浙江大学
ZHEJIANG UNIVERSITY

CSML TALK • 2026

从低秩表征到生成式先验

物理场稀疏重构中的表征学习路径

方榭楷 Shikai Fang

浙江大学百人计划研究员 · 信息与电子工程学院

低秩函数 → 生成分布 → 在线状态



个人介绍



方楒楷 博士

百人计划研究员 | 博士生导师

单位 信息与电子工程学院

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

PROBLEM

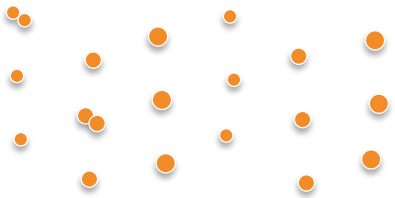
从稀疏离网观测恢复连续动态场

输入：观测集合

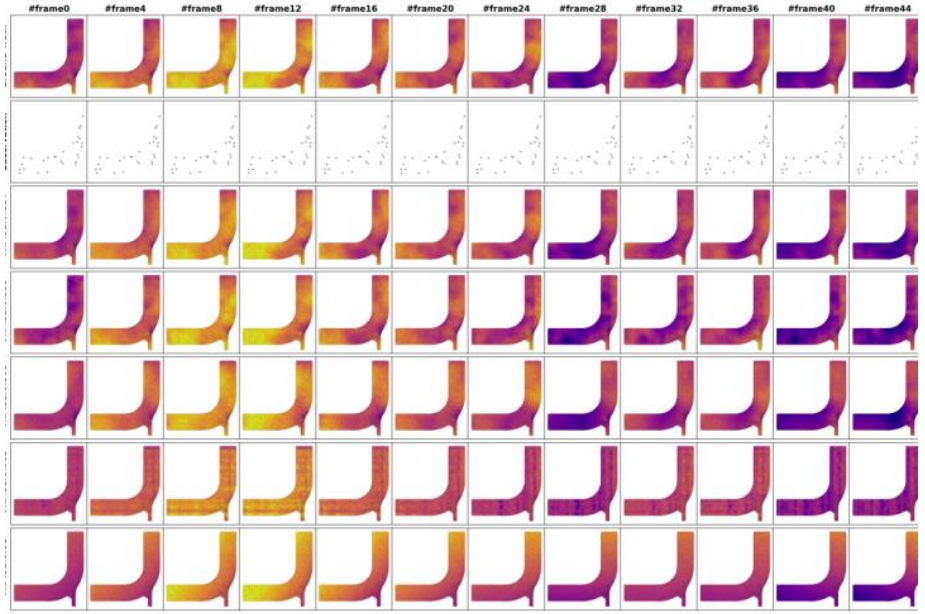
$$\mathcal{O}_t = \{(x_n, y_n, t_n)\}$$

有限　噪声

不规则　持续到达



conditional inference



re 4: Reconstruction of *Turbulent Flow* dynamics under uniform sampling pattern with $\rho =$

连续・高维・动态・不确定性可表达

观测几何直接影响可辨识性与重构精度

输出：任意坐标、任意时刻可查询的完整物理场

ROADMAP

浙江大学 ZHEJIANG UNIVERSITY

表征决定压缩、先验与推断

观测 o

1 · 连续化

FunBaT

低秩潜在函数

field as functions

2 · 生成化

SDIFT

core 序列的分布

dynamics as distribution

3 · 流式化

StreamPhy

持续更新的状态

history as state

一条主线：从表示一个场，到学习一类场，再到让推断随数据流动。

张量分解：低秩表征的离散原型

- Multi-dim array for high-order structural data

Entry: (index1, index2..)-> value ⇔ Interaction of multiple objects

integer coordinates!

Social Network



(user,user, message)

Online Ads



(item, group, site, device)

Traffic Flow



(city, road, population, period)

一个 entry 由多个 mode 的对象共同定位

mode 患者 / 医生 / 医院 / 时间

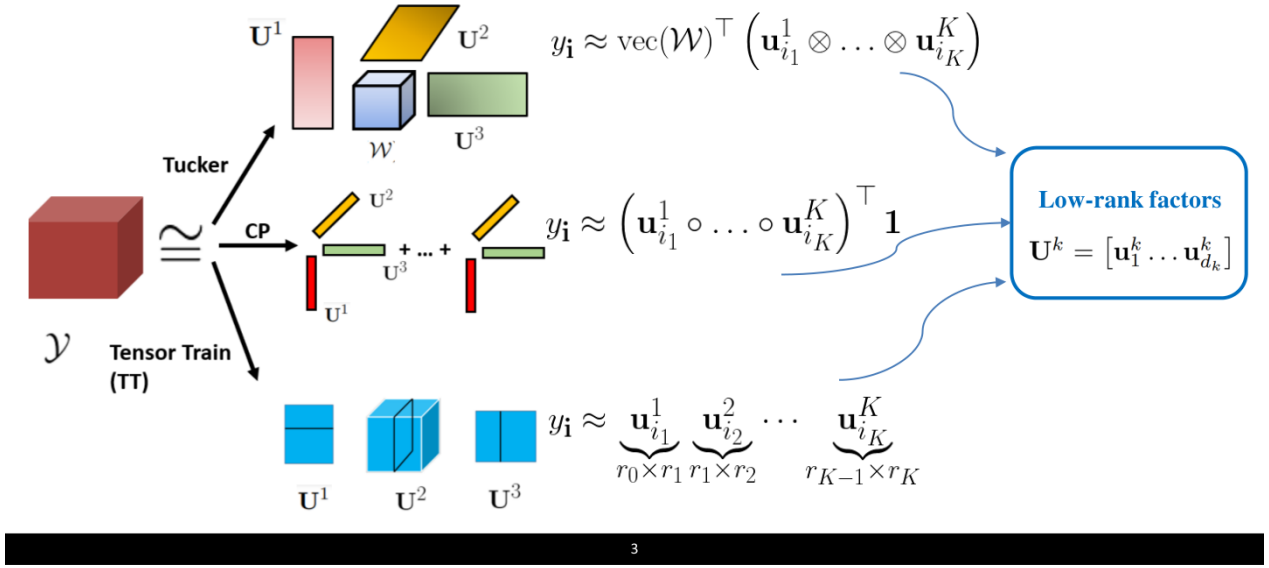
index 离散整数：1, 2, …, I_k

entry 多对象交互产生的数值

低秩假设：少量因子刻画高维交互。

Tucker, Some Mathematical Notes on Three-Mode Factor Analysis, Psychometrika, 1966.

Tucker 分解：因子表示对象，core 表示交互



Tucker 在容量、灵活性与可解释性之间提供平衡

$$y_i \approx \text{vec}(W)^T (u^1{}_{i1} \otimes \cdots \otimes u^K{}_i{}^K)$$

1. **factor matrices** 每个 mode 内对象的低维表示
2. **core tensor W** 跨 mode 的全局交互规则
3. **秩 R** 压缩率与表达力之间的旋钮

问题：当索引变为经纬度和时间等连续坐标，factor matrix 还够吗?

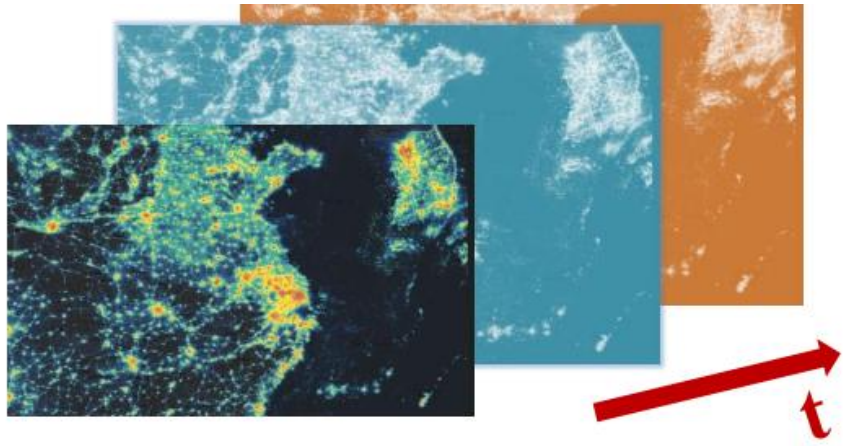
Tucker, Some Mathematical Notes on Three-Mode Factor Analysis, Psychometrika, 1966.

物理场的自然索引是连续坐标

Functional Bayesian Tucker Decomposition for Continuous-Indexed Tensor Data

Shikai Fang, Xin Yu, Zheng Wang, Shibo Li, Robert M. Kirby, Shandian Zhe · ICLR 2024

ICLR 2024



latitude · longitude · height · time

离散化的三个代价

- **细粒度信息丢失** — binning 先把连续结构压扁
- **分辨率需要人工选择** — 不同网格带来不同误差
- **新坐标不能自然查询** — 离网预测通常需要额外插值

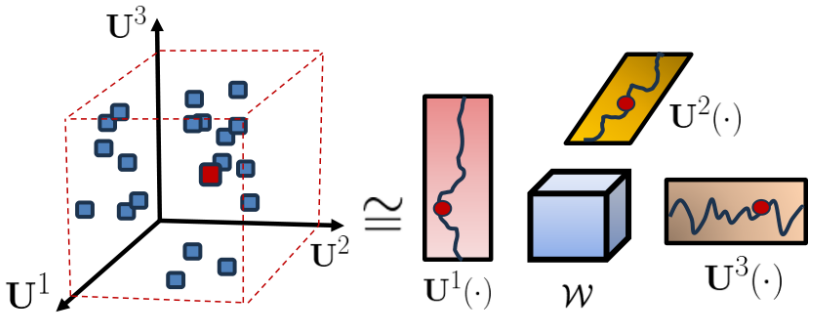
把每个 mode 的 factor 从向量提升为坐标函数。

[1] Functional Bayesian Tucker Decomposition for Continuous-Indexed Tensor Data, ICLR 2024.

Functional Tucker：将因子提升为坐标函数

Tensor entry ⇔ Tucker-form interaction of latent functions

$$f(\mathbf{i}) = f(i_1, \dots i_K) \approx \mathrm{vec}(\mathcal{W})^\top \left(\mathbf{U}^1(i_1) \otimes \dots \otimes \mathbf{U}^K(i_K)\right)$$



$$\mathbf{f(i) \approx vec(W)^T [U^1(i_1) \otimes \cdots \otimes U^K(i^K)]}$$

$\mathbf{U^k(i_k)}$ 沿第 k 个连续坐标轴变化的潜在函数

$\mathbf{W}$ 跨空间、时间等维度的低秩交互

$\mathbf{f(i)}$ 任意坐标 i 上直接定义的场值

连续坐标直接进入潜在函数。

[1] Functional Bayesian Tucker Decomposition for Continuous-Indexed Tensor Data, ICLR 2024.

Gaussian Process：函数先验、核与后验不确定性

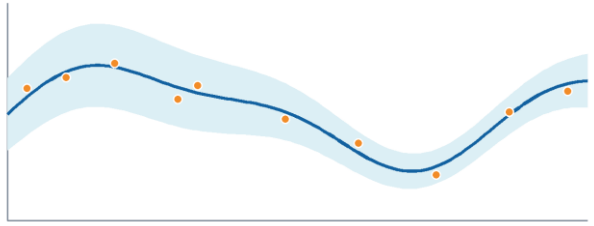
标准定义

$$f(x) \sim GP(m(x), k(x, x'))$$

$$y = f(X) + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I)$$

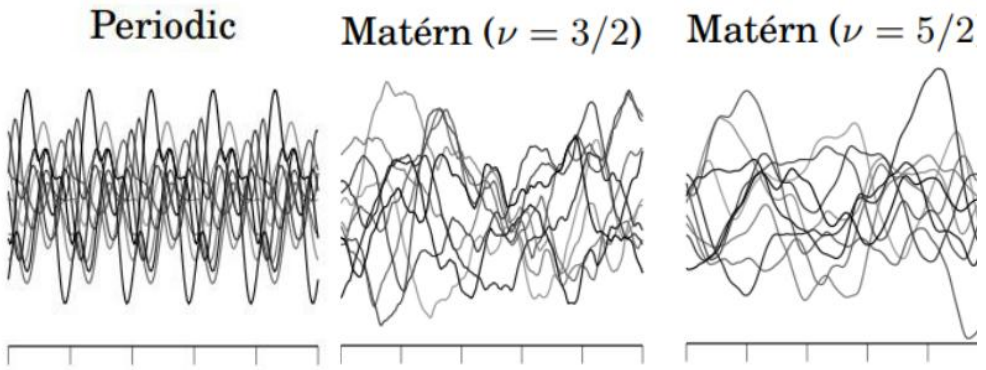
$$\mu(x^*) = K(x^*,X)[K(X,X) + \sigma^2 I]^{-1} y$$

$$\Sigma(x^*) = K(x^*,x^*) - K(x^*,X)[K(X,X) + \sigma^2 I]^{-1} K(X,x^*)$$



观测点 · 后验均值 · 预测不确定性

核函数编码结构先验



- **Periodic** 周期结构
- **Matérn 3/2** 较低光滑度、局部变化更快
- **Matérn 5/2** 更平滑的函数轨迹

Rasmussen & Williams, Gaussian Processes for Machine Learning, MIT Press, 2006; [1] FunBaT, ICLR 2024.

SS-GP：从 GP 核到线性状态空间模型

Linear-Cost GP with Chain-Structure

Temporal GPs

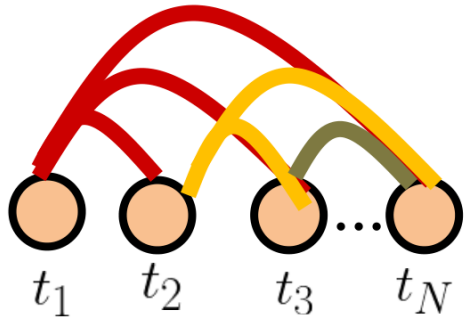
Temporal States:

LTI-SDE

$$\frac{\mathrm{d}\gamma_{\mathbf{r}}(t)}{\mathrm{d}t} = \mathbf{F}\gamma_{\mathbf{r}} + \mathbf{L}\xi(t)$$

discrete form

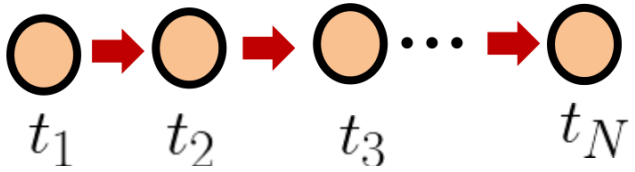
State Space Model



Space: $\mathcal{O}(N^2)$

Time: $\mathcal{O}(N^3)$

$$p\left(\gamma_{\mathbf{r}}\left(t_{n+1}\right) \mid \gamma_{\mathbf{r}}\left(t_n\right)\right) = \mathcal{N}\left(\gamma_{\mathbf{r}}\left(t_{n+1}\right) \mid \mathbf{A}_n\gamma_{\mathbf{r}}\left(t_n\right), \mathbf{Q}_n\right)$$



Space: $\mathcal{O}(N)$

Time: $\mathcal{O}(N)$

[1] Functional Bayesian Tucker Decomposition for Continuous-Indexed Tensor Data, ICLR 2024.

FunBaT：条件期望传播（CEP）与状态空间推断

Algorithm 1 FunBaT

Input: Observations $\mathcal{D}$ of a K-mode continuous-indexed tensor , kernel hyperparameters, sorted unique indexes set $\{\mathcal{I}_k\}$ of each mode.
Initialize approx. posterior $q(\tau)$, $q(\mathcal{W})$, $\{q(\mathbf{Z}^k)\}$ and message factors for each likelihood.
repeat
  for $k = 1$ **to** K **do**
    Approximate messages factors $\{f_n(\mathbf{Z}^k(i_k^n)), f_n(\tau), f_n(\mathcal{W})\}_{n=1}^N$ in parallel with CEP (13)
    Update the approximated posterior $q(\tau)$, $q(\mathcal{W})$ by merging the message (14)(15).
    Update $q(\mathbf{Z}^k)$ sequentially by KF and RTS smoother based on (16)
  end for
until Convergence
Return: $q(\tau), q(\mathcal{W}), \{q(\mathbf{Z}^k)\}_{k=1}^K$

时间复杂度 O(NKR)

N：观测数 · K：维度数 · R：各 mode 的秩

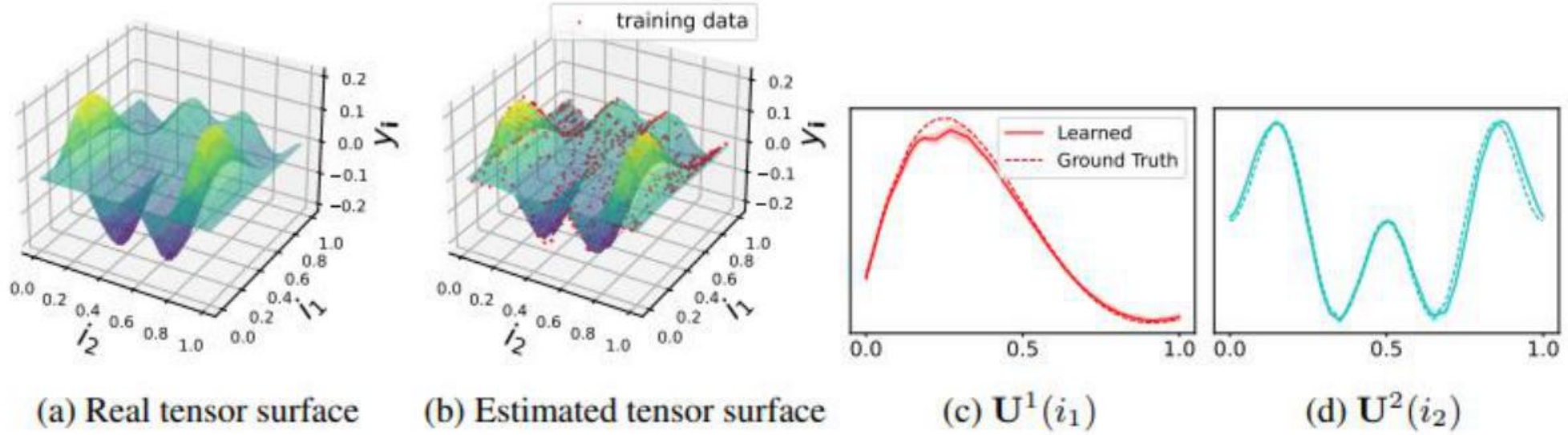
Algorithm 1：消息 → 全局参数 → 各 mode 状态链

1. **CEP 并行构造消息** — $f_n(Z^k)$, $f_n(W)$, $f_n(\tau)$
2. **合并全局消息** — 更新 $q(W)$ 与 $q(\tau)$
3. **沿坐标链顺序推断** — Kalman filter + RTS smoother
4. **重复至收敛** — 返回 $q(W)$, $q(\tau)$, $\{q(Z^k)\}$

观测 → 高斯消息 → 各 mode 的状态空间链

[1] Functional Bayesian Tucker Decomposition for Continuous-Indexed Tensor Data, ICLR 2024.

合成实验：场、潜在函数与后验不确定性



(a) Real tensor surface (b) Estimated tensor surface (c) $\mathbf{U}^1(i_1)$ (d) $\mathbf{U}^2(i_2)$

Figure 1: Results of Synthetic Data

[1] Functional Bayesian Tucker Decomposition for Continuous-Indexed Tensor Data, ICLR 2024.

真实数据实验：重构误差（RMSE / MAE）

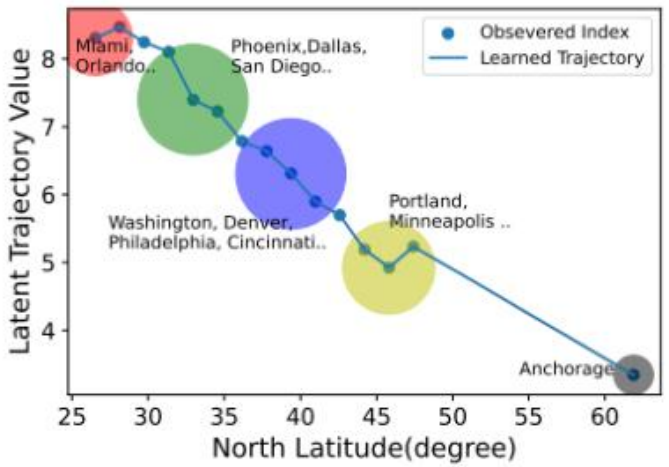
BeijingAir：2014–2017 年北京空气质量记录；分别预测 PM2.5、PM10 与 SO_2。

	RMSE			MAE		
Datasets	*PM2.5*	*PM10*	*SO2*	*PM2.5*	*PM10*	*SO2*
Resolution: $50 \times 50 \times 150$						
P-Tucker	0.805 ± 0.017	0.787 ± 0.006	0.686 ± 0.02	0.586 ± 0.003	0.595 ± 0.005	0.436 ± 0.011
Tucker-ALS	1.032 ± 0.049	1.005 ± 0.029	0.969 ± 0.027	0.729 ± 0.016	0.741 ± 0.007	0.654 ± 0.034
Tucker-SVI	0.792 ± 0.01	0.8 ± 0.026	0.701 ± 0.08	0.593 ± 0.01	0.605 ± 0.019	0.423 ± 0.031
Resolution: $100 \times 100 \times 300$						
P-Tucker	0.8 ± 0.101	0.73 ± 0.021	0.644 ± 0.023	0.522 ± 0.011	0.529 ± 0.013	0.402 ± 0.008
Tucker-ALS	1.009 ± 0.027	1.009 ± 0.026	0.965 ± 0.023	0.738 ± 0.01	0.754 ± 0.007	0.68 ± 0.011
Tucker-SVI	0.706 ± 0.011	0.783 ± 0.067	0.69 ± 0.086	0.509 ± 0.008	0.556 ± 0.031	0.423 ± 0.031
Methods using continuous indexes						
FTT-ALS	1.020 ± 0.013	1.001 ± 0.013	1.001 ± 0.026	0.744 ± 0.007	0.755 ± 0.007	0.696 ± 0.011
FTT-ANOVA	2.150 ± 0.033	2.007 ± 0.015	1.987 ± 0.036	1.788 ± 0.031	1.623 ± 0.014	1.499 ± 0.018
FTT-cross	0.942 ± 0.025	0.933 ± 0.012	0.844 ± 0.026	0.566 ± 0.018	0.561 ± 0.011	0.467 ± 0.033
RBF-SVM	0.995 ± 0.015	0.955 ± 0.02	0.794 ± 0.026	0.668 ± 0.008	0.674 ± 0.014	0.486 ± 0.026
BLR	0.998 ± 0.013	0.977 ± 0.014	0.837 ± 0.021	0.736 ± 0.007	0.739 ± 0.008	0.573 ± 0.009
FunBaT-CP	0.296 ± 0.018	0.343 ± 0.028	$\mathbf{0.386 \pm 0.009}$	$\mathbf{0.18 \pm 0.002}$	0.233 ± 0.013	0.242 ± 0.003
FunBaT	$\mathbf{0.288 \pm 0.008}$	$\mathbf{0.328 \pm 0.004}$	$\mathbf{0.386 \pm 0.01}$	0.183 ± 0.006	$\mathbf{0.226 \pm 0.002}$	$\mathbf{0.241 \pm 0.004}$

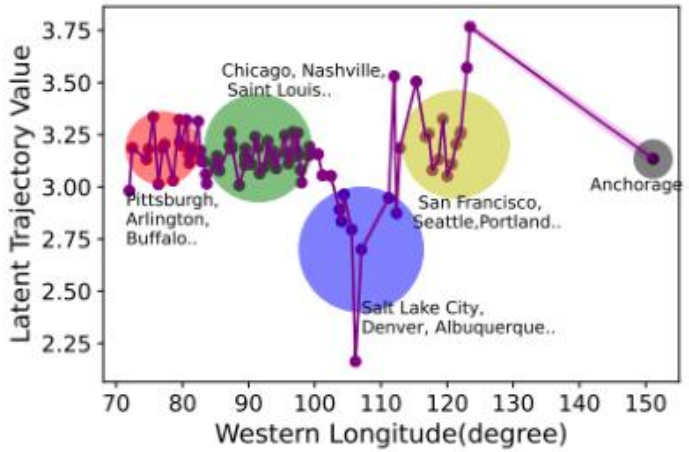
[1] Functional Bayesian Tucker Decomposition for Continuous-Indexed Tensor Data, ICLR 2024.

可解释性：潜在函数的地理与气候结构

US-TEMP：248 个城市，1750–2016 年气温记录；三个连续 mode 分别对应纬度、经度和时间。



(a) $\mathbf{U^1}(i_1)$: Mode of latitude



(b) $\mathbf{U^2}(i_2)$: Mode of longitude

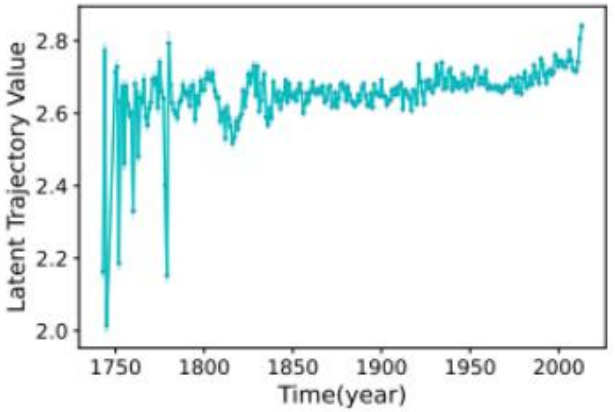


Figure 3: $\mathbf{U^3}(i_3)$: Mode of time

[1] Functional Bayesian Tucker Decomposition for Continuous-Indexed Tensor Data, ICLR 2024.

TRANSITION

从直推式重构到归纳式生成建模

FunBaT：直推式学习

transductive

稀疏样本 1 —— 初始化 + 迭代 —— 重构

稀疏样本 2 —— 初始化 + 迭代 —— 重构

稀疏样本 3 —— 初始化 + 迭代 —— 重构

仅针对当前实例推断后验；新实例需从头推断。

SDIFT：归纳式学习

inductive

训练场集合 —— 生成式先验 $p\theta(W_{1:t})$

新稀疏观测 —— 条件后验采样 完整场演化

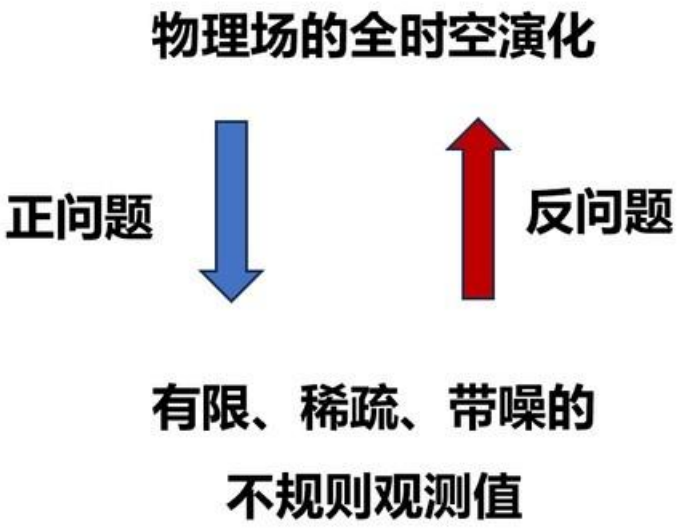
学习跨场景共享分布；新样本通过观测条件化。

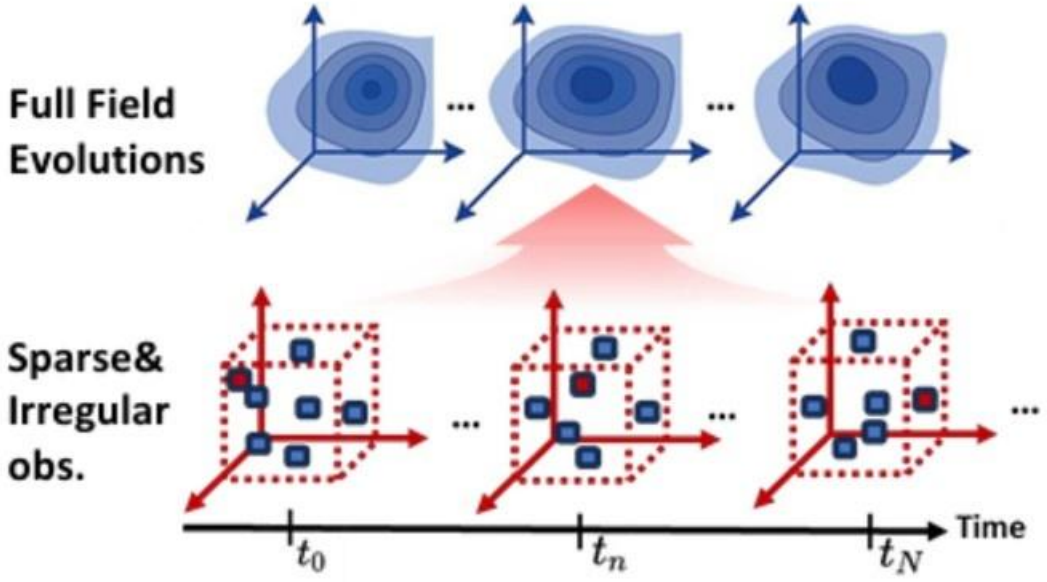
[1] Functional Bayesian Tucker Decomposition (ICLR 2024); [2] Generating Full-field Evolution (NeurIPS 2025).

• 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

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





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

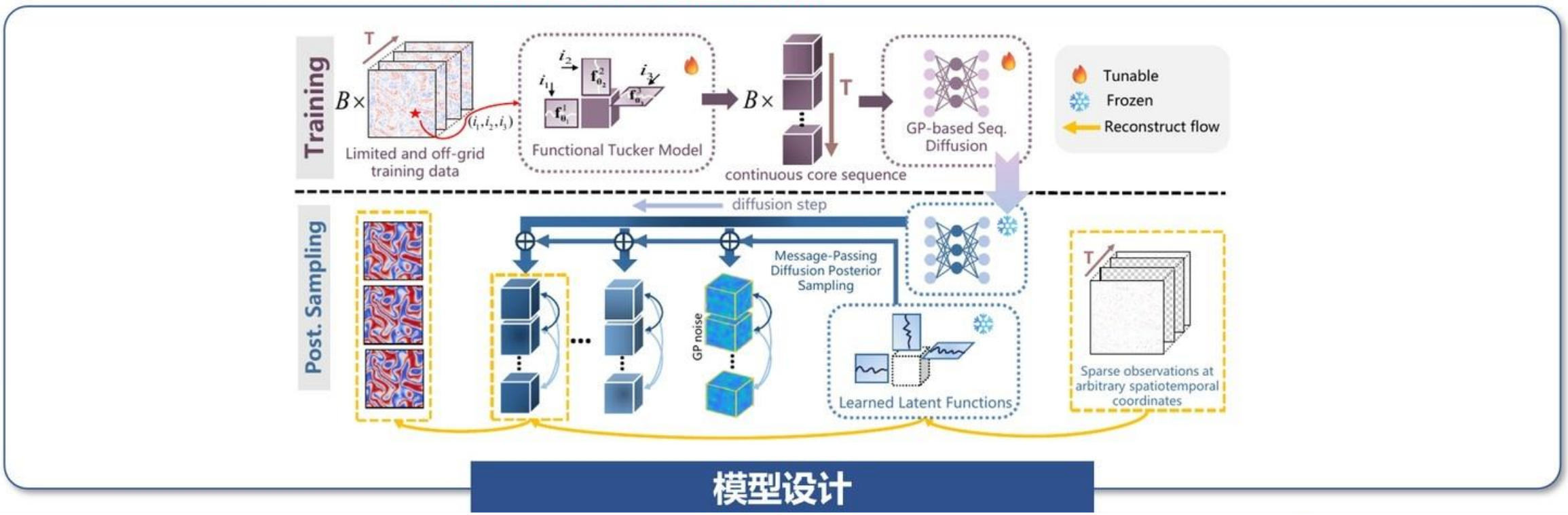
[2] Generating Full-field Evolution of Physical Dynamics from Irregular Sparse Observations, NeurIPS 2025.

• 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

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

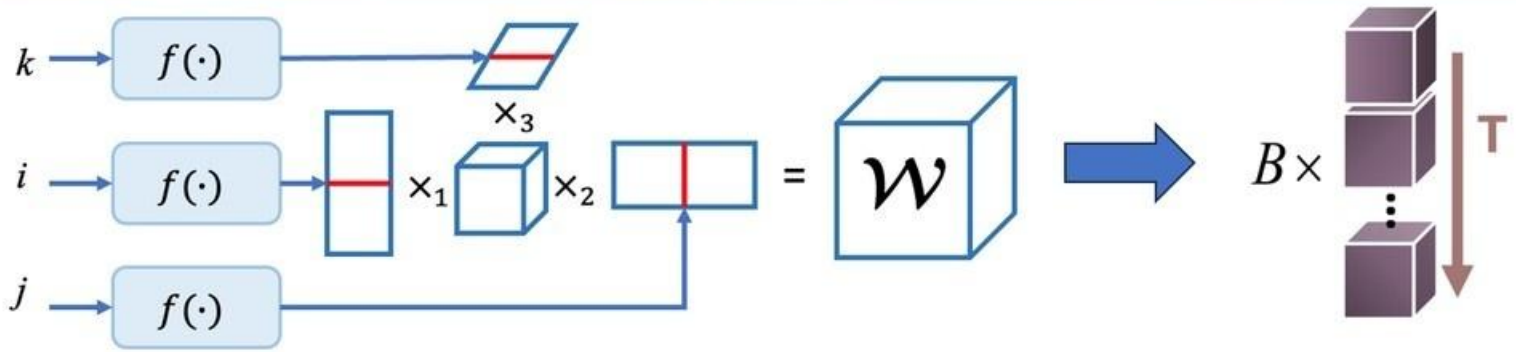
[2] Generating Full-field Evolution of Physical Dynamics from Irregular Sparse Observations, NeurIPS 2025.

• 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

泛函张量Tucker编码器

将高维连续物理场表征为隐空间中的张量核序列



$$y_{\mathbf{i},t_m} \approx \mathrm{vec}(\mathcal{W}_{t_m})^\top \left(\mathbf{f}^1_{\boldsymbol{\theta}_1}(i_1) \otimes \cdots \otimes \mathbf{f}^K_{\boldsymbol{\theta}_K}(i_K)\right).$$

泛在逼近性证明

Theorem 1. *(UAP of FTM) Let $X_1, \cdots, X_K$ 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 $\{R_1 > 0, \cdots, R_K > 0\}$, coefficients $\{a_{r_1,\cdots,r_K}\}_{r_1,\cdots,r_K}^{R_1,\cdots,R_K}$ and neural networks $\{\{f_{r_k}^k\}_{r_k}^{R_k}\}_k^K$ such that*

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

[2] Generating Full-field Evolution of Physical Dynamics from Irregular Sparse Observations, NeurIPS 2025.

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

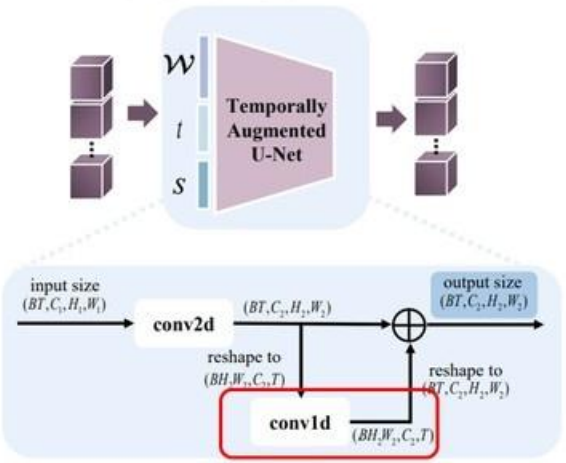
浙江大学 ZHEJIANG UNIVERSITY

高斯过程驱动的隐空间序列扩散模型

高斯过程噪声源

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

时间维度增强U-Net



在隐空间中一次生成整个序列！

目标函数

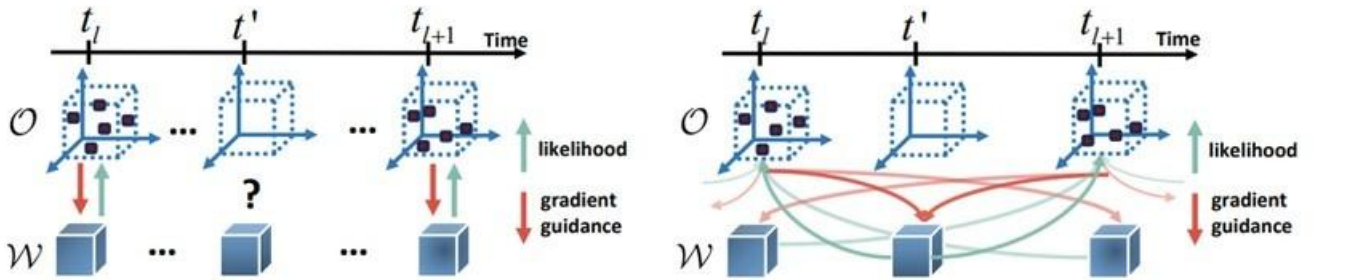
$$\mathcal{L}_{\text{GPSD}} = \mathbb{E}_{\sigma(s), \boldsymbol{\mathcal{W}}_{t_m} \sim \mathcal{W}, \boldsymbol{\mathcal{E}}_{t_m} \sim \mathcal{E}, t \sim \mathcal{T}} \left[\lambda(\sigma(s)) \left\| D_\theta(\boldsymbol{\mathcal{W}}_{t_m} + \boldsymbol{\mathcal{E}}_{t_m}; \sigma(s), t_m) - \boldsymbol{\mathcal{W}}_{t_m}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

[2] Generating Full-field Evolution of Physical Dynamics from Irregular Sparse Observations, NeurIPS 2025.

• 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

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:

来自其他时刻观测的聚合信息

$$\nabla_{\boldsymbol{\mathcal{W}}^s_{t_n}} \log p(\mathcal{O}|\boldsymbol{\mathcal{W}}^s_{t_n}) \approx \underbrace{\mathbf{1}_{\mathcal{T}_{\text{obs}}}(t_n) \cdot \nabla_{\boldsymbol{\mathcal{W}}^s_{t_n}} \log p(\mathcal{O}_{t_n}|\boldsymbol{\mathcal{W}}^s_{t_n})}_{\text{来自当前观测的后验信息}} + \sum_{t_l \in \mathcal{T}_{\text{obs}} \setminus t_n} \mathcal{G}^s_{t_l,t_n},$$

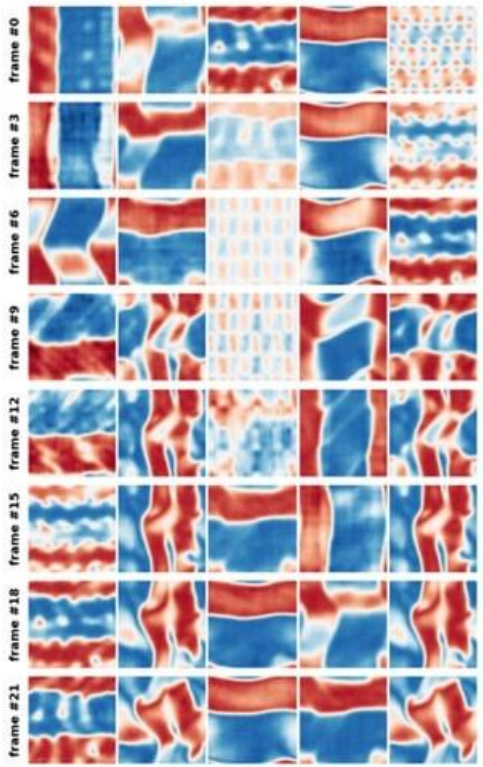
似然函数具有二次项结构，梯度计算效率高！

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

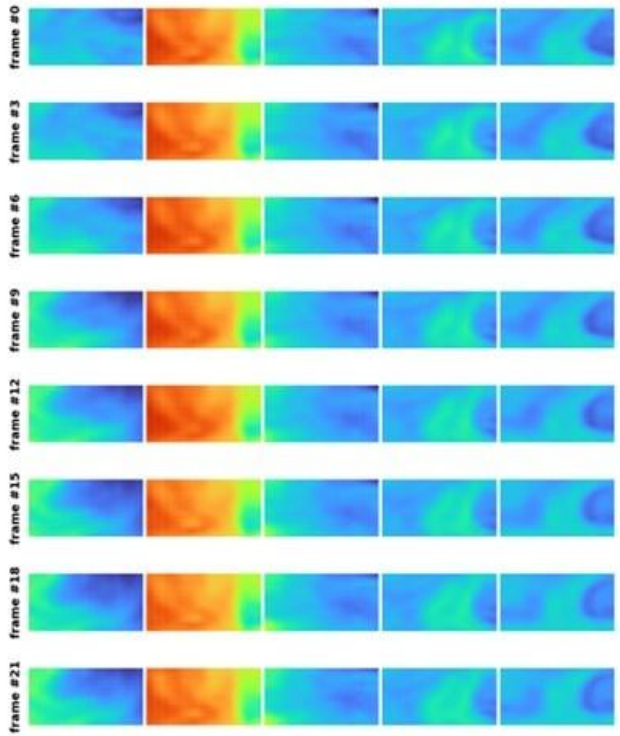
[2] Generating Full-field Evolution of Physical Dynamics from Irregular Sparse Observations, NeurIPS 2025.

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

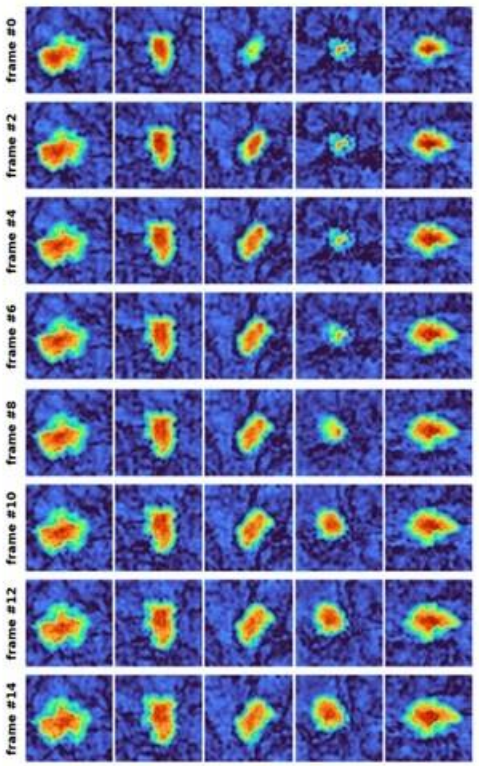
实验验证：从毫米到光年的三个动态物理场



毫米尺度：
活性物质



千米尺度：
海洋声速场演化



光年尺度：
超新星爆炸

[2] Generating Full-field Evolution of Physical Dynamics from Irregular Sparse Observations, NeurIPS 2025.

真实数据实验：重构误差（VRMSE）

Methods	*Supernova Explosion* $\rho = 1\%$	*Supernova Explosion* $\rho = 3\%$	*Ocean Sound Speed* $\rho = 1\%$	*Ocean Sound Speed* $\rho = 3\%$	*Active Matter* $\rho = 1\%$	*Active Matter* $\rho = 3\%$
Observation setting 1 : $\mathcal{T}_{obs} = \mathcal{T}_{tar}$						
Tensor-based						
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
DEMOTE [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
Attention-based						
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
Diffusion-based						
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	**0.283 ± 0.026**	**0.272 ± 0.025**	**0.146 ± 0.046**	**0.108 ± 0.043**	**0.215 ± 0.068**	**0.156 ± 0.046**
Observation setting 2 : $\lvert\mathcal{T}_{obs}\rvert = \frac{1}{2}\lvert\mathcal{T}_{tar}\rvert$						
Tensor-based						
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
DEMOTE [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
Attention-based						
Senseiver [18]	Not capable	-	-	-	-	-
Diffusion-based						
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	**0.433 ± 0.163**	**0.335 ± 0.122**	**0.181 ± 0.084**	**0.165 ± 0.041**	**0.296 ± 0.096**	**0.256 ± 0.087**

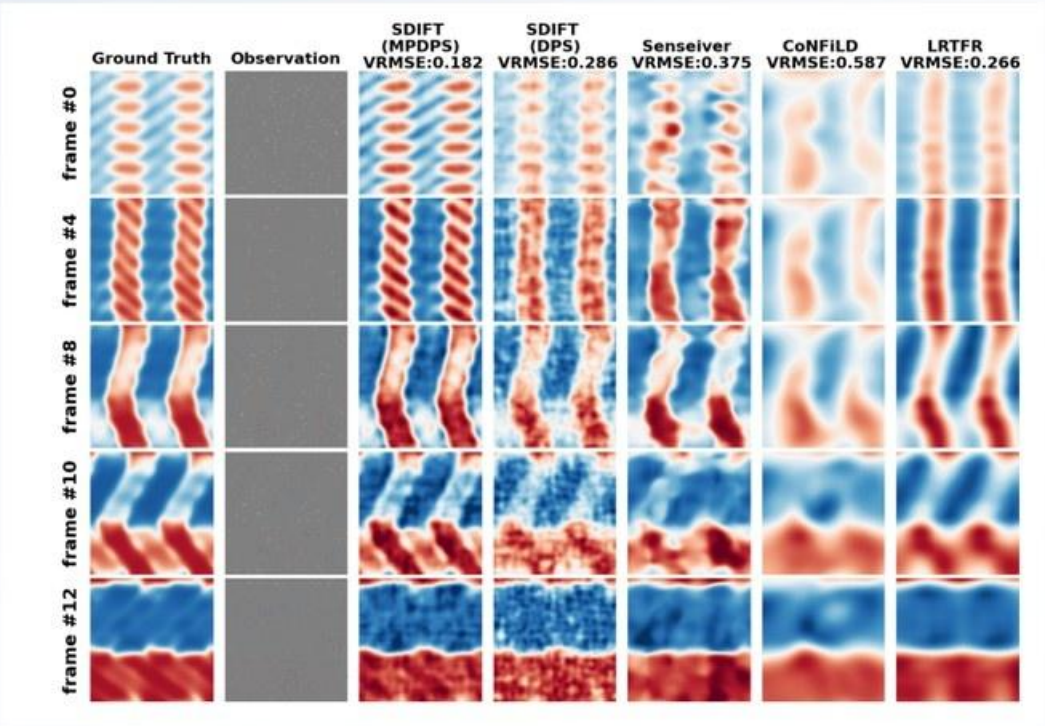
[2] Generating Full-field Evolution of Physical Dynamics from Irregular Sparse Observations, NeurIPS 2025.

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

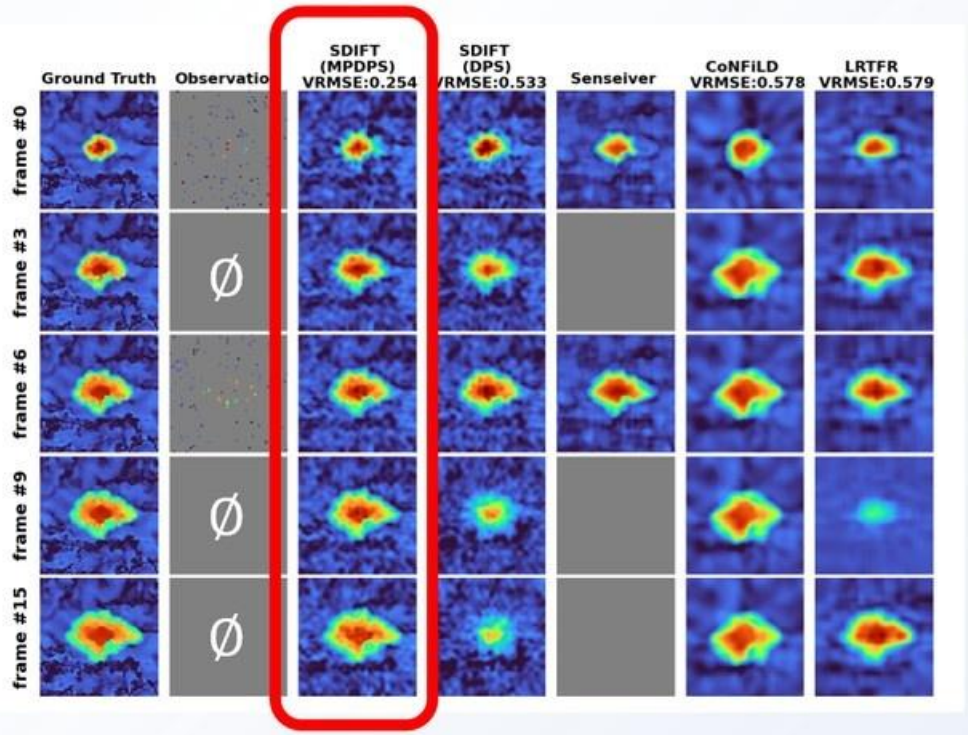
浙江大学 ZHEJIANG UNIVERSITY

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）



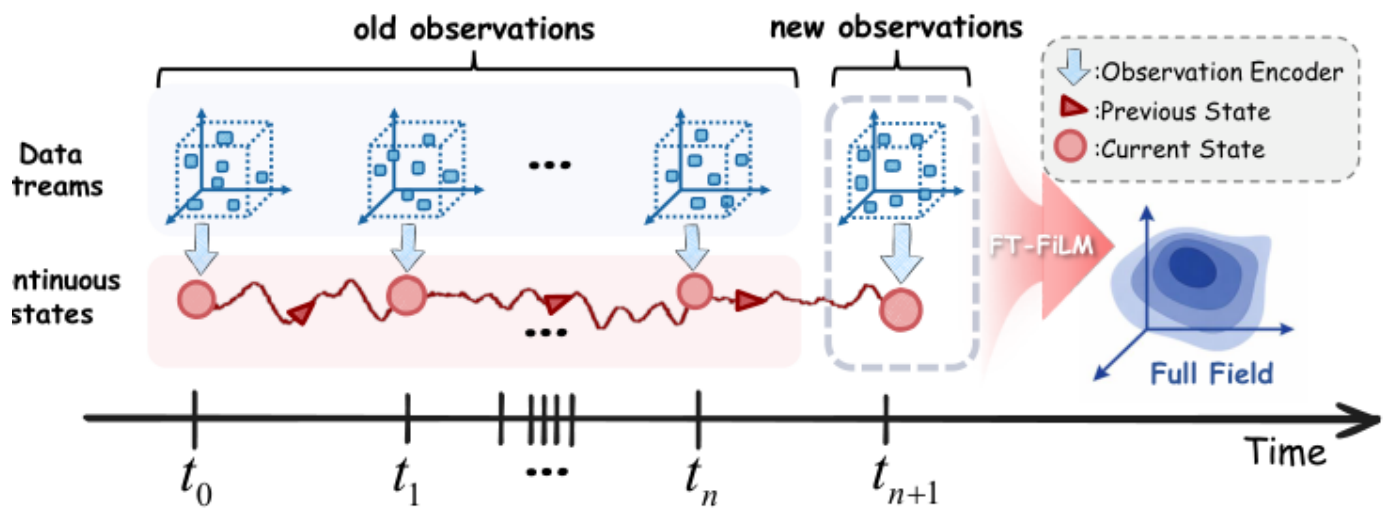
[2] Generating Full-field Evolution of Physical Dynamics from Irregular Sparse Observations, NeurIPS 2025.

StreamPhy：观测到达一次，状态更新一次

StreamPhy: Streaming Inference of High-Dimensional Physical Dynamics via State Space Models

Panqi Chen, Yifan Sun, Shikai Fang, Xiao Fu, Lei Cheng · arXiv:2605.07384 (2026)

STREAMING / ONLINE



old observations → state → new observation → updated full field

SDIFT

预先给定时间窗口
迭代后验采样

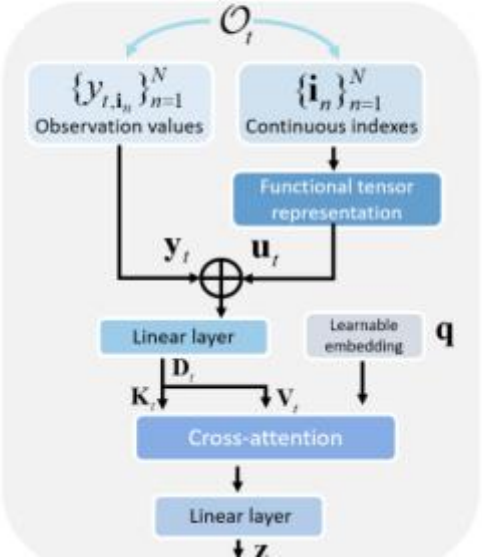
StreamPhy

单步递推更新 · 任意时刻立即查询

$\mathcal{O}_t + X_{t-1} \rightarrow X_t \rightarrow$ full field

[3] StreamPhy: Streaming Inference of High-Dimensional Physical Dynamics via State Space Models, arXiv:2605.07384, 2026.

自适应观测编码 + HiPPO-SSM：历史的在线压缩

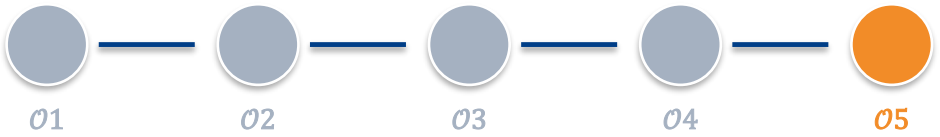


Observation encoder

坐标特征 + 数值 → cross-attention → 定长输入

Streaming state update

masking



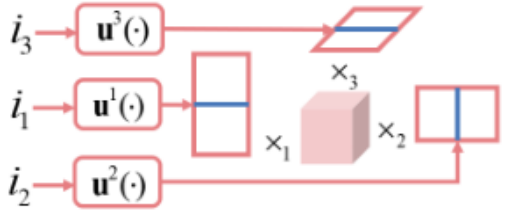
随机掩码训练 →
适应未见观测几何

$$X_t = A(\Delta t)X_{t-1} + B(\Delta t)e_t$$

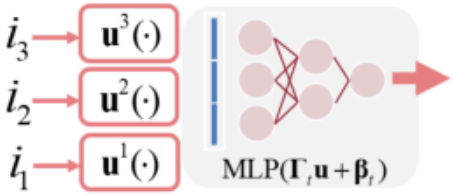
HiPPO 以有限维状态近似保留长程历史；Δt 显式处理非均匀到达。

[3] StreamPhy: Streaming Inference of High-Dimensional Physical Dynamics via State Space Models, arXiv:2605.07384, 2026.

FT-FiLM：状态调制的高表达力连续场解码器



(a) Functional Tucker model [14, 16].



(b) Functional tensor FiLM.

Figure 3: Illustrative comparison of FTM and FT-FiLM (K=3).

FTM：多线性乘法交互

FT-FiLM：状态产生缩放 / 偏置，再非线性读取

FTM 固定秩下受多线性结构限制

FT-FiLM 严格包含 FTM 的函数类

较小潜状态仍可解码复杂动力场。

[3] StreamPhy: Streaming Inference of High-Dimensional Physical Dynamics via State Space Models, arXiv:2605.07384, 2026.

StreamPhy 的在线推断机制

在线状态必须兼顾可递推更新与连续场可解码性。

① 不规则观测编码

cross-attention 将数量、位置均变化的点集映射为定长输入；随机掩码覆盖更多稀疏模式。

② 历史递推压缩

HiPPO 将输入历史投影到有限基系数；状态按 Δt 更新，无需保存并重算完整历史。

③ 状态解码为场

状态生成逐层缩放 / 偏置，与坐标特征非线性耦合；任意位置可直接查询。

不规则观测集合

$$\mathcal{O}_t = \{(i_n, y_n)\}$$

set encoder

有限维历史状态

$$X_t$$

FT-FiLM

连续完整物理场

$$\hat{u}_t(x)$$

$$X_t = A(\Delta t)X_{t-1} + B(\Delta t)e_t \quad \rightarrow \quad \hat{u}_t(x) = Dec\varphi(x; X_t)$$

共享的状态更新规则与连续场解码器，使每次新观测仅需一次前向递推。

[3] StreamPhy: Streaming Inference of High-Dimensional Physical Dynamics via State Space Models, arXiv:2605.07384, 2026.

在线重构实验：平均 VRMSE

	Uniform Sampling						Slab Sampling					
	Turbulent Flow		*Ocean Sound Speed*		*Active Matter*		*Turbulent Flow*		*Ocean Sound Speed*		*Active Matter*	
	$\rho = 3\%$	$\rho = 5\%$	$\rho = 1\%$	$\rho = 3\%$	$\rho = 1\%$	$\rho = 3\%$	$\rho = 3\%$	$\rho = 5\%$	$\rho = 1\%$	$\rho = 3\%$	$\rho = 1\%$	$\rho = 3\%$
Tensor-based												
LRTFR [14]	0.5633	0.3505	0.3453	0.2176	0.3021	0.2582	0.6517	0.4897	0.4075	0.3757	0.9988	0.9537
OFTD [12]	0.5416	0.4810	0.2571	0.1470	0.5542	0.3289	0.5034	0.4901	0.5348	0.4847	0.7387	0.6721
CATTE [13]	0.6192	0.5916	0.0941	0.0850	0.1581	0.1567	0.6563	0.6257	0.1225	0.1187	0.1960	0.1920
Attention-based												
Senseiver [30]	0.2779	0.2361	0.1553	0.1248	0.2555	0.2020	0.3052	0.2578	0.2008	0.1768	0.2762	0.2430
Diffusion-based												
SDIFT [16]	0.3248	0.2821	0.1466	0.1084	0.2158	0.1564	0.4070	0.4029	0.2713	0.1452	0.4231	0.3985
SSM-based												
StreamPhy (SH)	0.1283	0.0954	0.0862	0.0760	0.0895	0.0833	0.2094	0.1979	0.1324	0.1045	0.1452	0.1278
StreamPhy (MH)	**0.0935**	**0.0696**	**0.0628**	**0.0562**	**0.0752**	**0.0732**	**0.1067**	**0.0978**	**0.0692**	**0.0634**	**0.0905**	**0.0866**

Table 1: Average VRMSEs under two sampling patterns across datasets and observation ratios.

[3] StreamPhy: Streaming Inference of High-Dimensional Physical Dynamics via State Space Models, arXiv:2605.07384, 2026.

重构可视化、模块消融与单步推理时间

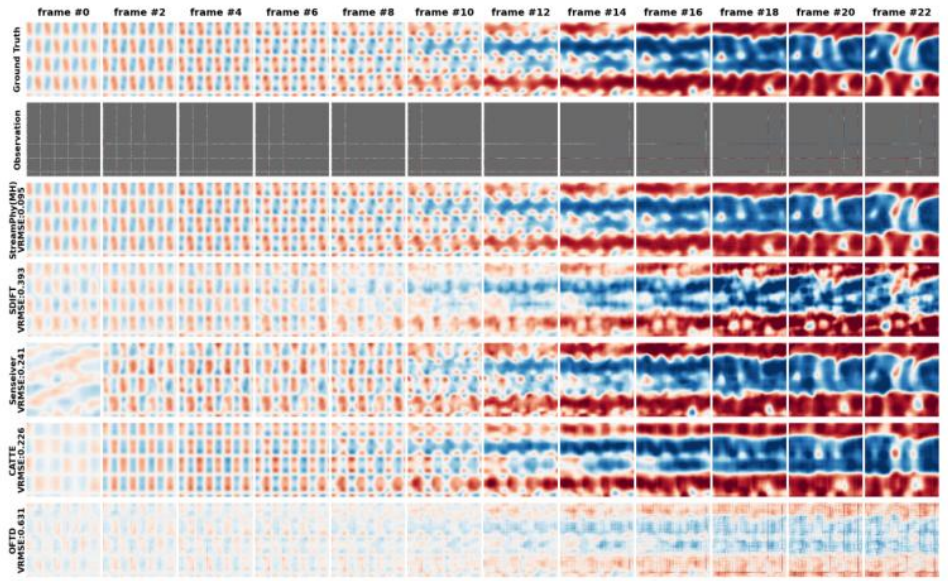


Figure 5: Reconstruction of *Active Matter* dynamics under slab sampling with $\rho = 3\%$.

消融实验

	Uniform Sampling						Slab Sampling					
	Turbulent Flow		*Ocean Sound Speed*		*Active Matter*		*Turbulent Flow*		*Ocean Sound Speed*		*Active Matter*	
	$\rho = 3\%$	$\rho = 5\%$	$\rho = 1\%$	$\rho = 3\%$	$\rho = 1\%$	$\rho = 3\%$	$\rho = 3\%$	$\rho = 5\%$	$\rho = 1\%$	$\rho = 3\%$	$\rho = 1\%$	$\rho = 3\%$
StreamPhy	**0.0935**	**0.0696**	**0.0628**	**0.0562**	**0.0752**	**0.0732**	**0.1067**	**0.0978**	**0.0692**	**0.0634**	**0.0905**	**0.0866**
StreamPhy w.o. SSM	0.8537	0.8501	0.3935	0.3923	0.9023	0.8687	0.8598	0.8556	0.3966	0.3954	0.9115	0.9107
StreamPhy w.o. mask	0.2420	0.1789	0.0899	0.0786	0.0898	0.0786	0.2841	0.2643	0.2030	0.1768	0.2401	0.1511
StreamPhy with FTM	0.1435	0.1134	0.1054	0.1034	0.1279	0.1229	0.1731	0.1574	0.1202	0.1154	0.1522	0.1454

单步推理时间

Methods	*Turbulent Flow* $\rho = 3\%$	$\rho = 5\%$	#Para.	*Ocean Sound Speed* $\rho = 1\%$	$\rho = 3\%$	#Para.	*Active Matter* $\rho = 1\%$	$\rho = 3\%$	#Para.
StreamPhy (MH)	**0.0419s**	**0.0448s**	10.3M	**0.0325s**	**0.0434s**	9.94M	**0.0479s**	**0.0498s**	11M
SDIFT [16]	5.14s	5.21s	15M	0.84s	0.89s	15M	1.31s	1.42s	12M

[3] StreamPhy: Streaming Inference of High-Dimensional Physical Dynamics via State Space Models, arXiv:2605.07384, 2026.

浙江大学 ZHEJIANG UNIVERSITY

一条路径的三个阶段：物理场表征与状态建模

比较维度	FunBaT	SDIFT	StreamPhy
物理场表征	潜在函数 $U^k(i_k)$ + 张量核 W	张量核序列 $W_{1:t}$ 的生成分布	在线状态 X_t + 连续解码器
学习 / 推断范式	**直推式推断 实例级后验**	**归纳式学习 跨场景生成先验**	**归纳式学习 共享状态更新规则**
物理状态	**当前场的 低秩函数参数**	**动力轨迹的 随机潜变量序列**	**历史观测的 有限维递推摘要**
新观测与输出	CEP + 滤波 / 平滑 → 任意坐标连续场	观测引导后验采样 → 完整时间窗口	单步更新 X_t → 当前时刻完整场

实例级后验 → 跨场景生成先验 → 在线流式状态推断

[1] FunBaT (ICLR 2024); [2] SDIFT (NeurIPS 2025); [3] StreamPhy (arXiv:2605.07384).

浙江大学 ZHEJIANG UNIVERSITY

重构只是起点：Physical AI 需要怎样的物理世界表征?

FunBaT
函数 · 一个场

SDIFT
分布 · 一类场

StreamPhy
状态 · 观测流

representation path

主流视觉世界模型

image / video

- 规则像素网格上的稠密观测
- 视觉预测与交互式环境生成
- 以图像 / 视频为主要观测载体

稀疏观测驱动的物理世界模型

world models from sparse observations

异构观测 点 / 切片 / RF / 声学 / 异步传感器

连续物理 守恒约束 / 跨尺度动力学 / 可校准不确定性

跨域泛化 cross-field / cross-law / cross-operator

关键：适合物理世界的潜在表征是什么?

重构与逆问题是入口；目标是从稀疏观测中理解状态、动力与规律。

[4] NVIDIA, Cosmos World Foundation Model Platform for Physical AI, 2025; [5] Assran et al., V-JEPA 2, 2025; [6] Google DeepMind, Genie 2, 2024.

浙江大学
ZHEJIANG UNIVERSITY

Q&A

函数 → 分布 → 状态

寻找适合物理世界的表征

xuangu-fang.github.io

FunBaT · ICLR 2024
SDIFT · NeurIPS 2025
StreamPhy · arXiv:2605.07384

