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

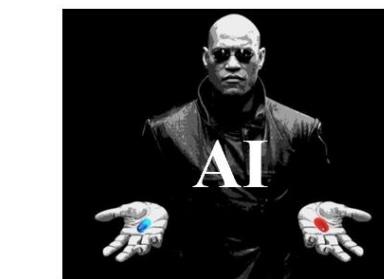
NeurIPS 2025

Group Meeting 12/18

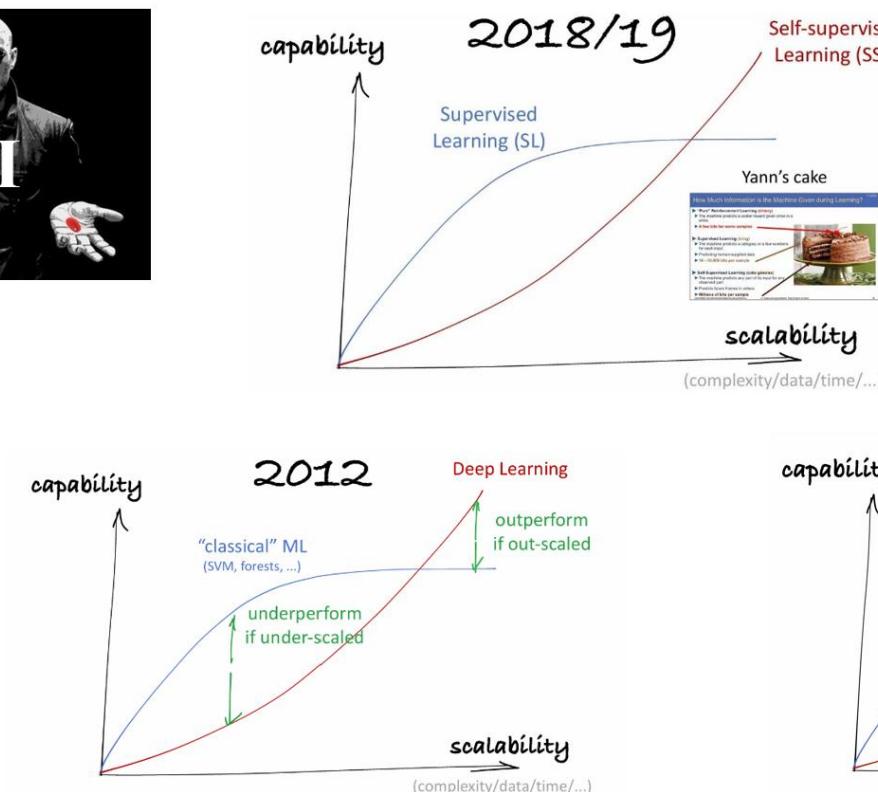
研究背景



- 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...

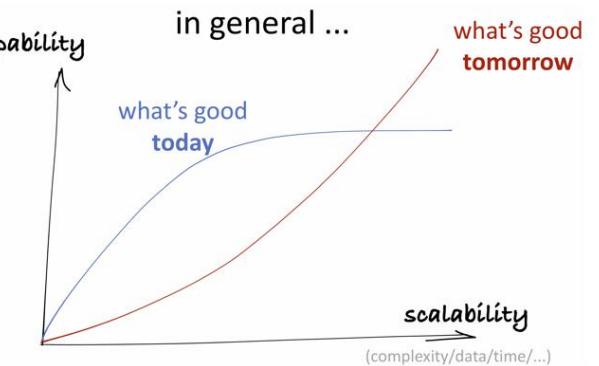


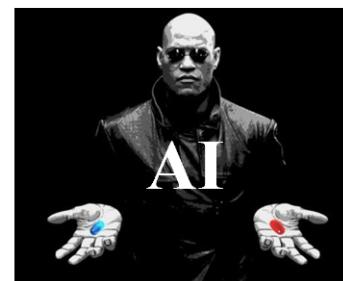
Fig. credits: Kaiming He's talk on NeurIPS 2024

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



Yes! But...

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, 对真实世界影响力有限

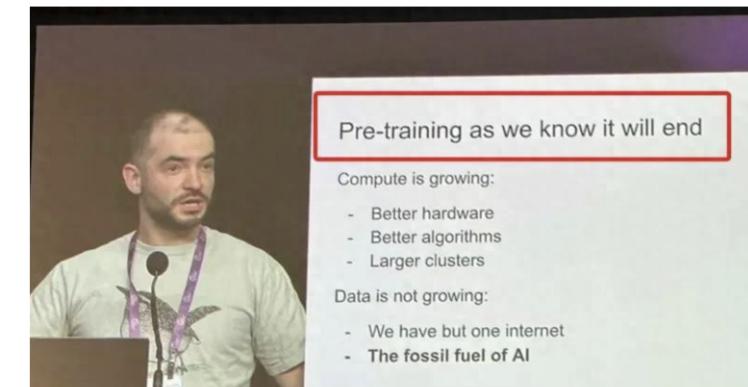
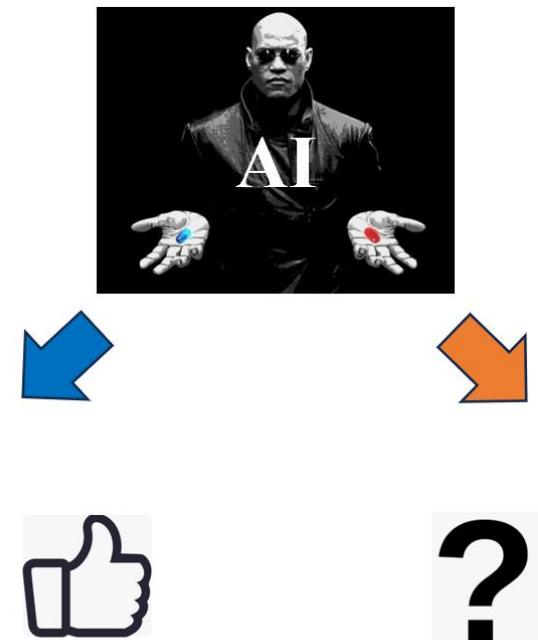


Fig. credits: Ilya's talk on NeurIPS 2024

研究背景

• 下一个风口：物理世界信号驱动的智能建模



Action / Sensor / Signal based

Vision / Language / Internet based

“Welcome to
the real world! ”

- 更广跨的数据来源：
真实物理世界的演化&观测信号
- 更广泛的应用场景：
真实物理世界的科学与工程场景

研究挑战

• 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！

研究挑战



稀疏张量：极高成本 + 超高维度 + 极稀疏信号 + 噪声

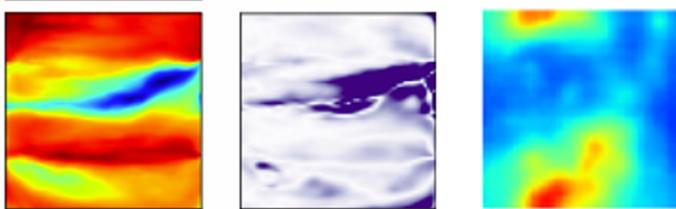
Tensor-Datasets	Tensor-Size	#Observed entries	Observed Ratio
SG (climate)	2321*5596*1600	105,764	0.0005%
ACC (cloud comp.)	3000*150*30000	1,220,000	0.1%

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...）

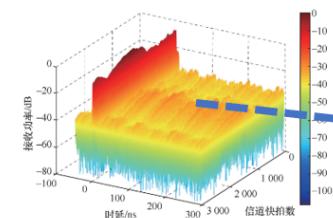
AI在CV/NLP任务上的发展出的新范式，很难直接迁移到物理世界信号的建模中！

研究挑战

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



难以支撑 多源、多物理耦合的复杂建模



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

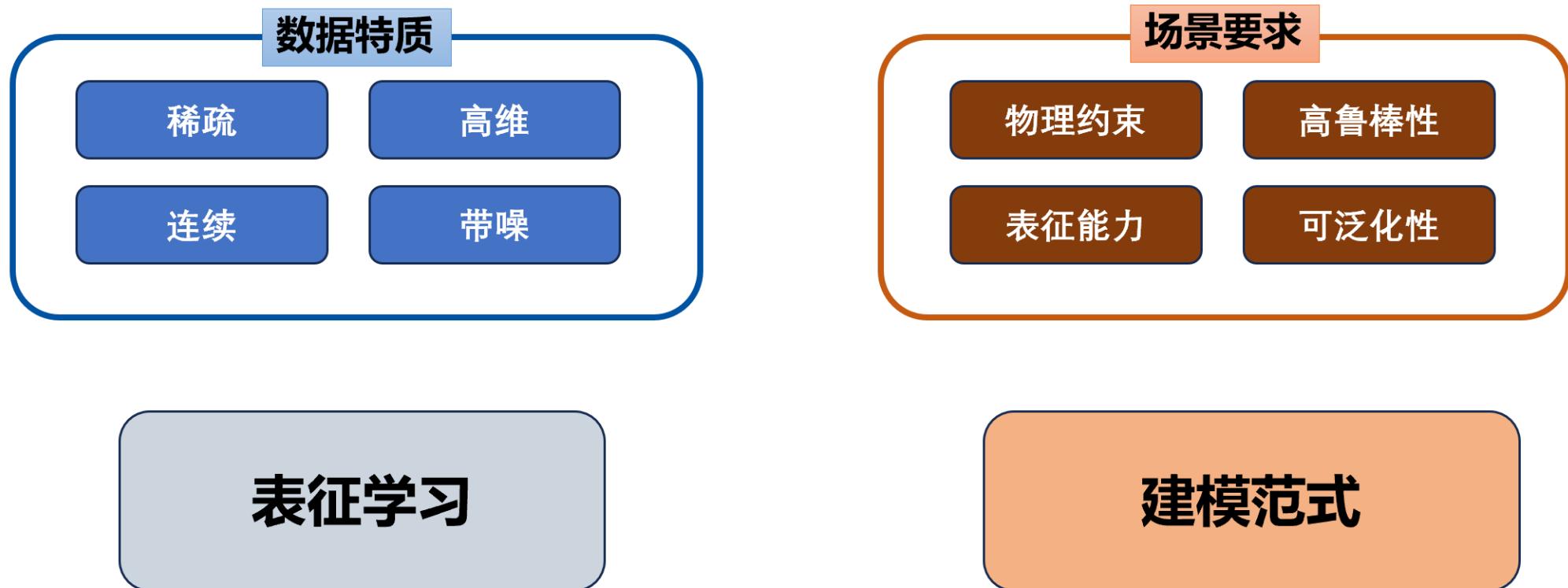
数据驱动+表征学习

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！

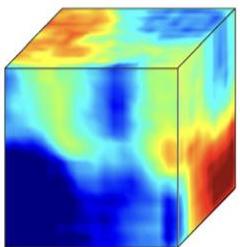
研究思路

对一个以“**物理世界观测信号**”为对象的AI模型，都需要解决的两个核心问题：

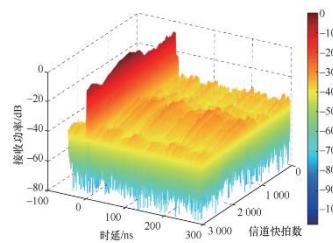


- 海洋中的声速场、陆地上的信道响应、工程流体中的速度场

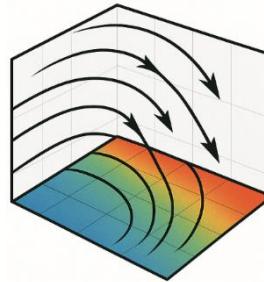
物理信号 在空间、时间、频率/谱域、观测模态（如声学、射频、电磁或流体量测）、传感器方向等维度上耦合变化。



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。



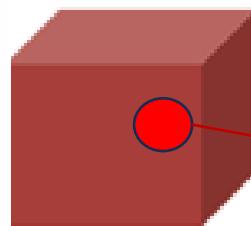
城市数据



社交网络



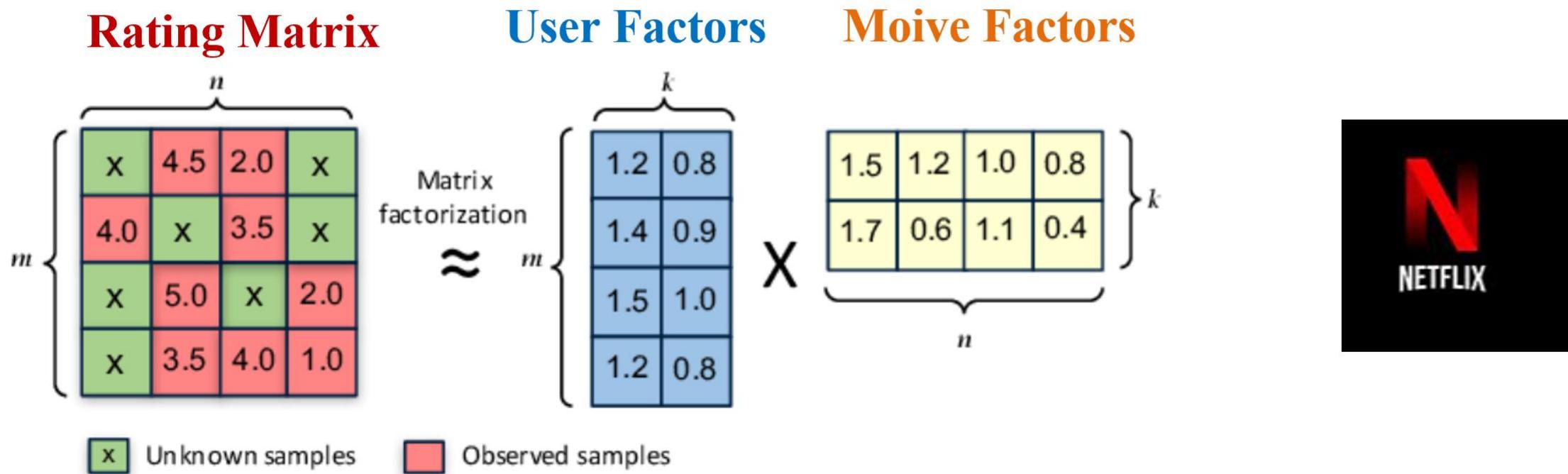
能源系统



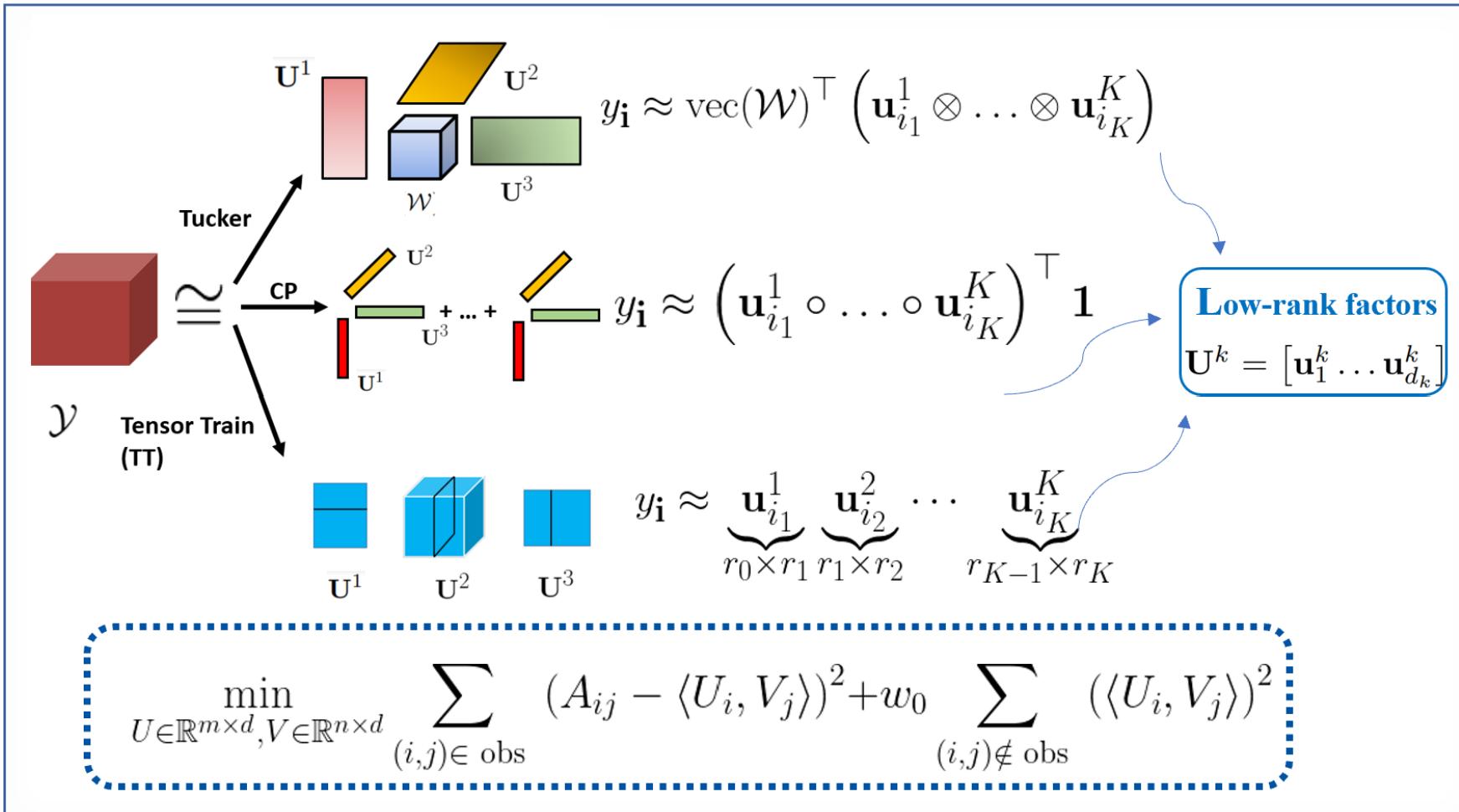
天然呈现多维 张量 结构

格式: (index1, index2..)-> value

- Learn **low-rank** factors of **high-order** tensor
- 2-D case: **Collaborative Filtering** (Matrix Factorization)

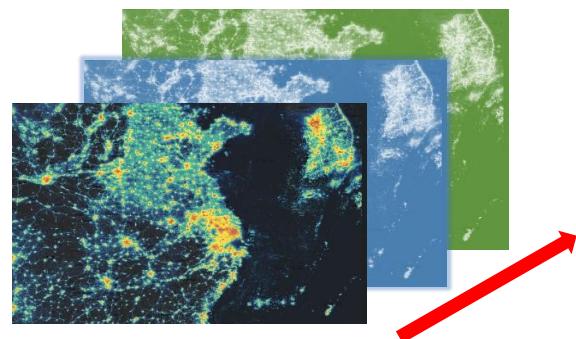


□ 经典张量模型：CP, Tucker, Tensor-Train



真实应用场景的挑战：

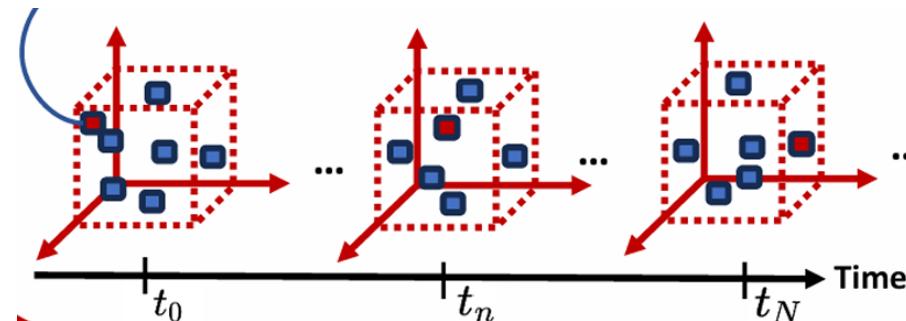
- 时变张量



$(\text{region}, \text{site}, \text{weather}) \times \text{time}$

张量结构随时间连续演化

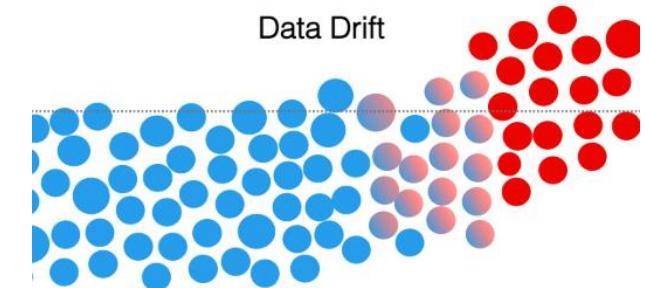
- 泛函张量



$(\text{latitude}, \text{longitude}, \text{height}, \text{time})$

索引皆为连续实数的时空张量

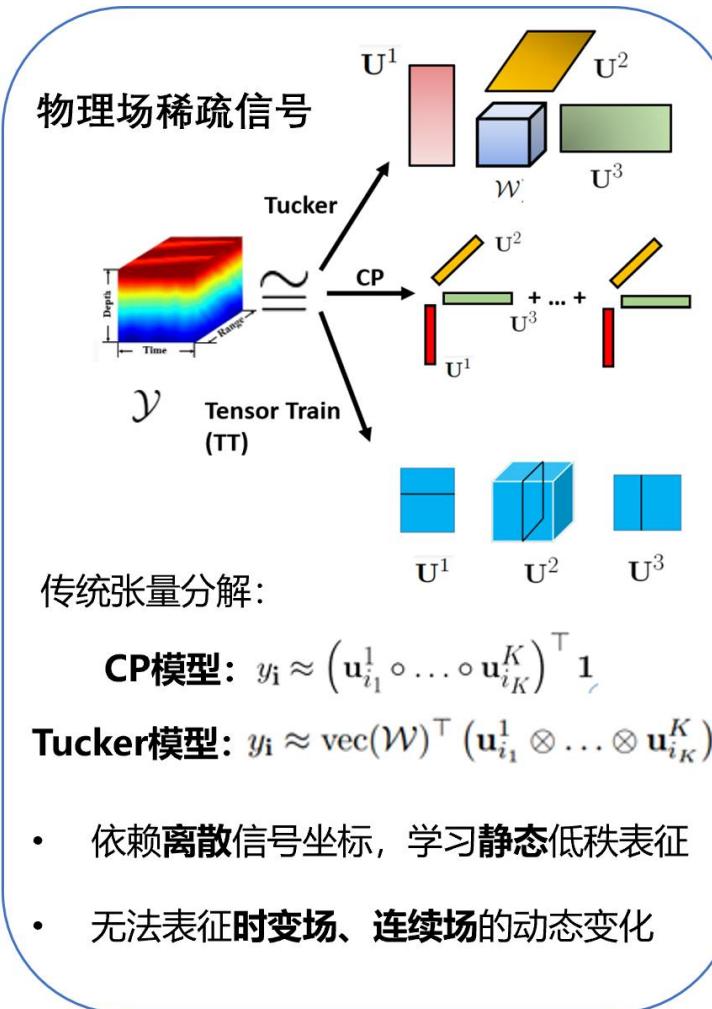
- 流式推断



大批量观测数据实时更新，
分布实时变化

动态张量学习：时变张量和泛函张量

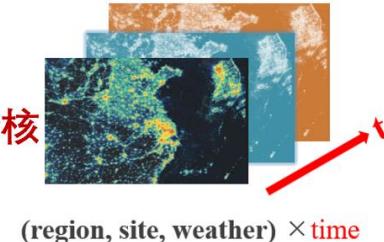
广义动态张量学习系列 代表成果



- 首次提出 **连续时间 Tucker 分解模型** [ICML 2022 oral]

$$y_i(t) \approx \text{vec}(\mathcal{W}(t))^T (\mathbf{u}_{i_1}^1 \otimes \dots \otimes \mathbf{u}_{i_K}^K)$$

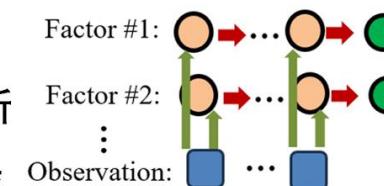
- 核心创新：随机动力系统-> **连续时变Tucker核**
- 应用场景：多维动态交互的时变物理场建模



- 首次提出 **流式因子轨迹张量模型** [NeurIPS 2023 spotlight]

$$y_i(t) \approx \text{vec}(\mathcal{W})^T (\mathbf{u}_{i_1}^1(t) \otimes \dots \otimes \mathbf{u}_{i_K}^K(t))$$

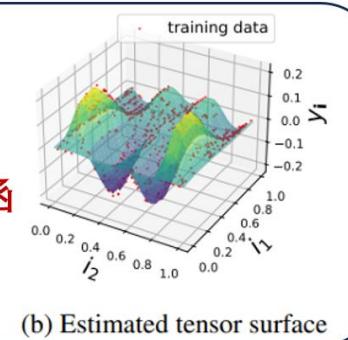
- 核心创新：隐空间滤波->**多因子时空轨迹推断**
- 应用场景：对动态信号中时变因子的实时响应



- 首次提出 **贝叶斯泛函张量模型** [ICLR 2024]

$$y_i \approx \text{vec}(\mathcal{W})^T (\mathbf{U}^1(i_1) \otimes \dots \otimes \mathbf{U}^K(i_K))$$

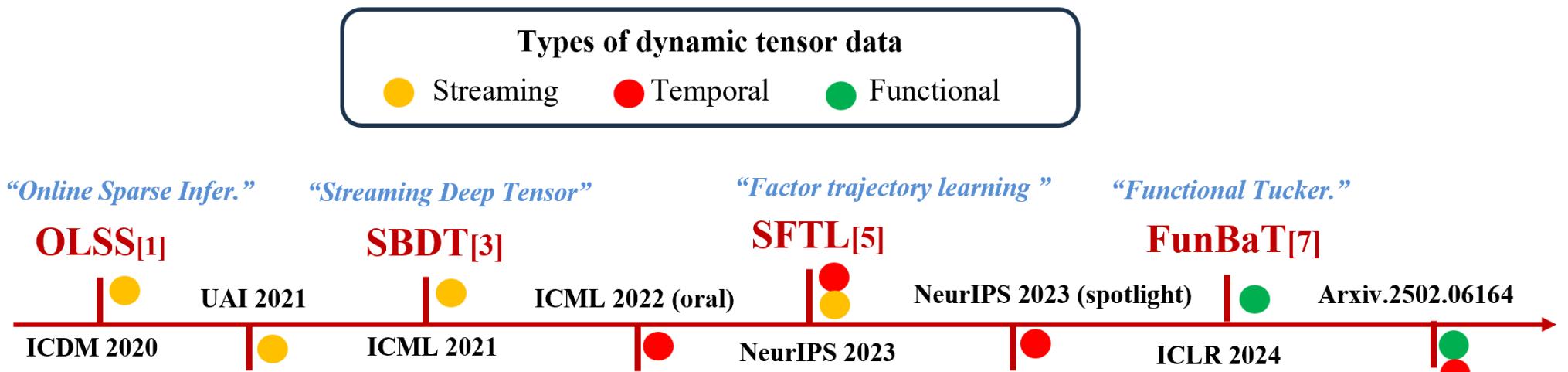
- 核心创新：高斯过程+似然解耦->**维度连续泛函**
- 应用场景：多维连续时空场的重构和反演



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



[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**

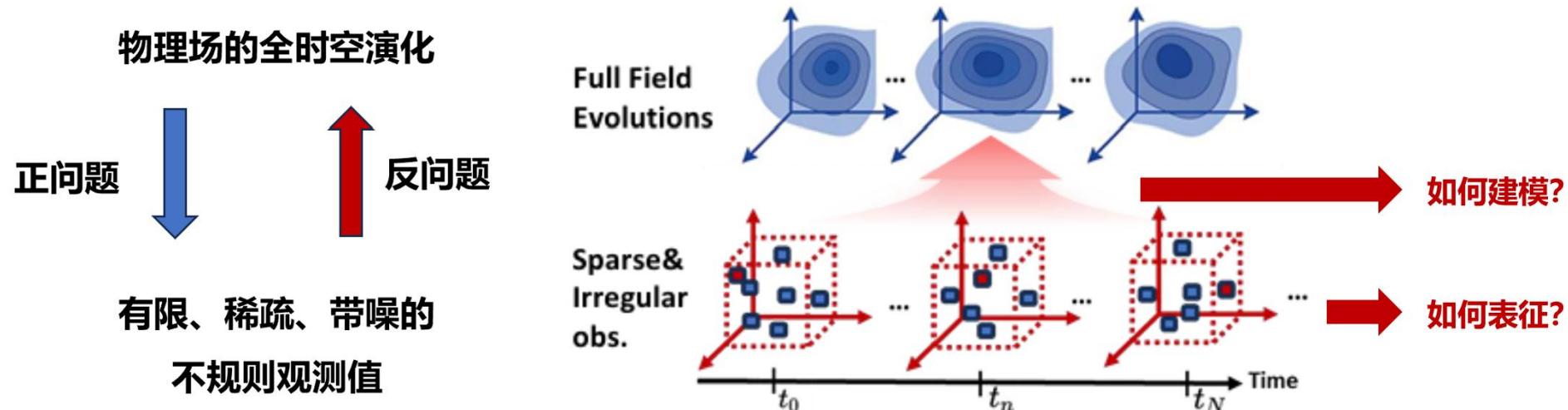
[5]: Fang et al., "Streaming Factor Trajectory Learning for Temporal Tensor Decomposition", NeurIPS 2023

[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

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



Problem Statement: We consider a K -mode physical full field tensor with an extra time mode \mathcal{Y} , whose spatiotemporal coordinate is denoted as $(\mathbf{i}, t) = (i_1, \dots, i_K, t)$. Without loss of generality, suppose we have M off-grid observation timesteps $\mathcal{T} = \{t_1, \dots, t_M\}$. At each t_m , the observation set $\mathcal{O}_{t_m} = \{(\mathbf{i}_{n_m}, t_m, y_{\mathbf{i}_{n_m}, t_m})\}_{n_m=1}^{N_m}$ contains N_m samples, where $y_{\mathbf{i}_{n_m}, t_m}$ is the entry value of \mathcal{Y} indexed at (\mathbf{i}_{n_m}, t_m) . We denote the entire collection of observations by $\mathcal{O} = \{\mathcal{O}_{t_m}\}_{m=1}^M$.

We assume the patterns of observations at each time step is varying. The goal is to infer the value of underlying function $y(\mathbf{i}, t) : \mathbb{R}^K \times \mathbb{R}_+ \rightarrow \mathbb{R}$ at arbitrary continuous spatiotemporal coordinates. If we treat the problem as a function approximation task, it can be solved by numerical methods [13, 14] and functional tensor methods [11, 12].

如何定义问题?

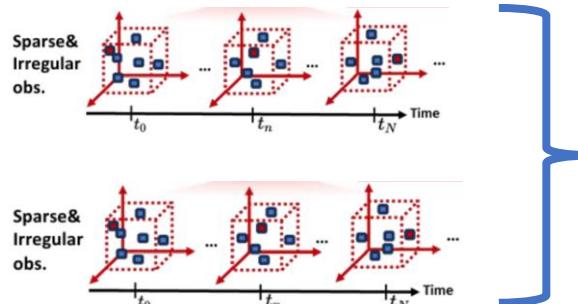
□ 当做传统图形/视频 补全任务?

- CNN/VIT (image) -> image-target
- 问题: 空时连续 (无限分辨率) + 稀疏观测

□ 当做传统张量分解任务?

- 表征能力不够
- 传统张量方法 只 利用一个张量样本 中的观测值, 信息量不够, 泛化性不够

Sample 1:



每一条轨迹样本 的生成: 服从同一的物理规律, 但初始值/边界值/参数 不一样

Sample N:

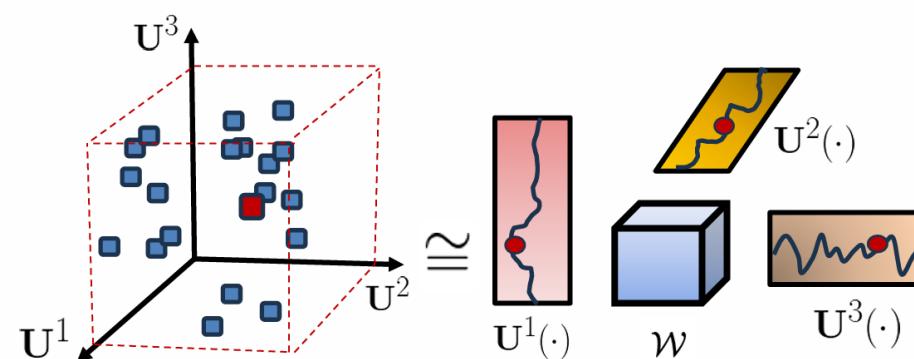
每一条时空张量数据、都需要训练一个自己的张量分解模型! (泛化性差)

□ 先解决数据表征的问题，

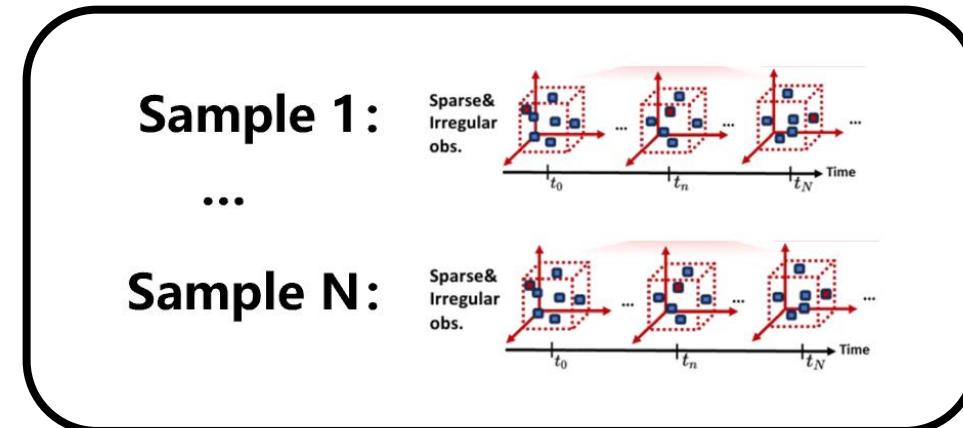
□ 利用泛函张量的形式：任何的稀疏不规则观测张量 -> 编码连续坐标的隐函数 + Compact Tucker Core

Tensor entry \Leftrightarrow Tucker-form interaction of latent functions

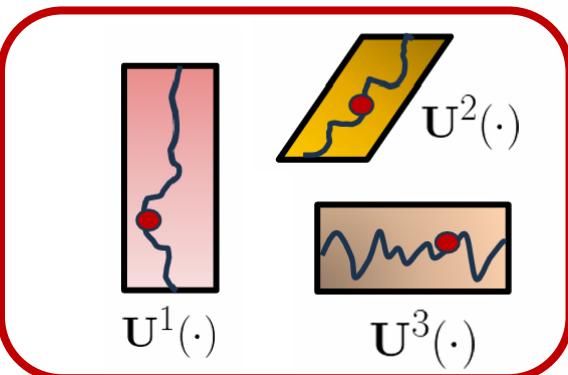
$$f(\mathbf{i}) = f(i_1, \dots, i_K) \approx \text{vec}(\mathcal{W})^\top (\mathbf{U}^1(i_1) \otimes \dots \otimes \mathbf{U}^K(i_K))$$



□ 每一条 时空张量序列 样本 -> 共享的坐标编码 隐函数 + Time-Dependent 的 Tucker Core



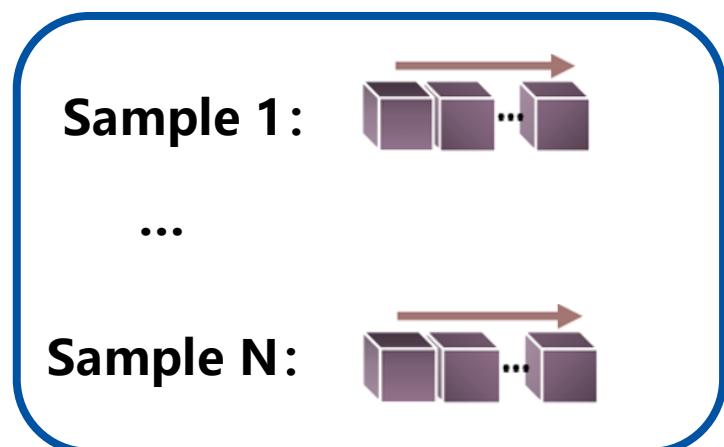
所有数据共享的“空间坐标” 编码器



To train the latent functions $\{f_{\theta_k}\}_{k=1}^K$ and obtain batches of core sequences $\{\mathcal{W}^b\}_{b=1}^B$, we propose to minimize following objective function over the entire training dataset $\mathcal{D} = \{\mathcal{O}^b\}_{b=1}^B$:

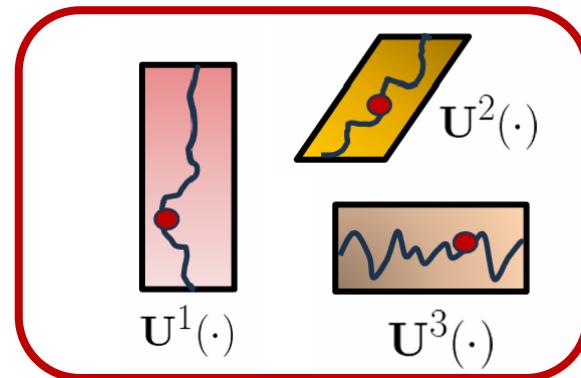
$$\mathcal{L}_{FTM} = \mathbb{E}_{(i, t_m, y_{i, t_m}) \sim \mathcal{D}} \left[\|y_{i, t_m} - \text{vec}(\mathcal{W}_{t_m})^\top (f_{\theta_1}^1(i_1) \otimes \cdots \otimes f_{\theta_K}^K(i_K))\|_2^2 + \beta \text{TV}(\mathcal{W}) \right], \quad (8)$$

每条时空张量轨迹-> Seq of Tucker Core



核心思路：重构 as 生成

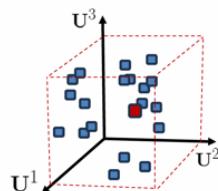
所有数据共享的“空间坐标”编码器



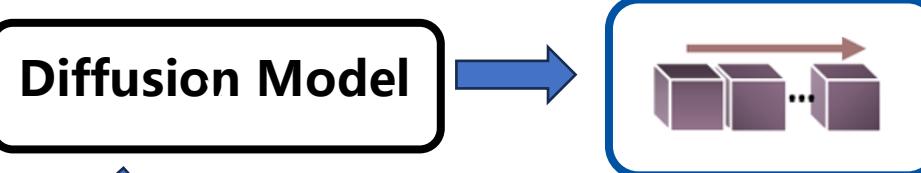
Diffusion Model



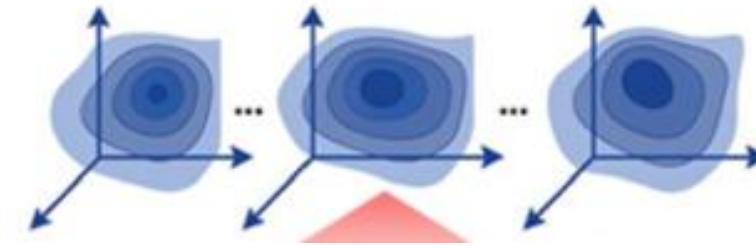
隐空间中生成的“任意时间截”的Core



Seq of Tucker Core



Full Field Evolutions



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

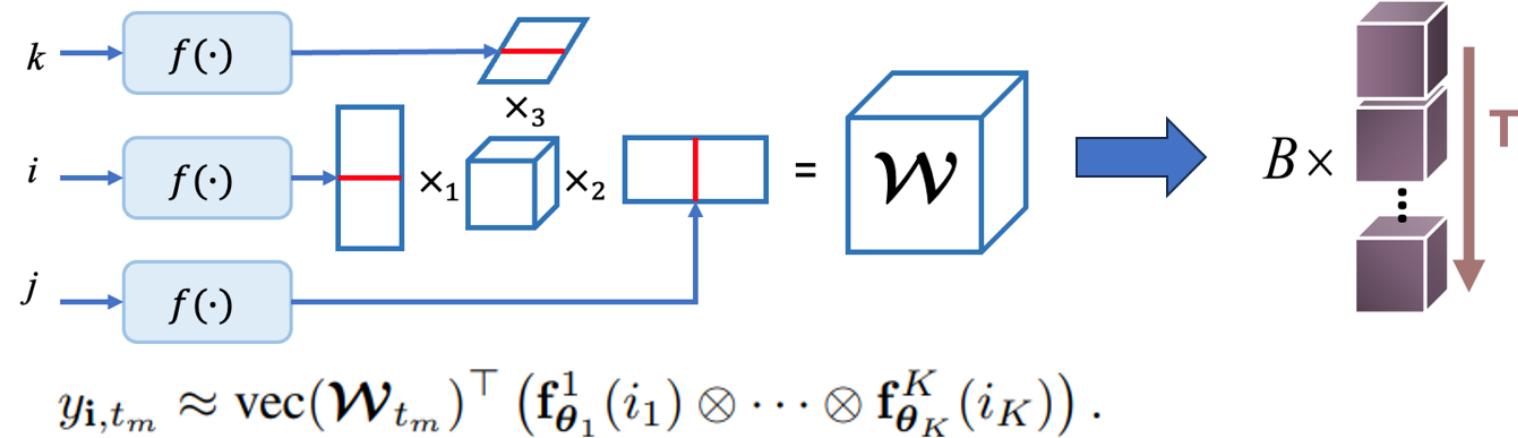
任意时空坐标下的值

• 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

将稀疏高维观测信号表征为隐空间中的张量核序列！



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 $u \in L^2(X_1 \times \dots \times X_K)$. 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

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

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



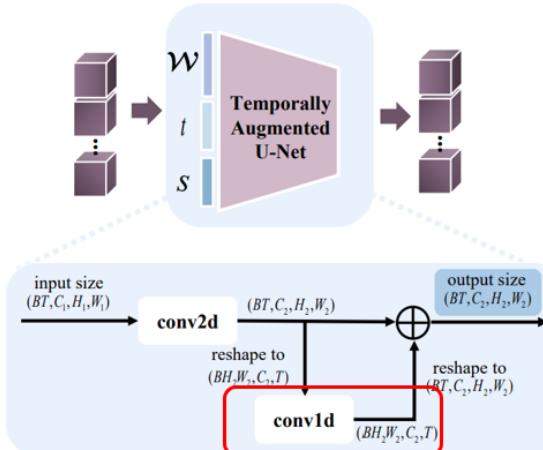
高斯过程驱动的隐空间序列扩散模型

高斯过程噪声源

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

允许对连续时间建模！

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

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

• 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

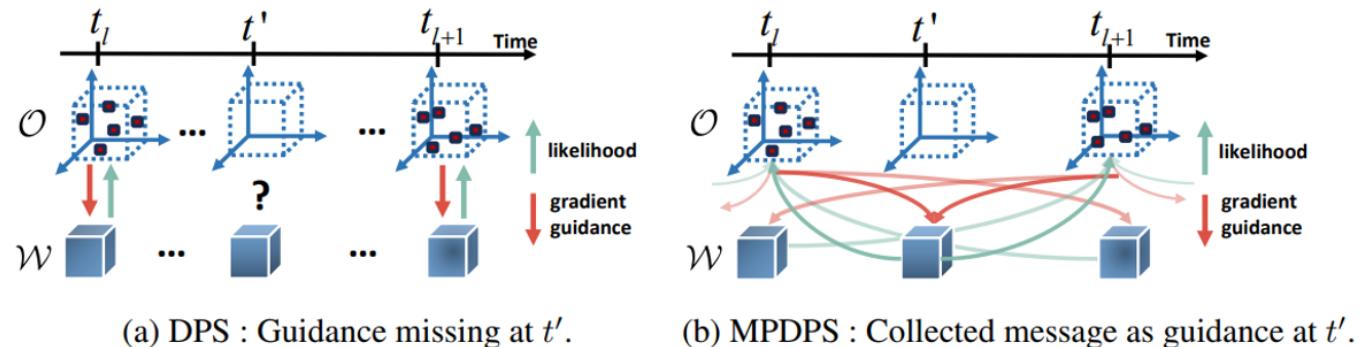


Figure 2: Illustration of DPS and MPDPS to handle the case of no observations at t' .

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

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

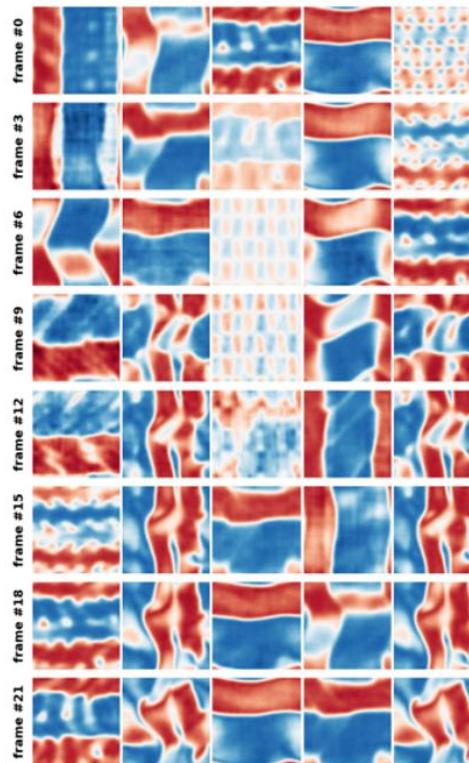
来自当前观测的后验信息

似然函数具有二次项结构，梯度计算效率高！

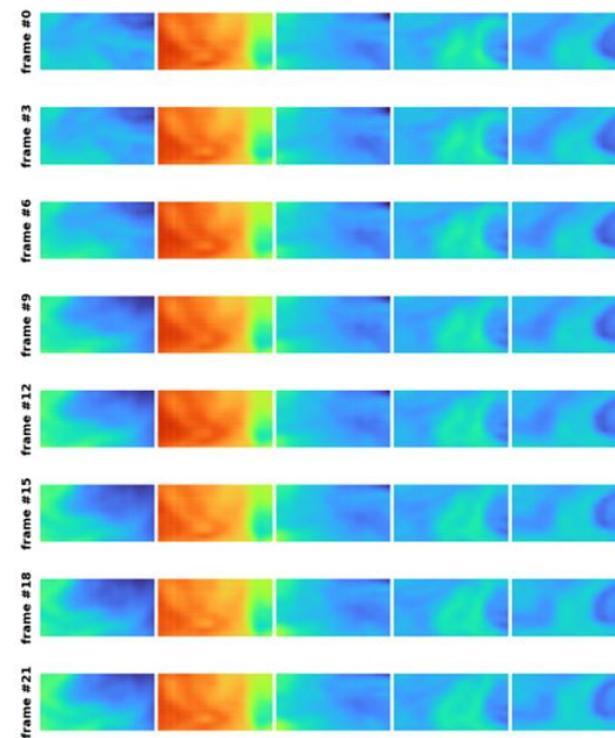
$$\{\mathcal{G}_{t_l, t_m}^s\}_{t_m \in \mathcal{T}_{\text{tar}} \setminus t_l} := \nabla_{\mathcal{W}_{\setminus t_l}^s} \log p(\mathcal{O}_{t_l} | \mathcal{W}_{\setminus t_l}^s) = \nabla_{\mathcal{W}_{\setminus t_l}^s} \log \int p(\mathcal{O}_{t_l} | \mathcal{W}_{t_l}^0) p(\mathcal{W}_{t_l}^0 | \mathcal{W}_{\setminus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

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

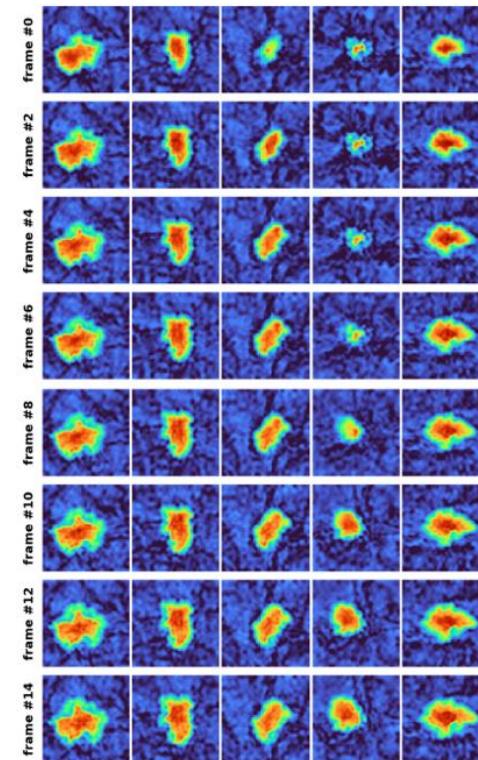
实验验证：从厘米到光年的三个动态物理场



厘米尺度：
生物质流体



千米尺度：
海洋声速场演化



光年尺度：
超新星爆炸

• 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

不同观测比例下的补全误差(VRMSE)

Methods	Supernova Explosion		Ocean Sound Speed		Active Matter	
	$\rho = 1\%$	$\rho = 3\%$	$\rho = 1\%$	$\rho = 3\%$	$\rho = 1\%$	$\rho = 3\%$
Observation setting 1 : $\mathcal{T}_{obs} = \mathcal{T}_{tar}$						
<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
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
<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	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 : $\mathcal{T}_{obs} = \frac{1}{2} \mathcal{T}_{tar}$						
<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
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
<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	0.433 ± 0.163	0.335 ± 0.122	0.181 ± 0.084	0.165 ± 0.041	0.296 ± 0.096	0.256 ± 0.087

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.
CoNFiLD	31.3s	44.0s	23M	15.9s	20.3s	10M
SDIFT w/ MPDPS	2.23s	5.43s	26M	0.84s	0.89s	15M

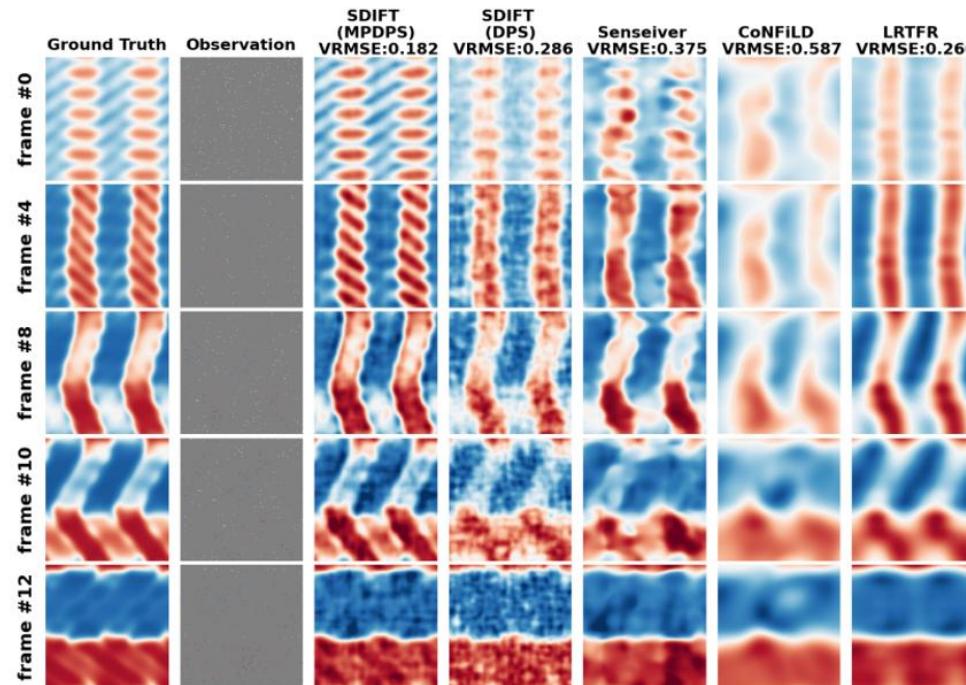
Table 2: Average sampling speed for reconstruction with different observation ratios on observation setting 1.

**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**

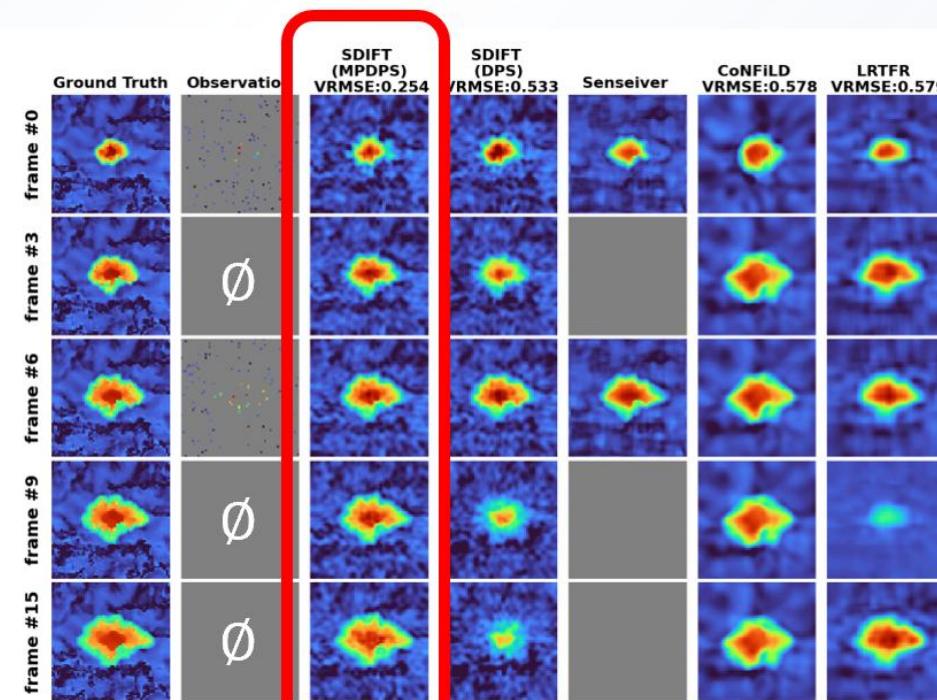
• 代表成果：SDIFT 生成稀疏物理场的全域演化 (NeurIPS 2025)

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)

对噪声的鲁棒性

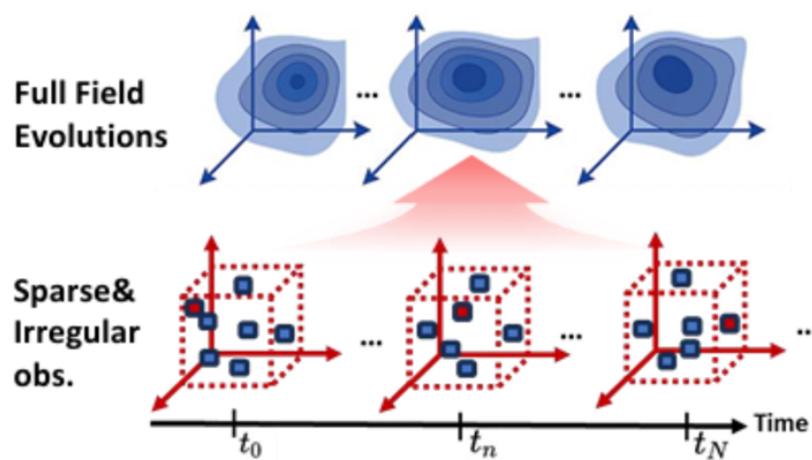
利用物理连续性，“跨时刻传递” MPDPS 机制来的稳健性

Noise configurations	SDIFT w/ DPS $\rho = 1\%$	SDIFT w/ MPDPS $\rho = 1\%$
Gaussian noise ($\sigma = 0.1$)	0.207 ± 0.066	0.156 ± 0.052
Gaussian noise ($\sigma = 0.3$)	0.210 ± 0.069	0.164 ± 0.049
Laplacian noise ($\sigma = 0.1$)	0.200 ± 0.075	0.170 ± 0.054
Laplacian noise ($\sigma = 0.3$)	0.224 ± 0.054	0.177 ± 0.051
Poisson noise ($\sigma = 0.1$)	0.186 ± 0.081	0.168 ± 0.048
Poisson noise ($\sigma = 0.3$)	0.214 ± 0.070	0.171 ± 0.053

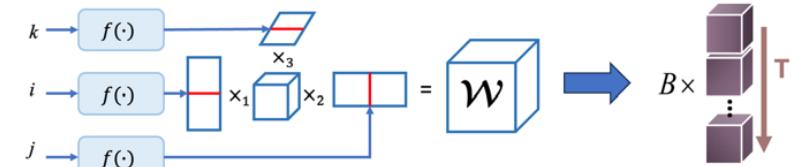
Table 4: VRMSE of reconstruction of our proposed method on *Ocean Sound Speed* datasets over varying noise with observation setting 1 and $\rho = 1\%$.

SDIFT 总结

稀疏物理场重构任务



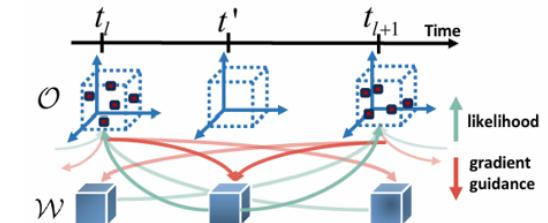
表征学习



泛函Tucker 张量分解

建模范式

隐空间序列生成

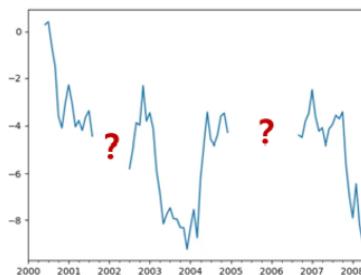


消息传递-后验采样机制

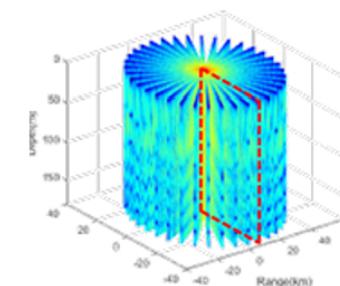
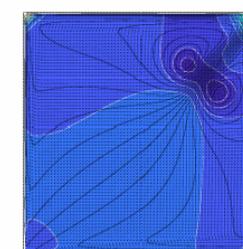
Top of my mind

为什么在 AI4Sci. 中，我们需要更多的生成式建模？

- 传统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

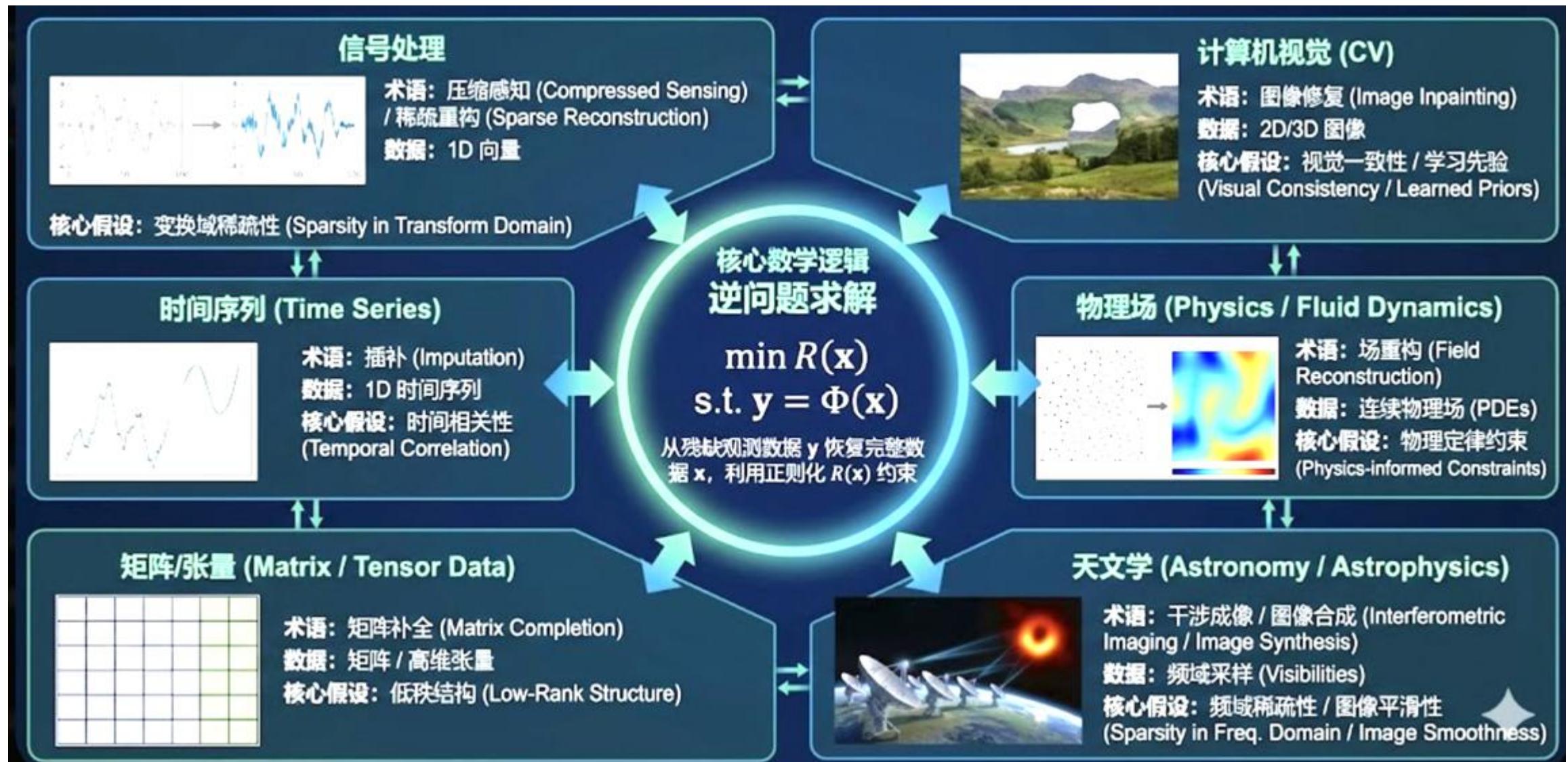


One for One...



One for All !

- 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



核心思路：

