



## Towards Intelligence Driven by Physical-World Signals

# 物理世界信号驱动的智能建模

方楷楷

浙江大学 百人计划研究员

2025年11月13日

华为松鼠会分享

# 个人介绍



浙江大学  
ZHEJIANG UNIVERSITY



## 方楷楷

博士

百人计划研究员 | 博士生导师

✉ 邮箱 fsk@zju.edu.cn

🎓 研究方向 · 机器学习

- 生成式物理仿真和信号处理
- LLM多智能体系统
- 大规模高维时空序列

单位 信息与电子工程学院



Morgan Stanley

達摩院  
ALIBABA DAMO ACADEMY



Microsoft

# Research

微软亚洲研究院

# 目录



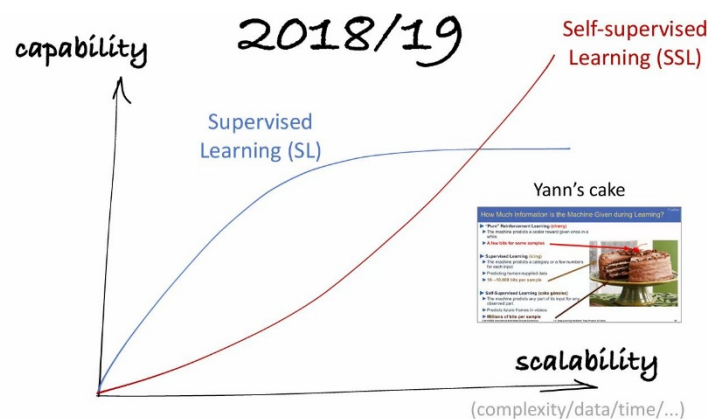
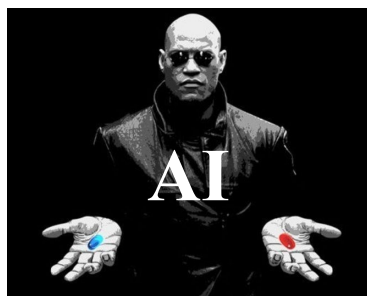
- **研究背景与挑战**
- 相关成果
- 未来方向

# 研究背景

- 过去十年AI 领域 重大范式转移: **深度网络, 表征学习, 生成模型, 自监督+预训练**



Vision / Language / Internet based



Yes! But...

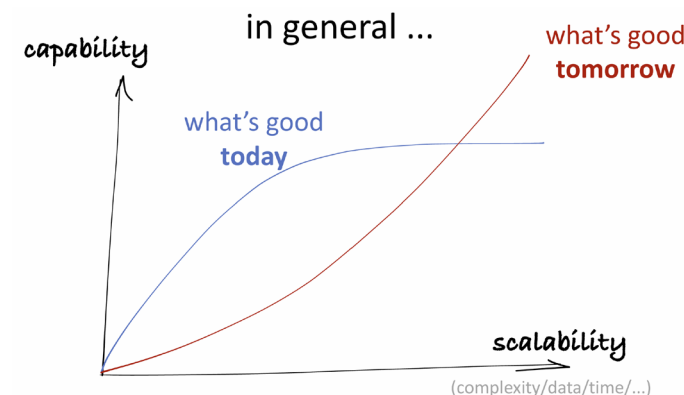
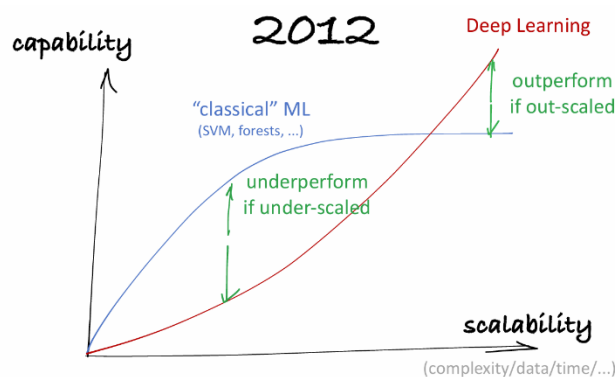
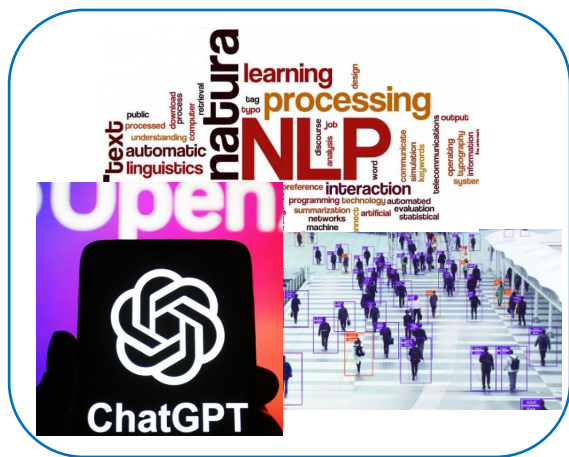


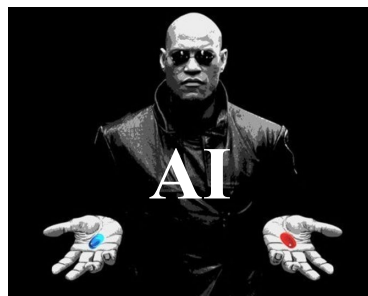
Fig. credits: Kaiming He's talk on NeurIPS 2024

# 研究背景

- 过去十年AI 领域 重大范式转移：**深度网络，表征学习，生成模型，自监督+预训练**



Vision / Language / Internet based



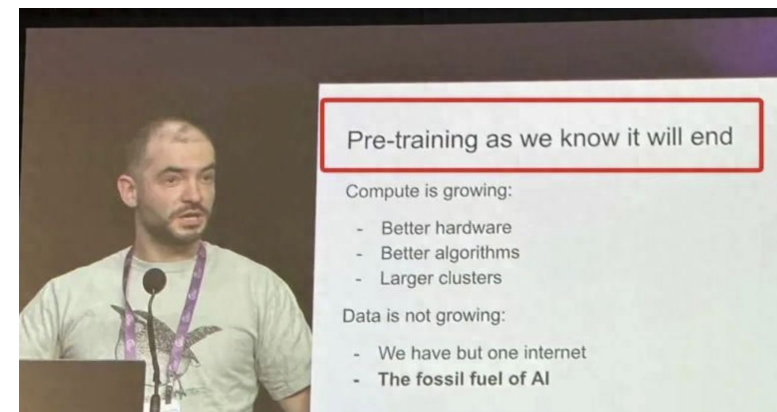
## 潜在瓶颈：

### - 数据端：

“数据墙”临近，依赖“人”/互联网的虚拟数据快被用完了\*

### - 应用端：

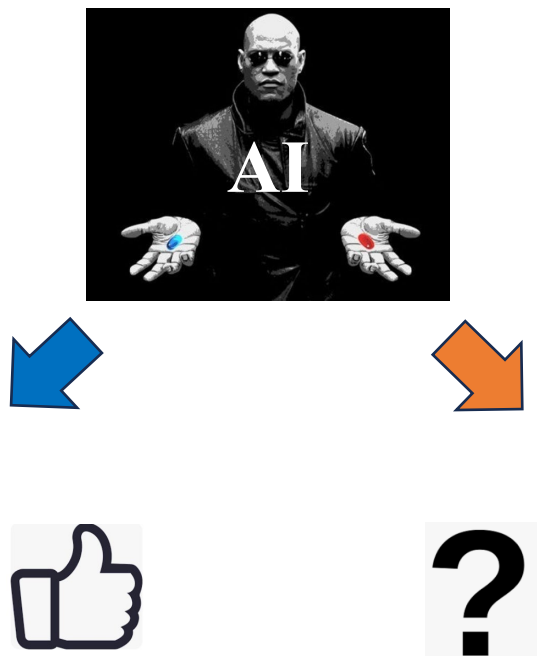
chatbot, 图文生成等“务虚”场景影响力有限



# 研究背景

- 下一个风口：**物理世界信号驱动的智能建模**

## AI 向实！！



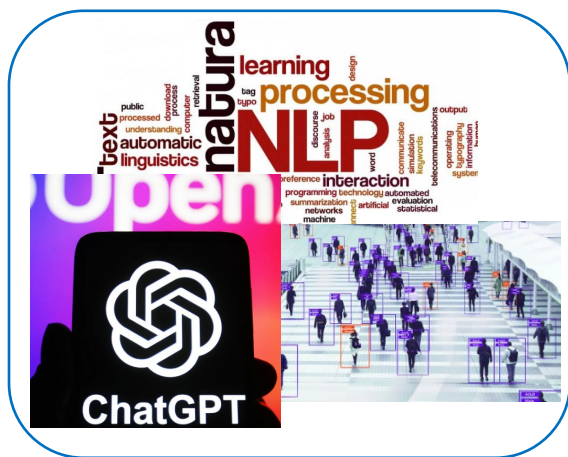
- 数据来源：  
**真实物理世界的演化&观测信号**
- 应用场景：  
**真实物理世界的科学与工程场景**

Vision / Language / Internet based

Action / Sensor / Signal based

“Welcome to  
the real world!”

- 下一个风口：**物理世界信号驱动的智能建模**



数据特质

数据量大&获取成本低，  
信噪比高，场域/尺度单一

数据量小&获取成本高，高阶&低秩，  
信噪比低（稀疏+多噪），多场/域/源混合

建模范式

数据驱动+表征学习

第一性原理为主（滤波，时频变换，动力系统...）

AI在CV/NLP任务上的发展出的新范式，**很难直接迁移到物理世界信号的建模中！**

## 稀疏张量：极高成本 + 超高维度 + 极稀疏信号 + 噪声

Tensor-Datasets	Tensor-Size	#Observed entries	Observed Ratio
SG (climate)	2321*5596*1600	105,764	0.0005%
ACC (cloud comp.)	3000*150*30000	1,220,000	0.1%

### 数据特质

数据量大&获取成本低，  
信噪比高，场域/尺度单一

数据量小&获取成本高，高阶&低秩，  
信噪比低（稀疏+多噪），多场/域/源混合

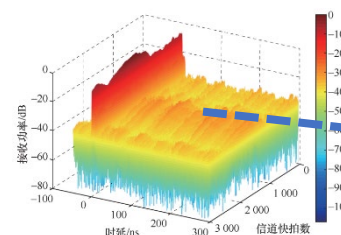
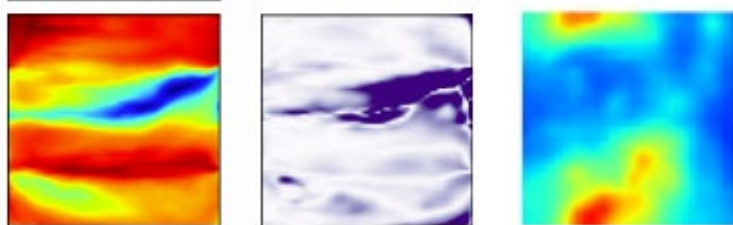
### 建模范式

数据驱动 + 表征学习

第一性原理为主（滤波，时频变换，动力系统...）

AI在CV/NLP任务上的发展出的新范式，很难直接迁移到物理世界信号的建模中！

例子：环境场感知信号的温度、盐度、风场 多系统高度非线性耦合



例子：对信道场中的小尺度、高度非线性特征建模

难以支撑 多源、多物理耦合的复杂建模

难以支撑多维精细化信号感知

模型表征能力不足

数据特质

数据量大&获取成本低，  
信噪比高，场域/尺度单一

数据量小&获取成本高，高阶&低秩，  
信噪比低（稀疏+多噪），多场/域/源混合

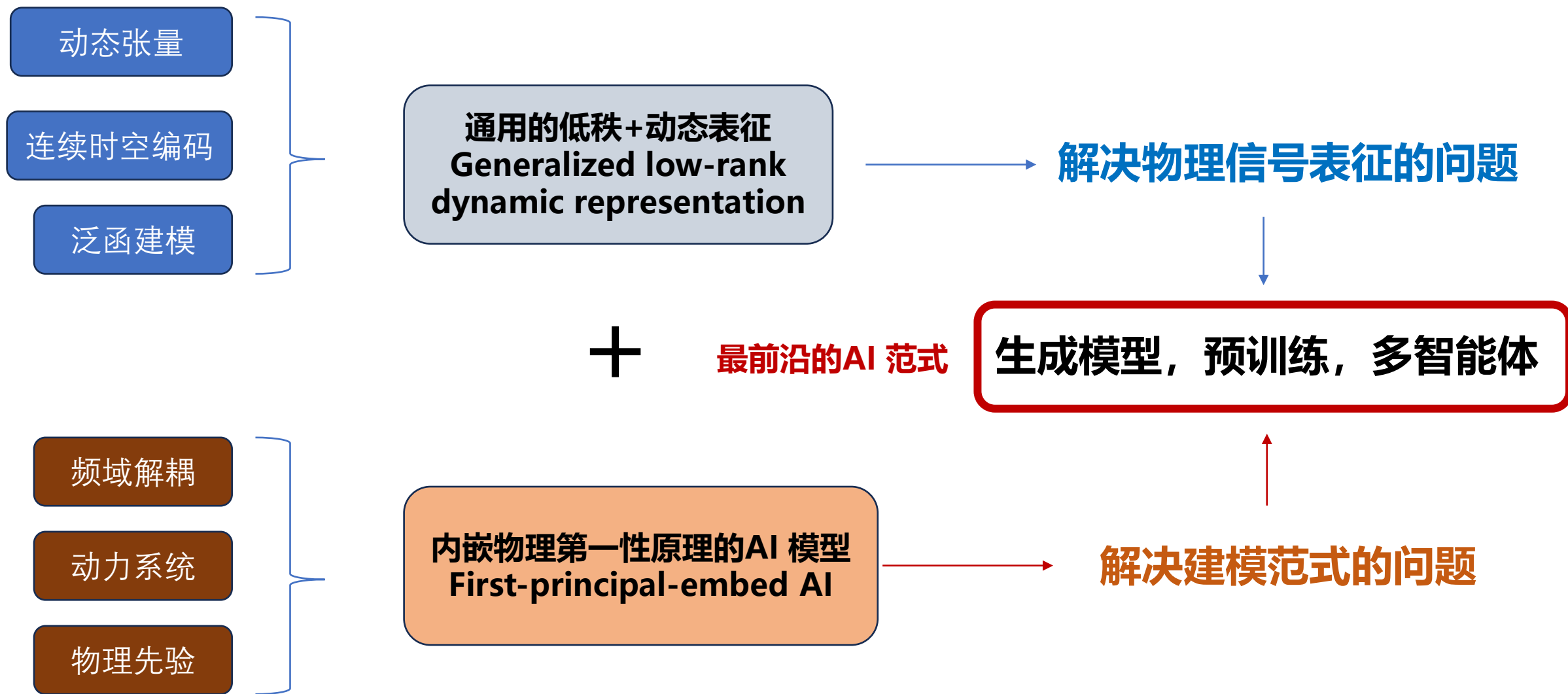
建模范式

数据驱动+表征学习

第一性原理为主（滤波，时频变换，动力系统...）

AI在CV/NLP任务上的发展出的新范式，很难直接迁移到物理世界信号的建模中！

# 研究思路



# 目录

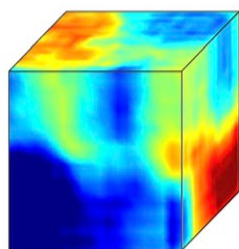


- 研究背景与挑战
- **相关成果:**
  - **广义动态张量学习**
  - 生成式物理场仿真
  - Agent 与产业基座模型
- 未来方向

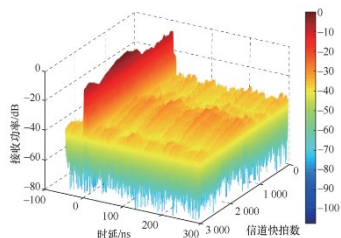
# 研究成果系列1：广义动态张量学习

海洋中的声速场、陆地上的信道响应、工程流体中的速度场

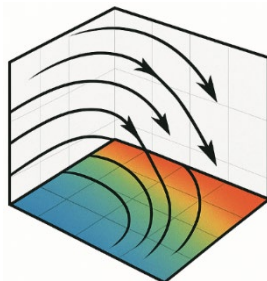
物理信号在空间、时间、频率/谱域、观测模态（如声学、射频、电磁或流体量测）、传感器方向等维度上耦合变化。



声速场



信道场



流速场

城市群的经济指标、社交网络中的交互、能源系统中的负荷

结构化交互信号 反应特定坐标-特定设备-特定模式-特定条件.. 等高阶多模动态交互的观测结果。



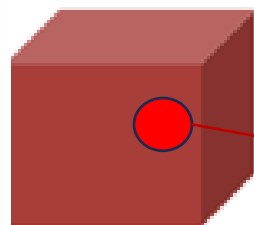
城市数据



社交网络



能源系统



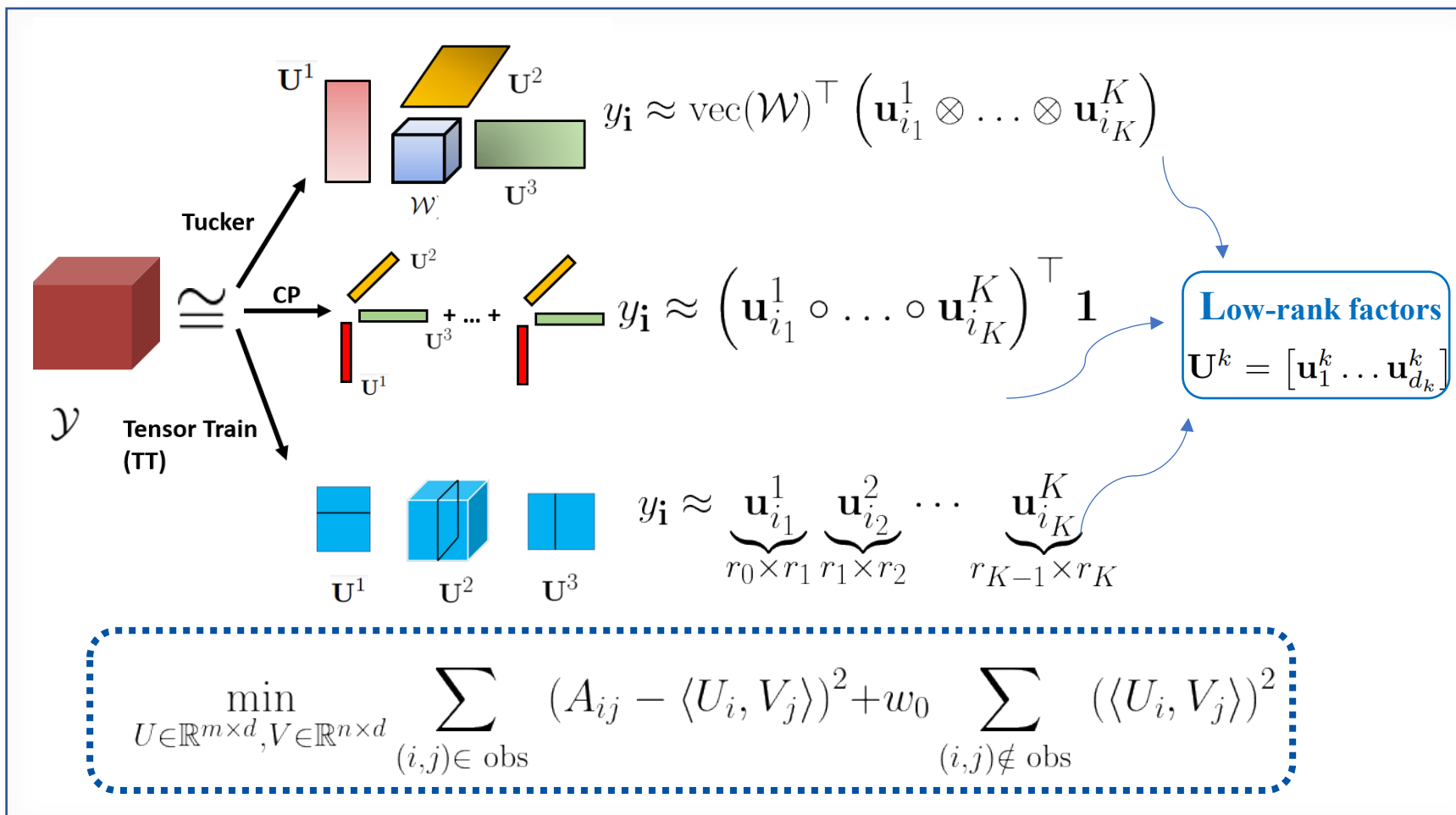
天然呈现多维 **张量** 结构

格式: (index1, index2.. )-> value

多维张量的**精准建模与重构**，直接决定了探测灵敏度、定位精度、通信容量与链路可靠性等关键性能指标

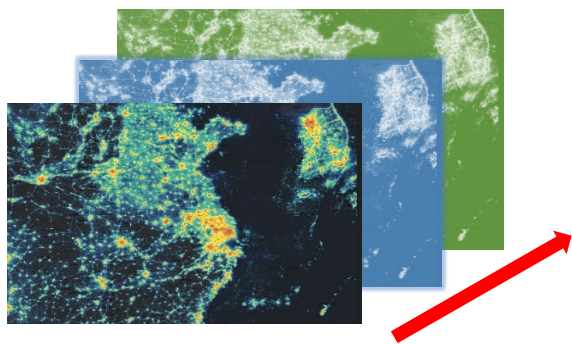
# 广义动态张量学习：背景

- 经典张量学习方法：稀疏观测 -> 学习静态低秩因子表征 -> 离散网格上重构 & 补全



## 真实应用场景的挑战：

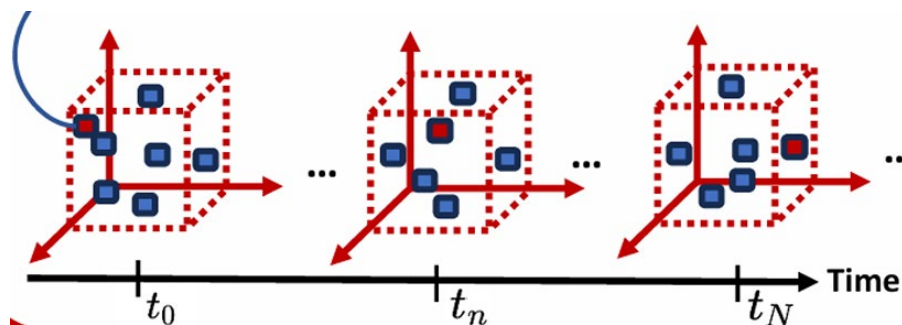
- 时变张量



(region, site, weather)  $\times$  time

张量结构随时间连续演化

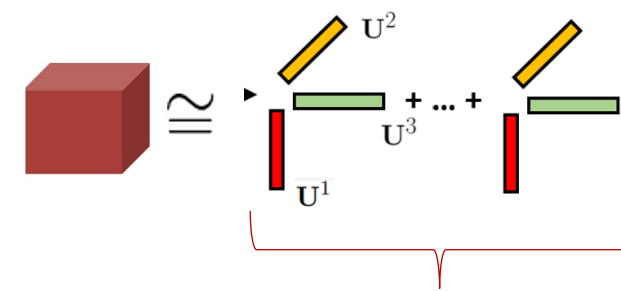
- 泛函张量



(latitude, longitude, height, time)

索引皆为连续实数的时空张量

- 分解秩确定

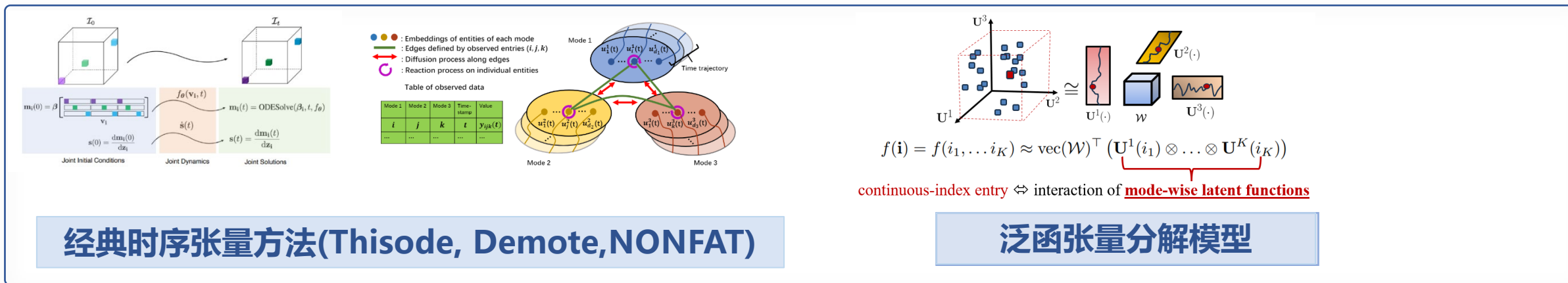


R: ?

最优分解秩的确定

# 代表成果: CATTE 自适应泛函时序张量分解 (NeurIPS 2025)

现有泛函张量模型难以有效建模**时空连续**数据, 且需手动设定模型复杂度。



经典时序张量方法(Thisode, Demote, NONFAT)

泛函张量分解模型

现有  
问题

- 经典时序张量只能处理空间离散时序数据, 无法处理**空时连续索引**的数据。
- 现有泛函张量模型**无法有效建模**时空连续数据。
- 现有泛函张量模型**无法自动确定**合适的模型复杂度。

解决  
方法

隐式神经表示

神经常微分方程

贝叶斯变分推断

**科学问题**

如何设计泛函时序张量模型，并且在训练过程中自适应学习模型复杂度？

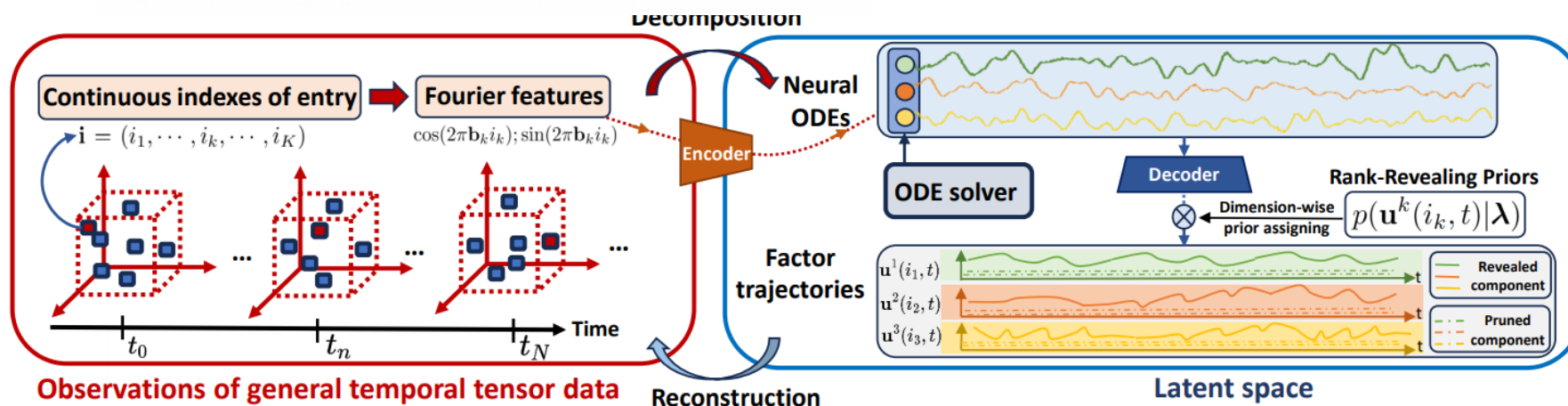
**研究方法**

□ 采用隐式神经表示建模因子，并通过神经常微分方程建模其随时间的演化。

□ 采用深度变分贝叶斯对模型进行稀疏建模，并通过变分推断构建证据下界损失。

◆ 泛函时序张量表征:  $y_i(t) = \mathbf{F}(i_1, \dots, i_K, t) \approx \mathbf{1}^T [\otimes_k \mathbf{u}^k(i_k, t)]$ , 编码连续索引的动态隐因子轨迹

◆ 稀疏先验:  $p(\mathbf{u}^k(i_k, t) | \lambda) = \mathcal{N}(\mathbf{u}^k(i_k, t) | \mathbf{0}, \Lambda^{-1}), \forall k$ , 推理中自适应调整分解秩



# 代表成果: CATTE 自适应泛函时序张量分解 (NeurIPS 2025)



## 先验与似然函数

$$p(\mathbf{u}^k(i_k, t) | \boldsymbol{\lambda}) = \mathcal{N}(\mathbf{u}^k(i_k, t) | \mathbf{0}, \boldsymbol{\Lambda}^{-1}), \forall k,$$

$$p(\boldsymbol{\lambda}) = \prod_{r=1}^R \text{Gamma}(\lambda_r | a_r^0, b_r^0),$$

$$p(y_n | \mathcal{U}, \tau) = \mathcal{N}(y_n | \mathbf{1}^T [\underset{k}{\otimes} \mathbf{u}^k(i_k^n, t_n)], \tau^{-1})$$

## 后验建模

$$q(\mathcal{U}) = \prod_{n=1}^N \prod_{k=1}^K \mathcal{N}(\mathbf{u}^k(i_k^n, t_n) | \mathbf{g}^k(i_k^n, t_n), \sigma^2 \mathbf{I}),$$

$$q(\boldsymbol{\lambda}) = \prod_{r=1}^R \text{Gamma}(\lambda_r | \alpha_r, \beta_r), q(\tau) = \text{Gamma}(\tau | \rho, \iota),$$

$$\text{KL}(q(\mathcal{U}, \boldsymbol{\lambda}, \tau) \| p(\mathcal{U}, \boldsymbol{\lambda}, \tau | \mathcal{D}))$$

## 变分证据下界

## 误差拟合项

## 稀疏正则项

$$\text{ELBO} = \underbrace{\mathbb{E}_{q(\mathcal{U}, \boldsymbol{\lambda}, \tau)} [\ln p(\mathcal{D} | \mathcal{U}, \boldsymbol{\lambda}, \tau)]}_{\text{误差拟合项}} + \underbrace{\mathbb{E}_{q(\mathcal{U}, \boldsymbol{\lambda})} [\ln \frac{p(\mathcal{U} | \boldsymbol{\lambda})}{q(\mathcal{U})}]}_{\text{稀疏正则项}} - \text{KL}(q(\boldsymbol{\lambda}) \| p(\boldsymbol{\lambda})) - \text{KL}(q(\tau) \| p(\tau)).$$

$$\mathbb{E}_{q(\mathcal{U}, \boldsymbol{\lambda}, \tau)} [\ln p(\mathcal{D} | \mathcal{U}, \boldsymbol{\lambda}, \tau)] = -\frac{N}{2} \ln(2\pi) + \frac{N}{2} (\psi(\rho) - \ln \iota) - \frac{1}{2} \sum_{n=1}^N \frac{\rho}{\iota} \{ y_n^2 - 2y_n \{ \mathbf{1}^T [\underset{k}{\otimes} \mathbf{g}^k(i_k^n, t_n)] \} + \mathbf{1}^T [\underset{k}{\otimes} \text{vec}(\mathbf{g}^k(i_k^n, t_n) \mathbf{g}^k(i_k^n, t_n)^T + \sigma^2 \mathbf{I})] \},$$

$$\mathbb{E}_{q(\mathcal{U}, \boldsymbol{\lambda})} [\ln \frac{p(\mathcal{U} | \boldsymbol{\lambda})}{q(\mathcal{U})}] = -\text{KL}(q(\mathcal{U}) \| p(\mathcal{U} | \boldsymbol{\lambda} = \mathbb{E}_q(\boldsymbol{\lambda}))) = - \sum_{n=1}^N \sum_{k=1}^K \sum_{r=1}^R \frac{1}{2} \{ \ln(\frac{\beta_r}{\alpha_r \sigma^2}) + \frac{\alpha_r}{\beta_r} \{ \sigma^2 + [\mathbf{g}_r^k(i_k^n, t_n)]^2 \} - 1 \},$$

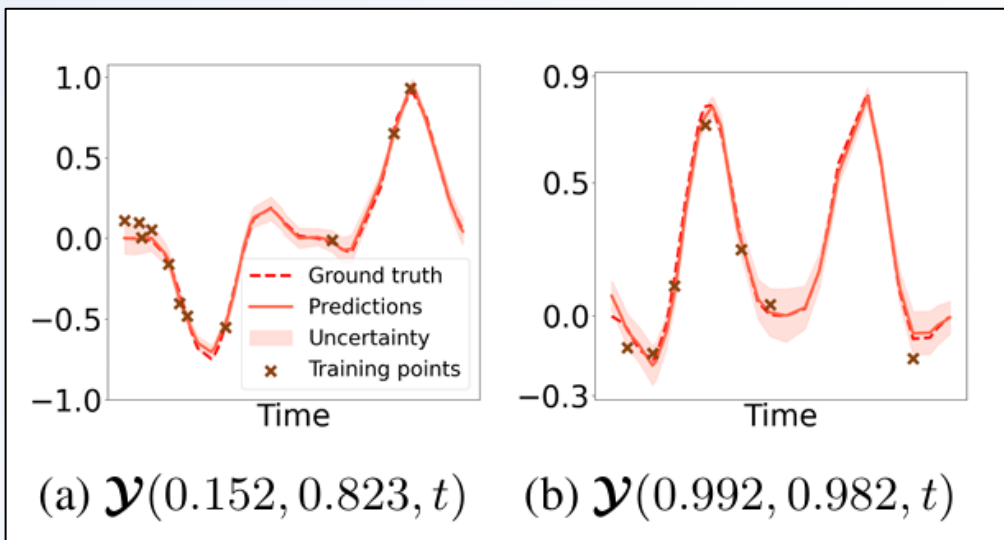
## 仿真数据实验

- **Rank1**

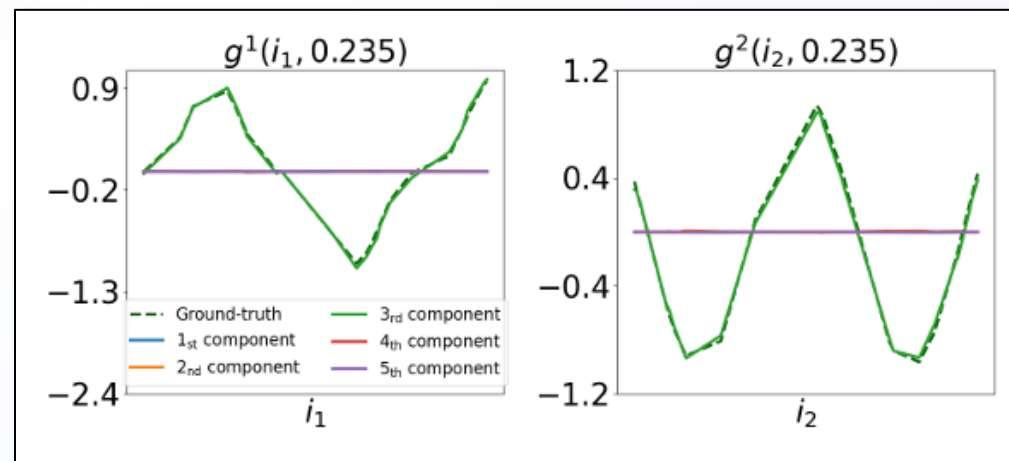
$$\mathcal{Y}(i_1, i_2, t) = \mathbf{1}^T [\mathbf{u}^1(i_1, t) \otimes \mathbf{u}^2(i_2, t)],$$

where  $\mathbf{u}^1(i_1, t) = -\cos^3(2\pi t + 2.5\pi i_1)$ ;  $\mathbf{u}^2(i_2, t) = \sin(3\pi t + 3.5\pi i_2)$ .

- 随机采样6250个带噪的不规则的实数索引值对 $(i, y)$ 用于训练



- 补全效果, 随机实数索引点的时间轨迹



- 所学的因子轨迹 (唯一性)

## 实际数据实验

### 补全误差(RMSE, MAE)

Datasets	RMSE			MAE		
	CA Traffic	Server Room	SSF	CA Traffic	Server Room	SSF
$R = 3$						
THIS-ODE	$0.672 \pm 0.002$	$0.132 \pm 0.002$	$2.097 \pm 0.003$	$0.587 \pm 0.002$	$0.083 \pm 0.002$	$2.084 \pm 0.003$
NONFAT	$0.504 \pm 0.010$	$0.129 \pm 0.002$	$9.796 \pm 0.010$	$0.167 \pm 0.009$	$0.078 \pm 0.001$	$8.771 \pm 0.043$
DEMOTE	$0.447 \pm 0.001$	$0.131 \pm 0.001$	$9.789 \pm 0.001$	$0.118 \pm 0.002$	$0.090 \pm 0.0015$	$8.757 \pm 0.001$
FunBaT-CP	$0.563 \pm 0.025$	$0.425 \pm 0.003$	$0.696 \pm 0.047$	$0.244 \pm 0.025$	$0.308 \pm 0.001$	$0.549 \pm 0.038$
FunBaT-Tucker	$0.584 \pm 0.009$	$0.498 \pm 0.058$	$0.730 \pm 0.201$	$0.189 \pm 0.014$	$0.381 \pm 0.053$	$0.614 \pm 0.128$
LRTFR	$0.379 \pm 0.042$	$0.151 \pm 0.004$	$0.595 \pm 0.018$	$0.187 \pm 0.022$	$0.110 \pm 0.002$	$0.464 \pm 0.0165$
$R = 5$						
THIS-ODE	$0.632 \pm 0.002$	$0.132 \pm 0.003$	$1.039 \pm 0.015$	$0.552 \pm 0.001$	$0.083 \pm 0.002$	$1.032 \pm 0.002$
NONFAT	$0.501 \pm 0.002$	$0.117 \pm 0.006$	$9.801 \pm 0.014$	$0.152 \pm 0.001$	$0.071 \pm 0.004$	$8.744 \pm 0.035$
DEMOTE	$0.421 \pm 0.002$	$0.105 \pm 0.003$	$9.788 \pm 0.001$	$0.103 \pm 0.001$	$0.068 \pm 0.003$	$8.757 \pm 0.001$
FunBaT-CP	$0.547 \pm 0.025$	$0.422 \pm 0.001$	$0.675 \pm 0.061$	$0.204 \pm 0.052$	$0.307 \pm 0.002$	$0.531 \pm 0.051$
FunBaT-Tucker	$0.578 \pm 0.005$	$0.521 \pm 0.114$	$0.702 \pm 0.054$	$0.181 \pm 0.005$	$0.391 \pm 0.097$	$0.557 \pm 0.041$
LRTFR	$0.376 \pm 0.016$	$0.167 \pm 0.006$	$0.532 \pm 0.036$	$0.182 \pm 0.012$	$0.121 \pm 0.005$	$0.418 \pm 0.003$
$R = 7$						
THIS-ODE	$0.628 \pm 0.007$	$0.154 \pm 0.016$	$1.685 \pm 0.009$	$0.548 \pm 0.006$	$0.089 \pm 0.002$	$1.674 \pm 0.008$
NONFAT	$0.421 \pm 0.016$	$0.128 \pm 0.002$	$9.773 \pm 0.015$	$0.137 \pm 0.006$	$0.077 \pm 0.002$	$8.718 \pm 0.035$
DEMOTE	$0.389 \pm 0.005$	$0.094 \pm 0.006$	$9.790 \pm 0.002$	$0.091 \pm 0.001$	$0.062 \pm 0.006$	$8.753 \pm 0.006$
FunBaT-CP	$0.545 \pm 0.009$	$0.426 \pm 0.001$	$0.685 \pm 0.049$	$0.204 \pm 0.037$	$0.307 \pm 0.001$	$0.541 \pm 0.039$
FunBaT-Tucker	$0.587 \pm 0.011$	$0.450 \pm 0.041$	$0.642 \pm 0.037$	$0.195 \pm 0.022$	$0.330 \pm 0.026$	$0.507 \pm 0.029$
LRTFR	$0.365 \pm 0.042$	$0.156 \pm 0.012$	$0.502 \pm 0.033$	$0.161 \pm 0.014$	$0.118 \pm 0.009$	$0.392 \pm 0.028$
Functional Automatic Rank Determination						
CATTE (Ours)	<b><math>0.284 \pm 0.016</math></b>	<b><math>0.078 \pm 0.001</math></b>	<b><math>0.373 \pm 0.003</math></b>	<b><math>0.085 \pm 0.004</math></b>	<b><math>0.047 \pm 0.003</math></b>	<b><math>0.288 \pm 0.003</math></b>
CATTE w.o. FARD	$0.301 \pm 0.020$	$0.091 \pm 0.008$	$0.402 \pm 0.013$	$0.094 \pm 0.010$	$0.0657 \pm 0.005$	$0.310 \pm 0.010$

Table 1: Predictive errors and standard deviation. The results were averaged over five runs.

### 超参学习曲线

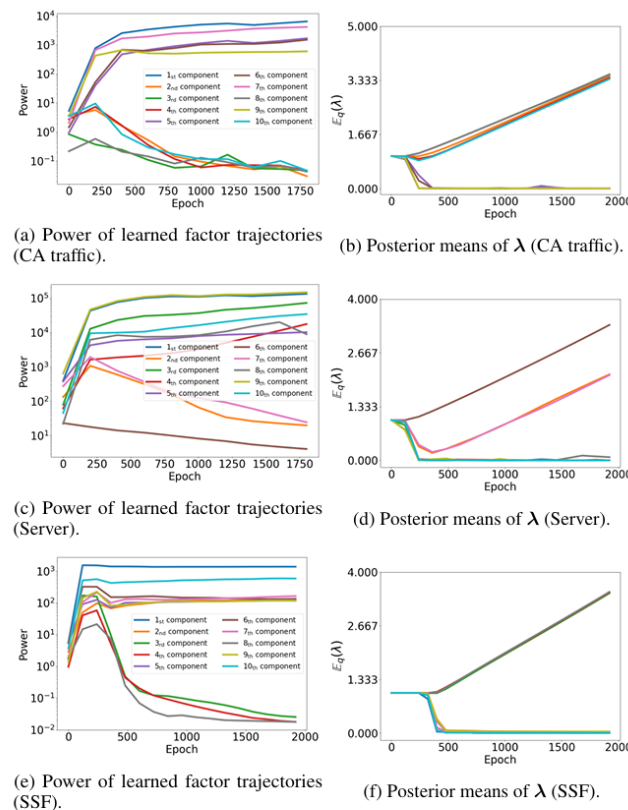
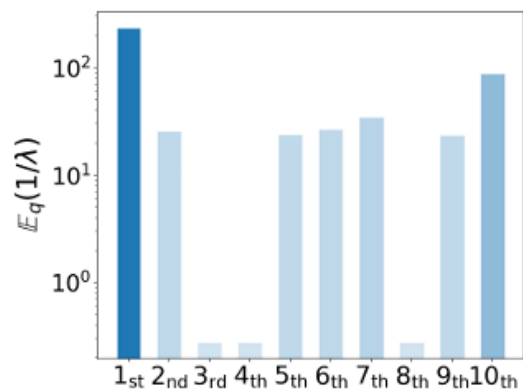
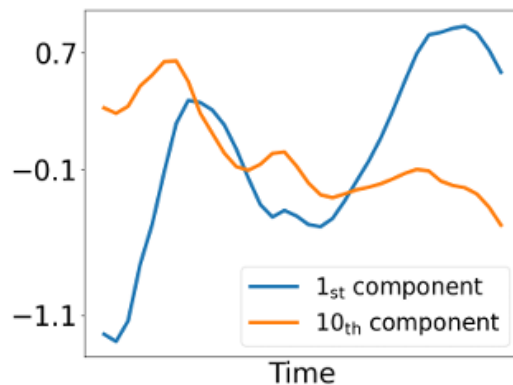


Figure 6: Rank learning curves of three datasets.

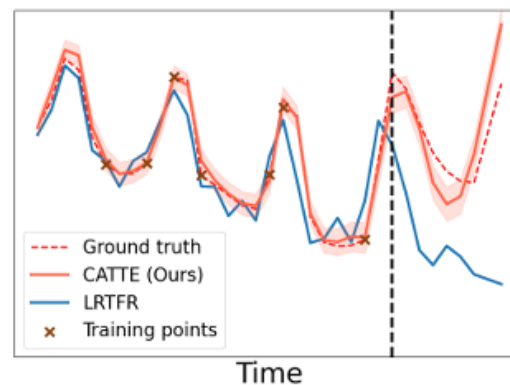
## 可解释性和可拓展性



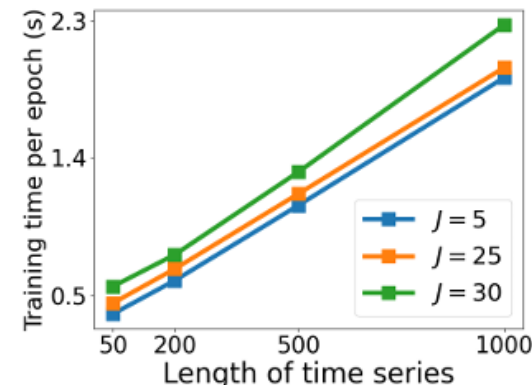
(a) Posterior mean of  $\frac{1}{\lambda}$



(b) Factor trajectories



(c) Entry predictions



(d) Scalability



$$p(\mathbf{u}^k(i_k, t) | \boldsymbol{\lambda}) = \mathcal{N}(\mathbf{u}^k(i_k, t) | \mathbf{0}, \boldsymbol{\Lambda}^{-1}), \forall k,$$

● 外推

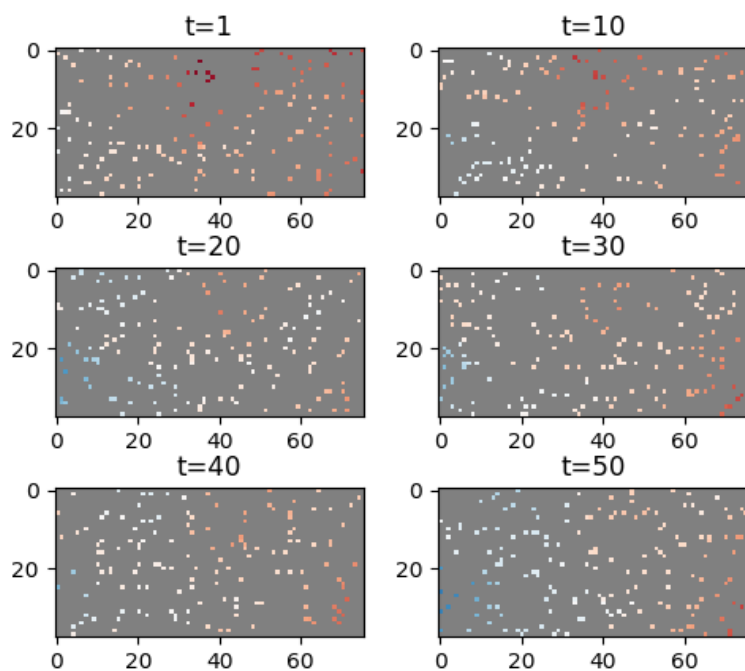
● 可拓展性

# 代表成果: CATTE 自适应泛函时序张量分解 (NeurIPS 2025)



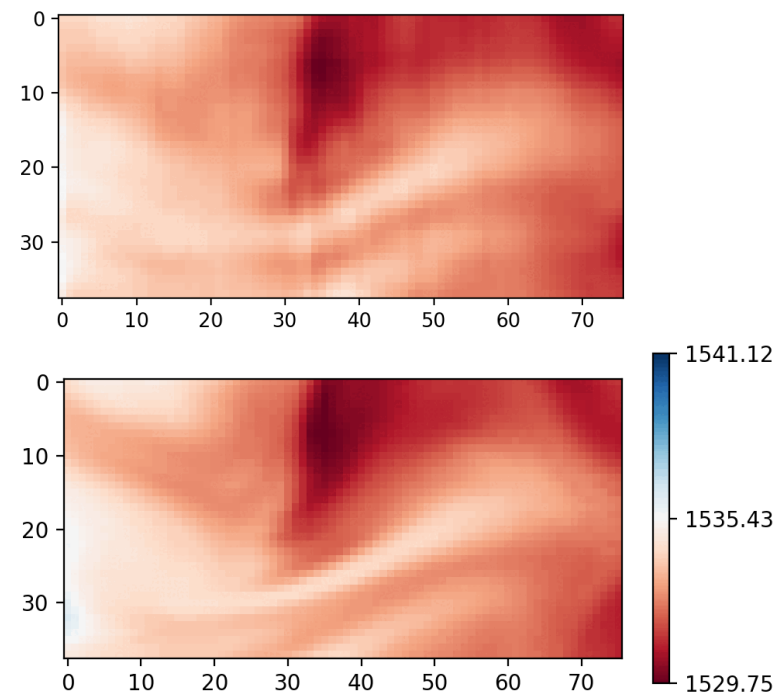
浙江大学  
ZHEJIANG UNIVERSITY

## 单源观测下初步验证泛函稀疏时变建模的可行性



5%空时观测

## 重构结果



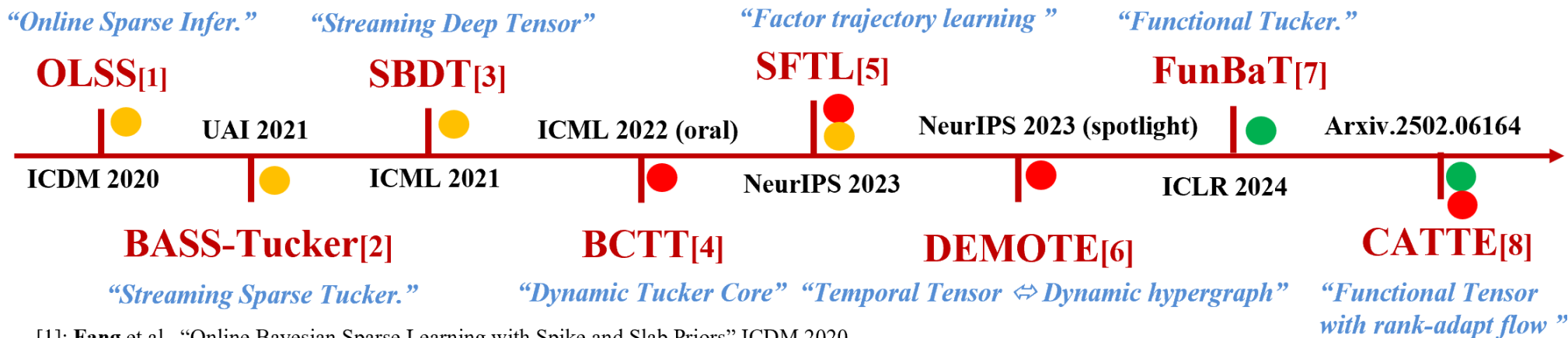
真实结果

# 广义动态张量学习：总结

基于**广义动态张量**的数据**表征学习**，能很好地应对物理信号的**高维-稀疏-动态-带噪-不规则-流式响应**等挑战

## Types of dynamic tensor data

● Streaming   ● Temporal   ● Functional



[1]: Fang et al., "Online Bayesian Sparse Learning with Spike and Slab Priors" ICDM 2020

[2]: Fang et al., "Bayesian Streaming Sparse Tucker Decomposition", UAI 2021

[3]: Fang et al., "Streaming Bayesian Deep Tensor Factorization" ICML 2021

[4]: Fang et al., "Bayesian Continuous-Time Tucker Decomposition", ICML 2022, oral

[5]: Fang et al., "Streaming Factor Trajectory Learning for Temporal Tensor Decomposition", NeurIPS 2023

[6]: Wang\* & Fang\* et al., "Dynamic Tensor Decomposition via Neural Diffusion-Reaction Processes", NeurIPS 2023, spotlight

[7]: Fang et al., "Functional Bayesian Tucker Decomposition for Continuous-indexed Tensor", ICLR 2024

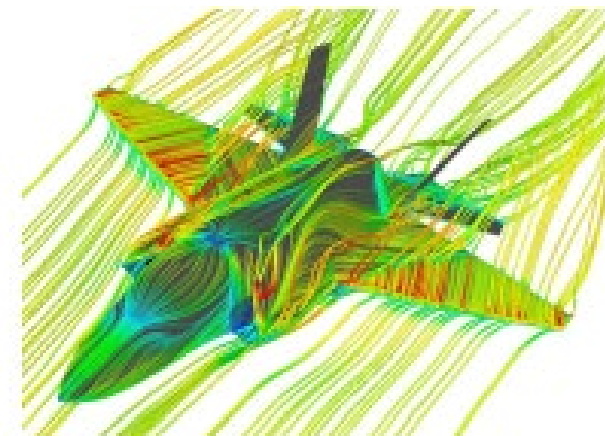
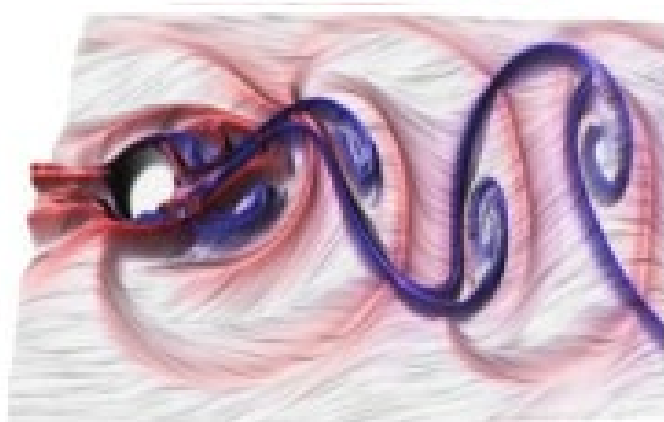
[8]: Chen et al., & Fang, "Functional Complexity-adaptive Temporal Tensor Decomposition". NeurIPS 2025

# 目录



- 研究背景与挑战
- **相关成果:**
  - 广义动态张量学习
  - **生成式物理场仿真**
  - Agent 与产业基座模型
- 未来方向

# 生成式物理场仿真：背景

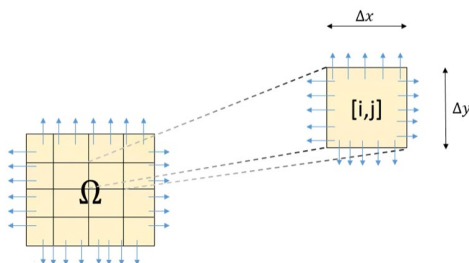


$$\mathcal{F}[u](\mathbf{x}) = f(\mathbf{x}) \quad (\mathbf{x} \in \Omega), \quad u(\mathbf{x}) = g(\mathbf{x}) \quad (\mathbf{x} \in \partial\Omega),$$

传统仿真：完全基于第一性原理

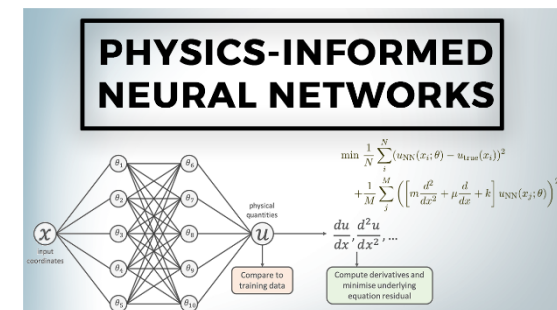
有限元、数值求解

- 高精度
- 低效率 + 泛化差



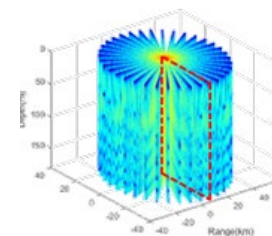
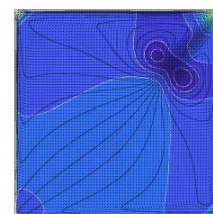
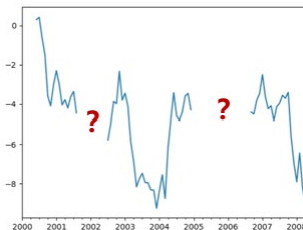
第一性原理+ 数据驱动：DNN 代理模型  
PINN, Operator Learning

- 高效率+泛化强
- 低精度



# 生成式物理场仿真：背景

- 更进一步：**生成** as **仿真**



- 物理场建模和仿真中的 补全、预测、去噪、超分等不同任务，需要设计单独的算法模块

One for One...



One for All !

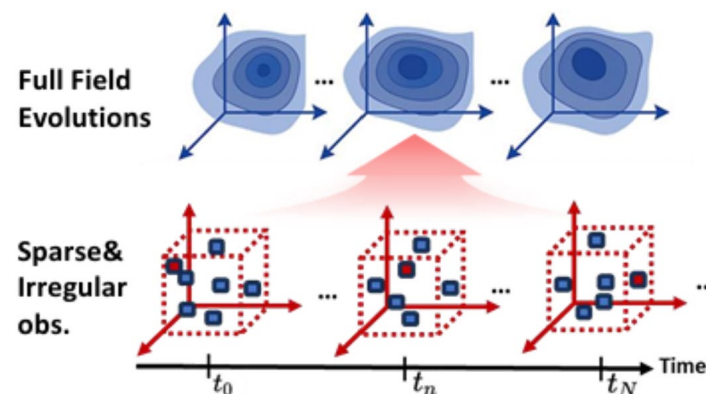
Why?

- 补全、预测、差值、去噪、超分... -> 统一被建模成 **“条件生成”** 任务
- 更灵活注入**第一性原理和物理约束**作为生成条件
- 天然的不确定性度量
- 更好的**Scaling** 能力和 **跨任务泛化**

# • 代表成果: SDIFT 生成稀疏物理场的全域演化 (NeurIPS 2025)



Task: 从稀疏、带噪、不规则的观测值中, **生成** 高维物理场 时空全域 演化过程



难点: □ 基于Diffusion和Transformer的模型无法建模**空时连续索引 + 高维** 数据。

□ 基于张量的模型缺乏生成能力, 表征能力、泛化性不够, 难以应对**极其稀疏**的缺失数据。

Solution:

泛函张量分解模型编码器

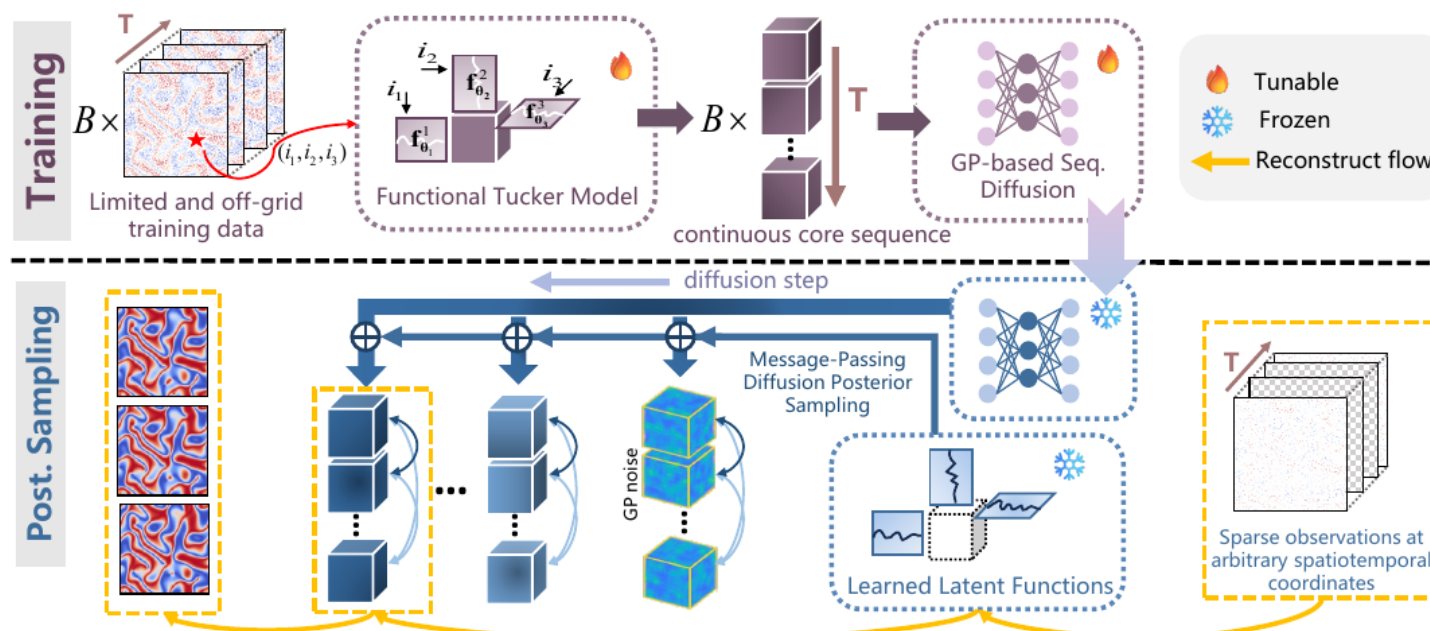
高斯过程噪声驱动的隐空间  
序列扩散模型

消息传递扩散后验采样

# 代表成果: SDIFT 生成稀疏物理场的全域演化 (NeurIPS 2025)

## 研究方法

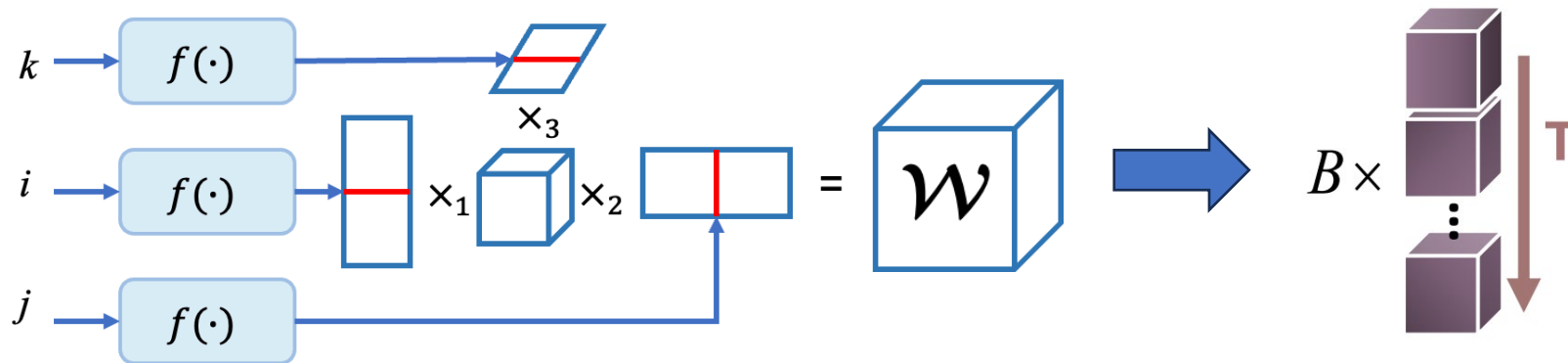
- 采用**泛函张量Tucker模型**作为编码器，将空间连续索引的数据映射至隐空间
- 利用**高斯过程噪声驱动的序列扩散模型**，学习信号场隐空间的动态演化分布
- 提出基于**消息传递的扩散后验采样**，实现多维信号场任意时间戳的连续补全



## 模型设计

• 代表成果: SDIFT 生成稀疏物理场的全域演化 (NeurIPS 2025)

泛函张量Tucker编码器



$$y_{\mathbf{i}, t_m} \approx \text{vec}(\mathcal{W}_{t_m})^\top (\mathbf{f}_{\theta_1}^1(i_1) \otimes \cdots \otimes \mathbf{f}_{\theta_K}^K(i_K)).$$

泛在逼近性证明

**Theorem 1.** (UAP of FTM) Let  $X_1, \dots, X_K$  be compact subsets of  $\mathbb{R}^K$ . Choose  $u \in L^2(X_1 \times \cdots \times X_K)$ . Then, for arbitrary  $\epsilon > 0$ , there exists sufficiently large  $\{R_1 > 0, \dots, R_K > 0\}$ , coefficients  $\{a_{r_1, \dots, r_K}\}_{r_1, \dots, r_K}^{R_1, \dots, R_K}$  and neural networks  $\{\{f_{r_k}^k\}_{r_k}^{R_k}\}_k^K$  such that

$$\left\| u - \sum_{r_1}^{R_1} \cdots \sum_{r_K}^{R_K} [a_{r_1, \dots, r_K} \prod_{k=1}^K f_{r_k}^k] \right\|_{L^2(X_1 \times \cdots \times X_K)} < \epsilon. \quad (7)$$

# 代表成果: SDIFT 生成稀疏物理场的全域演化 (NeurIPS 2025)

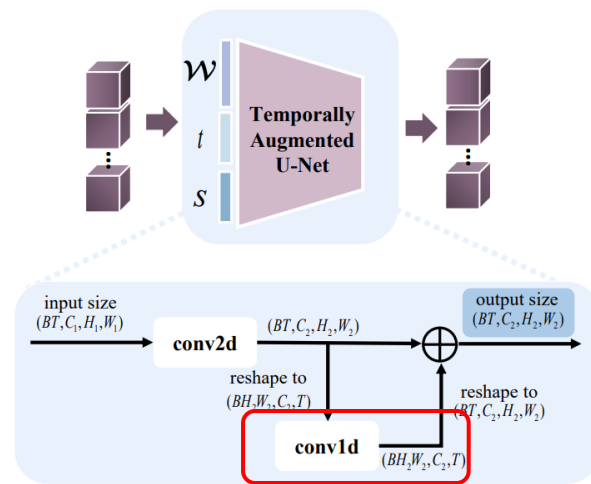


## 高斯过程驱动的隐空间序列扩散模型

### 高斯过程噪声源

$$\mathcal{E} = \{\mathcal{E}_{t_1}, \dots, \mathcal{E}_{t_M}\} \sim \mathcal{GP}(0, \kappa(t_i, t_j))$$

### 时间维度增强U-Net

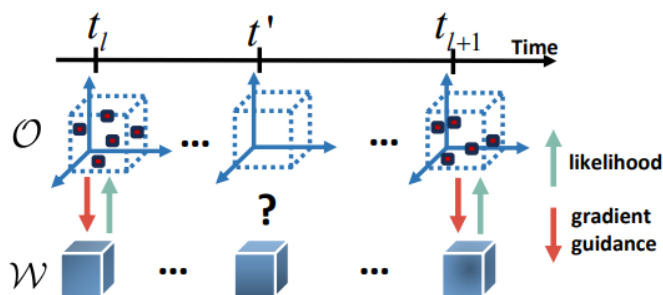


### 目标函数

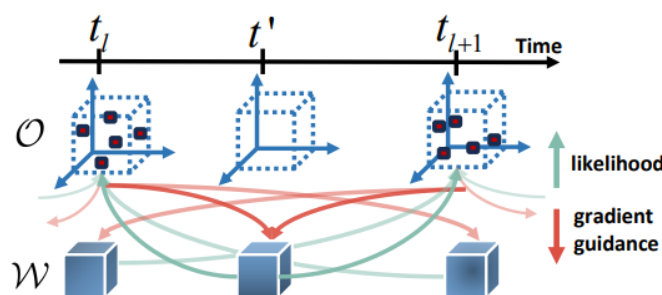
$$\mathcal{L}_{\text{GPSD}} = \mathbb{E}_{\sigma(s), \mathcal{W}_{t_m} \sim \mathcal{W}, \mathcal{E}_{t_m} \sim \mathcal{E}, t \sim \mathcal{T}} \left[ \lambda(\sigma(s)) \|D_{\theta}(\mathcal{W}_{t_m} + \mathcal{E}_{t_m}; \sigma(s), t_m) - \mathcal{W}_{t_m}\|_2^2 \right]$$

# 代表成果: SDIFT 生成稀疏物理场的全域演化 (NeurIPS 2025)

## 消息传递扩散后验采样器



(a) DPS : Guidance missing at  $t'$ .



(b) MPDPS : Collected message as guidance at  $t'$ .

Figure 2: Illustration of DPS and MPDPS to handle the case of no observations at  $t'$ .

允许观察信号/物理约束

“跨时刻传递”

来生成全域演化!

所提出的后验梯度引导信息:

$$\nabla_{\mathbf{W}_{t_n}^s} \log p(\mathcal{O} | \mathbf{W}_{t_n}^s) \approx \underbrace{\mathbf{1}_{\mathcal{T}_{\text{obs}}}(t_n) \cdot \nabla_{\mathbf{W}_{t_n}^s} \log p(\mathcal{O}_{t_n} | \mathbf{W}_{t_n}^s)}_{\text{来自当前观测的后验信息}} + \sum_{t_l \in \mathcal{T}_{\text{obs}} \setminus t_n} \overbrace{\mathcal{G}_{t_l, t_n}^s}^{\text{来自其他时刻观测的聚合信息}}$$

来自当前观测的后验信息

来自其他时刻观测的聚合信息

似然函数具有二次项结构, 梯度计算效率高!

$$\{\mathcal{G}_{t_l, t_m}^s\}_{t_m \in \mathcal{T}_{\text{tar}} \setminus t_l} := \nabla_{\mathbf{W}_{\setminus t_l}^s} \log p(\mathcal{O}_{t_l} | \mathbf{W}_{\setminus t_l}^s) = \nabla_{\mathbf{W}_{\setminus t_l}^s} \log \int p(\mathcal{O}_{t_l} | \mathbf{W}_{t_l}^0) p(\mathbf{W}_{t_l}^0 | \mathbf{W}_{\setminus t_l}^s) d\mathbf{W}_{t_l}^0$$

# 代表成果: SDIFT 生成稀疏物理场的全域演化 (NeurIPS 2025)



浙江大学  
ZHEJIANG UNIVERSITY

## 实际数据实验

### 不同观测比例下的补全误差 (VRMSE)

Methods	Supernova Explosion		Ocean Sound Speed		Active Matter	
	$\rho = 1\%$	$\rho = 3\%$	$\rho = 1\%$	$\rho = 3\%$	$\rho = 1\%$	$\rho = 3\%$
<b>Observation setting 1 : <math>\mathcal{T}_{obs} = \mathcal{T}_{tar}</math></b>						
<i>Tensor-based</i>						
LRTFR [11]	0.558 ± 0.044	0.429 ± 0.043	0.345 ± 0.036	0.217 ± 0.066	0.302 ± 0.104	0.258 ± 0.022
DEMOTÉ [38]	1.285 ± 0.102	1.213 ± 0.217	0.358 ± 0.127	0.314 ± 0.086	0.950 ± 0.486	0.871 ± 0.497
NONFAT [27]	1.229 ± 0.127	1.197 ± 0.204	0.402 ± 0.090	0.330 ± 0.101	0.921 ± 0.457	0.867 ± 0.413
<i>Attention-based</i>						
Senseiver [18]	0.446 ± 0.041	0.349 ± 0.023	0.264 ± 0.037	0.2005 ± 0.031	0.345 ± 0.094	0.264 ± 0.076
<i>Diffusion-based</i>						
CoNFILD [5]	0.561 ± 0.082	0.427 ± 0.037	0.201 ± 0.034	0.145 ± 0.012	0.529 ± 0.087	0.5075 ± 0.830
SDIFT w/ DPS	0.339 ± 0.116	0.291 ± 0.033	0.194 ± 0.073	0.160 ± 0.035	0.298 ± 0.065	0.174 ± 0.043
SDIFT w/ MPDPS	<b>0.283 ± 0.026</b>	<b>0.272 ± 0.025</b>	<b>0.146 ± 0.046</b>	<b>0.108 ± 0.043</b>	<b>0.215 ± 0.068</b>	<b>0.156 ± 0.046</b>
<b>Observation setting 2 : <math> \mathcal{T}_{obs}  = \frac{1}{2} \mathcal{T}_{tar} </math></b>						
<i>Tensor-based</i>						
LRTFR [11]	0.783 ± 0.416	0.813 ± 0.296	0.610 ± 0.323	0.508 ± 0.297	0.620 ± 0.484	0.598 ± 0.527
DEMOTÉ [38]	1.351 ± 0.209	1.223 ± 0.397	0.549 ± 0.181	0.533 ± 0.198	1.261 ± 0.614	1.277 ± 0.603
NONFAT [27]	1.278 ± 0.214	1.254 ± 0.2785	0.465 ± 0.153	0.420 ± 0.189	1.126 ± 0.514	1.270 ± 0.485
<i>Attention-based</i>						
Senseiver [18]	Not capable	-	-	-	-	-
<i>Diffusion-based</i>						
CoNFILD [5]	0.757 ± 0.199	0.6575 ± 0.148	0.310 ± 0.054	0.2615 ± 0.038	0.8265 ± 0.167	0.779 ± 0.161
SDIFT w/ DPS	0.659 ± 0.057	0.6450 ± 0.054	0.412 ± 0.156	0.407 ± 0.136	0.674 ± 0.153	0.637 ± 0.113
SDIFT w/ MPDPS	<b>0.433 ± 0.163</b>	<b>0.335 ± 0.122</b>	<b>0.181 ± 0.084</b>	<b>0.165 ± 0.041</b>	<b>0.296 ± 0.096</b>	<b>0.256 ± 0.087</b>

Table 1: VRMSEs of the reconstruction results for all methods across three datasets, evaluated under two observation settings with different observation ratios.

Methods	Supernova Explosion			Ocean Sound Speed			Active Matter		
	$\rho = 1\%$	$\rho = 3\%$	#Para.	$\rho = 1\%$	$\rho = 3\%$	#Para.	$\rho = 1\%$	$\rho = 3\%$	#Para.
CoNFILD	31.3s	44.0s	23M	15.9s	20.3s	10M	27.6s	31.5s	10M
SDIFT w/ MPDPS	<b>2.23s</b>	<b>5.43s</b>	26M	<b>0.84s</b>	<b>0.89s</b>	15M	<b>1.31s</b>	<b>1.42s</b>	12M

Table 2: Average sampling speed for reconstruction with different observation ratios on observation setting 1.

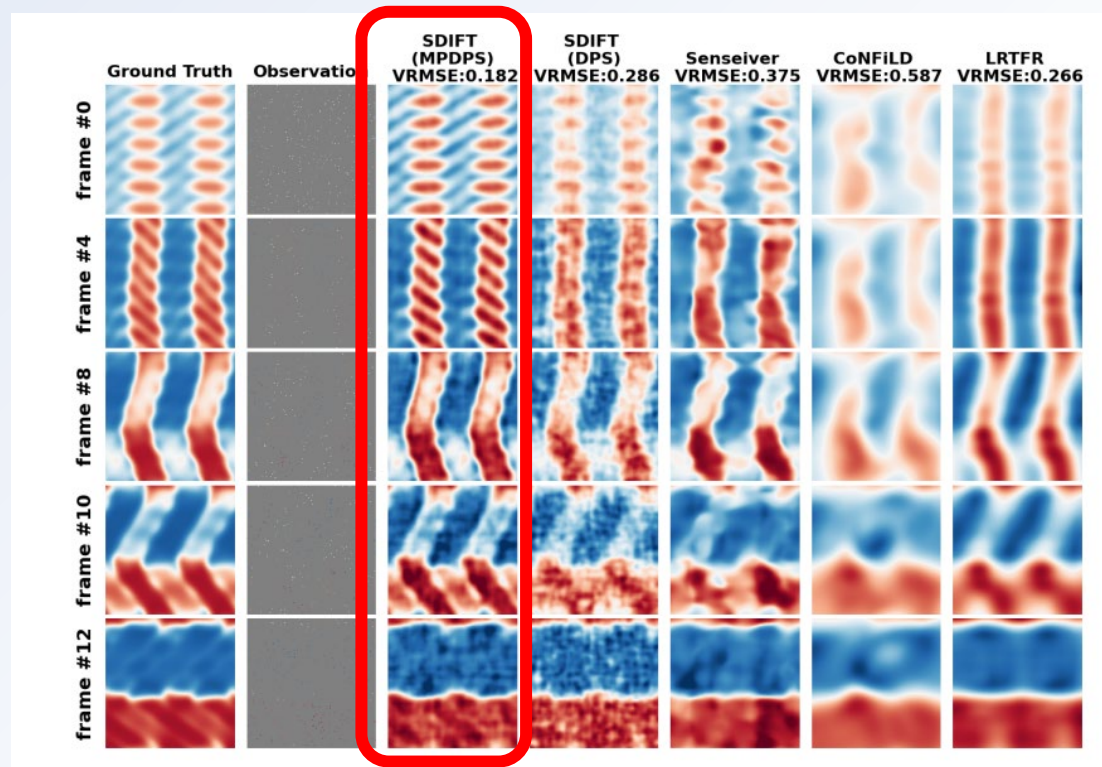
**声速场空时补全**  
**1%观测, 相较于CoNFILD**  
**(Nature Communications**  
**2024)**  
**补全误差下降46%**  
**补全速度提升20倍**

# 代表成果: SDIFT 生成稀疏物理场的全域演化 (NeurIPS 2025)

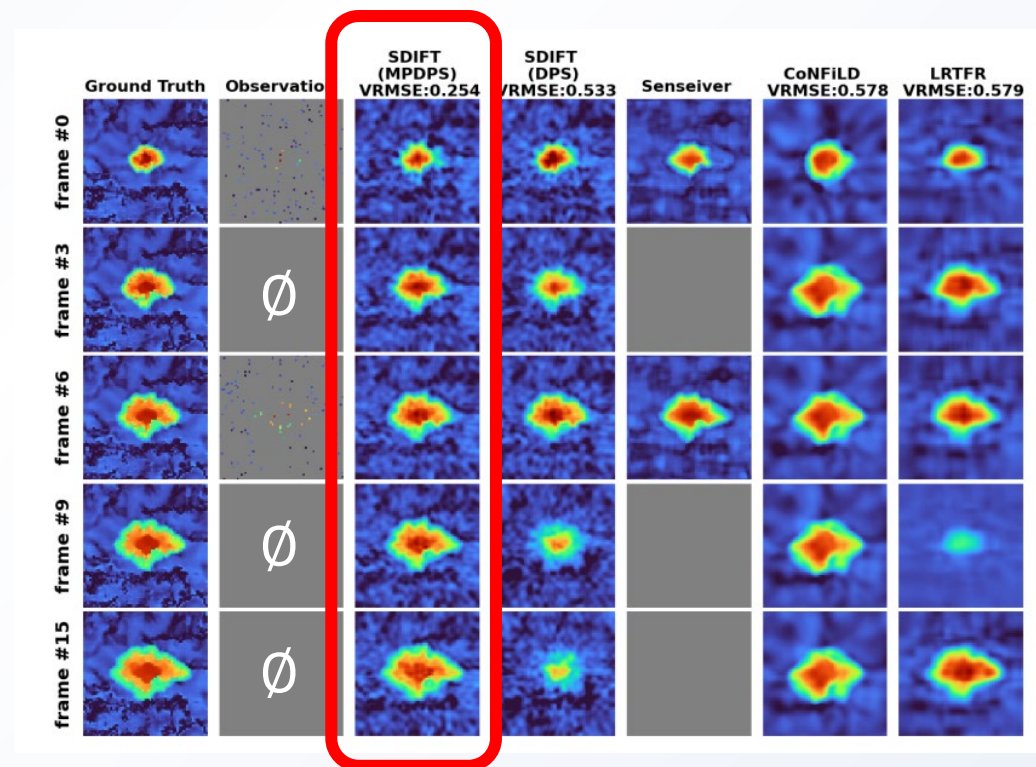


## 实际数据实验可视化

### 1%重建效果 (每个时刻都有观测)



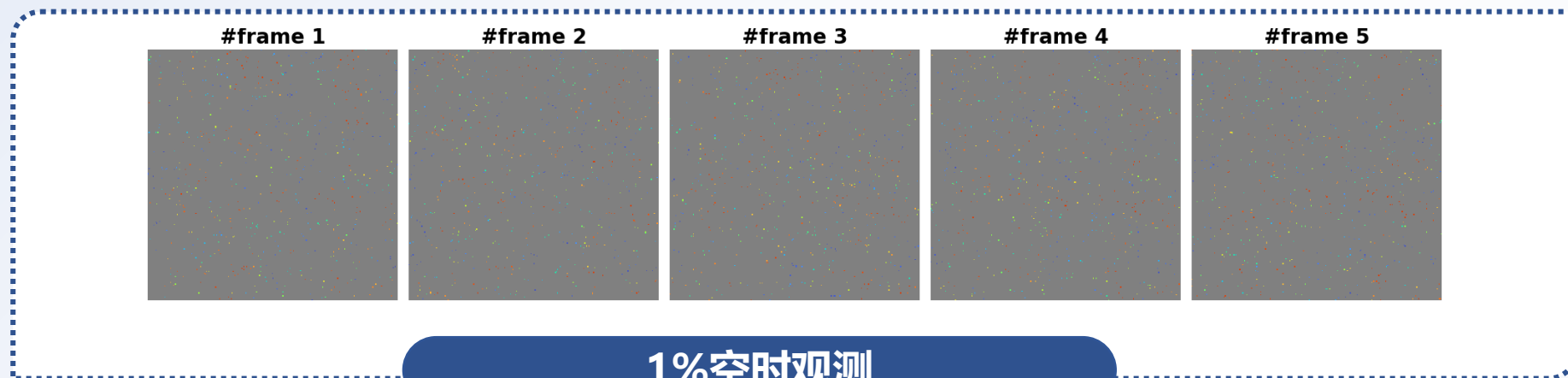
### 1%重建效果 (间隔时刻有观测)



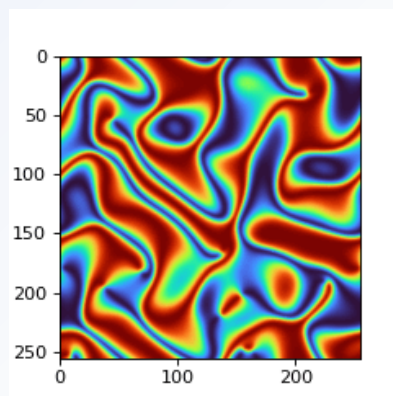
# 代表成果: SDIFT 生成稀疏物理场的全域演化 (NeurIPS 2025)



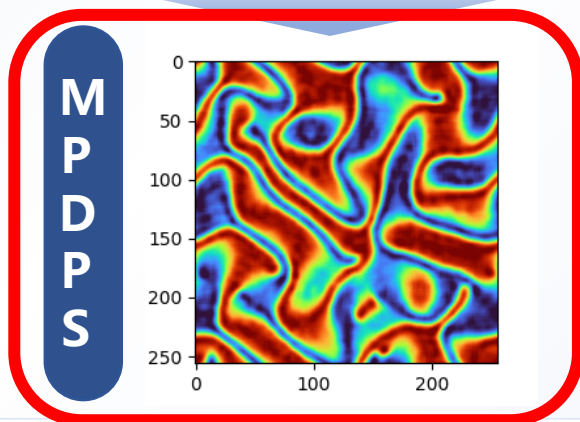
## 时序可视化效果



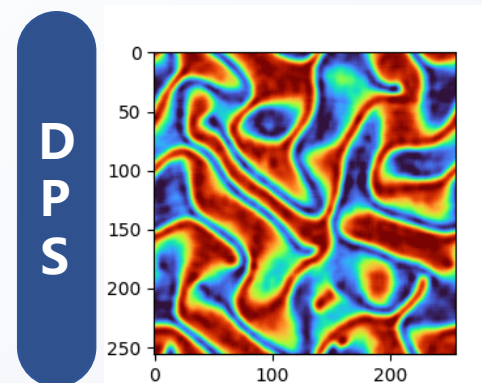
真实值



M  
P  
D  
P  
S



D  
P  
S



# • 生成式物理场仿真：总结

- 基于**生成范式**, 注入**物理约束**, 是下一代物理场仿真乃至**数字孪生**最有想象力的路线之一



# 目录



- 研究背景与挑战
- **相关成果:**
  - 广义动态张量学习
  - 生成式物理场仿真
  - **Agent 与产业基座模型**
- 未来方向

## • Agent 与产业基座模型



### 金融信号 原生基座模型：MarS



# MarS

[Paper](#) + [Project Website](#)

build pass license MIT version 1.0.0 python 3.11 | 3.12 platform linux PRs welcome docs latest

MARS: A FINANCIAL MARKET SIMULATION ENGINE POWERED BY GENERATIVE FOUNDATION MODEL

Junjie Li<sup>‡</sup>, Yang Liu<sup>‡</sup>, Weiqing Liu<sup>‡</sup>, Shikai Fang, Lewen Wang, Chang Xu & Jiang Bian  
Microsoft Research Asia  
{junli, yangliu2, weiqing.liu, fangshikai, lewen.wang, chanx, jiang.bian}@microsoft.com

### 工业级 R&D 自动化智能体：R&D Agent



# R&D-Agent

[Live Demo](#) | [Demo Video](#) | [YouTube](#) | [Documentation](#) | [Tech Report](#) | [Papers](#)

CI passing CodeQL passing Dependabot Updates passing Lint pull request title passing Release passing  
platform Linux pypi v0.8.0 python 3.10 | 3.11 release v0.8.0 license MIT pre-commit enabled mypy checked Ruff  
chat discord docs passing Read the Docs Pull Request Preview passing arXiv 2505.14738

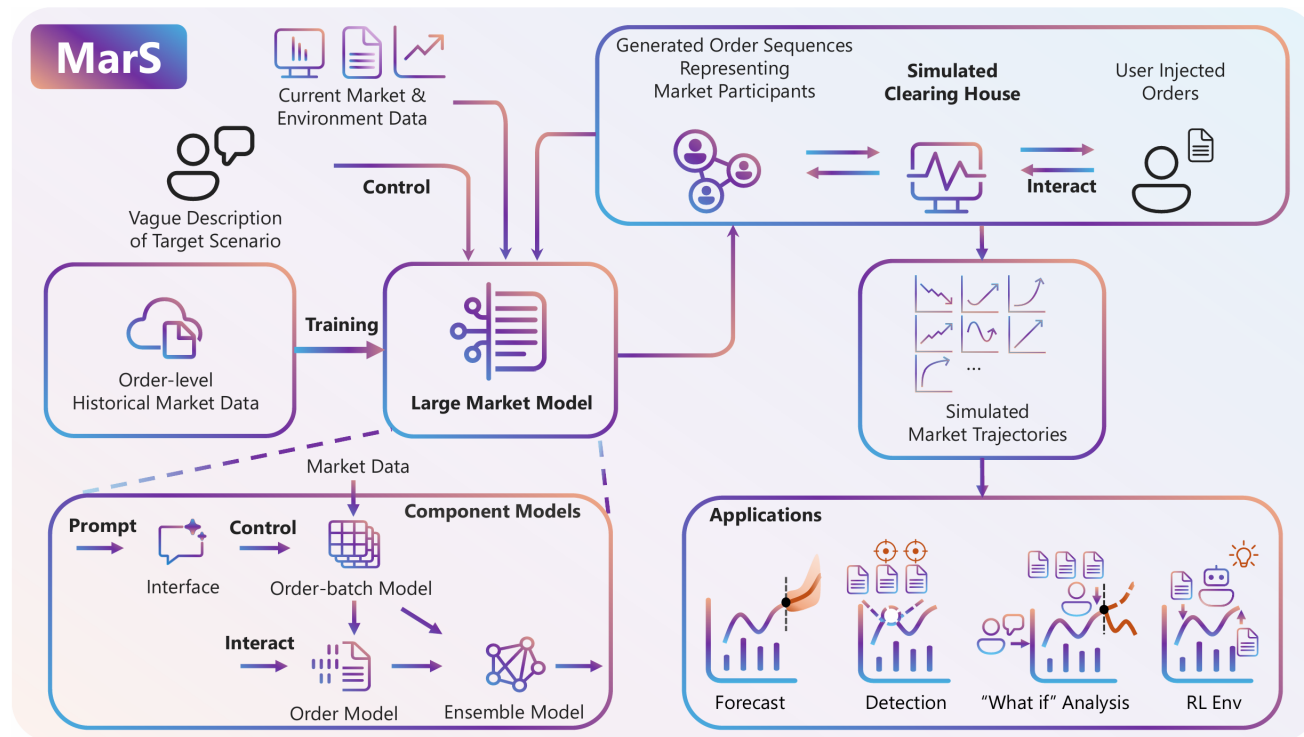
R&D-AGENT: AN LLM-AGENT FRAMEWORK TOWARDS AUTONOMOUS DATA SCIENCE

Xu Yang<sup>\*,†</sup>, Xiao Yang<sup>\*,†</sup>, Shikai Fang<sup>‡</sup>, Yifei Zhang<sup>‡</sup>, Jian Wang<sup>‡</sup>, Bowen Xian<sup>‡</sup>,  
Qizheng Li<sup>‡</sup>, Jingyuan Li<sup>‡</sup>, Minrui Xu<sup>‡</sup>, Yuante Li<sup>‡</sup>, Haoran Pan<sup>‡</sup>,  
Yuge Zhang<sup>‡</sup>, Weiqing Liu<sup>†,‡</sup>, Yelong Shen<sup>§</sup>, Weizhu Chen<sup>§</sup>, Jiang Bian<sup>‡</sup>  
<sup>‡</sup>Microsoft Research Asia, <sup>§</sup>Microsoft GenAI

# 金融场景原生基座模型：MarS

## 搭建产业原生基座模型的灵魂四问

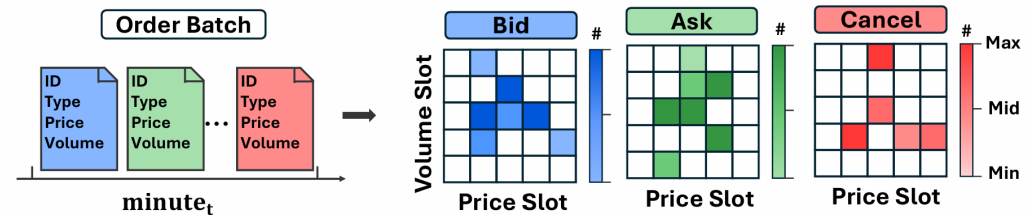
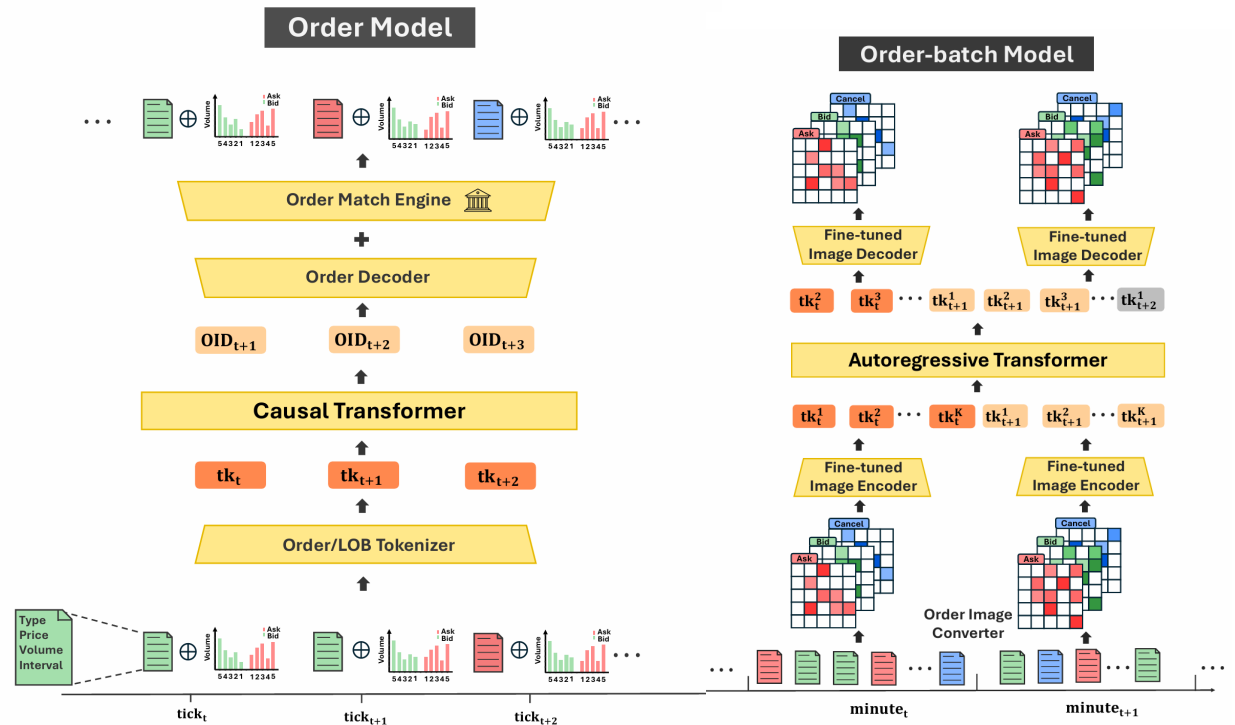
- 原子化的数据单元是什么？  
Tokenizer 怎么做？
- 能够Scaling吗？
- 自回归的范式如何对接下游任务？



# 金融场景原生基座模型：MarS

## 搭建产业原生基座模型的灵魂四问

- 原子化的数据单元是什么？  
Tokenizer 怎么做？
- 能够Scaling吗？
- 自回归的范式如何对接下游任务？

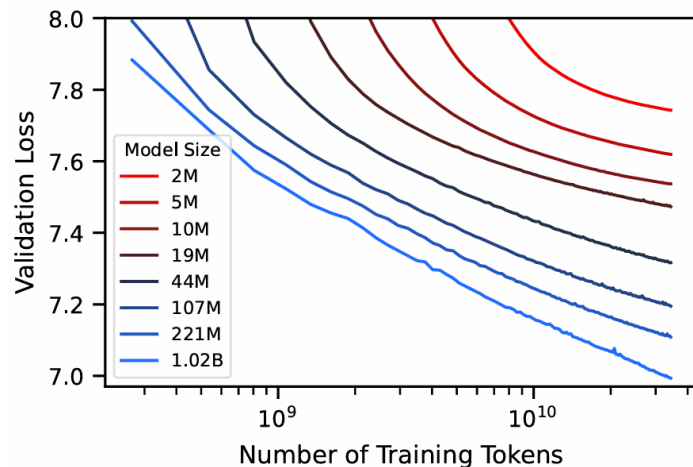


# 金融场景原生基座模型：MarS

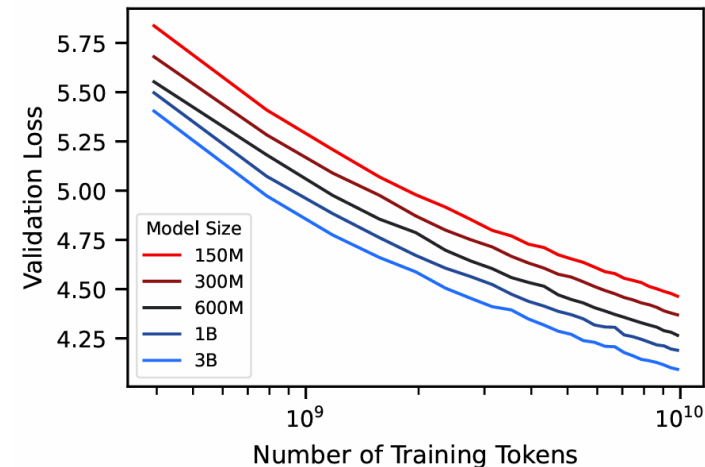


## 搭建产业原生基座模型的灵魂四问

- 原子化的数据单元是什么？  
Tokenizer 怎么做？
- 能够Scaling吗？
- 自回归的范式如何对接下游任务？



(a) Order Model

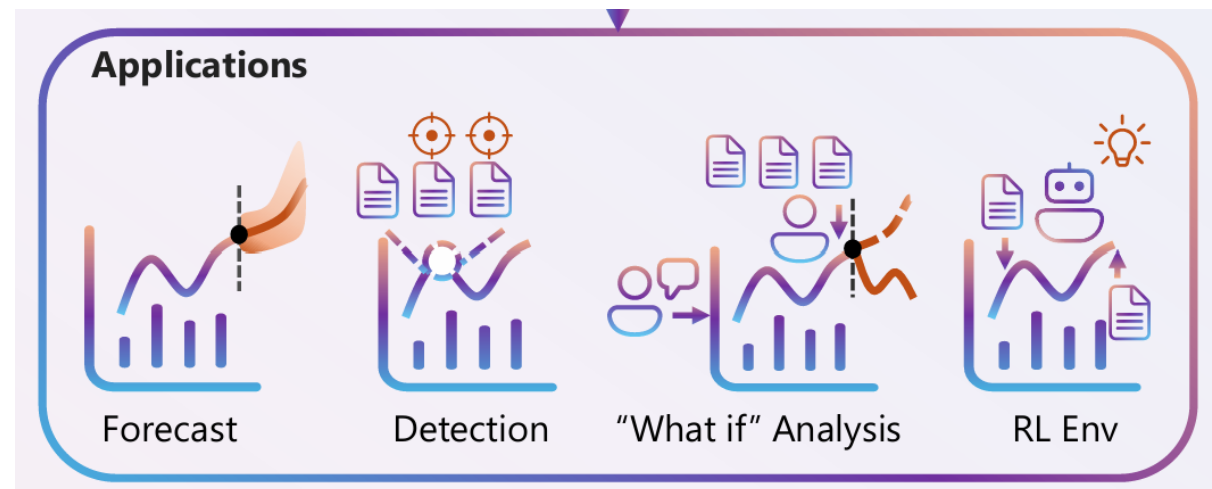


(b) Order-Batch Model

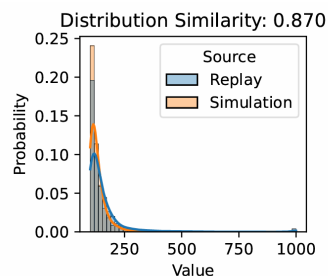
# 金融场景原生基座模型：MarS

## 搭建产业原生基座模型的灵魂四问

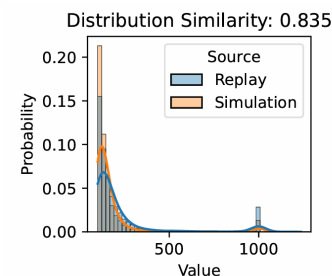
- 原子化的数据单元是什么？  
Tokenizer 怎么做？
- 能够Scaling吗？



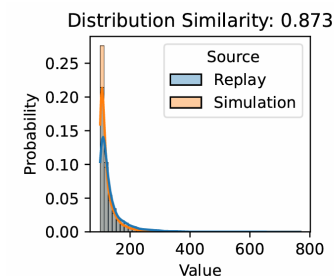
- 自回归的范式如何对接下游任务？



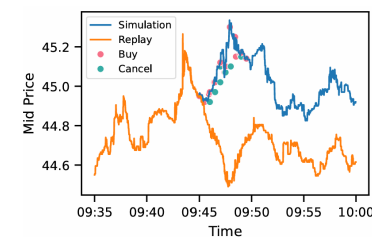
(a) Pre-manipulation



(b) Manipulation period



(c) Post-manipulation

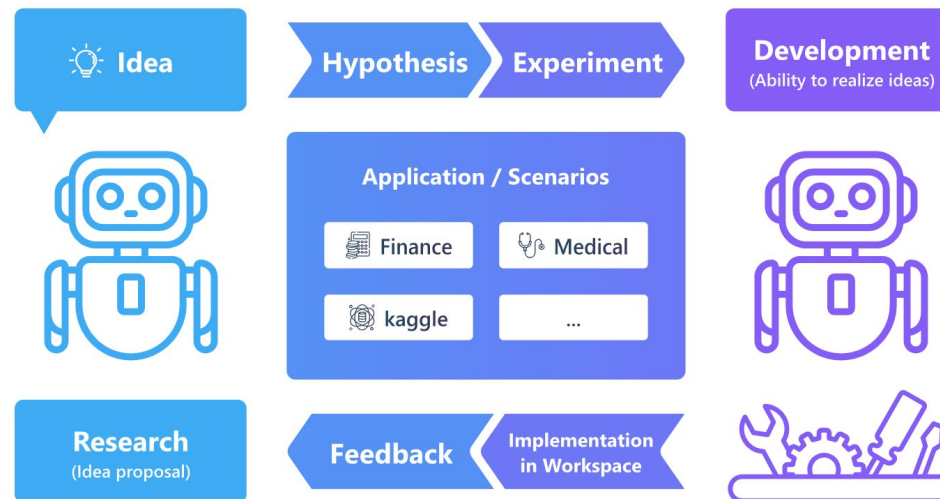


(a) Synthetic market interaction

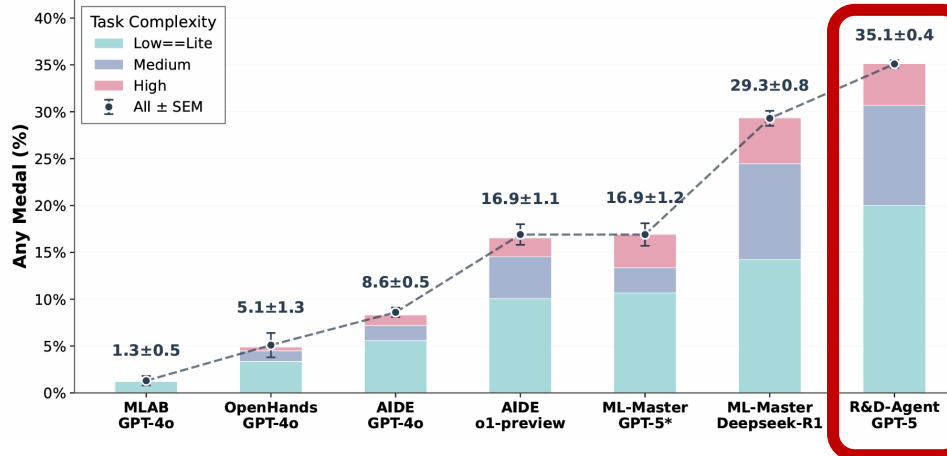
# 工业级 R&D 自动化智能体: R&D Agent

## 搭建工业级 Agent 的灵魂三问:

- 为什么需要Agent? 有清晰的 Input & Feedback 吗?
- 能够并发/高通量吗? 能够 Adaptive吗?
- Agent 怎么协作? 怎么平衡 探索-利用?



## OpenAI-MLE-Bench 上最好的开源模型!



# 工业级 R&D 自动化智能体: R&D Agent



## 搭建工业级 Agent 的灵魂三问:

- 为什么需要Agent? 有结构化的 Input & Feedback 吗?
- 能够并发/高通量吗? 能够 Adaptive吗?
- Agent 怎么协作? 怎么平衡 探索-利用?

### Hypotheses 🏆

	RD-Agent proposes the hypothesis ↓	because the reason ↓	based on the observation ↓	Knowledge ↓ gained after practice	action
1	Introducing new features derived from the raw features, such as calculating distances and ratios, will improve the model's performance by providing additional meaningful information.	Introducing derived features can reveal complex relationships and interactions, potentially improving model performance.	The dataset contains various cartographic variables that can be transformed into new meaningful features.	If new features are derived from existing variables, they may capture additional meaningful information and improve model performance.	Feature engineering

#### Structured Experiment Analysis — Memory Context Component

You are an advanced assistant analyzing results in data-driven R&D. Your task is to analyze the current experiment's hypothesis, implementation (code and its changes), and results, explicitly comparing them with previous best SOTA result step by step.

##### Step-by-step Analysis Process:

1. **Verify Submission Format:** Check format compliance and validity
2. **Evaluate Alignment with Competition Requirements:** Assess consistency with evaluation protocol
3. **Analyze Experimental Results:** Compare performance with SOTA and validate hypothesis

##### Key Analysis Components:

- **SOTA Comparison:** Direct comparison with historical best performance

25

- **Code Change Analysis:** Analyze implementation differences via diff
- **Performance Evaluation:** Score comparison with metric-aware reasoning
- **Hypothesis Validation:** Whether experimental results support or refute the hypothesis

##### Memory Integration Guidelines:

1. **Historical Context:** Reference previous similar attempts and their outcomes
2. **Pattern Recognition:** Identify recurring issues or successful strategies
3. **Knowledge Transfer:** Extract reusable insights for future experiments
4. **Risk Assessment:** Evaluate potential pitfalls based on historical failures

#### Competition Analysis Prompt — Planning Component

You are a data science assistant that extracts structured information from unstructured text. The user will provide you a Kaggle competition description, and you need to extract specific details from it.

Please answer in json format with the following schema:

- **“Task Type”:** The type of competition task, e.g., ‘Classification’, ‘Regression’, ‘Time-Series Forecasting’
- **“Data Type”:** The type of competition data, e.g., ‘Tabular’, ‘Time Series’, ‘Text’, ‘Image’, ‘Audio’
- **“Brief Description”:** A brief description of the competition
- **“Dataset Description”:** The dataset structure based on processed data folder description
- **“Submission Specifications”:** The submission specification & sample submission file descriptions
- **“Metric Evaluation Description”:** A precise explanation of how submissions are scored
- **“Metric Name”:** The name of the metric which this competition uses for scoring
- **“Metric Direction”:** True or False as True means bigger metric number is better
- **“Longer time limit required”:** True or False, whether the scenario requires a longer time limit

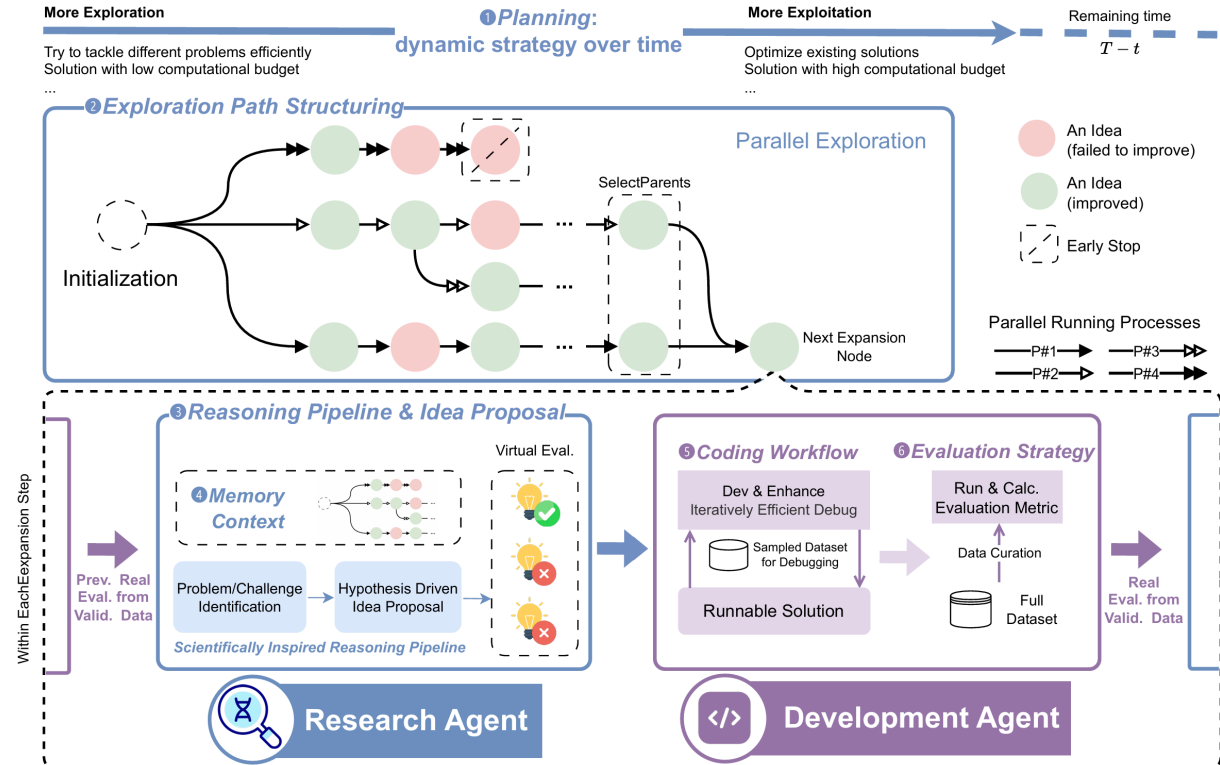
# 工业级 R&D 自动化智能体: R&D Agent



## 搭建工业级 Agent 的灵魂三问:

- 为什么需要Agent? 有结构化的 Input & Feedback 吗?
- 能够并发/高通量吗? 能充分 adaptive & Config 吗?
- Agent 怎么协作? 怎么平衡 探索-利用?

Phase	Design Aspect	Agent Designs				Frameworks	
		AIDE Jiang et al., 2025	ML-Master Liu et al., 2025	KompeteAI Kulibaba et al., 2025	MLE-STAR Nam et al., 2025	AIRA Toledo et al., 2025	R&D-Agent (Ours)
Research	Planning	/	/	/	/	/	⚙️ Dynamic
	Path Structuring	Tree (Greedy)	Tree + MCTS	Tree + Merging	Chain	⚙️ Tree + MCTS	⚙️ Adaptive
	Memory Context	/	Sibling	Sibling	Sibling	⚙️ Sibling	⚙️ Collab. Comm.
	Reasoning Pipeline	One step	One step	One step	One step	One step	⚙️ Scientific multi-step
Development	Coding Workflow	Node Debug	Node Debug	Debug	Debug	/	⚙️ Eff. & Iter. Debug
	Evaluation Strategy	/	/	/	/	/	⚙️ Aggregated



# 目录



- 研究背景与挑战
- 相关成果：
  - 广义动态张量学习
  - 生成式物理场仿真
  - Agent 与产业基座模型
- 未来方向

# • 未来方向?

• 未来的 **物理信号** 主流场景 = 更广泛的传感器收集到的 **“真实世界信号”** + 更多端侧智能体/基座模型交互的 **“智能互动信号”**



- **“自然语言”**会是最高效的通信媒介吗?

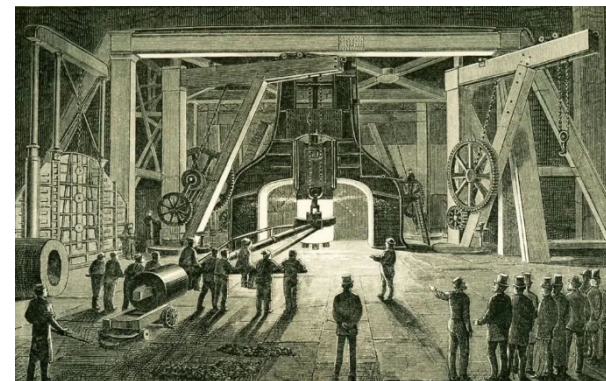
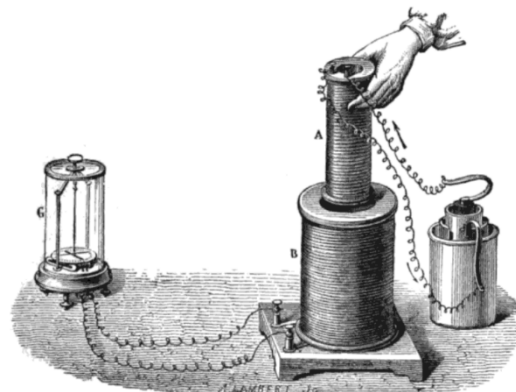
- **数据采集、存储、计算、建模**一体

- **多智能体的建模、协同、控制**一体

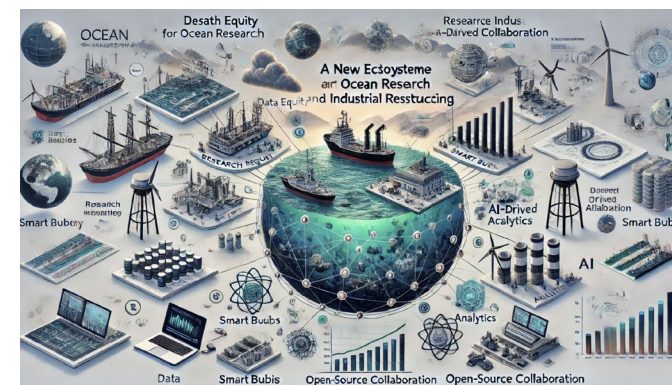
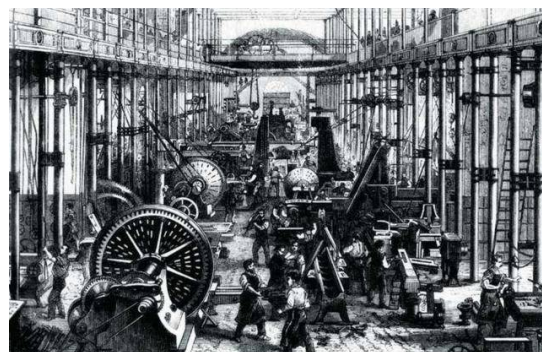
# 总结

现阶段的 Agent, 产业基座模型, 还处于技术路线未固定、需求场景未定型的阶段

- 工程 引领 学术
- 需求 引领 技术



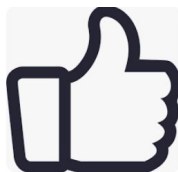
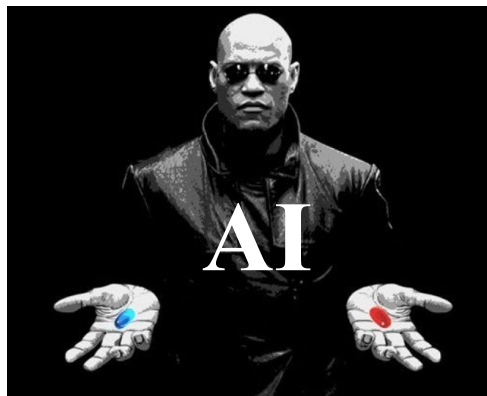
新的蒸汽机、内燃机、发电机的原型机刚刚确定, 但对应新的技术标准、工程机器、产业模式、乃至社会模式都还在诞生的前夜



# AI 向实：可上九天揽月，可下五洋捉鳖！

“Why we choose to go to the moon?”

“because they are hard!”



# Reference:

---



浙江大學  
ZHEJIANG UNIVERSITY

## **CATTE :Functional Complexity-adaptive Temporal Tensor Decomposition**

Panqi Chen, Lei Cheng†, Jianlong Li, Weichang Li, Weiqing Liu, Jiang Bian, **Shikai Fang†**

The Thirty-Ninth Annual Conference on Neural Information Processing Systems (**NeurIPS 2025**)

## **Generating Full-field Evolution of Physical Dynamics from Irregular Sparse Observations**

Panqi Chen, Yifan Sun, Lei Cheng†, Yang Yang, Weichang Li, Yang Liu, Weiqing Liu, Jiang Bian, **Shikai Fang†**

The Thirty-Ninth Annual Conference on Neural Information Processing Systems (**NeurIPS 2025**)

## **MarS: a Financial Market Simulation Engine Powered by Generative Foundation Model**

Junjie Li, Yang Liu, Weiqing Liu, **Shikai Fang**, Lewen Wang, Chang Xu, Jiang Bian

The 13th International Conference on Learning Representations (**ICLR 2025**)

## **- R&D-Agent: Automating Data-Driven AI Solution Building Through LLM-Powered Automated Research, Development, and Evolution**

Xu Yang\*, Xiao Yang\*,**Shikai Fang**, Bowen Xian, Yuante Li, Jian Wang, Minrui Xu, Haoran Pan, Xinpeng Hong, Weiqing Liu, Yelong Shen, Weizhu Chen, Jiang Bian



**感谢聆听!**