfied. The trial and error way of selecting the network structure, which may lead to either an inadequate or a highly redundant structure, can be avoided. In this paper, the proposed adaptive wavelet network is applied first to a class of nonlinear dynamical systems with a partially known model and an affine-in-input structure. Then, the adaptive wavelet network is applied to a class of nonlinear nonaffine dynamical systems.

**Index Terms**—Adaptive control, multiresolution, online structural tuning, orthonormality, wavelet network.

### I. INTRODUCTION

MANY control systems in practice have either partially unknown or completely unknown nonlinearities. Neural network (NN) is found to be an effective tool to deal with the control problems of such nonlinear uncertain systems [1]. In the early stage, NN is used to approximate unknown static nonlinearities. Using optimization technique (gradient method), the parameters (weights) of NN are tuned offline to minimize the approximation error. A limitation of this approach is its offline nature. In real time control, it would be hard to continuously tune NN so as to improve the system performance. Recently, adaptive neural control was introduced into the control design for nonlinear systems [2]–[10]. The parameters of NN are tuned

online to approximate the unknown nonlinear model with an appropriately chosen adaptive mechanism. Via the Lyapunov’s direct method, it shows that adaptive neural control can guarantee the stability of closed-loop systems.

It is shown in [3] that the tracking error produced by adaptive neural control schemes is a direct result of the function approximation error. Therefore, an accurate function approximation will be the decisive factor for achieving the desired control performance. In the design of an NN, the precision of the function approximation depends on the selection of the NN structure. It is proven that an NN is able to approximate any nonlinear function defined on a compact set to a prespecified accuracy [11]. However, it is still an open problem in choosing a suitable NN structure to meet a desired precision specification. Often, a trial and error method is employed in the structure selection for adaptive neural control. It is highly desirable, if possible, to tune the NN structure in an online fashion, so as to improve the precision of the function approximation. Unfortunately, it is a difficult task for such an online structure adjustment of NN. Any change of an NN structure will cause the change of the optimal parameter set that achieves the universal approximation. Note that an adaptive neural controller continuously tunes its adjustable parameters to approach the unknown optimal parameter set. Obviously, the adaptive neural controller will have to retune its parameters from the moment the unknown system parameter set varies. This paper will propose an adaptive wavelet controller (AWC), which can partially overcome the previous difficulty associated with adaptive neural control.

The AWC design is based on the promising function approximation capability of wavelet networks—the multiresolution approximation [12], [13]. A wavelet network is a one-layer network consisting of orthonormal “father wavelets” and “mother wavelets” and widely used as a kind of building blocks for the approximation of unknown functions. In [14]–[17], the properties of wavelet networks were discussed and applied to nonlinear system identification and nonparametric estimation. It was proven that a wavelet network is a universal approximator that can approximate any functions with arbitrary precision by a linear combination of father wavelets and mother wavelets [12]. Comparing with NN, a wavelet network has its own properties. First, wavelet networks are one-layer networks and linear in the parameter space. This feature makes the wavelet training much easier compared to multilayer NNs. Second, unlike NNs with radial basis functions, wavelet networks use orthonormal basis in the sequence of subspace of  $L^2(\mathcal{R}^d)$ . In particular, the coefficients needed for reconstructing any function are *fixed* and *unique*. Therefore, it is more suitable for online learning as the coefficient of each wavelet can be tuned independently (regardless of the updating process of other wavelets). The adaptive control approach can be easily applied to wavelet

Manuscript received July 2, 2004; revised October 28, 2005. J.-X. Xu is with Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576, Singapore (e-mail: elecexjx@nus.edu.sg).

Y. Tan is with Department of Electrical and Electronics Engineering, The University of Melbourne, Parkville, VIC 3010, Australia (e-mail: y.tan@ee.unimelb.edu.au).

Digital Object Identifier 10.1109/TNN.2006.886759

## 核心思想

利用正交小波基的正交性 + 多分辨率分析，实现构造性增长

## 增长策略

从粗分辨率开始 → 若精度不达标 → 逐层增加分辨率

## 核心贡献

首次将正交小波与构造性学习相结合，证明可行性

新增分辨率不影响已调好的网络（正交性的价值）

避免了传统“试错法”选网络结构的困境

**Xu & Tan 的工作证明了正交小波网络用于构造性学习的可行性，是CWNN领域的奠基性工作。**

# »» 存在的局限

## 维数灾难与计算爆炸

- ✘ 盲目增加节点以捕获高维非线性特征
- ✘ 节点激增导致指数级计算负担

## 闲置时频局部化优势

- ✘ 仅依赖时域残差，闲置时频局部化优势
- ✘ 缺乏主导频率的能量评估

## 繁杂的超参数依赖

- ✘ 节点增删与学习率高度依赖人为设定
- ✘ 调参耗时且强烈削弱了网络的自适应初衷

## 在线控制的闭环失稳风险

- ✘ 只增不减导致冗余积压与权重剧烈波动
- ✘ 节点激增导致训练不充分

**现有CWNN的增长策略是“按分辨率增长”  
即在空间中逐层加密，而非根据被逼近函数的实际频率能量分布进行“精准构建”**

# 频率-能量驱动自构建小波网络框架

## 核心创新

引入频率估计器，分析未知函数的空间频率能量分布

扩展到非线性动态系统，利用在线测量数据进行频域分析

## 关键机制

根据能量分布选择初始小波基的位置和尺度，而非从粗层开始逐层加满

网络可动态地增减小波基，配有Lyapunov稳定性分析

## 预期效果

用更少的小波基达到更高的逼近精度，计算成本显著降低

适用于实时控制、轨迹跟踪等动态场景



Dunsheng Huang, Dong Shen, Lei Lu, Ying Tan. Optimizing Basis Function Selection in Constructive Wavelet Neural Networks and Its Applications. **IEEE Transactions on Artificial Intelligence**, vol. 7, no. 5, pp. 2648-2663, 2026.



Dunsheng Huang, Dong Shen, Lei Lu, Ying Tan. Adaptive Frequency-Based Constructive Wavelet Neural Network for Nonlinear Dynamic Systems. **IEEE Transactions on Neural Networks and Learning Systems**, vol. 37, no. 6, pp. 2929-2943, 2026.

# OUTLINE



**研究背景**

Background



**自构建小波**

Self-Constructive WNN



**自适应构建**

Adaptive Construction



**展望与拓展**

Perspective & Extension

# 自构建小波神经网络

考虑如下非线性静态映射:

$$y = f(x),$$

$$x \in \mathbf{R}^d$$

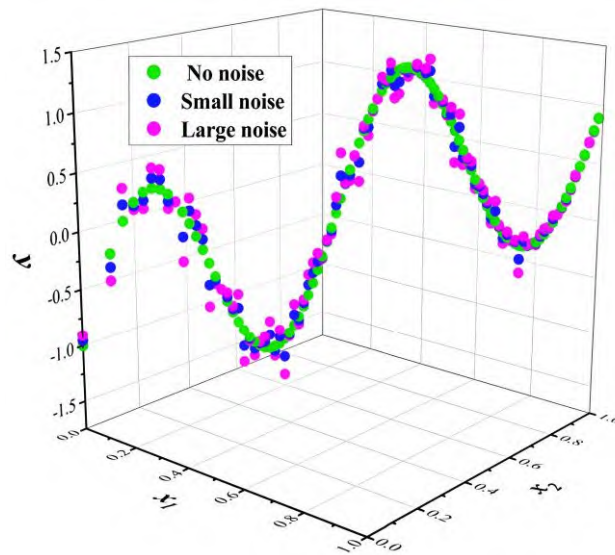
映射输入

$$d$$

输入  $x$  的维度

$$f(\cdot): \mathbf{R}^d \rightarrow \mathbf{R} \text{ 局部Lipschitz (LLC)}$$

假设1. 假设  $f(\cdot)$  满足  $f \in L^2(\mathbf{R}^d)$



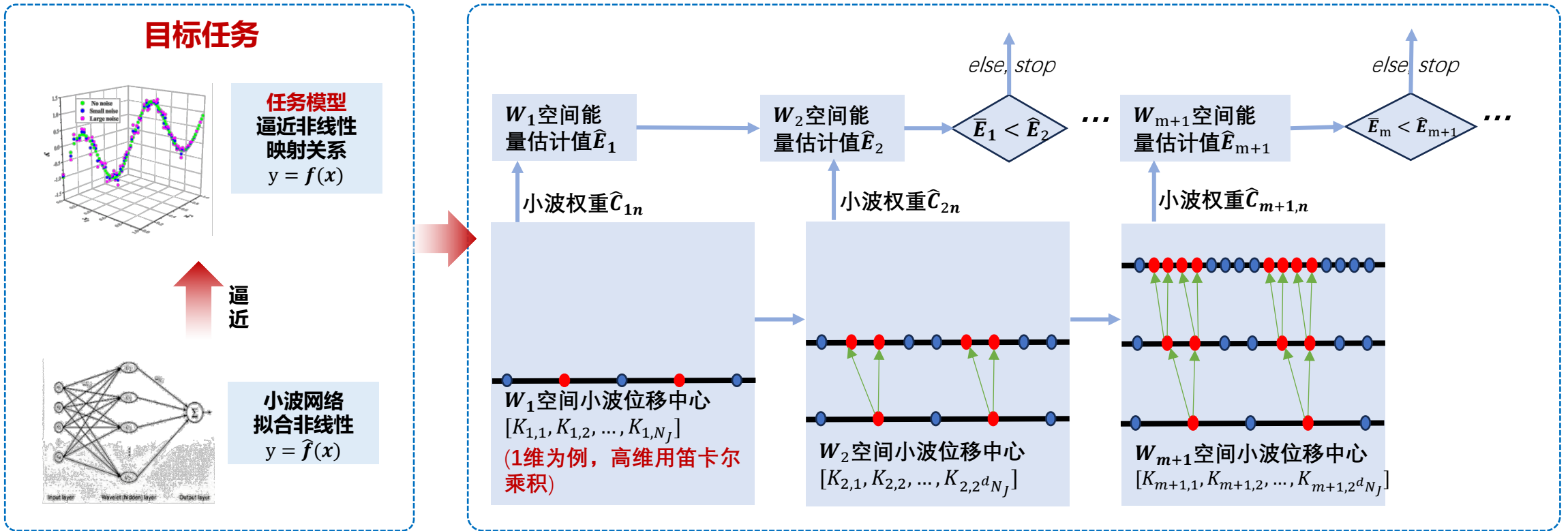
## 核心思想

针对当前所面临的挑战, 我们提出一种基于“频率—能量”驱动自构建小波神经网络 (FBCWNN), 通过能量分布自适应初始化结构, 并以高能量优先策略进行增量扩展, 实现逼近性能与结构复杂度的平衡, 具体由两个串联模块组成:

**Step 1.** 通过频率分析估计非线性映射在不同子空间(频率空间)的能量分布 —— 初始小波基频率估计 (确定初始小波空间)

**Step 2.** 通过能量分析估计同一小波子空间中不同小波基能量大小 —— 高能量基优先原则 (确定小波基扩增办法)

# Step 1: 初始小波基频率估计



## 初始频率估计思路

函数  $f(\cdot)$  可以被投影到小波子空间上，并由每个小波子空间的基进行近似。通过从测量数据中识别能够捕获  $f(\cdot)$  大部分能量的子空间，我们可以有效地逼近未知的非线性映射，随后可以确定初始频率。

## 逐分辨率估计能量

采用轻量网络架构由低频（低分辨率）空间到高频空间逐级训练并估计得到能量分布，从而定位主能量空间，以此作为网络的初始基底。

# Step 1: 初始小波基频率估计

## 能量计算方法

设  $\hat{E}_m$  表示映射  $f(\cdot)$  投影到小波子空间  $W_m$  上的能量, 则  $\hat{E}_m$  可以由下式计算得:

$$\hat{E}_m = \sum_n \hat{C}_{mn}^2 \|\psi_{mn}(x)\|^2$$

其中,  $\hat{C}_{mn}$  表示小波基函数  $\psi_{mn}(x)$  的系数 (或权重)。

Note: 子空间  $W_m$  中的能量由该子空间内各小波基对应系数的平方加权累积得到, 因此可以用来衡量该分辨率层对原始映射  $f(\cdot)$  的贡献大小

## EMA平滑化能量估计

为了抑制局部波动, 稳定地识别主要量子空间并确定截断位置, 利用指数移动平均 (EMA) 平滑化方法, 引入能量评价指标  $\bar{E}_m$ :

$$\bar{E}_m = \frac{\alpha \bar{E}_{m-1} + (1 - \alpha) \hat{E}_m}{1 - \alpha^m}, m = 2, 3, \dots$$

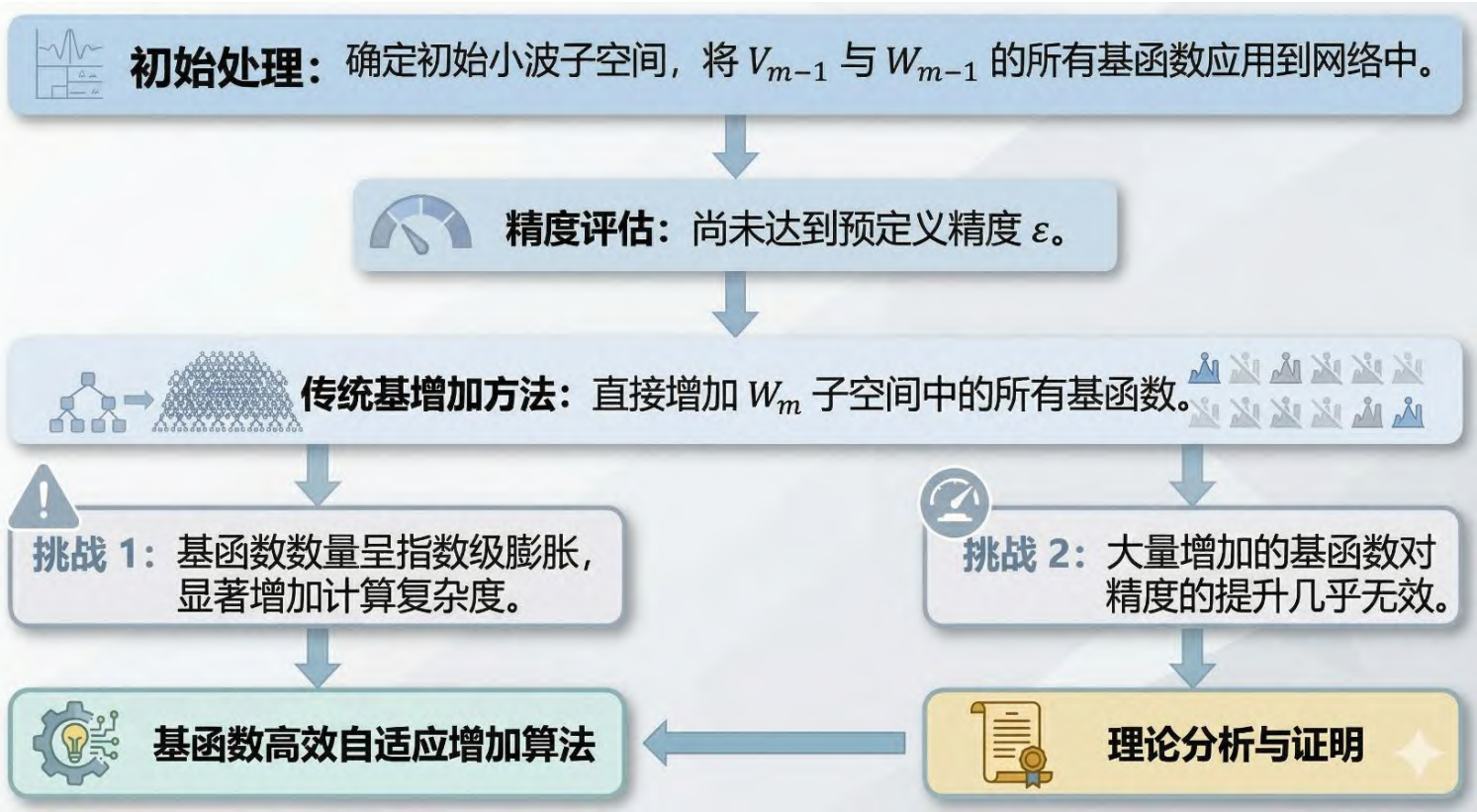
$\alpha$  是与具体问题相关的参数, 用于反映历史子空间能量的重要性。一般取  $0 < \alpha < 1$ 。

## 初始小波基频率判定准则

如果满足  $\hat{E}_{m+1} \leq \bar{E}_m$  则说明子空间  $W_{m+1}$  中的能量已经足够小, 此时可以终止 EMA 的递推计算, 并将  $2^{m-1}$  确定为初始频率。

Note: 直观上看, 若  $\hat{E}_{m+1} \leq \bar{E}_m$ , 则意味着后续子空间所包含的能量相较于当前平滑能量指标已经明显减弱, 因此  $W_{m+1}$  更可能处于低能量区域。相比于直接使用  $\hat{E}_m$ , 平滑后的评价指标  $\bar{E}_m$  更适合作为识别主要量子空间的判据

# Step 2: 高能量基优先原则



## 当前挑战

当模型未达到预设精度时, **基函数扩展**若采用一次性增加分辨率子空间全部基函数, 容易导致**基函数数量爆炸**, 引入大量未产生有效贡献却**消耗计算资源**的冗余基函数; 若采用逐个添加的方式, 使**收敛速度过于缓慢**

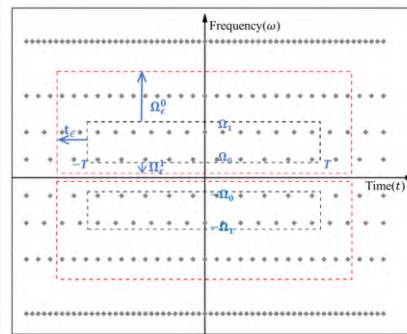
# Step 2: 高能量基优先原则

## 能量集中特性

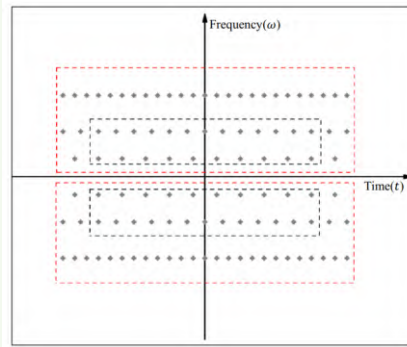
$$\begin{aligned} (Q_T f)(x) &= \chi_{[-T, T]}(x) f(x) \\ (P_\Omega \hat{f})(\omega) &= \chi_{[-\Omega, \Omega]}(\omega) \hat{f}(\omega) \end{aligned}$$

非线性映射的能量  $f(x)$  集中在一定范围的时频域区间内(图中的黑色虚线框), 小波的平移-尺度中心距离核心区越远, 其小波系数  $C_{mn}$  越小, 对逼近的贡献越弱。

## 距离衰减效应



由距离衰减效应, 对于给定的逼近精度  $\varepsilon$ , 只需从黑色虚线框的主能量小波基开始, 逐渐扩展边界 (红色虚线框) 内的小波基, 最大化提高逼近效率。



*Theorem 1:* Let the orthogonal wavelet functions

$$\psi_{mn}(x) = 2^{\frac{1}{2}md} \psi(2^m x - n), \quad (m \in \mathbf{Z}, n \in \mathbf{Z}^d)$$

defined in (3) form an orthonormal wavelet frame. Assume that Assumptions 1, 2 and 3 hold for nonlinear mapping  $f(\cdot) \in L^2(\mathbf{R}^d)$ . Let  $\varepsilon > 0$  be given. Then there exists a  $B_\varepsilon$  given by

$$\begin{aligned} B_\varepsilon &= B_\varepsilon(\mathbf{T}, \Omega_0, \Omega_1) \\ &= \{(m, n) | m_1 < m < m_0, |n|_s \leq 2^m \mathbf{T} + t_\varepsilon, \\ &\quad \mathbf{T}, t_\varepsilon \in \mathbf{Z}^d\}, \end{aligned}$$

where  $m_0, m_1, t_\varepsilon$  will be specified in the proof. For any  $\varepsilon > 0$ , the following inequality holds:

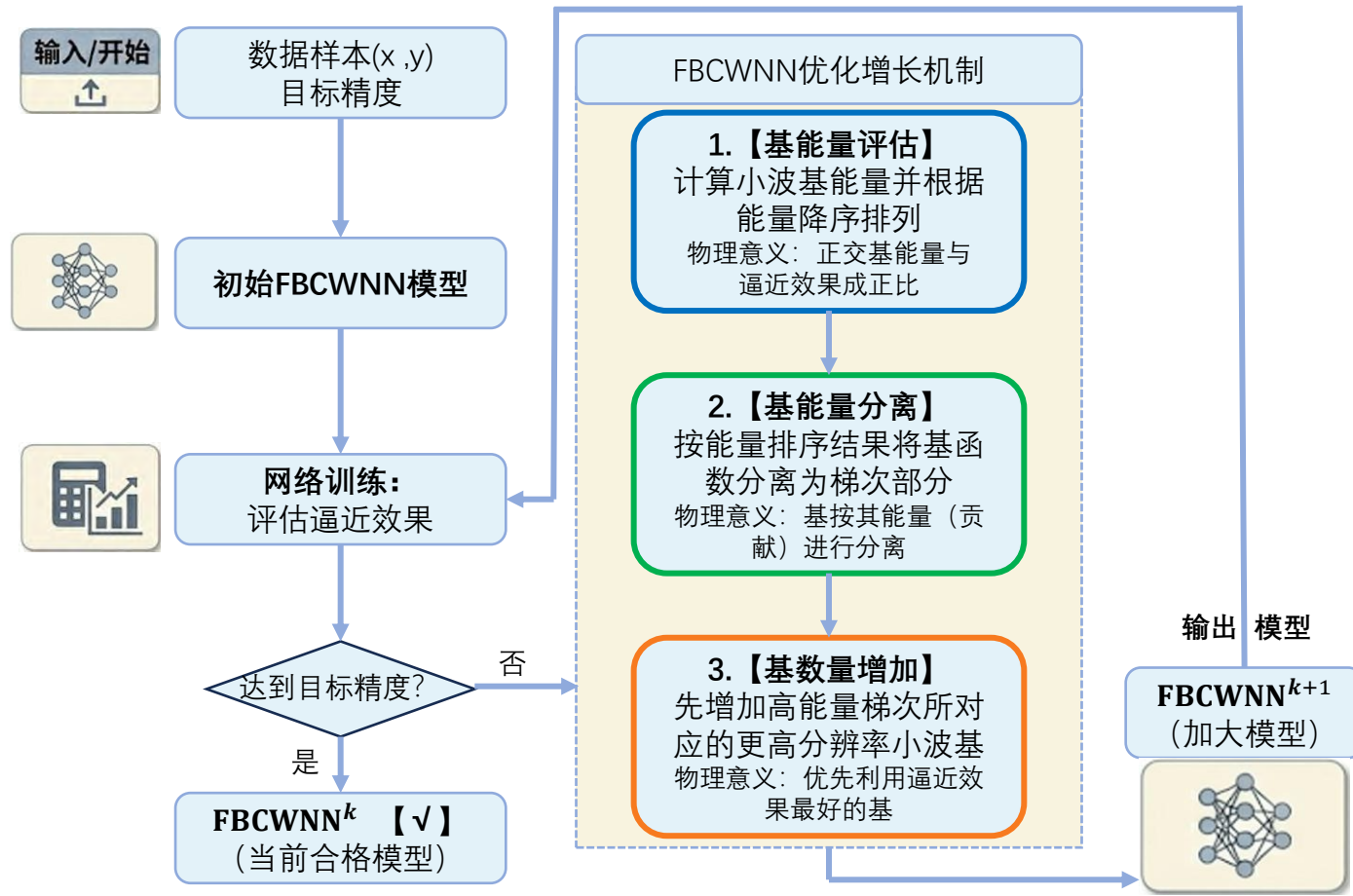
$$\begin{aligned} &\left\| f - \sum_{(m,n) \in B_\varepsilon} \langle \psi_{mn}, f \rangle \psi_{mn} \right\| \\ &\leq \| (1 - P_{\Omega_1} + P_{\Omega_0}) \hat{f} \| + \| (1 - Q_T) f \| + 2(2\pi)^{\frac{d}{2}} \varepsilon \| f \|. \end{aligned}$$

小波时频域的范围边界

给定边界下的逼近误差上界

# 自构建小波神经网络框架

## FBCWNN新增基函数算法框架



## 实验一（静态抗噪性验证）

在不同噪声水平下，所提出的网络能够精准拟合二维非线性映射，充分展示了其出色的噪声鲁棒性。

## 实验二（增量分批学习）

针对分批次流入的数据，网络表现出平滑的增量逼近能力，能够在不断变化的数据流中保持高效的学习。

## 实验三（时变自适应能力）

在非线性特征发生突变时，网络结构能够通过增加基函数实现在线自适应调整，并重新实现对新的特征的学习。

## 实验四（真实的高维工业数据应用）

在针对9维输入的刀具磨损实际训练任务中，所提方法在低于传统网络节点数的条件下，达到了目标训练精度。

# 实验一：离线静态映射逼近

## 实验数据

离线数据集  $\{x_i, y_i\}$  由非线性函数  $y = f(x)$  生成, 其中  $F(x_1)$  为噪声.

$$y = 0.5 + x_1 + x_2 + \sin(2\pi(x_1 + x_2)) + F(x_1)$$

TABLE II  
ACCURACY WITH TESTING DATASETS

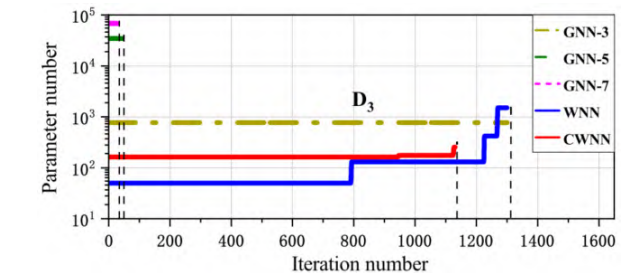
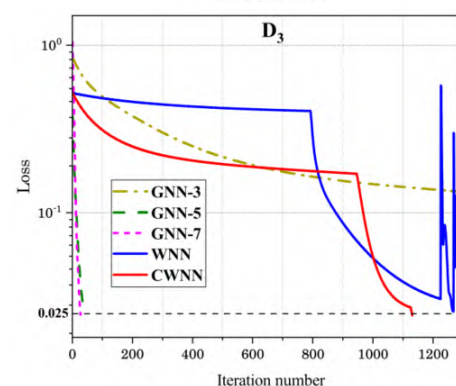
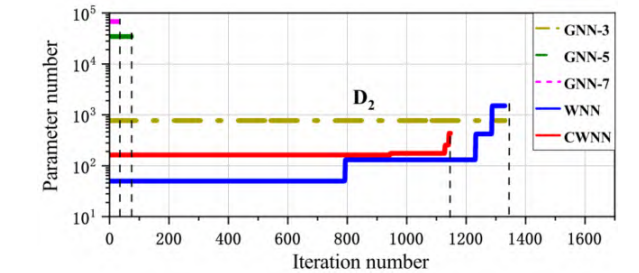
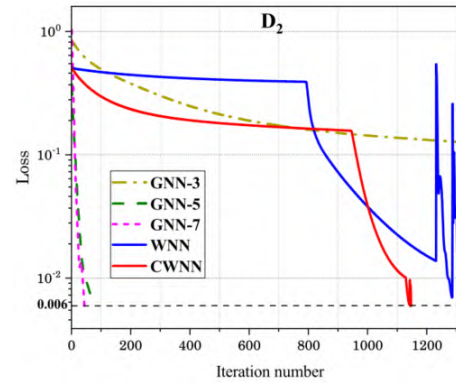
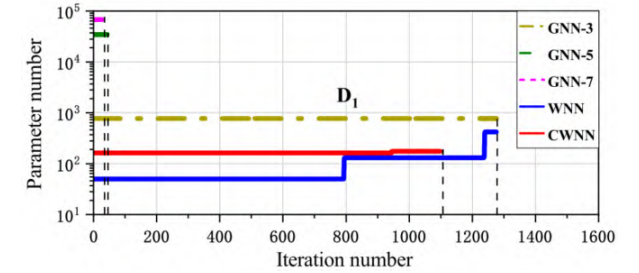
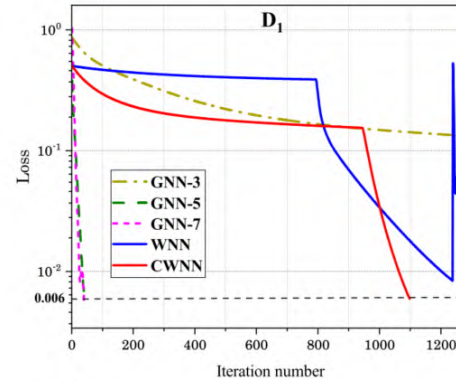
| Method         | GNN-3  | GNN-5  | GNN-7  | WNN           | CWNN          |
|----------------|--------|--------|--------|---------------|---------------|
| D <sub>1</sub> | 0.1417 | 0.1090 | 0.1696 | <b>0.0059</b> | 0.0068        |
| D <sub>2</sub> | 0.1607 | 0.0145 | 0.1652 | <b>0.0047</b> | 0.0058        |
| D <sub>3</sub> | 0.1929 | 0.1428 | 0.2264 | 0.0238        | <b>0.0154</b> |

**具有残差连接的全连接深度神经网络 (MLP系列)**: 训练收敛快, 但结构较浅的网络规模无法达到目标精度。在带有噪声的测试集上, MLP系列表现显著变差, 其中层数最深 (67,841个参数) 的网络出现了最严重的过拟合现象。

**标准 WNN**: 能够达到精度目标, 且对噪声具有鲁棒性。但参数消耗较大, 无噪声下需 420 个参数, 强噪声下飙升至 1,509 个参数。

**FBCWNN (图中的CWNN方法)**: 对噪声展现出极强的鲁棒性 (方差略低于标准 WNN)。同时参数量远低于其他方法, 在强噪声环境下仅需 254 个参数, 实现了相较于 WNN 超过 3.5 倍的参数缩减。

## 实验结果



# 实验二：分批增量学习

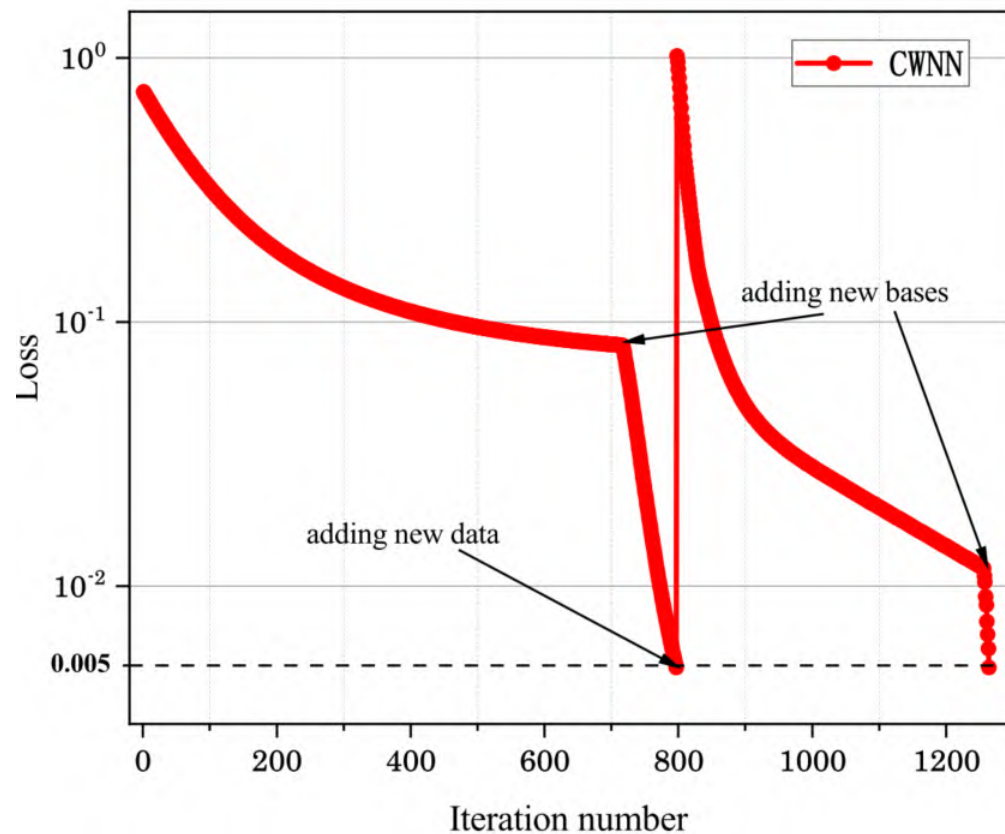
## 实验数据

数据集  $\{x_i, y_i\}$  由非线性函数  $y = f(x)$  生成，其中  $F(x_1)$  表示噪声函数。  
数据集  $\{x_i, y_i\}$  被划分为两个子数据集  $DS_1$  和  $DS_2$ 。网络首先在  $DS_1$  上进行训练，直到损失收敛至给定精度；随后将  $DS_2$  引入训练。通过 FBCWNN 框架新增基函数以重新达到目标精度。

$$y = 0.5 + x_1 + x_2 + \sin(2\pi(x_1 + x_2)) + F(x_1)$$

模型首先在数据集  $DS_1$  的区域训练至收敛。当突然加入全新区域  $DS_2$  的数据集时，误差短暂飙升，但 FBCWNN 机制自动触发，成功引入新的基函数，使网络迅速重新收敛至目标精度

## 实验结果



# 实验三：时变自适应能力

## 实验数据

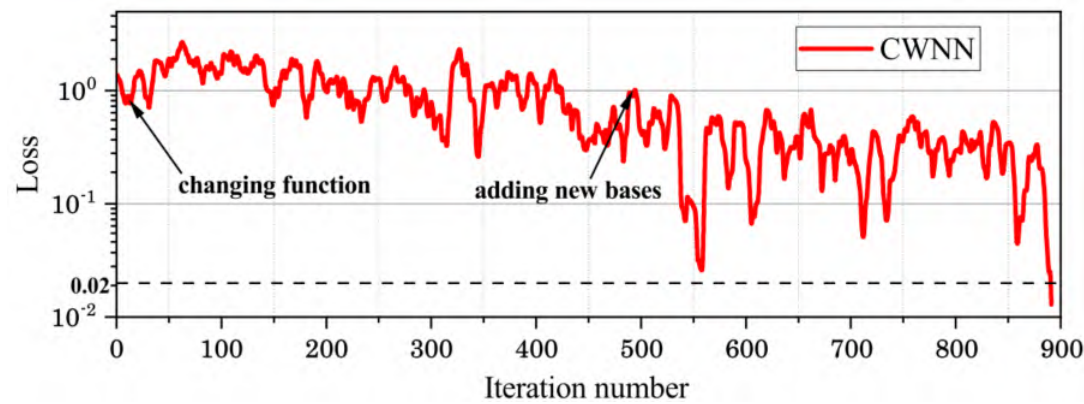
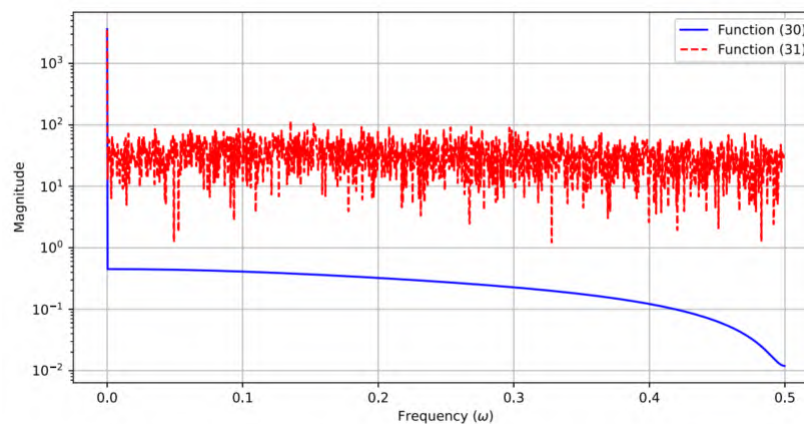
考虑非线性自回归模型(NAR), 非线性函数  $y = f(x)$  在训练过程中发生了突变, 由 (30) 变为 (31), 我们旨在通过增加新的基函数以捕获这个突变:

$$y_t = \arctan(\pi(y_{t-1}^2 + y_{t-2}^2))^{\frac{1}{2}} \quad (30)$$

$$y_t = \arctan(\pi(y_{t-1}^2 + y_{t-2}^2))^{\frac{1}{2}} + \cos(\pi(y_{t-1}^2 + y_{t-2}^2)) \quad (31)$$

当底层时间序列系统发生突变 (频率发生偏移) 时, 本研究的方法能够敏锐捕捉变化, 并自动增加新的小波基来适应新的映射关系, 展现出强大的在线时变适应力。

## 实验结果



# 实验四：真实高维工业数据应用

## 实验数据

该实验考虑9维输入的真实金属铣削机器数据集并进行训练和预测。  
非线性模型  $y = f(\cdot)$  满足下面关系：

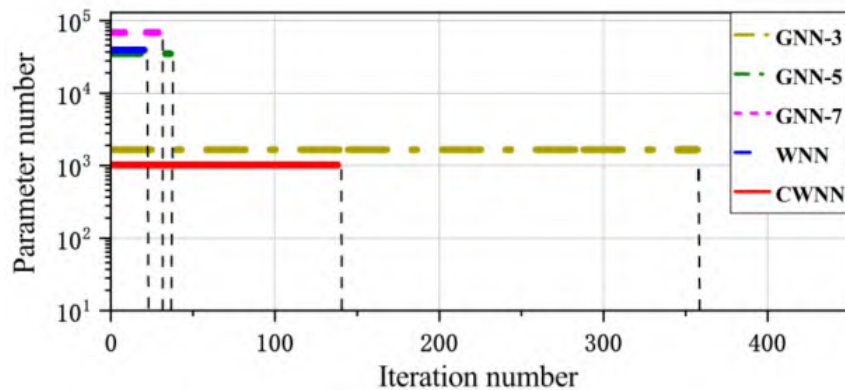
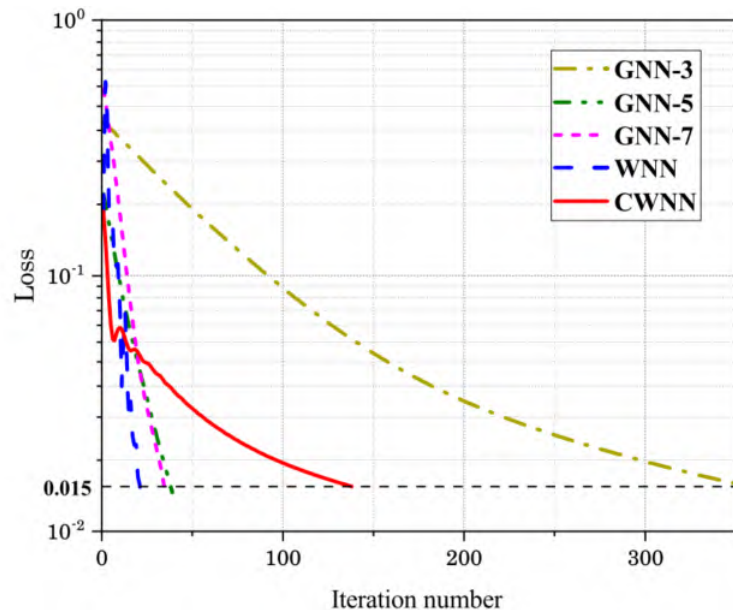
$$VB_t = f(VB_{t-1}, \mathbf{x}_t)$$

**具有残差连接的全连接深度神经网络 (MLP系列)**：均能达到预定精度，但代价是庞大的参数量（3层：1,665个；5层：35,201个；7层：68,737个参数），且7层网络在测试集上依然表现出过拟合迹象。

**标准 WNN**：虽然收敛所需的迭代次数最少（21个），但遭遇了严重的“维度灾难”，参数量暴增至 39,366 个。

**FBCWNN (图中的CWNN方法)**：展现出压倒性的效率优势，仅使用 1,510 个参数就达到了目标精度。参数量缩减至标准 WNN 的三十分之一以下，同时远低于其他MLP系列方法，完美平衡了“快速收敛”与“极低参数开销”。

## 实验结果



# OUTLINE



**研究背景**

Background



**自构建小波**

Self-Constructive WNN



**自适应构建**

Adaptive Construction



**展望与拓展**

Perspective & Extension

# 由静态非线性到动态系统

## 非线性动态系统

考虑如下包含未知非线性 $f(x)$ 的SISO动态系统:

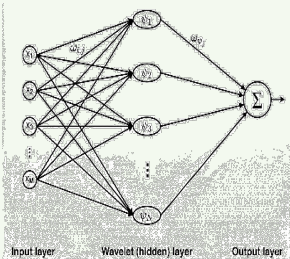
$$\begin{cases} \dot{x}_i = x_{i+1}, & i = 1, 2, \dots, d-1, \\ \dot{x}_d = f(x) + u, \\ y = x_1, \end{cases}$$

$x \in \mathbb{R}^d$ :  $x = [x_1, x_2, \dots, x_d]^T$  系统状态

$d$ :  $x$  的维度

$f(\cdot)$ :  $\mathbb{R}^d \rightarrow \mathbb{R}$ , 局部Lipschitz (LLC)且满足  $f \in L^2(\mathbb{R}^d)$

逼近系统未知非线性  
 $f(\cdot)$



自构建小波神经网络

$$y = \hat{f}(x)$$

## 关键挑战

面向复杂机械臂、机器人等非线性动态系统控制问题时, 想要将现有的自构建网络框架和传统自适应方法结合起来, 我们仍然面临着重要挑战:

- 现有的**FBCWNN框架**虽然解决了非线性映射能量分布问题 (确定初始分辨率) 和基函数扩展机理问题 (高能量准则), 但面对非线性系统控制时:
  - 没有离线数据、系统状态随时间剧烈演化**
  - 如果在控制闭环中**网络结构发生不当突变**, 会直接导致控制力矩发散, 系统彻底失稳!
  - 冗余节点累积与节点“只增不减”**
- 现有的**神经网络自适应控制框架**虽然解决了系统闭环稳定控制问题, 但当系统结构发生改变时, 很难保证原有的稳定性:
  - 网络结构的扩展, 原有训练好的参数可能受到影响**
  - 需要寻求合理的方法触发网络结构的改变并保证系统的稳定**

# 自适应构建型小波神经网络

## 核心思想

为了克服上述瓶颈问题，我们提出面向非线性系统控制的**自适应构建型小波神经网络 (AFBCWNN) 框架**

该框架包含一种以**频率-能量**为结构决策依据、在线控制过程中实现“**增—选—剪**”协同演化的自构建小波神经网络，通过**自适应地添加、筛选和裁剪**小波基函数，同时保证非线性动态系统的**逼近效率与闭环稳定**。

该框架由**三个核心机制**组成：

- 从静态映射扩展至非线性动态系统，**实现未知非线性函数  $f(x)$  在系统运行过程中在线逼近并实时补偿**；
- 基于基函数能量“**贡献度**”实施**新增—裁剪—冻结机制**——实现频率能量视角下对无用于空间的结构收缩；
- **严谨的 Lyapunov 稳定性保证**——基于 Lyapunov 函数的自适应参数更新律以及具有理论保证的基函数冻结与裁剪策略。

# 自适应控制与结构切换协同设计

## 设计控制律与自适应权重更新律

在结构扩展的同时，AFBCWNN 将小波网络输出直接嵌入控制律，并通过 Lyapunov 方法设计自适应参数更新律，使得参数学习、结构调整与系统稳定性三者形成统一闭环

$$u(t) = -\beta\delta - \hat{f}(x) + g(t, x)$$

$$\dot{\hat{v}}_{m_0} = -\phi_{m_0}(x)\delta,$$

$$\dot{\hat{w}}_m = -\psi_m(x)\delta, \quad m = m_0, m_0 + 1, \dots$$

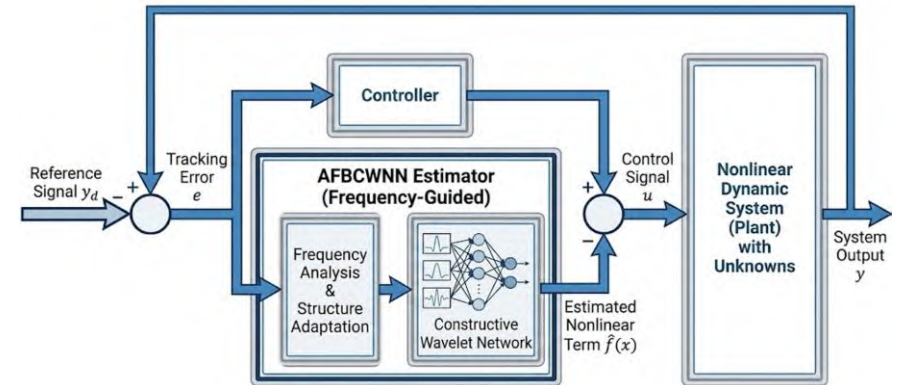
## 驻留时间内最大增广跟踪误差 (MATEDT)

结构变化并非连续发生，为避免频繁切换引发系统不稳定，引入驻留时间内最大增广跟踪误差 (MATEDT) 机制，对网络结构更新施加时间约束。

$$\bar{\delta}_i = \max_{t \in [(i-1)\Delta T, i\Delta T]} |\delta(t)|, \quad i = 1, 2, \dots$$

控制目标

$$\limsup_{t \rightarrow \infty} |\delta(t)| \leq \delta_{ac}$$



# 频率-能量引导的小波基扩展

## 频率-能量引导的小波基扩展

当系统误差指标  $\bar{\delta}(t)$  无法满足预设性能要求时，AFBCWNN 根据能量-频率机理新增更高频基函数：对所有小波基函数的能量进行计算并从高到低排序，并优先选取包含能量最高的子集；随后根据小波细分准则选择更高分辨率的小波空间中的基函数。

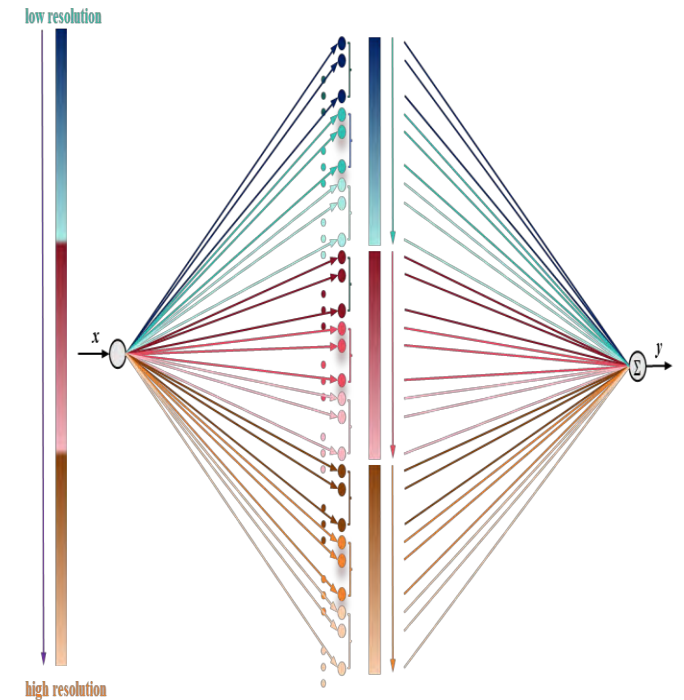
### 特点

小波函数的**正交性**——新增或裁剪某个基函数几乎不会影响其他节点

小波**时频域局部化定位性**——相同时域特征的小波基具有递进的分辨率（频率）

### 四个优势

- 新增基函数具有明确的物理与频谱意义
- 避免在低能量子空间中无效扩展
- 在同等逼近效率下降低网络增速
- 提供了一种基于交叉频谱分析和高能量基主导的研究思路



# 基于贡献度的裁剪与冻结机制

## 基于贡献度的裁剪与冻结机制

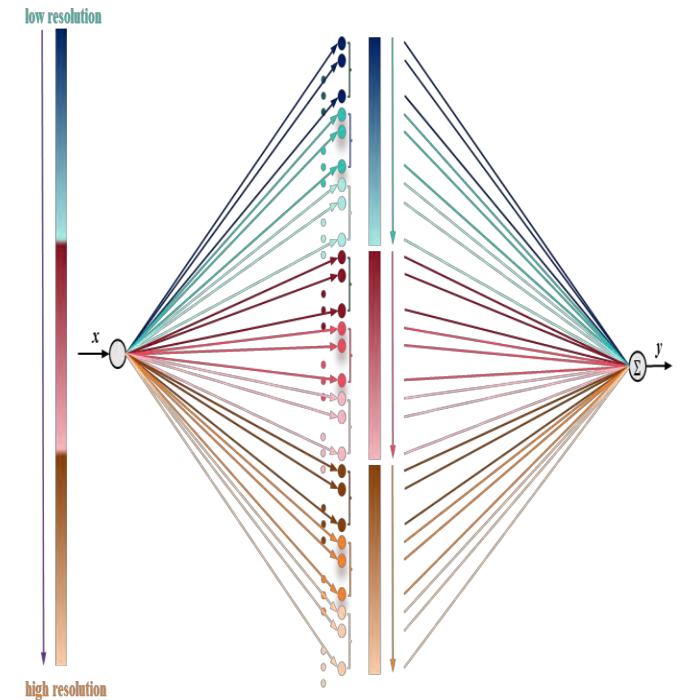
不同于FBCWNN等自构建网络“只增不减”的设计，AFBCWNN 在系统运行过程中持续评估各小波基函数对系统性能的实际贡献。当某些基函数在较长时间内权重几乎不变化或者对逼近的贡献度很低时，算法将对其进行

- **冻结**：停止参数更新，降低计算负担
- **裁剪**：将其从网络结构中移除

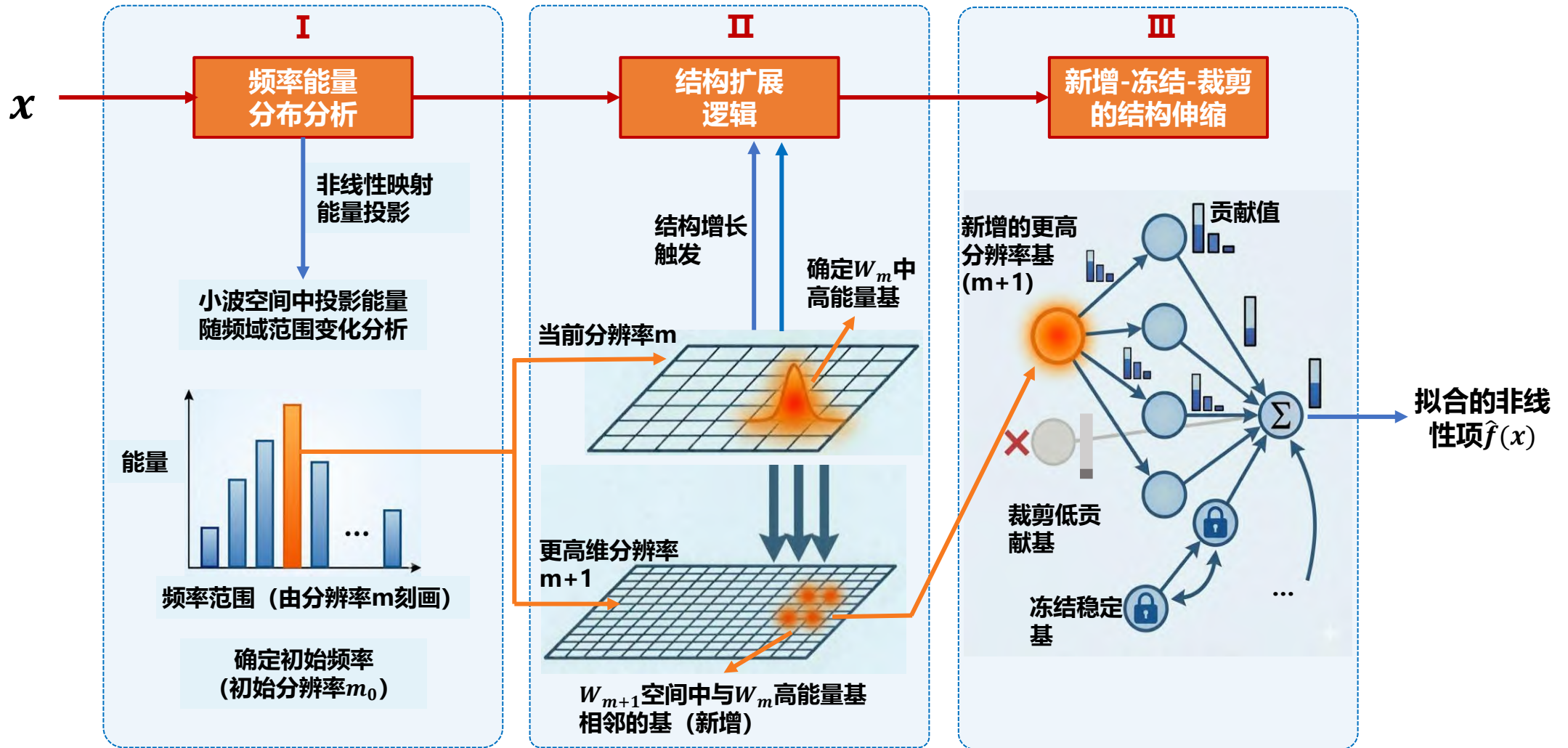
**作用**：在频率和能量视角下对长期无贡献的小波子空间进行的结构收缩  
**恢复机制**：当系统误差再次增大时，设计恢复机制维持网络的逼近能力

## 三个优势

- 网络结构始终保持紧凑与高效
- 计算复杂度随时间保持可控
- 在保证性能的同时，显著提升长期运行稳定性



# 框架结构



为了验证 AFBCWNN 在线环境下的控制效能与结构优越性，本研究将其应用于含未知动态的非线性系统轨迹跟踪任务中。实验设计既包括了自适应增益的关键参数影响分析，也涵盖了与非自适应方法、多种自适应神经网络（固定结构与动态结构）、以及前沿无模型强化学习方法的多维度基准对比测试。

仿真结果直观地展现了算法的三大核心优势：

**极简结构下的高精度轨迹跟踪：**在面临系统内部扰动与未知动态时，AFBCWNN 能够迅速实现在线收敛，将跟踪误差压制在极小的稳定边界内。与传统固定结构神经网络或常规 CWNN 相比，本方法在达到同等甚至更高跟踪精度的前提下，占用的网络节点数量呈数量级下降。

**网络拓扑的动态伸缩与抗膨胀：**仿真深刻揭示了裁剪与冻结机制的高效。在系统平稳运行期，算法自动清理无贡献节点，网络规模保持极度紧凑；当系统遭遇突发扰动时，网络迅速唤醒或扩展所需频率空间的节点以吸收扰动；扰动消退后，网络再次动态收缩。这种拓扑演化，有效改进了在线控制中的网络规模爆炸问题。

**结构突变下的控制力矩平滑与安全：**控制问题中担忧的“结构改变引发系统震荡”问题在仿真中被有效化解。得益于小波基的正交解耦特性以及基于 Lyapunov 的自适应律设计，即使在网络节点被瞬间裁剪或大量唤醒的时刻，系统输出的控制力矩依然保持平滑，未出现任何破坏闭环稳定性的危险。

# 实验一：算法与参数分析

## 实验数据

控制对象为一个包含未知复杂非线性函数  $f(\cdot)$  的二阶动态系统；控制目标为跟踪参考轨迹  $y_r(t) = 0.5 \sin^2(t)$ ；测试变量为四组自适应增益参数组合  $(\lambda, \beta) \in \{(10, 10), (10, 20), (20, 10), (20, 20)\}$ 。

$$\dot{x}_1(t) = x_2(t)$$

$$\dot{x}_2(t) = 8e^{-x_1^2} \sin\left(5\sqrt{x_1^2 + x_2^2}\right) + \sin(4x_2) + u(t)$$

$$y(t) = x_1(t)$$

在所有参数组合下，AFBCWNN 均能够驱动系统跟踪误差收敛至预设精度范围内，验证了算法的有效性与鲁棒性；增益  $\lambda, \beta$  具有双重影响，其中增大  $\beta$  可加快初始误差收敛速度、提升瞬态响应性能，并减少初始阶段所需新增的小波基数量，而适当减小  $\lambda$  有助于在结构裁剪前获得更优的稳态最大跟踪误差，因此适中的增益能在收敛速度与稳态精度之间取得平衡。

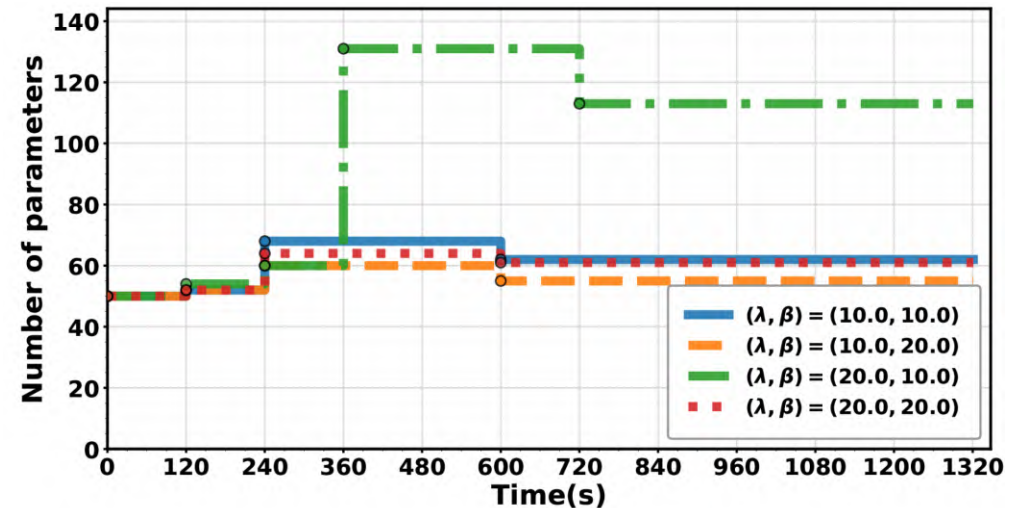
## 实验结果

TABLE I  
TRACKING PERFORMANCE  $\bar{\delta}_i$  WITH DIFFERENT  $(\lambda, \beta)$  FOR SYSTEM (22)

| $(\lambda, \beta) : (10.0, 10.0)$ |                  | $(\lambda, \beta) : (10.0, 20.0)$ |                  |
|-----------------------------------|------------------|-----------------------------------|------------------|
| Time $t(s)$                       | $\bar{\delta}_i$ | Time $t(s)$                       | $\bar{\delta}_i$ |
| $[0, \Delta T]$                   | 0.3706           | $[0, \Delta T]$                   | <b>0.1977</b>    |
| $[\Delta T, 2\Delta T]$           | 0.2033           | $[\Delta T, 2\Delta T]$           | <b>0.1145</b>    |
| $[2\Delta T, 3\Delta T]$          | 0.0362           | $[2\Delta T, 3\Delta T]$          | <b>0.0317</b>    |
| $[3\Delta T, 5\Delta T]$          | 0.0258           | $[3\Delta T, 5\Delta T]$          | <b>0.0231</b>    |
| $[5\Delta T, 7\Delta T]$          | 0.0251           | $[5\Delta T, 7\Delta T]$          | <b>0.0227</b>    |
| $[7\Delta T, 9\Delta T]$          | <b>0.0184</b>    | $[7\Delta T, 9\Delta T]$          | 0.0240           |
| $[9\Delta T, 11\Delta T]$         | 0.0204           | $[9\Delta T, 11\Delta T]$         | <b>0.0154</b>    |

TABLE II  
TRACKING ERROR  $\bar{\delta}_i$  WITH DIFFERENT  $(\lambda, \beta)$  FOR SYSTEM (22)

| $(\lambda, \beta) : (20.0, 10.0)$ |                  | $(\lambda, \beta) : (20.0, 20.0)$ |                  |
|-----------------------------------|------------------|-----------------------------------|------------------|
| Time $t(s)$                       | $\bar{\delta}_i$ | Time $t(s)$                       | $\bar{\delta}_i$ |
| $[0, \Delta T]$                   | 0.3572           | $[0, \Delta T]$                   | <b>0.1763</b>    |
| $[\Delta T, 2\Delta T]$           | 0.2185           | $[\Delta T, 2\Delta T]$           | <b>0.1231</b>    |
| $[2\Delta T, 3\Delta T]$          | 0.1461           | $[2\Delta T, 3\Delta T]$          | <b>0.0303</b>    |
| $[3\Delta T, 4\Delta T]$          | <b>0.0215</b>    | $[3\Delta T, 5\Delta T]$          | 0.0304           |
| $[4\Delta T, 6\Delta T]$          | 0.0263           | $[5\Delta T, 7\Delta T]$          | <b>0.0162</b>    |
| $[6\Delta T, 8\Delta T]$          | 0.0189           | $[7\Delta T, 9\Delta T]$          | <b>0.0173</b>    |
| $[8\Delta T, 11\Delta T]$         | 0.0303           | $[9\Delta T, 11\Delta T]$         | <b>0.0154</b>    |



# 实验二：多维基准对比

## 实验数据

控制对象为四维高阶非线性动态系统，控制目标为跟踪参考轨迹  $y_r(t) = 0.5\sin^2(t)$ ：

$$\begin{aligned}\dot{x}_i(t) &= x_{i+1}(t), \quad i = 1, 2, 3 \\ \dot{x}_4(t) &= -(x_1 + 2x_2 + 2x_3 + 2x_4) \\ &\quad + [1 - (x_1 + 2x_2 + x_3)^2](x_2 + 2x_3 + x_4) + u(t) \\ y(t) &= x_1(t)\end{aligned}$$

## 对比方法

- 基于模型的非自适应方法 FBCWNN (离线训练、固定权重)
- 基于模型的自适应方法：ACWNN (动态结构) 及 RBF-256、RBF-625 (固定结构)，在这个方法中我们进行了抗干扰测试，于  $t = 120s$  时引入外部扰动  $d(t) = 6\sin(4t)$ ，测试系统对外界扰动的鲁棒性
- 数据驱动的非模型方法 Transformer 与 SAC

# 实验二：多维基准对比

对比方法：基于模型的非自适应方法 FBCWNN（离线训练、固定权重）

## 实验结果

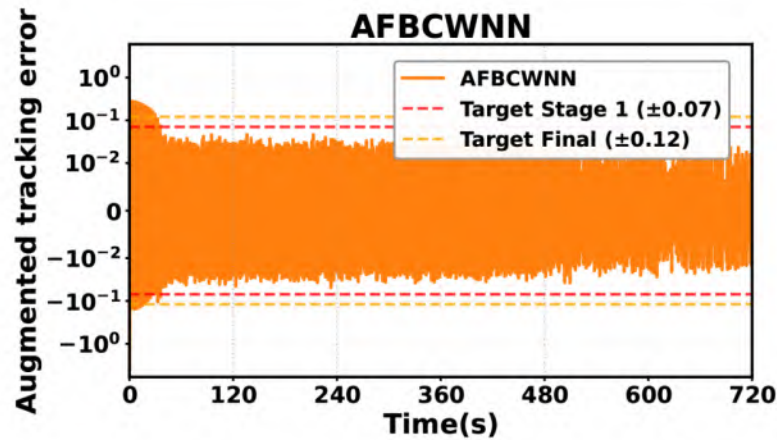
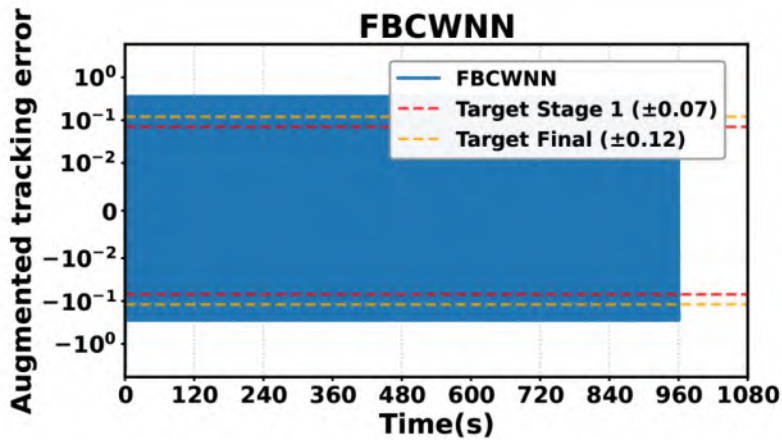


TABLE III  
TRACKING ERROR  $\bar{\delta}_i$  OF FBCWNN AND AFBCWNN FOR SYSTEM (23)

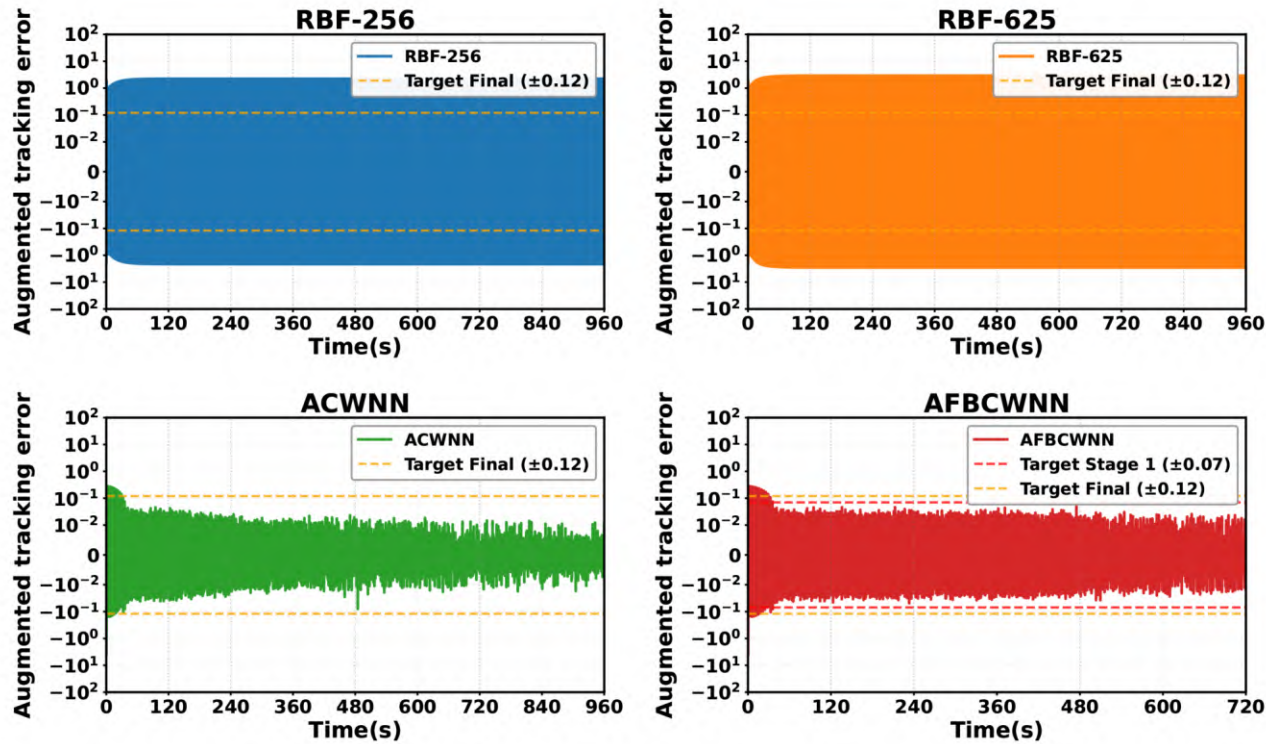
| FBCWNN                   |                  | AFBCWNN                  |                  |
|--------------------------|------------------|--------------------------|------------------|
| Time $t(s)$              | $\bar{\delta}_i$ | Time $t(s)$              | $\bar{\delta}_i$ |
| $[0, \Delta T]$          | 6.0000           | $[0, \Delta T]$          | <b>6.0000</b>    |
| $[\Delta T, 2\Delta T]$  | 0.3611           | $[\Delta T, 2\Delta T]$  | <b>0.0526</b>    |
| $[2\Delta T, 4\Delta T]$ | 0.3614           | $[2\Delta T, 4\Delta T]$ | <b>0.0542</b>    |
| $[4\Delta T, 6\Delta T]$ | 0.3615           | $[4\Delta T, 6\Delta T]$ | <b>0.0463</b>    |

非自适应方法FBCWNN 的稳态误差约为 0.36，而 AFBCWNN 依靠在线自适应能力在第二驻留阶段迅速收敛至 0.0526，体现了**在线自适应权重更新 + 结构调整**的必要性。

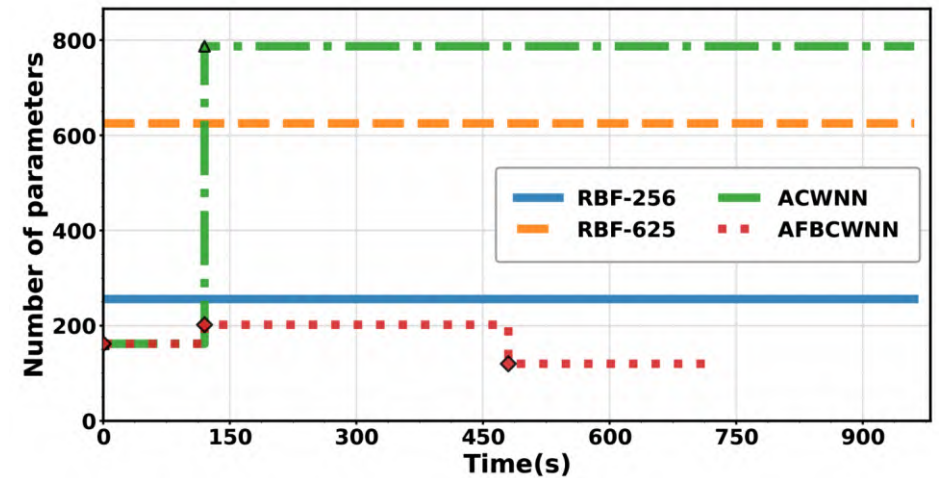
# 实验二：多维基准对比

对比方法：基于模型的自适应方法：ACWNN（动态结构）及 RBF-256、RBF-625（固定结构）

## 实验结果



无扰动情形

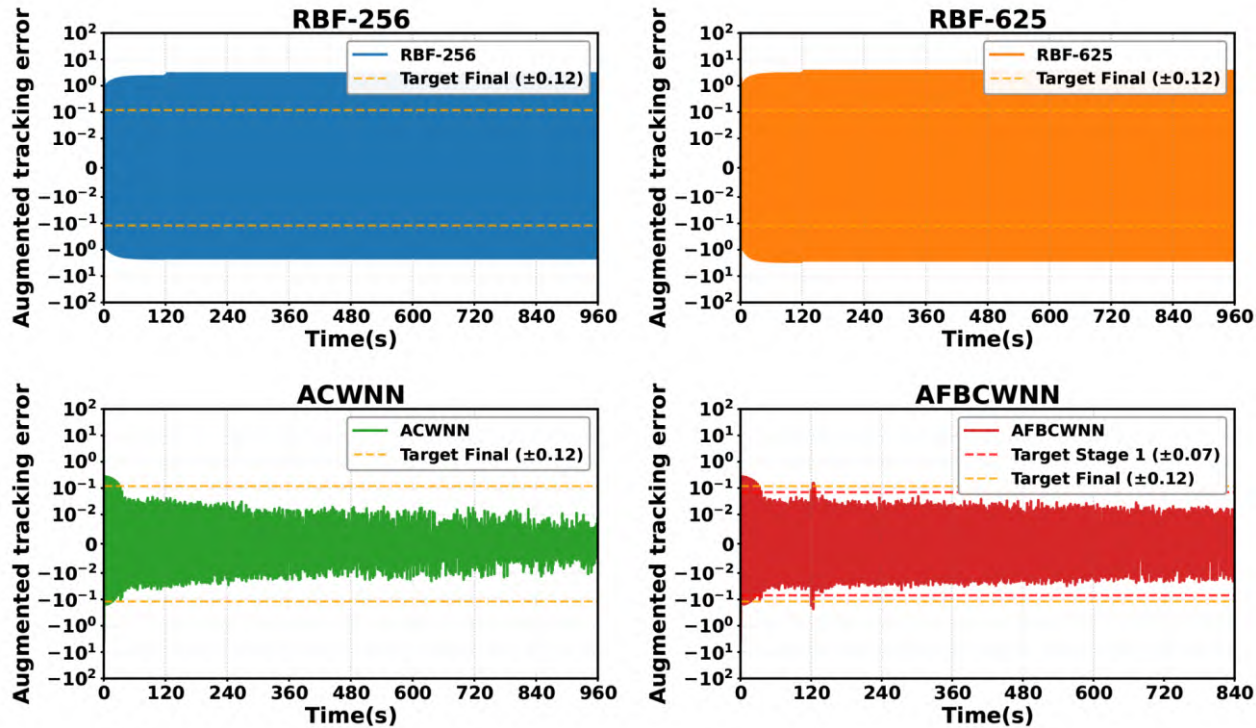


对于自适应框架下的神经网络，在**无扰动情形**下，RBF-256 与 RBF-625 的误差分别约为 2.3和 2.9，未达到预设精度  $\pm 0.12$ ，ACWNN 虽达到 0.0175但参数量约为 800，而 **AFBCWNN 仅用约 120-200 个参数即可达到相当精度**

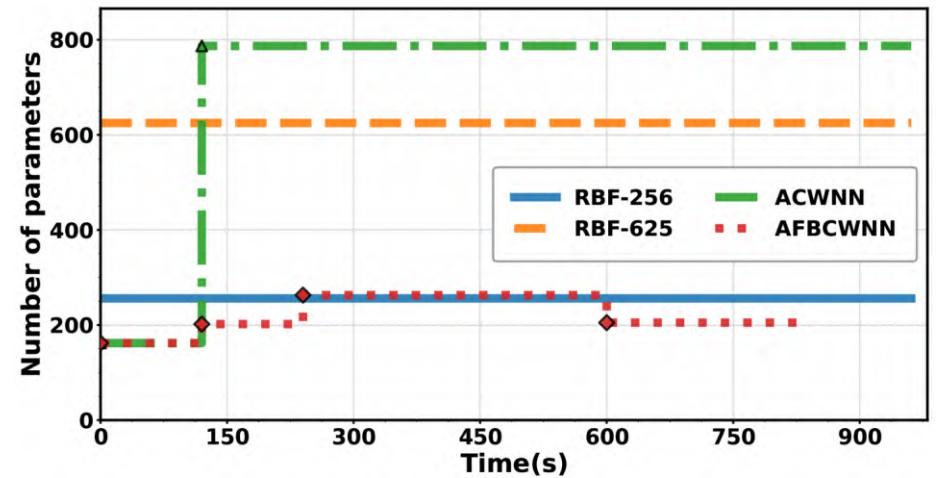
# 实验二：多维基准对比

对比方法：基于模型的自适应方法：ACWNN（动态结构）及 RBF-256、RBF-625（固定结构）

## 实验结果



有扰动情形



在扰动注入情形下，RBF 网络性能严重退化（误差超过 2.9 和 3.6），ACWNN 虽保持稳定但参数量仍约为 800，而 AFBCWNN 仅短暂波动后迅速恢复至 0.04，且参数量控制在 162-240，说明了其在高维系统下的高效与鲁棒性

# 实验二：多维基准对比

对比方法：基于模型的自适应方法：ACWNN（动态结构）及 RBF-256、RBF-625（固定结构）

## 实验结果

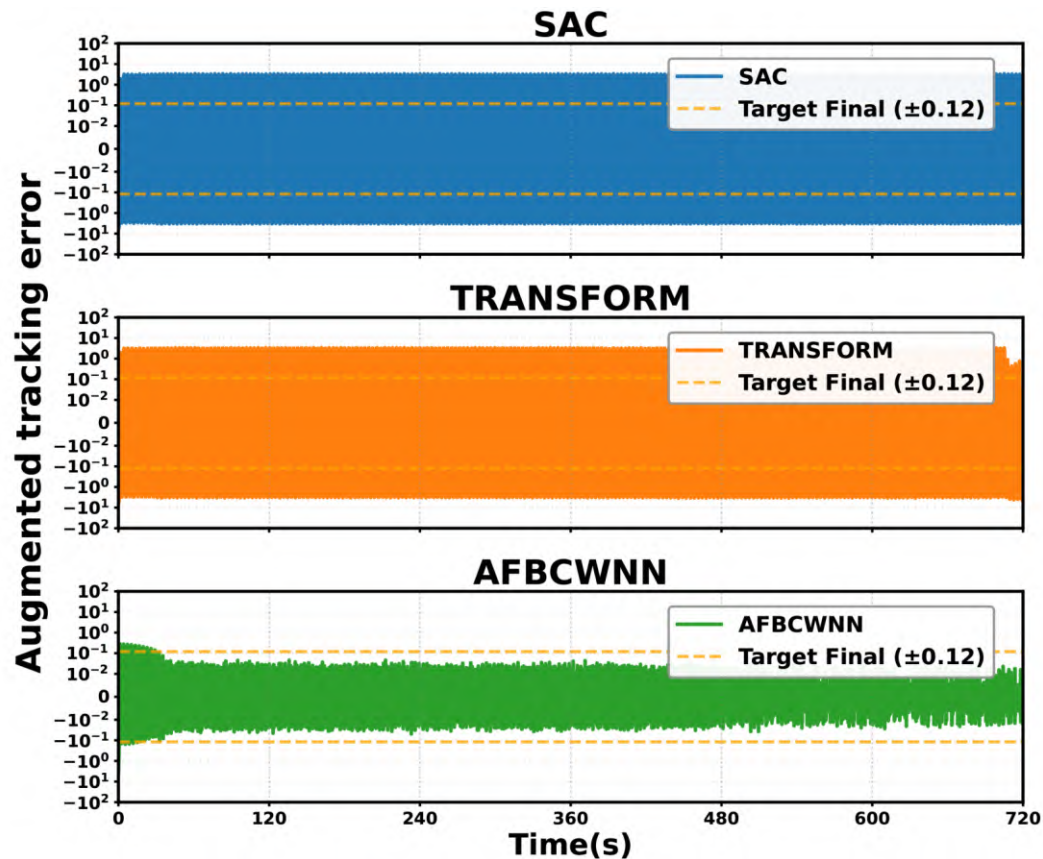


TABLE IV  
TRACKING ERROR  $\bar{\delta}_i$  UNDER MODEL-FREE METHODS FOR SYSTEM (23)

| Time $t(s)$              | SAC    | TRANSFORM | AFBCWNN |
|--------------------------|--------|-----------|---------|
| $[0, \Delta T]$          | 6.0000 | 6.0000    | 6.0000  |
| $[\Delta T, 2\Delta T]$  | 3.3984 | 3.4345    | 0.0526  |
| $[2\Delta T, 4\Delta T]$ | 3.3969 | 3.4358    | 0.0542  |
| $[4\Delta T, 6\Delta T]$ | 3.3972 | 3.4436    | 0.0463  |

相较无模型方法，在相同有限数据与安全约束下，Transformer 与 SAC 的误差约为 3 ~ 4，而 **AFBCWNN 仅用约 162 个参数即可实现快速稳定收敛，误差约为 0.0526**，表现出显著优势。该现象表明 Transformer 类模型和 SAC 方法在在线系统及低数据量场景下的性能受到一定限制，这主要源于其对大规模数据和相对稳定数据分布的依赖。

# OUTLINE



**研究背景**

Background



**自构建小波**

Self-Constructive WNN



**自适应构建**

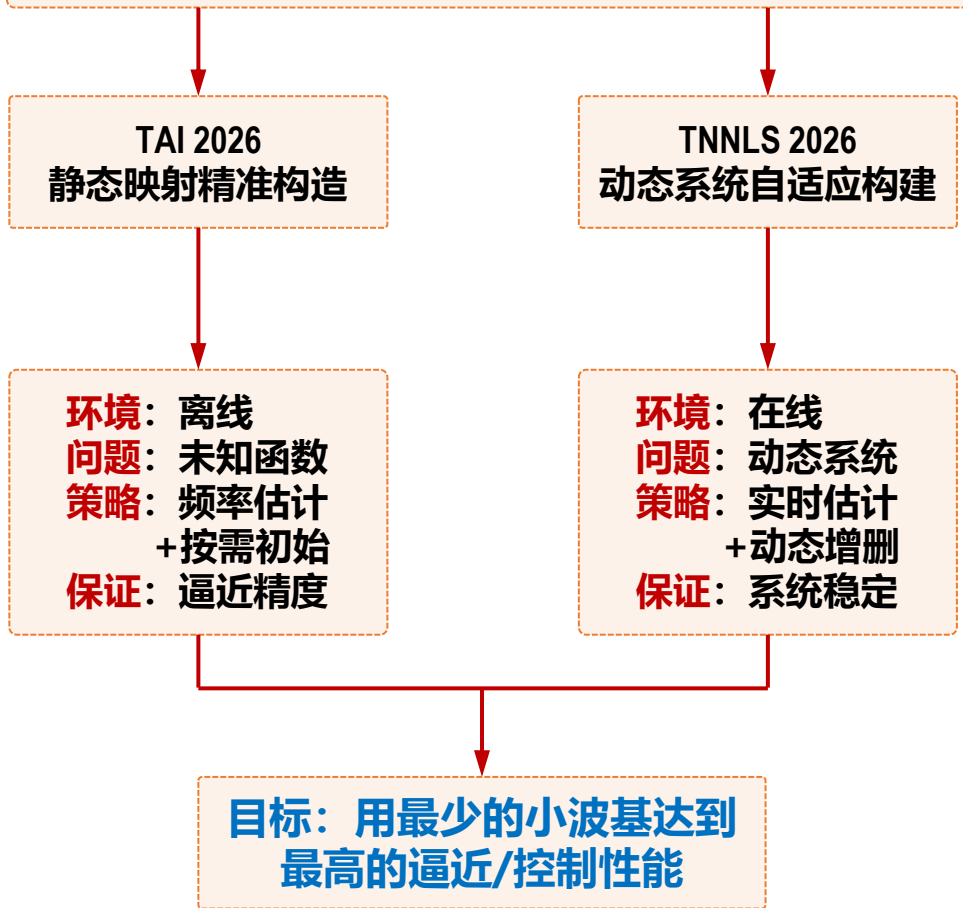
Adaptive Construction



**展望与拓展**

Perspective & Extension

## 核心思想 频率能量分析 → 指导小波基精准构造



| 对比维度 | 论文一 (TAI 2026)                    | 论文二 (TNNLS 2026)                           |
|------|-----------------------------------|--|
| 核心问题 | 未知 <b>静态非线性映射</b> 的基函数如何选择?       | <b>非线性动态系统</b> 如何在运行中在线学习与控制?              |
| 应用环境 | <b>离线</b> 数据驱动建模                  | <b>在线</b> 实时控制                             |
| 核心创新 | 引入 <b>频率估计器</b> 分析能量分布 → 指导初始小波选择 | <b>在线频域分析</b> + Lyapunov自适应律 + <b>动态剪枝</b> |
| 结构管理 | <b>高能量</b> 小波基优先                  | <b>持续调整</b> : 有增有减, 动态演进                   |
| 理论保证 | 给定精度的 <b>高维小波时频范围</b>             | <b>闭环系统一致最终有界</b> (UUB)                    |
| 计算效率 | 用 <b>更少基函数</b> 达到更高精度             | 剪枝策略 <b>降低在线计算成本</b>                       |
| 应用场景 | 静态非线性逼近/离线数据建模                    | 机器人控制、无人机姿态调节、电机驱动等实时控制系统                  |

论文一回答静态结构从哪来, 论文二回答在线结构怎么变, 给出了频率能量驱动的自构建小波神经网络从静态到动态的自适应学习理论。

## 方法贡献

首次将频域能量分析系统性地引入小波网络的构造过程，为基函数选择提供了数学指导而非经验试错。

## 算法贡献

构建了“频域感知 → 能量评估 → 结构演进”的完整算法体系。

## 理论贡献

静态场景：给出了高维小波时频范围的理论刻画

动态场景：提供了Lyapunov-based闭环稳定性严格证明

## 效率贡献

以更稀疏的网络结构、更低的计算成本实现高精度逼近/控制，为资源受限环境下的部署提供了可能。

## 自构建小波网络的理论探索和应用扩展

面向高维非线性逼近的  
并行自适应局部小波网络

基于同型小波神经元的  
参数自适应缩放策略

基于主副小波网络协同  
的抗扰动机制

# Thanks



<http://shendong.tech/>



DAILab\_RUC

## Acknowledgement

This work is supported by National Natural Science Foundation of China (62173333, 61673045, 61304085) and Beijing Natural Science Foundation (Z210002, 4152040).

