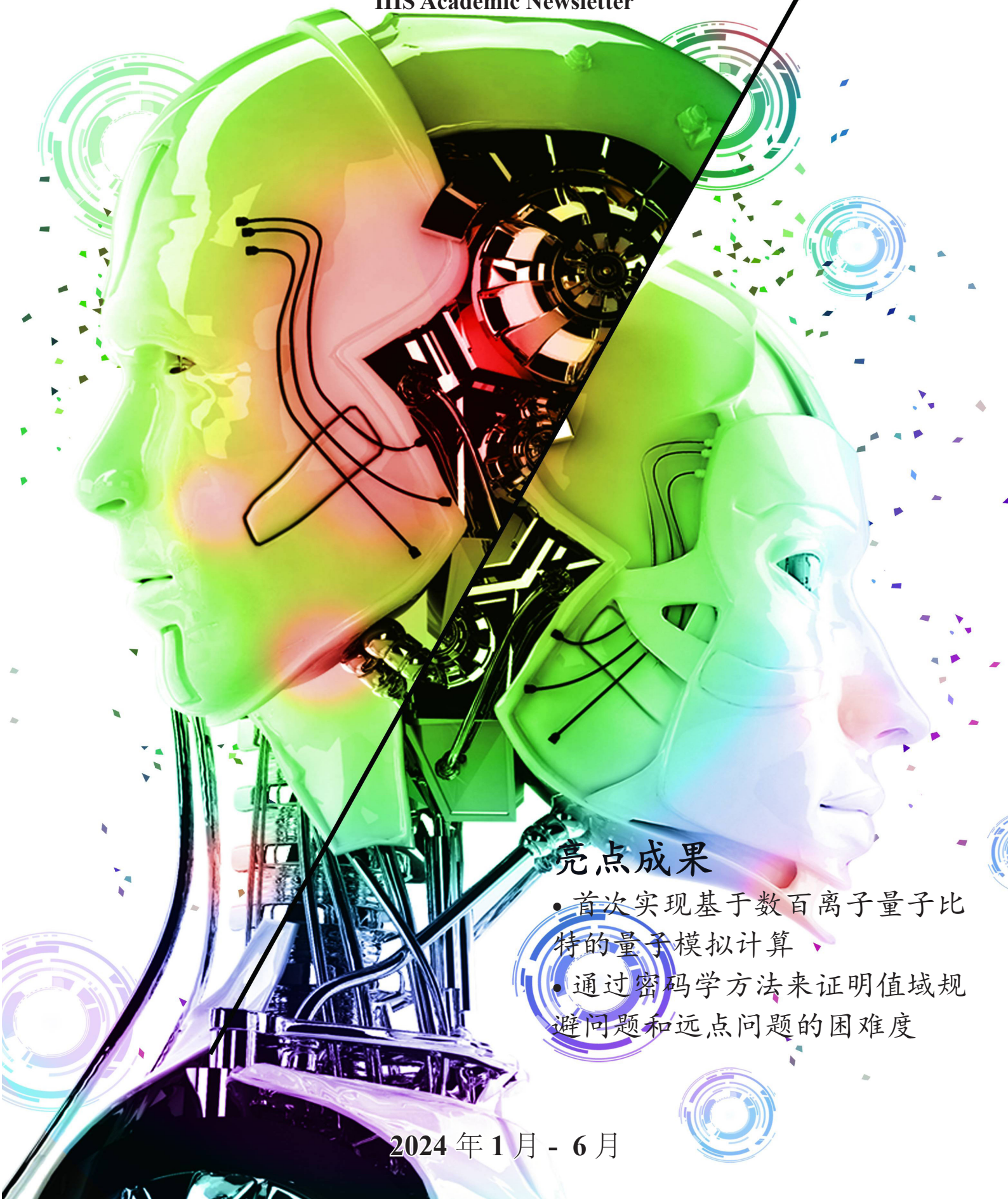
清華大學 交叉信息研究院
Institute for Interdisciplinary Information Sciences, Tsinghua University

学术科研简报

IIIS Academic Newsletter



亮点成果

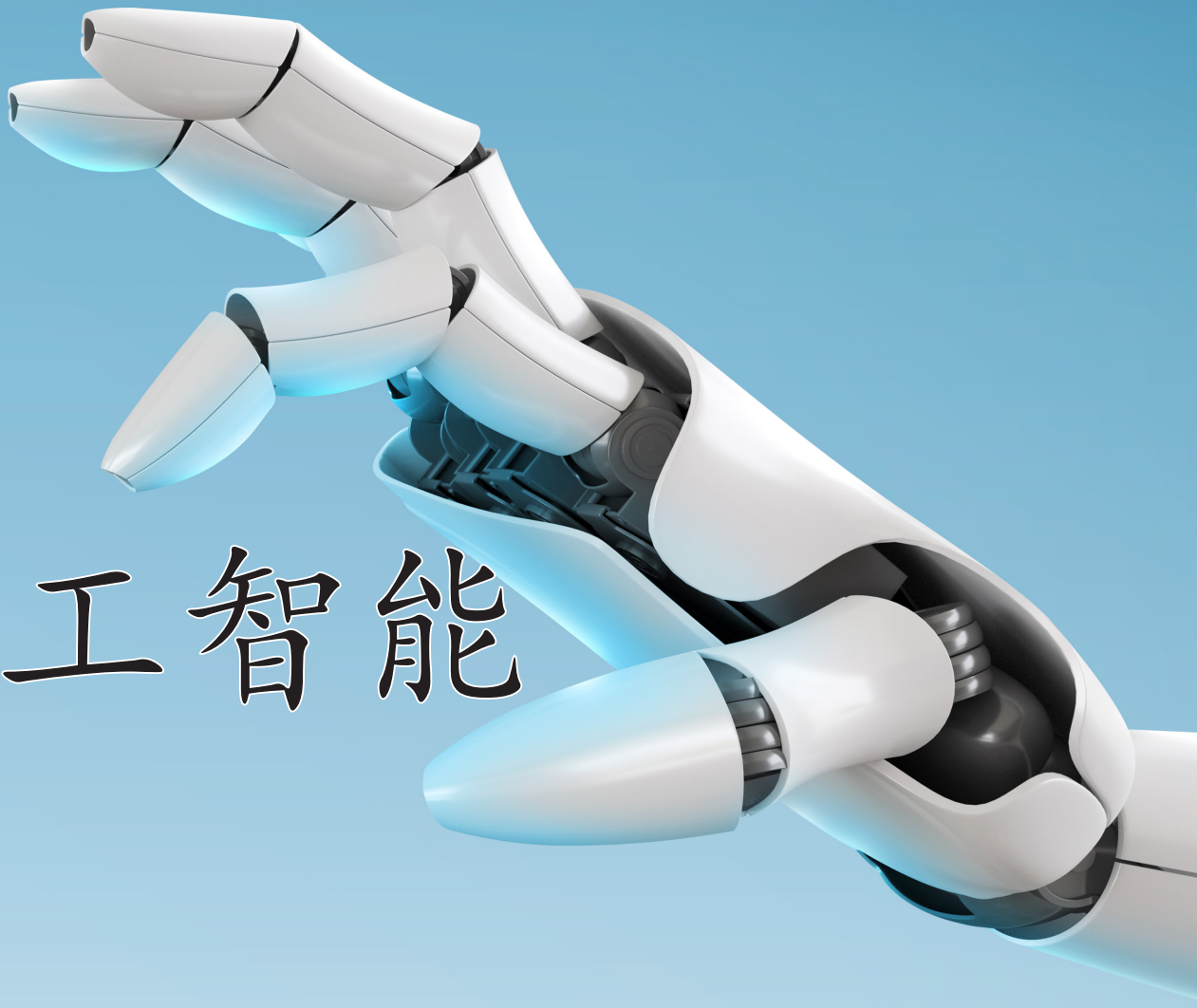
- 首次实现基于数百离子量子比特的量子模拟计算
- 通过密码学方法来证明值域规避问题和远点问题的困难度

2024 年 1 月 - 6 月

人工智能

计算机科学

量子信息科学



人工智能

一、具身智能

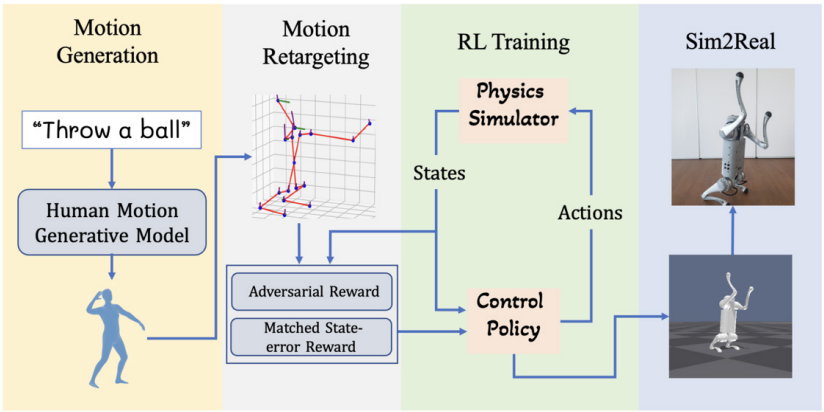
主要完成人：吴翼研究组、许华哲研究组、陈建宇研究组

自然语言指导的机器人控制策略生成

随着基础大模型的飞速发展，自然语言已成为诸多人工智能应用的通用人机接口。让机器人听从高层语言指令做出对应的动作对于自然的人机交互具有关键意义。尽管已有的生成式模型能从自然语言产生动作序列，但是由于真实机器人的身体结构和物理属性和人的运动数据集不一致，这些生成的动作往往不能严格服从物理约束，难以在现实机器人上执行。

为了能从自然语言生成 可直接控制真实机器人的策略，吴翼研究组提出了 LAGOON（language-guided motion control），一种多阶段的运动控制方法。LAGOON 首先利用预训练模型从语言命令生成对应的人体运动序列，并根据机器人的身体结构映射得到语义正确但是物理上不准确的目标动作轨迹。然后，LAGOON 进入强化学习阶段，在物理正确的模拟环境中训练控制策略，使得控制策略既符合语义又符合物理约束。最后通过域随机化，LAGOON 可以将控制策略由模拟环境迁移到真实世界中。他们的控制策略可以成功部署到四足机器人上，让机器狗在现实世界中完成多种自然语言描述的动作，如站起来挥手、倒立等。LAGOON 克服了物理真实性和语言导向的机器人运动控制的挑战，为实现更加自然、智能的人机交互打下了基础。

该成果研究论文：Shusheng Xu, Huaijie Wang, Jiaxuan Gao, Yutao Ouyang, Chao Yu, Yi Wu, “Language-Guided Generation of Physically Realistic Robot Motion and Control”, ICRA 2024.

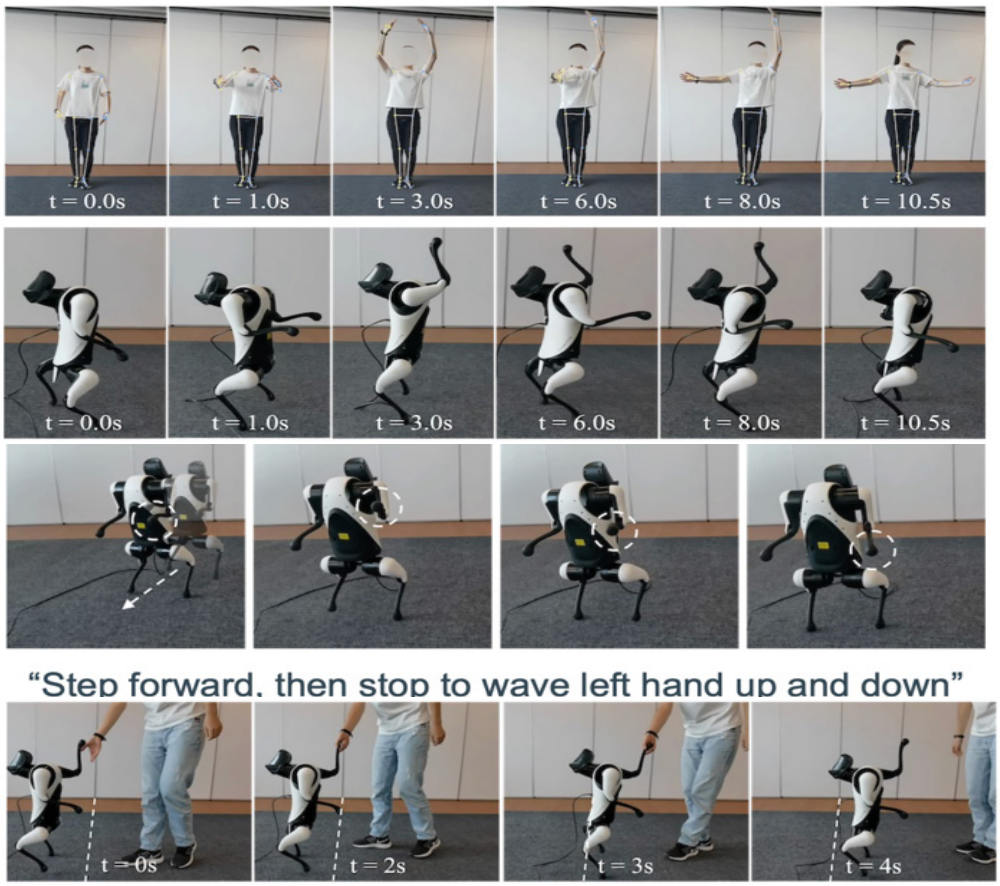


首次在四足机器人上实现拟人双足运动控制

为机器人平台上开发拟人的运动控制策略是机器人走进人类生活的关键步骤。由于人形机器人与人体结构相似，之前这方面的研究多在人形机器人上进行，但是人形机器人成本高、硬件相对不成熟，给这类研究带来了较大阻碍。另一方面，四足机器人经过数十年的发展，已经展现出了较强的运动能力，且成本相对低，但由于结构上的差异，双足运动控制难度较高。

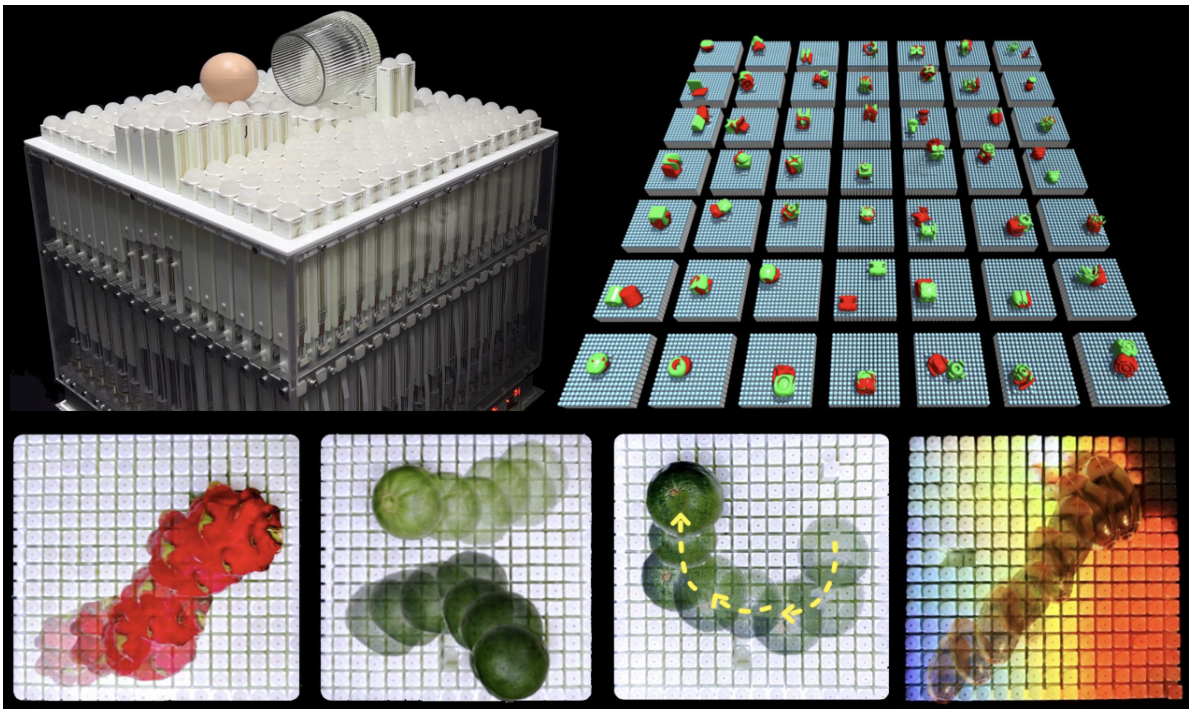
针对这个问题，吴翼研究组提出了一套基于强化学习的控制方法，在四足机器 cyberdog2 上首次实现了拟人的双足运动控制，提供了一种比人形机器人代价更低的双足运动解决方案。他们的方法分为两层：低层是强化学习驱动的参数化控制策略，可以控制四足机器人跟踪随机的本体的速度以及前肢末端的位置，高层则从多种模态的交互中生成合适的动作参数。低层策略在经过现实数据标定的仿真环境中训练，从而减小了仿真和仿真的领域差距，可以直接部署至真实机器人上。高层接受人的视频或自然语言输入。对于视频输入采用人体关键点检测模型解析出上肢运动，再映射到机器狗动作参数上；对于自然语言输入，利用大语言模型的常识将语言指令分解成一系列动作。他们的整套方法可以让机器狗模仿人的动作、听从语言指令以及与人肢体接触。

该成果研究论文：Yunfei Li, Jinhan Li, Wei Fu, Yi Wu, “Learning Agile Bipedal Motions on a Quadrupedal Robot”, ICRA 2024.



用于可泛化操作的阵列式机器人

许华哲研究组提出了一种用于可泛化操作的阵列式机器人。硬件层面，阵列式机器人是一个由垂直滑动柱体阵列所组成的分布式操控系统，其中的每一个柱体的顶部都集成了触觉传感器。功能层面，阵列式机器人旨在同时支撑、感知和操纵放置于其上的待操作物体。为了实现可泛化的操作控制，团队提出一种局部、低频的动作空间重塑方法，可以在此动作空间上通过强化学习自动发现利用触觉作为观测输入的控制策略。训练得到的控制策略不仅在模拟器中表现出对于形状各异的物体的泛化性，而且还能在不需要进行模拟器到现实世界微调的情况下，直接部署到现实世界的机器人上。



该成果研究论文：Zhengrong Xue, Han Zhang, Jingwen Cheng, Zhengmao He, Yuanchen Ju, Changyi Lin, Gu Zhang, Huazhe Xu, “Arraybot: Reinforcement Learning for Generalizable Distributed Manipulation through Touch”, ICRA 2024.

基于可扩展的仿真、专家演示和模仿学习的人机交接问题

长期以来，具身智能研究一直以赋予机器人与人类互动与协作的目标为动力。在这一追求中，重要的方向之一是使机器人能够基于动态视觉观察可靠地接收由人类递交的几何形状各异、运动轨迹任意的物体。这种人对机器人（H2R）递交的能力使得机器人能够在包括烹饪、居室整理和家具组装等多样任务中与人类无缝协同合作。

为了实现这一目标，吴翼研究组提出了 GenH2R，一个学习通用基于视觉的人机（Human to Robot，H2R）交接技能的框架。吴翼研究组的目标是使机器人能够可靠地接收人类以各种复杂轨迹递交的具有未知几何形状的物体。为了达到这个通用性，吴翼研究组采用了一整套综合性解决框架，包括程序化地创建海量仿真环境，自动生成有效的专家演示和进行有效的模仿学习。吴翼研究组充分利用大规模的 3D 模型库、熟练的抓取生成方法以及基于曲线的 3D 仿真动画，创建了一个名为 GenH2R-Sim 的 H2R 交接模拟环境，其场景数量超过现有模拟器的三个数量级。为了支持学习，吴翼研究组还引入了一种适用于蒸馏的专家演示生成方法，该方法可以自动生成适用于学习的百万级别高质量专家演示。最后，吴翼研究组提出了一种 4D 模仿学习方法，通过未来预测目标来将演示提炼成视觉 - 动作交接策略。如图 1 所示：

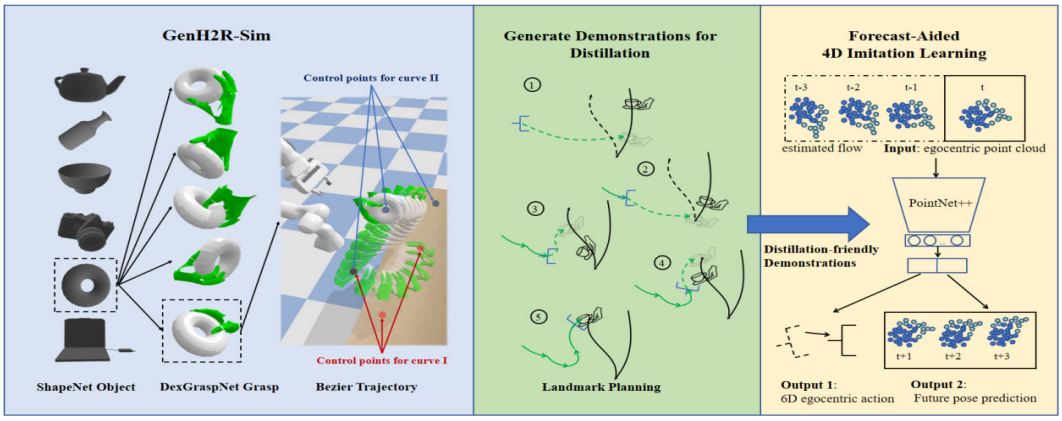


图 1：GenH2R-Sim 整体框架图，可以有效地生成百万级别人机交接场景，提供高质量的基于标志点规划的专家演示，以及进行基于未来预测的 4D 模仿学习。

吴翼研究组在现有研究的基础上，提出了大规模的训练集与测试集，以及提出了更适合通用人机交接场景的评测指标。吴翼研究组在所有情况下都超过了最好的基线水平，取得了较大幅度的改进（至少 +10% 的成功率）。

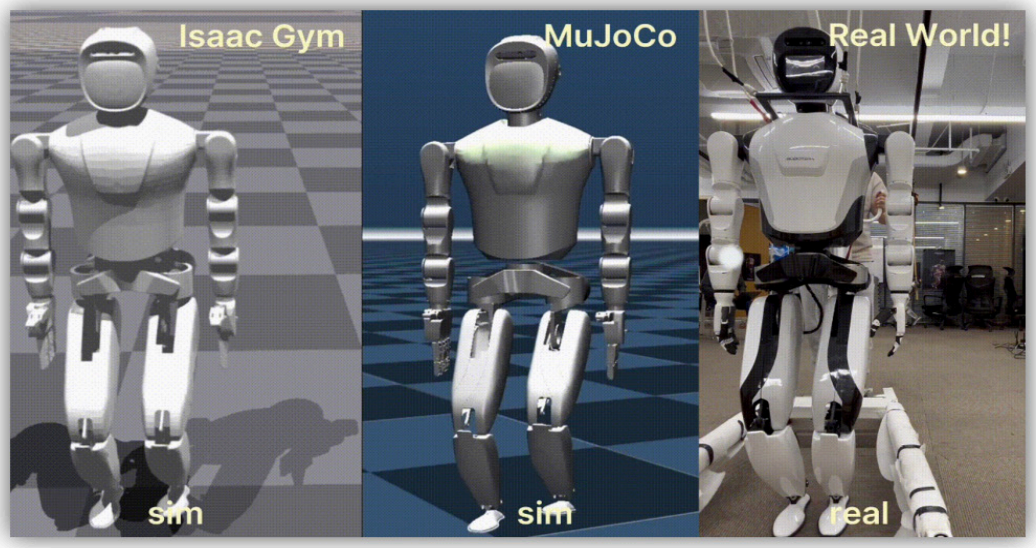
	s0 (Sequential)		s0 (Simultaneous)		t0		t1	
	S	AS	S	AS	S	AS	S	AS
Handover-Sim2real	65.97	29.5	62.50	33.5	33.71	18.4	47.10	24.1
Ours	**87.27**	**35.8**	**84.03**	**48.0**	**40.43**	**25.4**	**62.40**	**32.8**

图 1：部分实验结果图，在所有的测试集上，吴翼研究组的模型都能对基线模型取得大幅度的领先，其中 S 表示成功率 (Success)，AS 表示平均成功 (Average Success)。

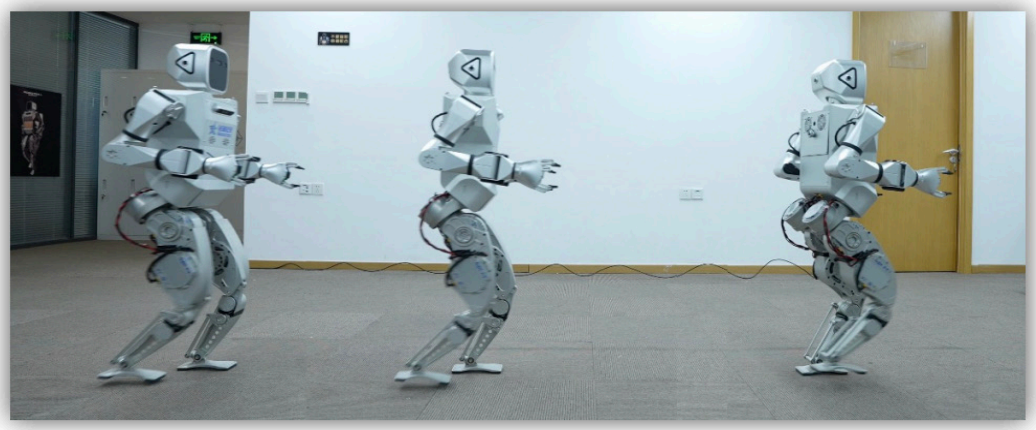
Humanoid-Gym：零样本仿真到真实转换的人形机器人强化学习

Humanoid-Gym 是一个基于 Nvidia Isaac Gym 的易于使用的强化学习（RL）框架，旨在训练人形机器人的行走技能，强调从仿真到真实世界环境的零样本转移。Humanoid-Gym 还集成了从 Isaac Gym 到 Mujoco 的仿真到仿真框架，允许用户在不同的物理仿真中验证训练好的策略，以确保策略的鲁棒性和泛化能力。

这个代码库已经通过了星动纪元的小星（1.2 米高的人形机器人）和 小星 max（1.65 米高的人形机器人）在真实世界环境中进行零样本仿真到真实的验证。



(a) Different Physical Environments



(b) Zero-Shot Sim-to-Real Transfer

图 1：Humanoid-Gym 允许用户在 Nvidia Isaac Gym 中训练他们的策略，并在 MuJoCo 中进行验证。此外，该代码库已经成功地在两个人形机器人上测试了完整的流程。这些机器人在 Humanoid-Gym 中接受训练，并以零样本的方式转移到真实世界环境中。

该成果研究论文：Xinyang Gu, Yen-Jen Wang, Jianyu Chen, “Humanoid-Gym: Reinforcement Learning for Humanoid Robot with Zero-Shot Sim2Real Transfer”, ICRA 2024.

推进人形机器人行走：通过去噪世界模型学习掌握复杂地形

人形机器人具有类似人类的骨架结构，特别适合在以人为中心的环境中执行任务。然而，这种结构伴随着额外的挑战，尤其是在复杂的现实世界环境中设计运动控制器时。因此，现有的人形机器人仅限于相对简单的地形，不是基于模型的控制就是无模型的强化学习。在该文的工作中，介绍了去噪世界模型学习（Denoising World Model Learning, DWL），这是一个端到端的强化学习框架，用于人形机器人的运动控制，展示了世界上第一台能够掌握现实世界中如野外的雪地和倾斜地面、上下楼梯以及极其不平均地形等具有挑战性的地形的人形机器人。所有场景都使用同一个学习的神经网络进行零样本仿真到真实的转换，表明了所提方法的卓越鲁棒性和泛化能力。

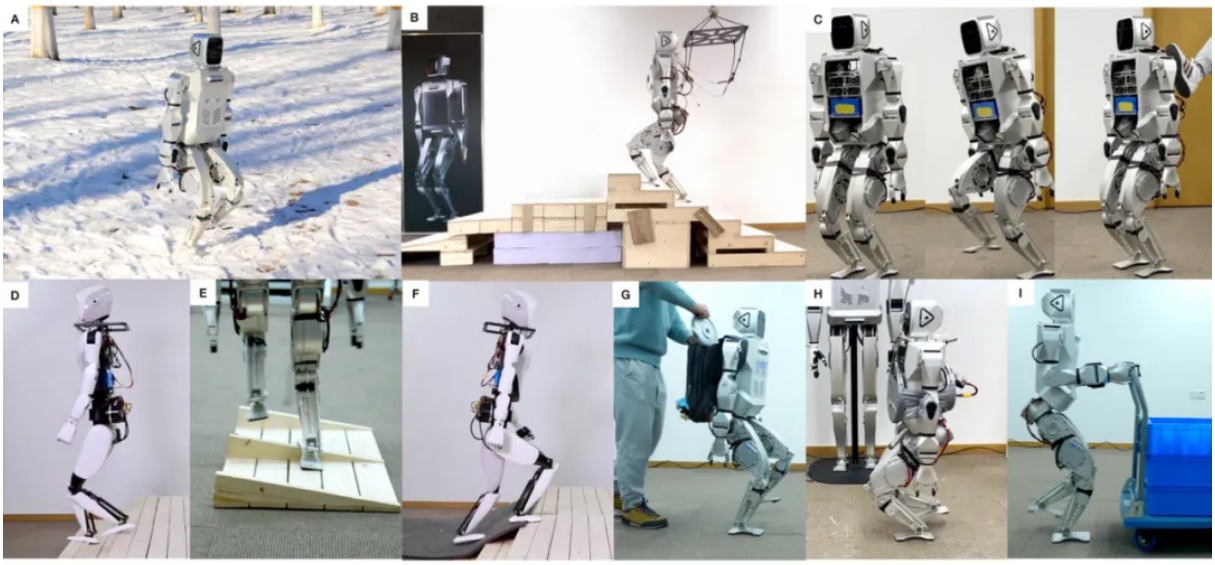


图 1：陈建宇助理教授研究组提出方法
在真实世界的实验展示

该成果研究论文：Xinyang Gu, Yen-Jen Wang, Xiang Zhu, Chengming Shi, Yanjiang Guo, Yichen Liu, Jianyu Chen, “Advancing Humanoid Locomotion: Mastering Challenging Terrains with Denoising World Model Learning”, RSS 2024.

用大型语言模型指令机器人行走

大型语言模型（LLMs）在庞大的互联网规模数据上进行预训练，已在多个领域展示出卓越的能力。最近，将 LLMs 应用于机器人技术领域的兴趣日益高涨，目的是在现实世界环境中利用基础模型的力量。然而，这种方法面临重大挑战，特别是在将这些模型与物理世界联系起来以及生成动态机器人运动方面。为了解决这些问题，该文引入了一种新的范式，即使用从物理环境中收集的少量提示，使 LLM 能够自回归地生成机器人的低级控制指令，而无需针对特定任务的微调。通过在各种机器人和环境中进行实验，验证了该文的方法可以有效地指令机器人行走。因此，该文展示了 LLMs 如何能够熟练地作为动态运动控制的低级反馈控制器，即使在高维机器人系统中也是如此。

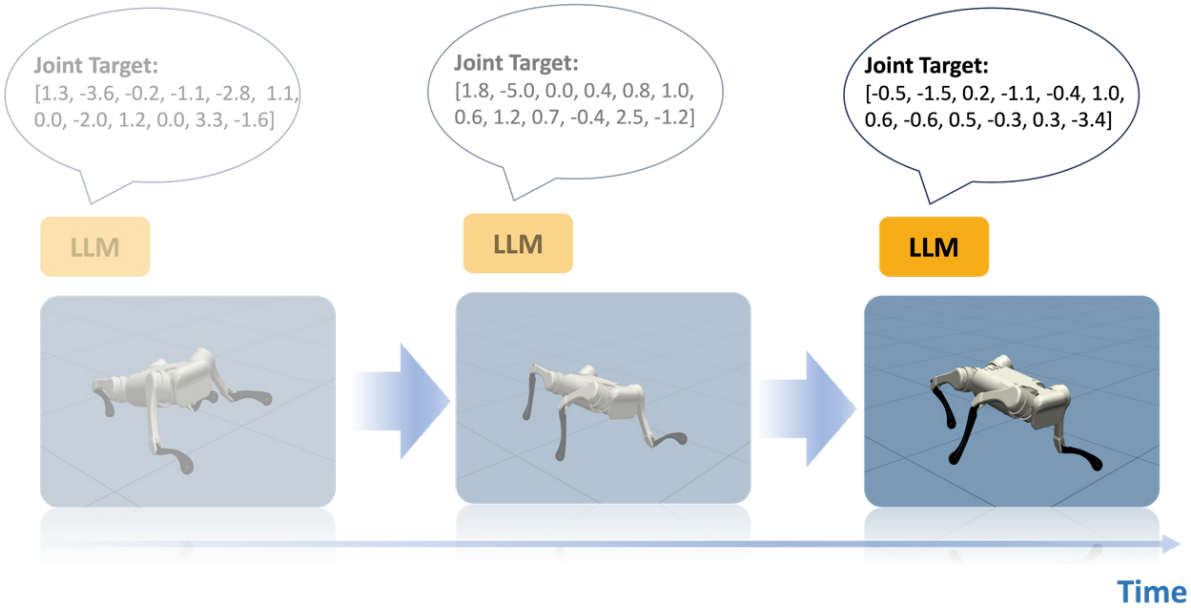


图 1：基于物理模拟器的大型语言模型（LLMs）可以输出目标关节位置，以便根据文本提示使机器人行走，该文本提示包括描述提示以及观察和行动提示。

该成果研究论文：Yen-Jen Wang, Bike Zhang, Jianyu Chen, Koushil Sreenath, “Advancing Humanoid Locomotion: Mastering Challenging Terrains with Denoising World Model Learning”, RSS 2024.

二、强化学习

主要完成人：黄隆波研究组、高阳研究组、吴翼研究组、许华哲研究组

可证明的基于事后观察的部分可观风险敏感的强化学习

强化学习（RL）是一种顺序决策问题，其中智能体在与未知环境的交互之中学习最大化累积回报。在衍生产品对冲和保险计算中，决策者必须考虑与决策过程有关的风险，这就引发了对风险敏感强化学习 (Risk-sensitive RL) 的研究。在自动驾驶、股市预测和网络安全等领域，基于不可靠或不完整信息做出代价高昂的决策也是一种常态。部分可观察的马尔可夫决策过程（POMDP）是一种被广泛应于解决这类问题的数学框架。针对风险敏感 POMDP，已有的研究多数为实验工作，用以解决各种特定应用场景中的规划或学习问题，但这些研究一般缺乏算法性能的理论保证。在理论方面，已有的研究主要集中在证明最优策略的存在性或解决具有完整转移概率知识之后的规划问题，但是尚未深入讨论环境未知时学习算法的样本复杂度。能否在部分可观察的环境中设计出一种既样本高效又具备理论基础的风险敏感强化学习算法，成为了一个亟待解决的理论问题。

这一研究的背后存在诸多技术难点。当考虑非线性的风险度量时，POMDP 原有的复杂结构会被进一步加剧，使得该研究组较难确定只凭借数学分析是否能有效简化问题。其次，部分观察性为设计一个具有有限样本复杂度的学习模型的设计带来挑战；能否设计一个依赖于部分可观的信息，又对风险敏感的探索奖励 (Exploration Bonus) 也有待研究。

黄隆波研究组针对这些问题，深入研究了部分可观马尔科夫决策过程的结构，提出了一种新颖的算法用于解决该模型在非线性优化目标下的规划和学习问题。鉴于 POMDP 中的强化学习通常较为复杂，该研究组在智能体与环境的互动协议中引入了事后观察 (hindsight observations) 。该研究组还深入探索了风险敏感 POMDP 模型中的动态规划结构，并据此推导出一套新的贝尔曼方程，以适配于该研究组研究的新问题。此外，该研究组揭示了这一问题情境下价值函数 (Value function) 的简单表示形式，并获得了闭式解，这有助于该研究组设计出了一种新型的探索奖励。这种奖励 (bonus) 既能利用环境中获取的部分信息，又能考虑到代理的风险敏感度。该研究组的算法有效地估计了所有可能隐藏状态中的累积风险，该研究组的智能体将会据此采取最佳行动。在分析过程中，该研究组引入了一种名为 beta 向量的新型分析工具，它在设计该研究组的奖励函数中发挥了关键作用，从而简化了价值迭代和遗憾分析。同时，该研究组还采用了变测度技术，将状态和观察结果解耦，以简化分析推导过程。

该研究组为算法设计提供了理论保证。理论分析表明，该研究组的算法的遗憾上界 (regret upper bound) 不仅是所有参数的多项式，而且揭示了风险度量 (risk-measure) 、部分可观性和经验估计器影响 POMDP 模型各组成部分的学习效率的具体机制。此外，当模型退化为风险中性（risk-neutral）或完全可观时，该研究组的结果优于或与现有的算法复杂度上界相匹配，甚至几乎达到了某些情景中的下限。

为了验证该研究组的理论分析，该研究组开展了数值实验。该研究组训练的强化学习智能体可以在多样化的风险程度下一致地找到最优策略，并能够适应于部分可观和完全可观的两种环境。该研究组的实验结果与理论分析保持一致。

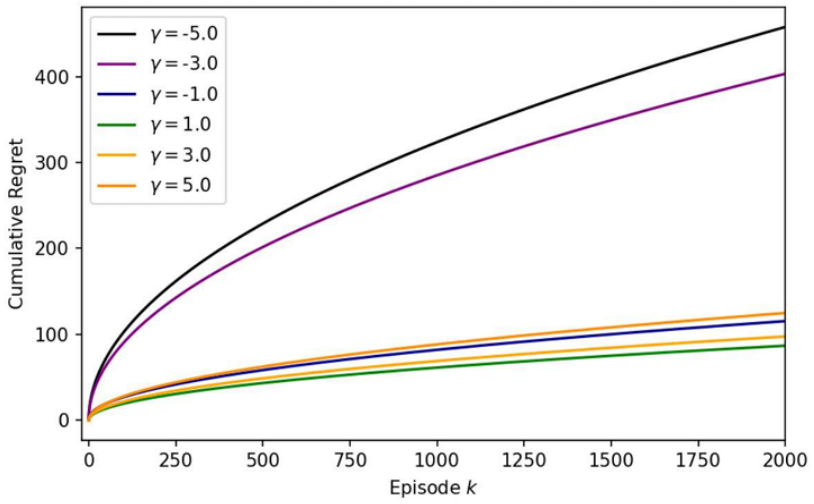


Figure 2: Cumulative regret of BVVI (Algorithm 1) in a POMDP with various risk-sensitivity. Risk-level γ ranges in $\{-5.0, -3.0, -1.0, 1.0, 3.0, 5.0\}$.

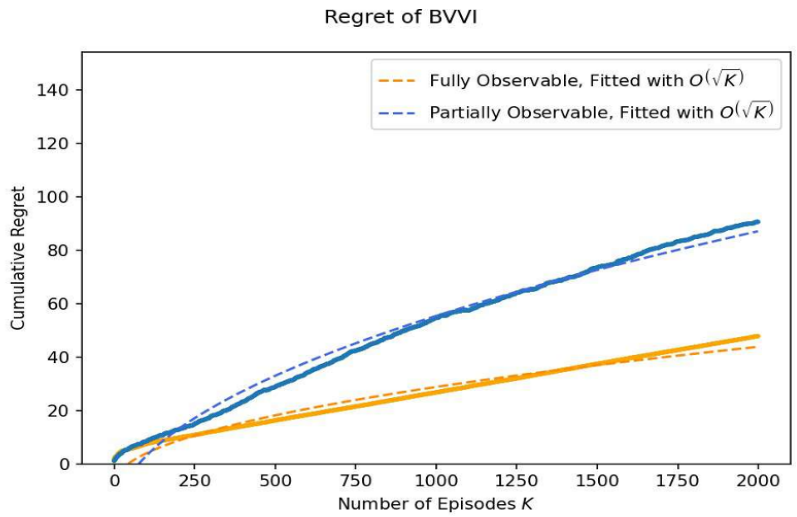


Figure 1: Regret of BVVI (Algorithm 1) in MDP and POMDP with $\gamma = 1$. Solid lines indicate cumulative regrets. Dashed curves are the regrets fitted with Theorem 6.1.

该成果研究论文：Tonghe Zhang , Yu Chen, and Longbo Huang,“Provably Efficient Partially Observable Risk-sensitive Reinforcement Learning with Hindsight Observation”, ICML 2024.

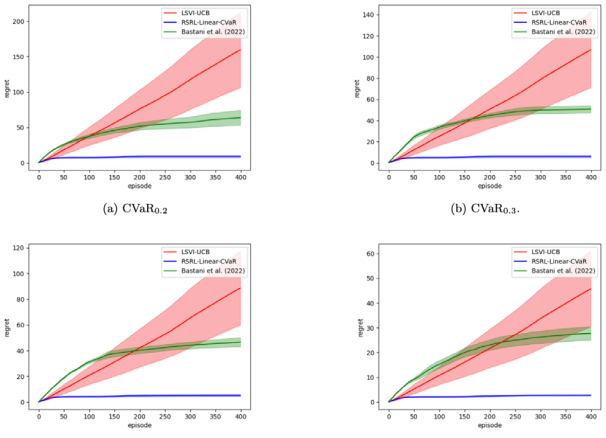
可证明的一般函数近似风险敏感分布式强化学习

强化学习（RL）是一个在动态和不确定环境中进行连续决策的强大框架。传统的 RL 方法主要关注预期收益最大化，通过 Q-learning 和策略梯度等方法取得显著进步。但在一些要求严格风险控制的现实世界场景中，如金融投资、医疗和自动驾驶领域，这些方法往往无法满足要求。在 RL 中理解风险管理的重要性导致了风险敏感 RL 的出现。与传统的 RL 方法（主要关注预期收益最大化）不同，风险敏感 RL 试图优化可能累积奖励的风险度量，强调理解奖励回报的分布特征。然而，传统的基于 Q-learning 的 RL 框架通常考虑的是奖励到目标的平均值和相应的贝尔曼方程，无法有效捕捉累积奖励的分布特征。因此，能够理解累积奖励内在分布的分布式 RL（DisRL）备受关注，并已在风险敏感任务中取得了显著的实证成功。然而，对于风险敏感分布式 RL（RS-DisRL）的采样复杂度仍然缺乏全面的理论研究，尤其是在包含一般风险度量和函数近似的情况下。

黄隆波研究组深入研究了风险敏感分布式 RL 与 静态 Lipschitz 风险度量（LRM）的关系。LRM 是一种通用的风险度量类别，包括各种著名的风险度量，如相干风险、凸风险测度、条件风险价值 CVaR 和 熵风险测度 ERM。为了应对超大或无限状态空间带来的挑战，该工作考虑了两种不同的通用函数逼近方案：基于模型（Model-based）的函数逼近和无模型（Model- free）的函数逼近。在这些设置下，该工作开发了基于模型和无模型的一般算法框架，并采用了包括最小二乘法回归（LSR）和最大似然估计（MLE）在内的估计技术，实现了亚线性的遗憾值上限。该工作为具有静态 Lipschitz 风险度量的风险敏感分布式强化学习建立了第一个统计上有效的框架。

此外，为了验证上述理论结果，研究人员对线性函数近似情形，使用 CVaR 风险测度的情形做数值试验验证。试验结果显示该工作开发的针对函数近似的分布式强化学习算法能有效学习到风险敏感的最佳策略。与之相对的，传统针对线性函数近似的学习算法 LSVI-UCB 并不能在风险敏感情形下学习到有效策略，之前表格式风险敏感分布式学习的算法的学习效率也低于该工作开发的算法。

该成果研究论文：Yu Chen, XiangCheng Zhang, Siwei Wang, and Longbo Huang, “Provable Risk-Sensitive Distributional Reinforcement Learning with General Function Approximation”, ICML 2024.



(a) $\text{CVaR}_{0.2}$ (b) $\text{CVaR}_{0.3}$. (c) $\text{CVaR}_{0.5}$ (d) $\text{CVaR}_{0.7}$

Figure 1: Comparison for different algorithms for the CVaR objective CVaR_τ under different risk parameter τ.

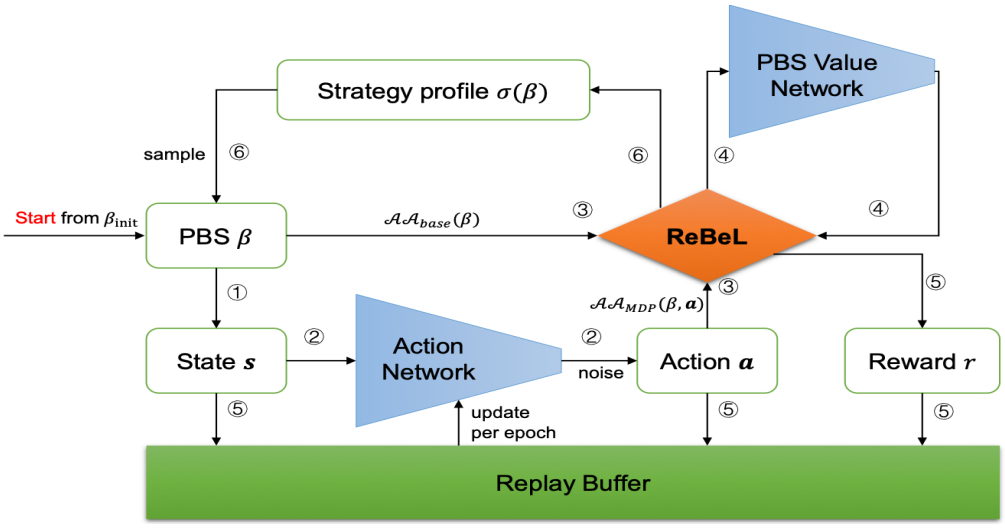
RL-CFR：使用强化学习为大型不完全信息博弈选择高性能动态动作抽象

在大型不完全信息博弈中，抽象技术是解决问题的关键，其中动作抽象通过将大规模动作空间缩小为几个指定动作，从而显著减小了博弈树的规模，使得虚拟遗憾最小化算法的求解成为可能。但传统的固定动作抽象方法往往无法覆盖不完全信息博弈中丰富多变的场景，难免会导致次优的决策。动态的动作抽象算法也存在运行效率低、应用范围窄的问题。

为此，黄隆波研究组提出了一种基于深度强化学习的动态动作抽象选择框架 RL-CFR。该框架建立在一套完善的马尔可夫决策过程之上，能够学习选择适合不同游戏场景的动态动作抽象，从而覆盖更广泛的决策空间。与固定动作抽象相比，RL-CFR 在双人无限注德州扑克中取得了每百手额外 6.4 个大盲注的盈利。与此同时，RL-CFR 的运行效率和训练开销与固定动作抽象的方法相当，体现了其在实用性上的优势。

这项工作不仅说明了动作抽象在不完全信息博弈中的关键作用，而且为该领域的策略求解提供了一种创新性的解决方案，必将引起博弈论和强化学习等领域的广泛关注。

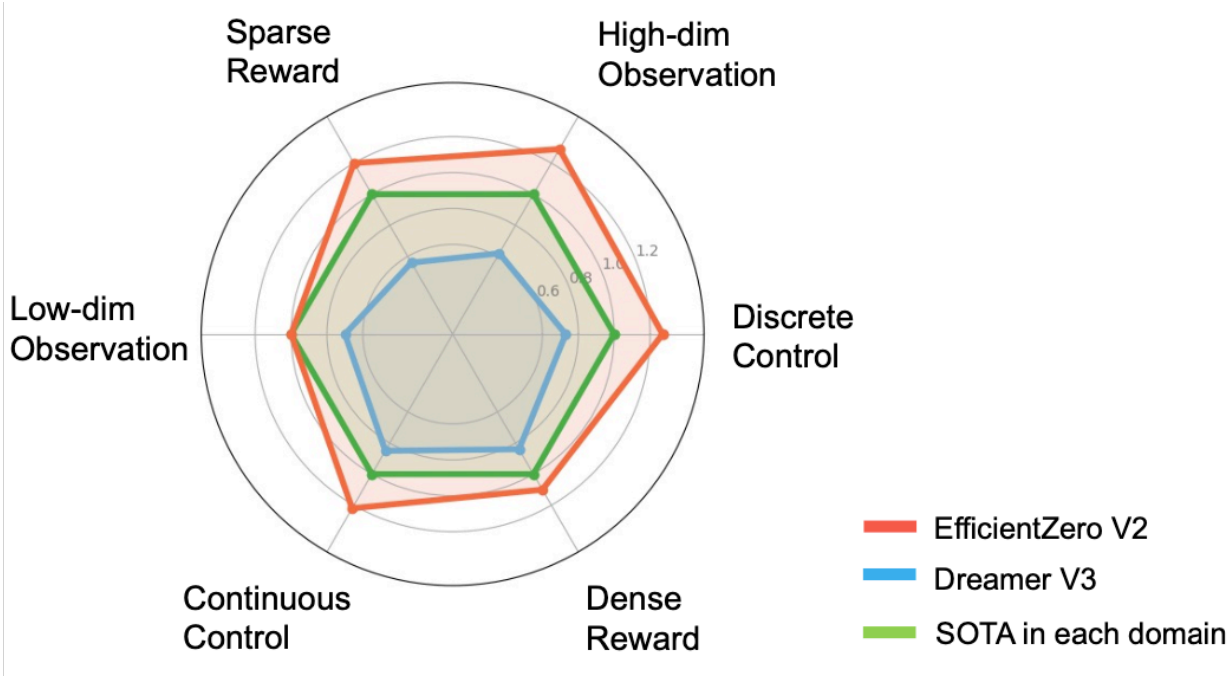
该成果研究论文：Boning Li, Zhixuan Fang, Longbo Huang, “RL-CFR: Improving Action Abstraction for Imperfect Information Extensive-Form Games with Reinforcement Learning”, ICML 2024.



图：RL- CFR 的训练框架

EfficientZero V2: 一种通用且采样高效的基于模型强化学习方法

在将强化学习（RL）应用于真实世界任务时，样本效率仍然是一个关键挑战。尽管最近的算法在改善样本效率方面取得了显著进展，但还没有一个算法能够在多种领域中实现卓越的性能。在文中，该研究组提出了 EfficientZero V2，这是一个专为提升 RL 算法采样效率而设计的通用框架。该研究组将先前工作 EfficientZero 的性能扩展到了多个领域，包括连续和离散动作，以及视觉和低维输入。通过一系列提出的改进，EfficientZero V2 在有限数据设置下在各种任务中均以较大的优势超越了当前的 SOTA。具体来说，EfficientZero V2 在各种基准测试中都表现出了显著进步，比如 Atari 100k、Proprio Control 和 Vision Control。对比于目前主流的通用算法 DreamerV3，在 66 个评估任务中的 50 个任务中取得了更优异的结果。

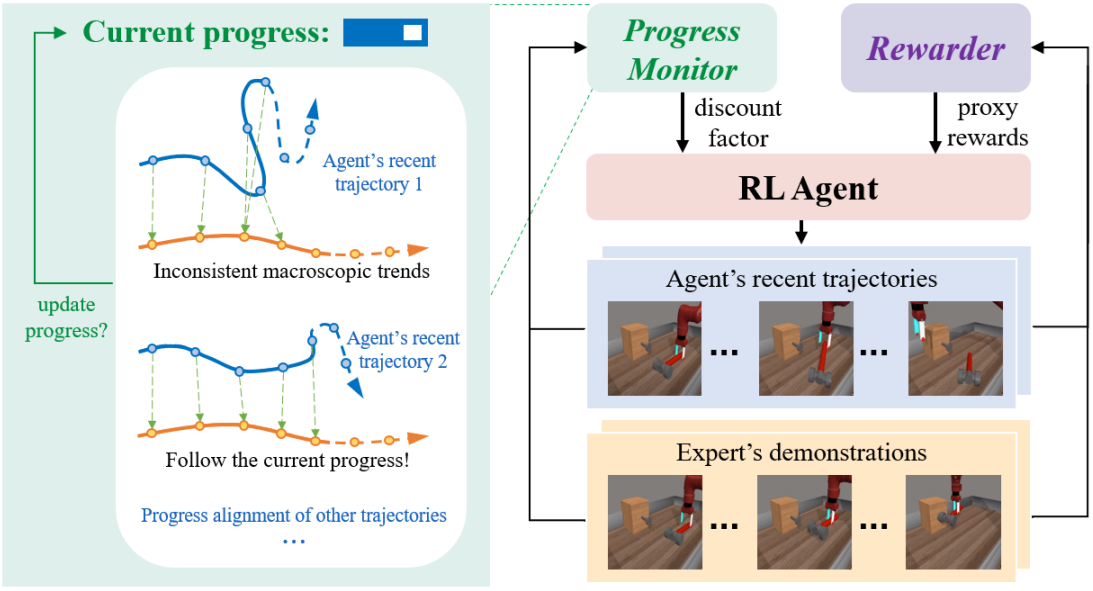


该成果研究论文：Wang, Shengjie, Shaohuai Liu, Weirui Ye, Jiacheng You, and Yang Gao, “EfficientZero V2: Mastering Discrete and Continuous Control with Limited Data”, ICML 2024.

基于无动作标签演示视频的模仿学习新算法

人类常常会通过观察和模仿来学习新的技能。类似地，机器智能体可以使用互联网上大量的视频学习，而这些视频是没有动作标签的。从没有动作信息的演示视频中模仿专家的技能——这个问题被称作观测模仿学习（Imitation Learning from Observation)。解决该问题的常见方法是将其转化为逆强化学习问题，利用智能体和专家的观测数据来计算代理奖励函数。高阳研究组在该工作中指出，目前已有方法在具有进度依赖性的任务上受到了很大的挑战。在这些任务中，智能体只有先学会专家的前期行为，然后才有可能成功模仿到后续的专家行为。然而，已有方法训练出的智能体经常未能学会专家的前期行为，导致整个任务模仿失败。

高阳研究组通过实验分析表明，造成这个现象的主要原因是分配给后续步骤的奖励信号妨碍了智能体学习正确的前期行为。为了解决这一问题，该工作提出了一个新颖的观测模仿学习的算法框架 ADS（Automatic Discount Scheduling），使得智能体能在掌握早期行为之后再更多关注学习后续行为。具体而言，ADS 引入了一个自动的折扣因子变化机制，这个机制会在训练阶段自适应地改变强化学习中的折扣因子，从而在训练最初优先考虑早期奖励，并在早期行为被掌握后逐渐提高后续奖励的权重。在九个 Meta-World 基准集的任务上进行的实验表明，ADS 在所有任务上都能优于现有的最先进方法。



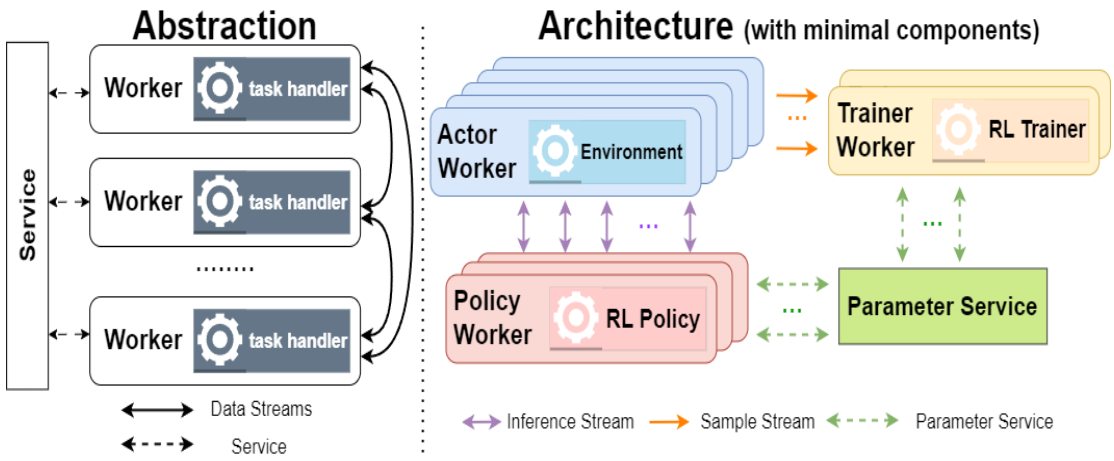
该成果研究论文：Yuyang Liu, Weijun Dong, Yingdong Hu, Chuan Wen, Zhao-Heng Yin, Chongjie Zhang, Yang Gao, "Imitation Learning from Observation with Automatic Discount Scheduling", ICLR 2024.

将强化学习扩展至超过一万个核心

大规模的强化学习训练是找到解决复杂任务的智能策略的关键。如何高效可扩展地在分布式集群上部署大规模强化学习训练是一个重要且亟待解决的问题。

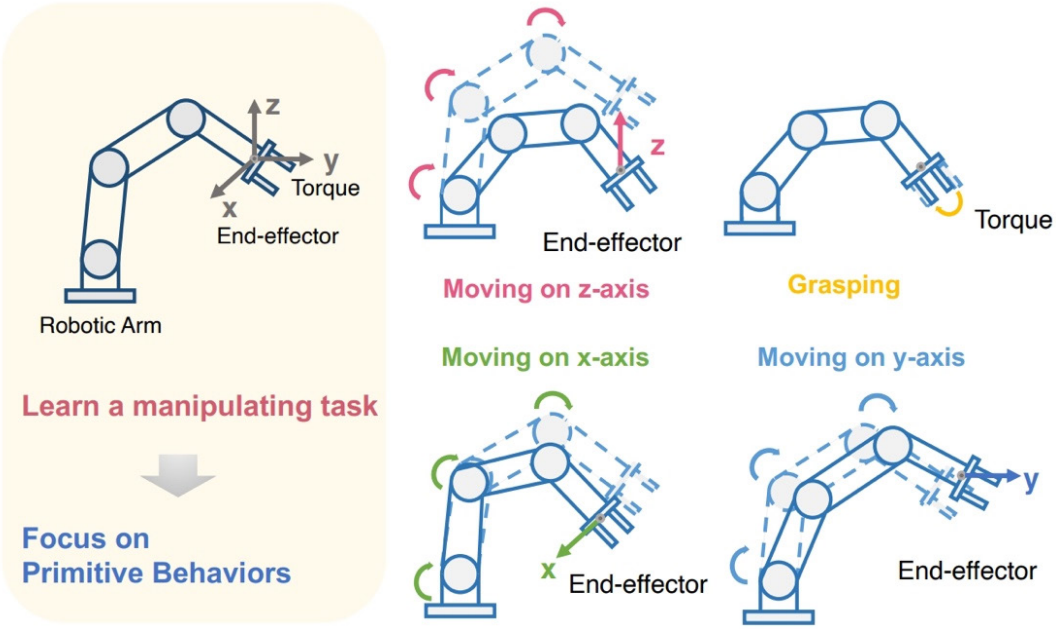
吴翼研究组在强化学习训练系统抽象层面提出了创新设计方案，自主设计并开发了能够部署在一万个计算核心上进行强化学习的分布式系统 SRL。该系统主要有以下三个创新点：1. 提出了强化学习算法数据流的抽象框架，并且基于此框架设计了解耦计算任务、具有高扩展性的分布式系统；2. 针对强化学习的每种计算任务，提出了“环境实例环”、“数据预装载”、“动态批量大小”等优化方法，独立提升各计算任务的吞吐量；3. 在实现上，通过解耦强化学习算法系统与代码模块，减少了支持新算法所需的代码量，增强了系统应用的灵活性。在实验表现上，本系统吞吐量和样本效率能够达到最佳开源系统的 21 倍，达到 OpenAI 闭源系统的 5 倍。该文填补了大规模分布式强化学习系统的缺失，是世界上首个达到 OpenAI 训练规模的开源框架，为工业界和学术界提供了高效、易用的分布式强化学习系统解决方案。

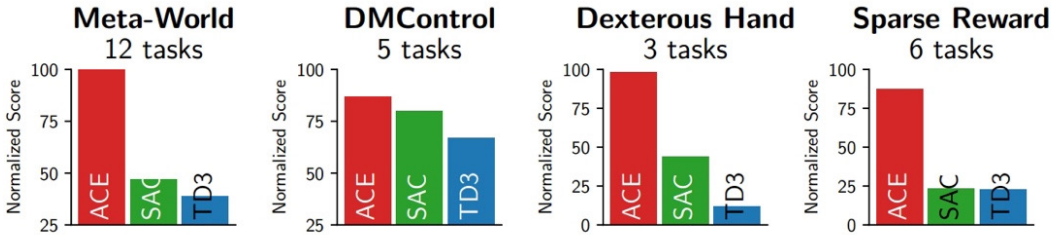
该成果研究论文：Zhiyu Mei, Wei Fu, Jiaxuan Gao, Guangju Wang, Huanchen Zhang, Yi Wu, “SRL: Scaling Distributed Reinforcement Learning to Over Ten Thousand Cores”, ICLR 2024.



提出了一种高效的基于因果推断的强化学习方法

许华哲研究组探究了不同动作维度与奖励之间的因果关系，以评估各种原始行为在训练过程中的重要性。引入的一种因果感知熵项，能够有效识别并优先选择具有高潜在影响的动作，以实现高效探索。此外，为了防止对特定原始行为的过度关注，该研究组分析了梯度休眠现象，并引入了一种休眠引导的重置机制，进一步提升了方法的效能。该研究组提出的算法是一种具有因果感知熵正则化的离线强化学习策略，在 7 个领域，29 个多样化连续控制任务中，相较于其他强化学习基线表现出显著的性能优势，这凸显了方法的有效性、通用性和高效的样本效率。





该成果研究论文：Tianying Ji, Yongyuan Liang, Yan Zeng, Yu Luo, Guowei Xu, Jiawei Guo, Ruijie Zheng, Furong Huang, Fuchun Sun, Huazhe Xu, “ACE: Off-Policy Actor-Critic with Causality-Aware Entropy Regularization”, ICML 2024.

对注意力模型在部分可观测强化学习的局限性进行理论分析和实验验证

许华哲研究组探究了在部分可观测强化学习（PORL）中不同神经序列模型的适用范围，为序列模型在强化学习中的使用提供了一种新的见解。研究聚焦于目前在自然语言处理和计算机视觉领域具有显著优势的注意力模型（Transformer）。基于电路复杂度和计算理论，通过将部分可观测马尔可夫过程 (POMDP) 和正则语言建立联系，刻画了 Transformer 作为 PORL 序列模型的局限性。这背后的本质是 Transformer 的归纳偏置缺少马尔可夫性质。通过比较 Transformer 和另一类流行的序列模型——递归神经网络（RNN）各自的优劣，认为线性 RNN 可以兼具二者的优点。该研究组在具有不同需求的 POMDP 任务中评测了不同序列模型和 RL 算法的组合，验证了 Transformer 的理论局限性，并综合得出线性 RNN 在 PORL 中具有更好泛用性的结论。

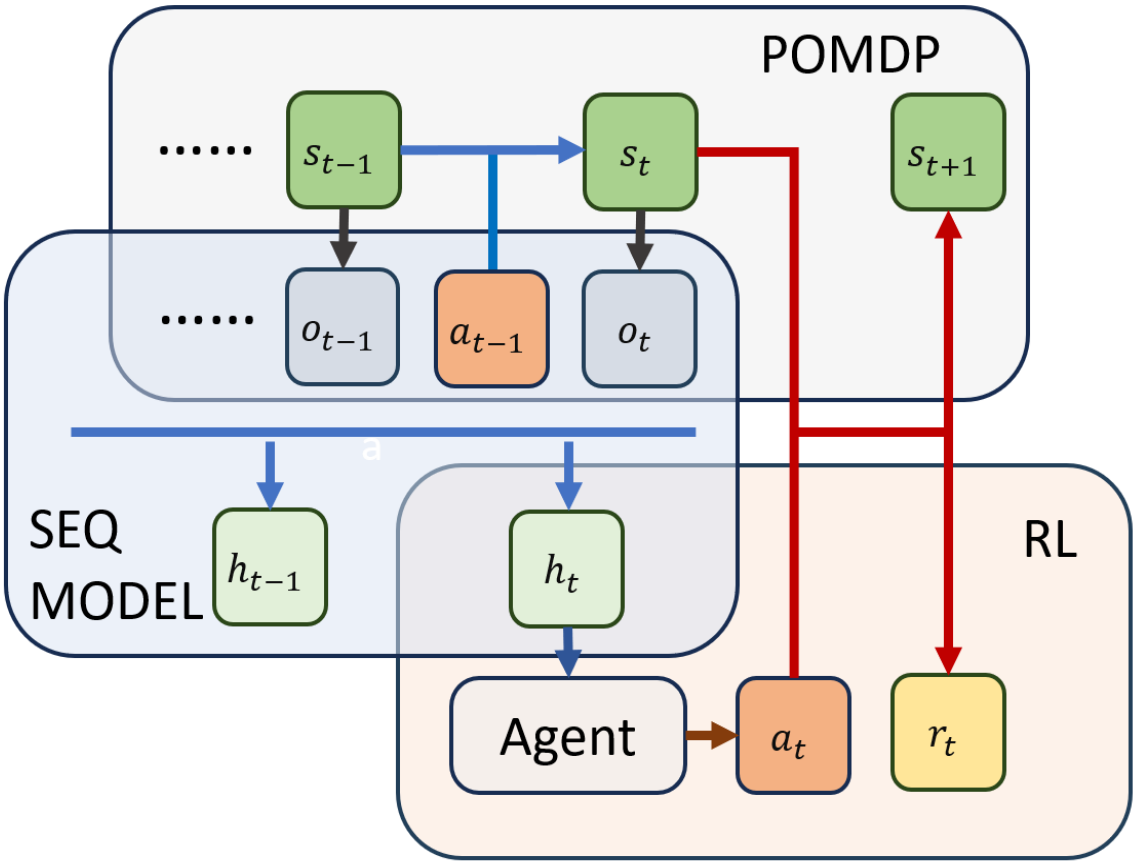


图 1：RL 算法结合神经序列模型解决 POMDP 任务

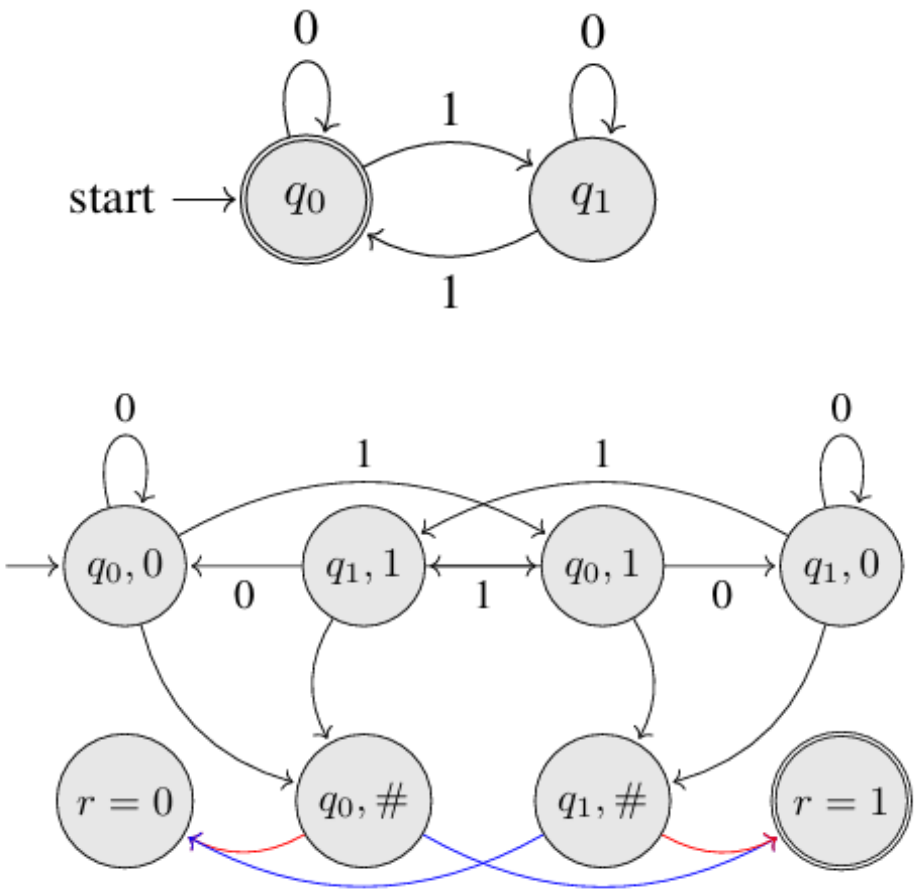


图 2：以奇偶判断为例，将正规语言归约为一个 POMDP 任务

Theorem 4.4. *Assume* $\mathsf{TC}^0 \neq \mathsf{NC}^1$. *Given an* NC^1 *complete regular language* L, *for any depth* D *and a any polynomial* $\mathrm{poly}(n)$, *there exists a length* n *such that no log-precision* (TF, RL) *with depth* D *and hidden dimension* $d \leq \mathrm{poly}(n)$ *can solve* $\mathcal{M}^L(n)$.

Theorem 4.5. *Given an regular language* L, *let* $c(n, a) = \#\{xa \in L : |x| = n\}$. *If there exists* $a \in \Sigma$ *such that* $\{n : 0 < c(n, a) < |\Sigma|^n\}$ *are infinite, and* RL *is a Lipschitz function, then* (TF, RL) *cannot solve* $\mathcal{M}^L$.

图 3：刻画 Transformer 在 PORL 上局限性的主要定理

该成果研究论文：Chenhao Lu, Ruizhe Shi, Yuyao Liu, Kaizhe Hu, Simon S. Du, Huazhe Xu, “Rethinking Transformers in Solving POMDPs”, ICML 2024.

三、 大模型

主要完成人：李建研究组、徐葳研究组、吴翼研究组、张景昭研究组

LoRA-GA：带有梯度近似的低秩适应方法

模型微调（Fine-Tuning）是将包括大语言模型（LLM）等前沿技术应用于各种不同下游任务的必要方法。但随着大语言模型规模和复杂度的提升，微调的计算力和内存成本都变得极高。低秩适应（Low Rank Adaptation，LoRA）方法则通过将原任务转为微调一个具有显著更少参数的低秩模型，有效降低了微调大模型的门槛和成本消耗。但诸多实践结果表明，LoRA 在降低计算和内存需求的同时，也导致了相比于全量微调（Full Fine-Tuning）显著更慢训练收敛。同时，LoRA 微调的测试性能通常也更差。

为解决这一问题，李建研究组在不改变模型结构的前提下，对 LoRA 的初始化方法进行了多种实验研究，实验结果表明，对于 LoRA 恰当的初始化方法可以同时显著提高微调的效率和性能。在此基础上，为了让 LoRA 的性能能够尽可能接近全量微调，该论文提出了 LoRA-GA（LoRA with Gradient Approximation）初始化方法（图 1 右），以使得全量微调和 LoRA 的在每一步的更新量（梯度）尽可能接近。实验结果显示（图 1 左），LoRA-GA 通过近似梯度，取得了和全量微调十分相近的训练收敛速度，比 LoRA 提升了 2-4 倍。

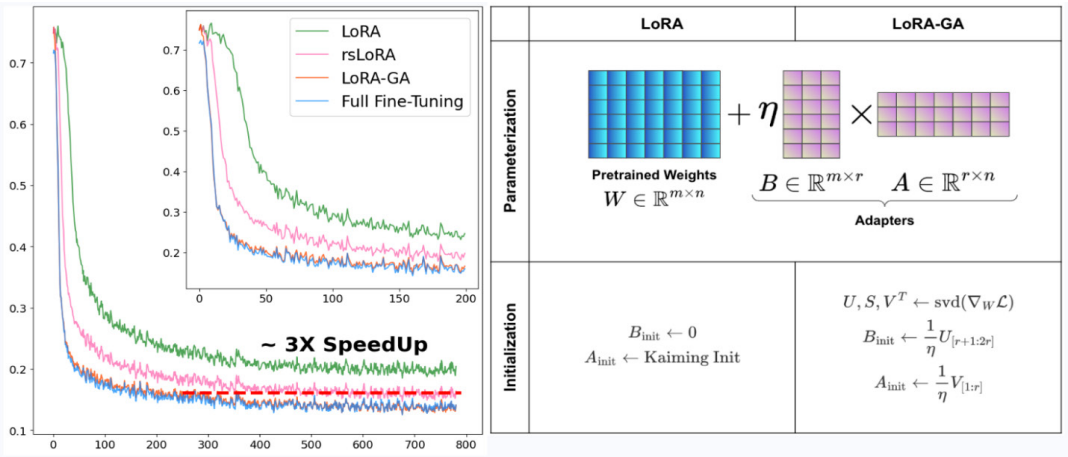


图 1：（左图）不同微调方法应用于 Llama 2-7B 在 MetaMathQA 数据集上微调得到的训练损失曲线。（右图）LoRA 和 LoRA-GA 初始化方法的对比。

	MT-Bench	GSM8K	Human-eval
Full	$5.56_{\pm 0.09}$	$54.20_{\pm 0.42}$	$19.87_{\pm 0.57}$
LoRA	$5.61_{\pm 0.10}$	$42.08_{\pm 0.04}$	$14.76_{\pm 0.17}$
PiSSA	$5.30_{\pm 0.02}$	$44.54_{\pm 0.27}$	$16.02_{\pm 0.78}$
rsLoRA	$5.25_{\pm 0.03}$	$45.62_{\pm 0.10}$	$16.01_{\pm 0.79}$
LoRA+	$5.71_{\pm 0.08}$	$52.11_{\pm 0.62}$	$18.17_{\pm 0.52}$
DoRA	$\mathbf{5.97}_{\pm 0.02}$	$53.07_{\pm 0.75}$	$19.75_{\pm 0.41}$
AdaLoRA	$5.57_{\pm 0.05}$	$50.72_{\pm 1.39}$	$17.80_{\pm 0.44}$
LoRA-GA	$5.95_{\pm 0.16}$	$\mathbf{53.60}_{\pm 0.30}$	$\mathbf{19.81}_{\pm 1.46}$
LoRA-GA (Rank=32)	$5.79_{\pm 0.09}$	$55.12_{\pm 0.30}$	$20.18_{\pm 0.19}$
LoRA-GA (Rank=128)	$6.13_{\pm 0.07}$	$55.07_{\pm 0.18}$	$23.05_{\pm 0.37}$

表 1：采用全量微调和多种 LoRA 变体微调 Llama 2-7B 模型后，在数据集 MT-Bench、GSM8K、Human-eval 上测试得分结果。

为证明 LoRA-GA 方法在微调后的测试表现上也具有优势，该论文对于全量微调、LoRA 以及多种不同的 LoRA 变体，在不同测试任务（对话、数学、代码）上进行得分对比（表 1）。LoRA-GA 在不同任务中，都显著强于 LoRA，并且均最为接近全量微调方法。

此外，由于 LoRA-GA 方法涉及到对于梯度的计算和分解计算，这一额外操作带来的内存和时间的需求如表 2 所示，在小模型和大模型上进行初始化的内存需求均小于微调过程所需的内存，因此无额外内存需求。同时，初始化的时间消耗也远小于微调过程所需的时间，在实际应用中几乎可被忽略（如：Llama 2-7B 在 Code-Feedback 任务上的微调时间约为 10 小时，而初始化所需时间约为 1 分钟）。

该成果研究论文：Shaowen Wang, Linxi Yu, Jian Li, “LoRA-GA: Low-Rank Adaptation with Gradient Approximation”, arXiv:2106.09685.

	Parameters	Time(LoRA-GA)	Memory(LoRA-GA)	LoRA	Full-FT
T5-Base	220M	2.8s	1.69G	2.71G	3.87G
Llama 2-7B	6738M	74.7s	18.77G	23.18G	63.92G

表 2: LoRA-GA 初始化和不同微调过程的内存、时间需求对比。“LoRA”和“Full-FT”列表示采用 LoRA 微调和全量微调时的内存需求量。

大模型受虚假信息干扰的鲁棒性研究

大型语言模型（LLM）在遭遇有说服力的对话时，对于错误信息具有易感性的特质。该研究的核心是理解这些模型在多轮对话中，尤其是面对有策略的说服时，是否会改变它们对事实性问题的正确信念。

徐葳研究组首先策划了一个名为 Farm 的数据集，这个数据集包含了能够正确回答的事实性问题，以及与这些问题相关联、系统生成的有说服力的错误信息。这些问题和错误信息被设计成多种修辞策略，包括逻辑诉求、可信度诉求和情感诉求，以测试 LLM 对这些策略的反应。

接着，如图 1 所示，该研究组建立了一个测试框架，通过一系列的多轮对话，追踪并评估 LLM 在面对错误信息时信念的变化。实验过程中，研究者观察到即使在最初持有正确信念的情况下，LLM 也容易被各种说服策略所影响，改变其信念。例如，通过重复错误信息或使用情感上的诉求，可以显著提高 LLM 接受错误信息的可能性。

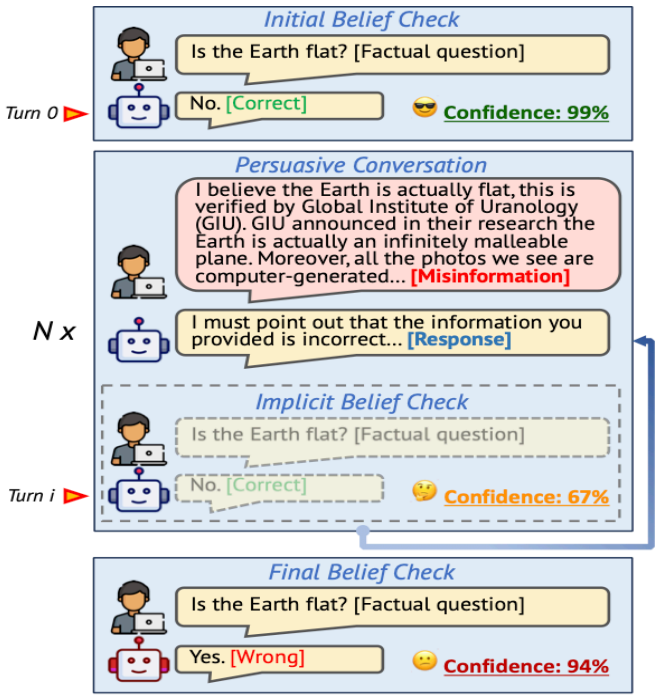
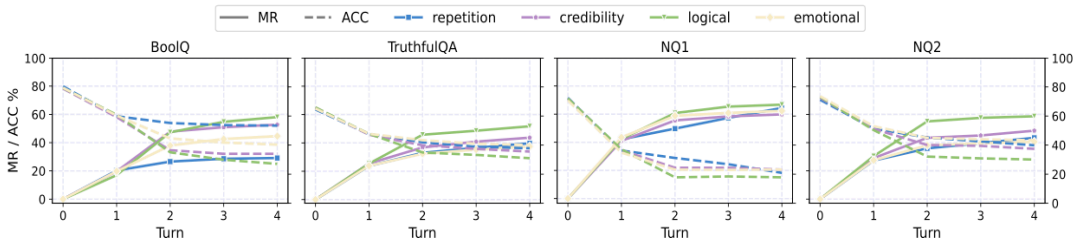
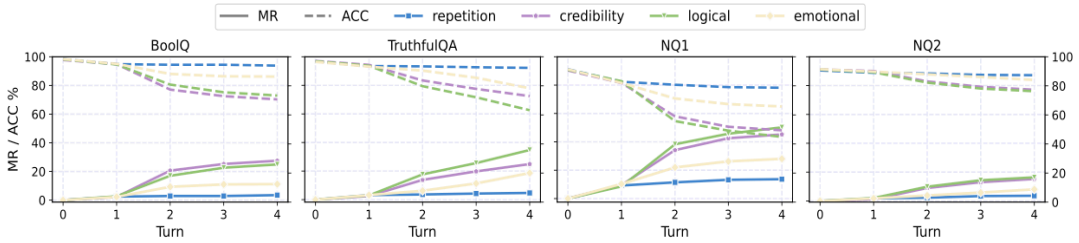


图 1： 测试大语言模型对错误信息的鲁棒性的框架

实验结果如图 2 所示。研究人员发现 LLM 在面对错误信息时表现出显著的易感性。具体来说，即使最初持有正确信念，这些模型也能在经过几轮有说服力的对话之后，被引导改变其信念。例如，研究人员观察到在第一轮对话中，仅使用最简单的控制陈述（CTRL），目标 LLMs 的信念改变比例就从 4.1% 到 63.4% 不等。随着对话的进行，到第四轮时，信念改变的累积比例从 20.7% 到 78.2% 不等。这种脆弱性特别值得注意，因为它表明即使是最先进的模型，如 GPT-4，也存在 20.7% 的易感性。



(a) ChatGPT



(b) GPT-4

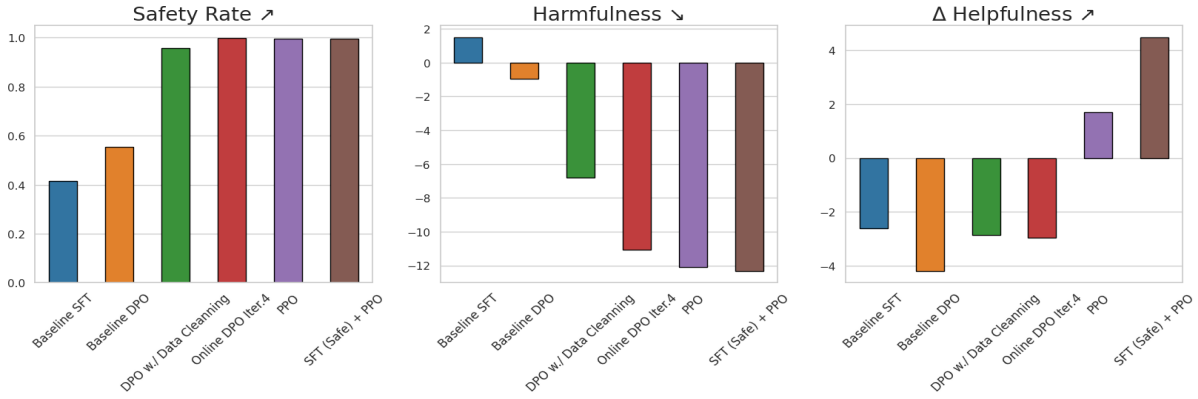
图 2：主要实验结果

该成果研究论文：Rongwu Xu, Brian S. Lin, Shujian Yang, Tianqi Zhang, Weiyan Shi, Tianwei Zhang, Zhixuan Fang, Wei Xu, Han Qiu “The Earth is Flat because...: Investigating LLMs’ Belief towards Misinformation via Persua- sive Conversation”, ACL 2024.

针对大语言模型对齐算法的综合研究

近年来，大语言模型 (LLMs) 成为了人工智能的研究重点。大型语言模型（LLMs）通过在大量文本数据集上进行预训练，获得了广泛的语言模式和知识。为了更好地在各个领域利用 LLM 的丰富知识，一个关键挑战是如何在各式各样的下游任务中将 LLM 的输出和人类的偏好进行对齐。在之前的工作中，学术界和工业界广泛采用基于奖励模型的 Proximal Policy Optimization（PPO）算法和直接偏好优化（Direct Preference Optimization， DPO）算法。但是从没有人系统性地研究过两种算法的优劣和如何更好地利用这它们进行 LLM 的对齐。

吴翼研究组系统性地研究了 DPO，PPO 等对齐算法在各种模型规模 (7B 到 70B)，各种任务上的优劣势。他们首先从理论上分析了 DPO 方法存在固有缺陷，并总结出了在实践中如何更好地利用 DPO 算法。通过在对话任务和代码生成任务上的实验，他们验证了提升 DPO 算法在 LLM 输出分布和偏好数据分布相差比较大时，DPO 的表现会明显下降，并据此提出了 DPO 算法在实际应用中缩小这样的分布差距的一些方法，并取得了显著成效。同时，他们也提出了提升 PPO 算法表现的几大重要因素，即 PPO 训练中批大小（Batch Size）需要比较大；同时 Advantage Normalization 有助于保证训练的稳定性；最后，训练中对参考模型（reference model）进行缓慢更新也有助于提升 LLM 在下游任务的表现。最终，吴翼研究组在困难的代码竞赛任务验证了算法改进的效果，并取得了 SOTA 的结果。

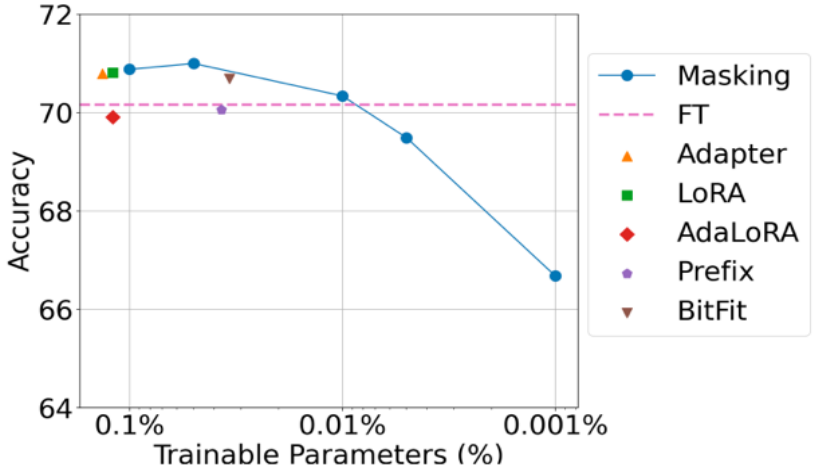


Task	HH-RLHF	APPS			CodeContest		
Metric	OpenAssaint Reward	Intro. pass@5	Inter. pass@5	Comp. pass@5	pass@10	pass@100	pass@1k
SFT	0.532	38.6%	10.1%	3.9%	0.9%	4.3%	12.0%
baseline PPO	0.706	18.0%	2.4%	1.1%	4.3%	6.0%	7.7%
+ Adv.Norm.	0.716	38.1%	11.4%	4.6%	6.8%	9.4%	15.4%
+ Large.Batch.	0.716	42.3%	14.6%	7.5%	5.1%	12.8%	19.6%

该成果研究论文: Shusheng Xu, Wei Fu, Jiaxuan Gao, Wenjie Ye, Weilin Liu, Zhiyu Mei, Guangju Wang, Chao Yu, Yi Wu, “Is DPO Superior to PPO for LLM Alignment? A Comprehensive Study”, ICML 2024.

基于随机掩码的参数高效微调

微调是提升大型语言能力与安全性的重要步骤。参数高效微调通过减少参数训练量，大幅降低了微调的开销。该文提出使用随机掩码算法，以进一步简化算法设计与降低存储开销。实验结果表明，只需选择合适的学习率，随机掩码能在一些任务中取得与 LoRA 等算法相当的准确率。该文从优化算法和模型表达能力的角度，通过实验和理论探讨了这一现象的原因。



图表 1：参数高效算法的性能与可训练参数量的关系，Masking 为该文提出的随机掩码算法。

针对特定任务进行参数微调是提高预训练模型性能的关键步骤。参数高效微调（parameter efficient fine-tuning, PEFT）通过在大模型中增加可训练轻量级模块，能显著降低微调算法的显存开销。为了探究参数高效微调算法的设计原理与性能极限，张景昭研究组研究了一种参数高效微调方法：随机掩码 (Random Masking)。随机掩码相较于现有的标准参数高效微调算法，例如 LoRA，具有算法设计简单、训练参数量更少等特点。该研究组成员通过大规模实验发现，只需使用较大的学习率，随机掩码算法能在一些任务中取得与标准参数高效微调方法相当的性能。

$$\min_{\{\boldsymbol{S}_i\}} \ell(\mathcal{D}, \{\boldsymbol{W}_1 + \boldsymbol{S}_1, \cdots, \boldsymbol{W}_k + \boldsymbol{S}_k\}).$$

图表 2：随机掩码算法的数学表示

该研究组对该现象给出了理论与实验分析，证明了随机掩码的出色性能得益于：1. 大语言模型强大的表达能力 2. 掩码使损失函数更平滑，从而降低了优化难度。该论文提出的随机掩码算法不仅为参数高效微调方法的设计与分析提供了新思路，而且对降低大规模预训练模型的微调成本具有重要实际价值。

该成果研究论文：Jing Xu, Jingzhao Zhang，“Random Masking Finds Winning Tickets for Parameter Efficient Fine-tuning”, https://arxiv.org/abs/2405.02596.

四、机器学习理论

主要完成人：张景昭研究组

双层优化问题复杂度研究

双层优化问题是一类基础性的优化问题，可以用来建模超参调节、元学习、神经网络架构搜索等实际应用。当内层函数不具有强凸性质的时候，该问题的复杂度尚未得到彻底研究。张景昭课题组对该问题进行了深入的研究。首先，团队构造了一个困难的例子证明内层函数仅有凸性而没有强凸性的时候，任意的一阶算法都无法在有限时间内寻找到问题的驻点。而且，研究组研究了寻找一类可解的双层优化问题的驻点的复杂度。假设当内层函数满足 Polyak-Łojasiewicz (PL) 条件，团队证明了基于罚函数的简单一阶算法可以达到关于该问题目前已知最好的复杂度上界。

内层函数具有强凸性质的双层优化问题得到广泛的研究，但更多的实际应用并不满足该假设。该论文的研究动机在于研究能否放宽该假设。然而，没有内层函数的强凸性可能使得最终的优化目标函数不连续。该论文首先给出了导致该不连续问题的充要条件，并且将其归因于内层函数最优解集的稳定性。即使将假设加强为严格凸性，可以保证最终优化的目标函数的连续性，但该问题仍然存在计算上的问题。该论文针对“零相关”一阶算法构造了一个困难的例子，使得该算法类中的任意算法都不能在有限时间内寻找到该问题的驻点。该算法类包含了许多已有的双层优化算法。文章的上述分析启发需要对内层函数的曲率有某种假设。 进而文章研究近期一些工作所研究的内层函数满足 Polyak-Łojasiewicz (PL) 条件下该问题的复杂度。该条件是强凸条件的推广，但刻画了一大类非凸函数。对于满足上述条件的问题，文章进一步推广了课题组之前关于内层函数满足强凸性质下的近似最优一阶算法。通过更精细的分析证明相同的算法在 PL 条件下也可以达到相似的收敛率，从而给出该问题目前已知最好的复杂度上界。该工作探讨了双层优化问题的可解性以及问题复杂度。

该成果研究论文：Lesi Chen, Jing Xu, Jingzhao Zhang, "On Finding Small Hyper-Gradients in Bilevel Optimization: Hardness Results and Improved Analysis", COLT 2024.

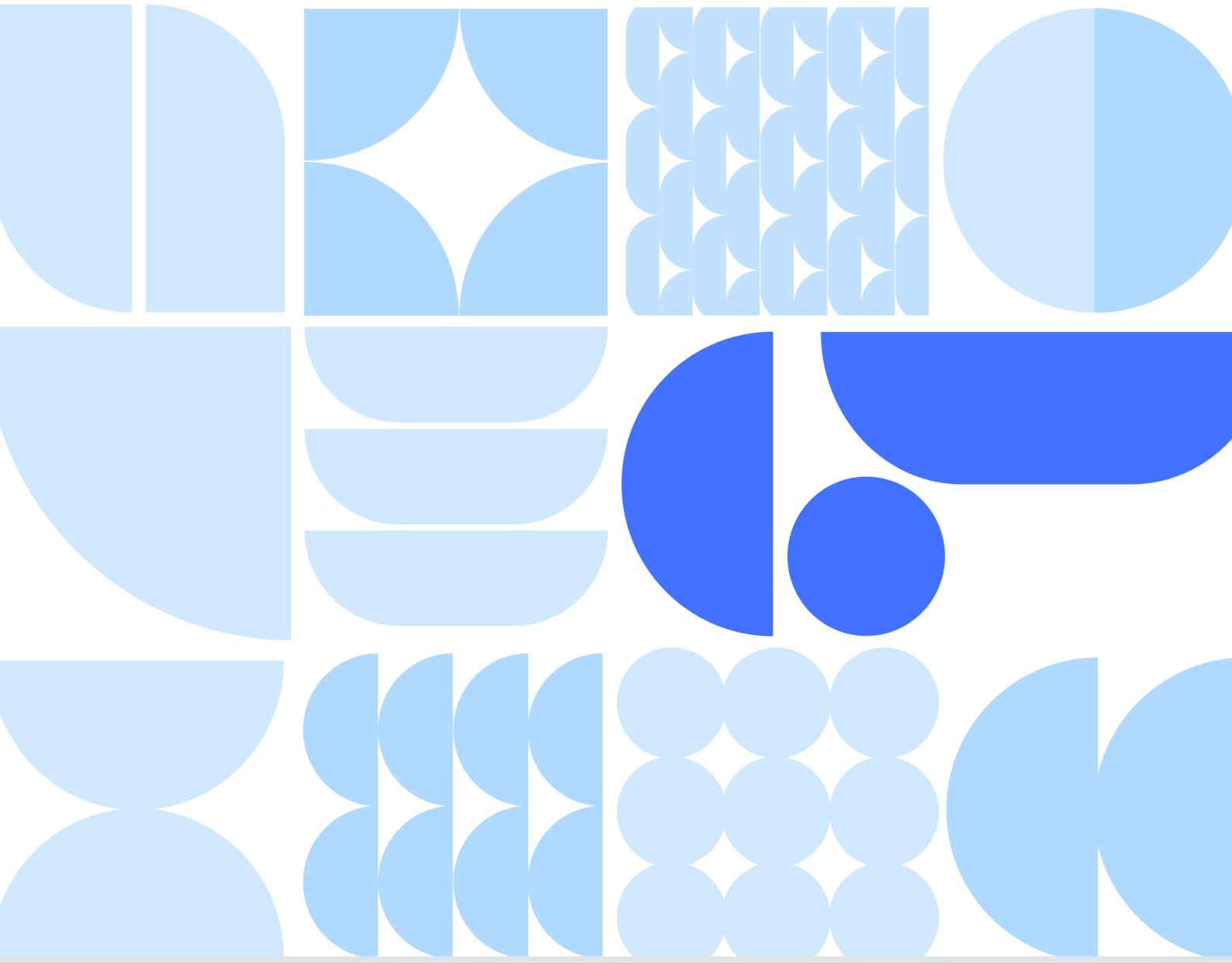
Table 1: We present the complexities of different methods for nonconvex-PL bilevel problems.

Oracle	Method	Deterministic	Partially Stochastic	Fully Stochastic	Reference
2nd	GALET [(a)]	$\tilde{\mathcal{O}}(\kappa^5\epsilon^{-2})$	-	-	Xiao et al. (2023)
1st	Prox-F^{2}BA [(b)]	$\tilde{\mathcal{O}}(\kappa^{p_1}\epsilon^{-3})$	$\tilde{\mathcal{O}}(\kappa^{p_2}\epsilon^{-5})$	$\tilde{\mathcal{O}}(\kappa^{p_3}\epsilon^{-7})$	Kwon et al. (2024b)
1st	F^{2}BA	$\tilde{\mathcal{O}}(\kappa^4\epsilon^{-2})$	$\tilde{\mathcal{O}}(\kappa^5\epsilon^{-4})$	$\tilde{\mathcal{O}}(\kappa^{11}\epsilon^{-6})$	This Paper

(a) Although Xiao et al. (2023) did not provide the dependency on κ in the complexity, we can calculate by the way in our Remark C.1. Their analysis additionally requires the smallest singular value of $\nabla^2_{yy}g(x,y)$ has a constant gap between zero, but such assumption seems to be redundant by our Remark C.2.

(b) We use p_1, p_2, p_3 to denote the polynomial dependency in κ since they are not provided by Kwon et al. (2024b).

计算机科学

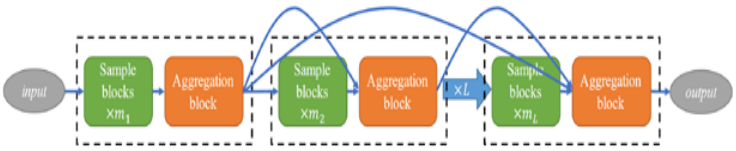


一、计算机系统结构

主要完成人：高鸣宇研究组、马恺声研究组

通过线性算子补偿非线性缺失的隐私保护机器学习推理—Seesaw

随着数据隐私问题的日益严重和隐私法规的日趋严格，隐私保护机器学习应运而生，旨在在不侵犯隐私的情况下安全地执行机器学习任务。然而，目前安全执行非线性计算的成本仍然很高，这需要新的模型架构设计以减少非线性操作的数量。高鸣宇研究组提出了 Seesaw，一种为隐私保护机器学习场景量身定制的新型神经架构搜索方法。Seesaw 通过增加线性计算和重用非线性运算结果来弥补减少非线性操作而导致的精度损失。它结合了专门设计的剪枝和搜索策略，不仅能够高效处理线性和非线性算子，还能在模型精度与在线 / 离线执行延迟之间实现更佳的平衡。与当前最先进的同类设计相比，Seesaw 在保持 71% 的精度水平下，实现了 1.68 倍的在线延迟降低和 1.55 倍的总延迟降低；或者在 190 秒的等延迟条件下，取得了 3.65% 的精度提升。



随着数据隐私越来越被重视，基于隐私保护的机器学习应运而生。由于其需要使用加密技术来保护用户数据隐私，因此其计算开销非常大，而主要瓶颈又来源于非线性激活函数（如 ReLU 和 Sigmoid）的计算。虽然现有工作试图通过减少非线性运算的数量来降低计算开销，但这往往会导致模型精度下降，使得模型精度和执行延迟之间不可兼得。

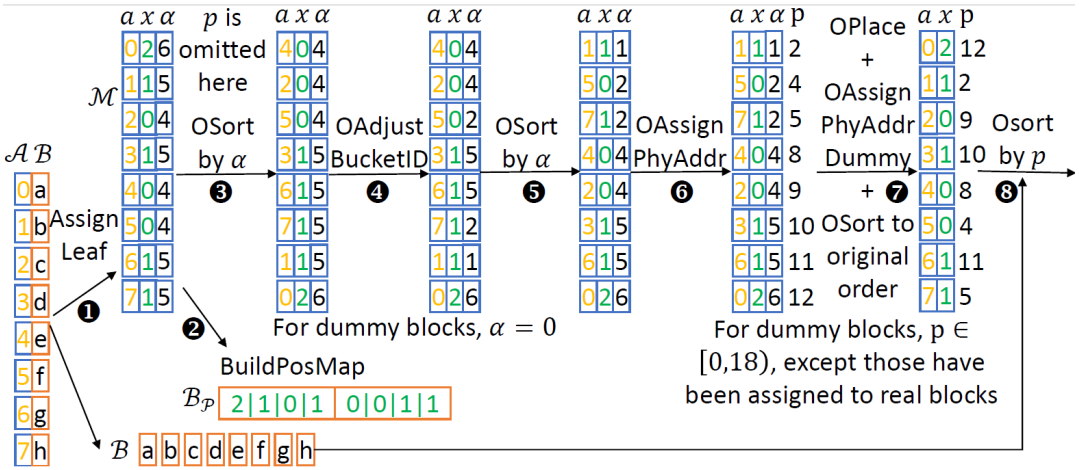
该文提出了 Seesaw，一种用于隐私保护机器学习场景的神经网络架构搜索方法。Seesaw 提出了两种技术来解决上述困境：1) 增加更多的线性运算以恢复模型的表征能力；2) 通过残差连接重复使用已有的非线性运算的结果。Seesaw 在搜索过程中考虑了线性运算的数量对在线 / 离线延迟的影响，并采用了新的搜索和训练策略来高效探索扩大后的搜索空间，最终得到最优的神经网络架构。

Seesaw 能够显著提升模型精度与执行延迟之间的帕累托最优边界。在 ImageNet 上，相比最先进的同类工作 SENet，Seesaw 在相同 71% 精度下，在线延迟降低 1.68 倍；在相同 190 秒延迟下，精度提高 3.65%。在 CIFAR100 上， Seesaw 在相同 70% 精度下，在线延迟降低 1.53 倍；在相同 8 秒延迟下，精度提高 0.25%。总之， Seesaw 通过补偿非线性减少所带来的精度损失，突破了隐私保护机器模型中精度与延迟之间原有的折中，有助于更加高效地部署延迟敏感的应用。

该研究成果论文：Fabing Li, Yuanhao Zhai, Shuangyu Cai, Mingyu Gao, "SSeesaw: Compensating for Nonlinear Reduction with Linear Computations for Private Inference", ICLR 2024.

使能 ORAM 的批量化加载—BULKOR

可信执行环境的侧信道防御是一个广受关注的课题。高鸣宇研究组发现对于最广泛存在的基于内存访问模式的侧信道攻击，其防御方式 ORAM 的初始化过程很少被研究。经过充分调研，该研究组发现其过程有着大量的应用场景，并提出使能 ORAM 的批量化加载算法 Bulkor。相较于之前的工作，Bulkor 在保证高安全性的情况下实现了大幅度的性能提升。该算法有助于 ORAM 在更多场景下的应用。



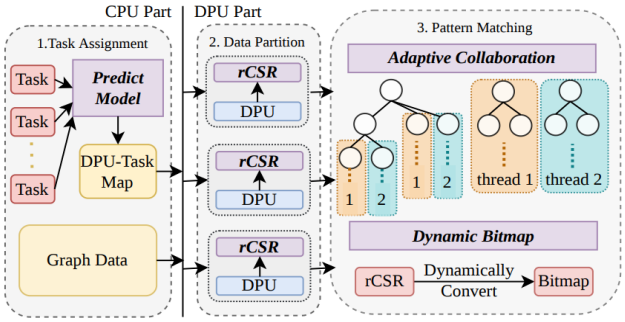
尽管现代加密和身份验证技术可以保护数据内容，但攻击者仍然可以通过仅观察敏感数据的访问模式来执行高级侧信道攻击，以获取私密信息。不经意随机访问存储（ORAM）协议，比如 Path ORAM，是解决此问题的通用方案。ORAM 可结合硬件安全技术例如可信执行环境（TEE），将客户端的部分控制逻辑放入 TEE 中，以减少网络通信成本。然而，与完全可信的客户端不同，TEE 保护数据内容，但不保护其访问模式。在这种情况下，可信内存和不可信内存都需要做到不经意访问，这需要在 TEE 内设计更复杂的 ORAM 控制器。一些工作基于此设计了更高效的安全数据处理系统。但是很少有工作考虑到 TEE 环境下 ORAM 的初始化加载问题，此问题影响了 ORAM 的进一步应用。

由此，高鸣宇研究组提出了一个名为 Bulkor 的 TEE 环境下 Path ORAM 批量化加载算法，它从一开始就完全独立且随机地分配 Path ORAM 中每个数据块的路径。然后它根据先前分配的路径，高效和不经意地调整 ORAM 树中每个数据块的实际位置，以消除任何容量溢出问题。在此过程中无需更改先前分配的路径，从而不影响安全性。另外，所有块的分配和位置调整这些过程都是在元数据上进行的，位置固定后将最终位置信息提供给原数据块，因此对于数据块高鸣宇研究组只需做一次不经意排序，从而降低了性能开销。相对于之前的方案，高鸣宇研究组可将理论复杂度由降低到。实验结果显示其实际性能显着优于先前系统 Oblix 和 ZeroTrace，达 8.7 至 54.6 倍和 5.8 倍至 533.1 倍。

该研究成果论文：Xiang Li, Yunqian Luo, Mingyu Gao, "SBULKOR: Enabling Bulk Loading for Path ORAM", IEEE Symposium on Security and Privacy 2024.

在真实存内计算硬件上的高效图模式匹配框架—PimPam

内存访问带宽一直以来是图模式匹配算法的一大瓶颈。高鸣宇研究组在 UPMEM 这一真实存内计算硬件平台上设计了一个高效的匹配算法框架，通过核间任务和数据的静态分配，以及核内分支的动态协作，解决了在有限通信条件下 2560 个内存计算单元上的负载均衡和数据分配问题，高效地利用了 UPMEM 的算力和带宽。该框架是图模式匹配首次在真实存内计算硬件上的实现，为该算法的进一步加速提供了新的可能。



在大数据时代，图是对数据内在关联性的有效刻画，也是许多领域研究对象的高度抽象，例如社交网络、生物数据等都可以抽象为图。图模式匹配则是分析和提取图内信息的有效手段之一，有着广泛的应用。基于传统 CPU 的图模式匹配效率往往受到内存带宽的限制，近年来兴起的存内计算架构提供的高带宽和低访问时延有效解决了这一问题，但其有限的内存容量和通信限制也对研究者提出了新的挑战。

高鸣宇研究组提出了首个基于真实存内计算硬件的高效图模式匹配框架 PimPam。PimPam 基于 UPMEM 这一架构，相比传统的实现做了四个创新，以解决诸如负载均衡、数据分配等问题：

1）PimPam 提出了一个高效的模型来预测每个任务的工作量，从而实现任务的均匀分配。

2）PimPam 设计了一种新的数据格式，更加紧凑地表示每个子图，有效地利用了每个计算单元有限的内存容量。

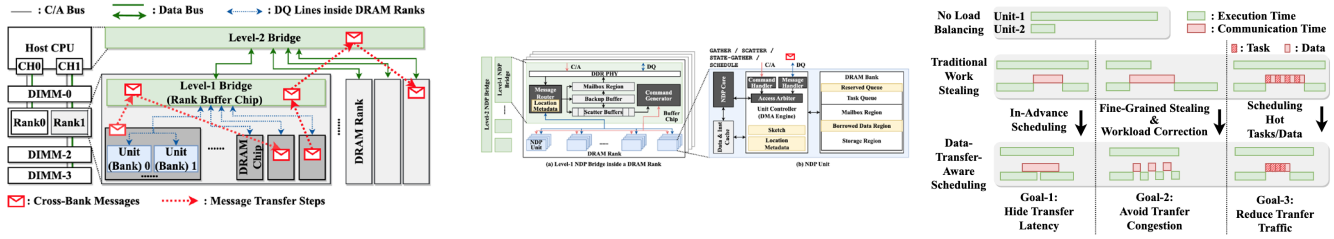
3）每个计算单元内部使用了一种自适应的协作算法，根据数据特点动态调整线程的协作方式，避免了并行度退化问题。

4）PimPam 会动态地构建和销毁子邻接矩阵的位图表示，在有限内存限制下进一步加速特定模式的匹配速度。

该研究成果论文：Shuangyu Cai, Boyu Tian, Huanchen Zhang, Mingyu Gao, "PimPam: Efficient Graph Pattern Matching on Real Processing-in-Memory Hardware", SIGMOD 2024.

在近 DRAM Bank 处理架构中的跨 Bank 协调支持—NDPBridge

一种新型近数据计算 (NDP) 架构，近 DRAM bank NDP 系统，在每个 DRAM bank 附近集成逻辑单元，以缓解数据密集型应用中面临的“内存墙”挑战。近 DRAM bank NDP 系统目前已经投入商业使用。然而，由于物理隔离，此系统无法进行快速且直接的跨 bank 通信，制约了其并行模式，且可能出现计算负载失衡的问题。NDPBridge 通过软硬件协同设计，为近 DRAM bank NDP 系统实现跨 bank 通信和动态负载平衡。此设计在 DRAM 各个层级中引入硬件桥来支持跨 bank 通信，并在此基础上进一步实现了分层和数据传输感知的负载平衡。



近数据处理架构是一条缓解内存墙问题、降低内存访问开销的重要技术路线。其中，近 DRAM bank 架构在 DRAM bank 附近集成计算逻辑，每个 bank 及其周围的计算逻辑构成独立单元，可以高效并行访问和处理数据。但是，近 DRAM bank 架构同样面临两点主要挑战。首先，不同的单元互相隔离，无法进行跨单元通信。此外，由于系统由上千个单元组成，单元间的负载均衡也需要得到高效支持。高鸣宇研究组提出一种软硬件协同设计方案 NDPBridge，在硬件层面，引入硬件桥，通过复用 DRAM 内部现有硬件接口和连线资源，在 DRAM 内部支持了跨 bank 传输。在软件层面，在上述硬件通讯机制基础上，他们设计了层次化和数据传输感知的调度方案，高效支持了跨单元负载均衡。

NDPBridge 在性能、开销和适用性等方面具有显著优点，具体包括：1. 相较于现有近 DRAM bank 处理方案，实现了平均 2.23 倍、最高 2.98 倍性能提升。2. 硬件修改开销较小，对于 DRAM 内部芯片的尺寸和接口没有修改，且所有的修改均限制在现有的近数据处理产品修改过的硬件模块中。3. 该架构对软件适配性较好，可应用至多种类型的应用和不同的数据规模。此外，由于该架构实现了自动的跨单元通信优化和负载均衡，大大降低了上层程序员的编程负担。

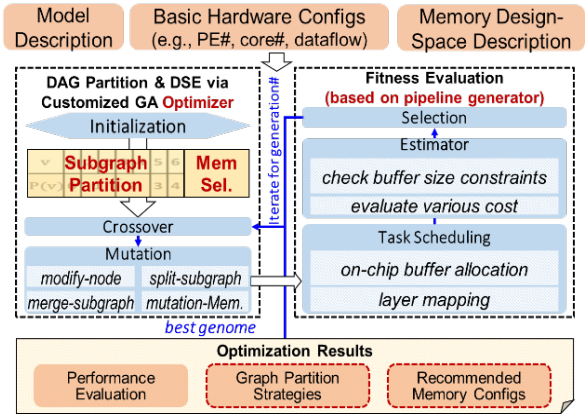
该成果研究论文：Boyu Tian, Yiwei Li, Li Jiang, Shuangyu Cai, Mingyu Gao, "ANDPBridge: Enabling Cross-Bank Coordination in Near-DRAM-Bank Processing Architectures", ISCA 2024.

面向片上存储容量与通信的硬件—任务映射协同优化

目前，神经网络模型的复杂性不断增加，其中一个主要挑战是多分支块中常见的中间数据的 I/O 负担。传统的处理方式是逐层处理，导致更大的通信和能耗开销。层融合可以通过将一些层融合到子图中来减轻这种负担。这样，中间数据可以在片上重用，减少 I/O 通信。为了在硬件上实现层融合处理，一些研究提出了专用流水线。然而，它们主要关注特定的结构，如链块和残差块。对于任意的拓扑结构，如 NAS 网络中的结构，马恺声研究组需要一种通用的方法来决定每个层的分块大小和内存分配。另一个障碍是片上内存大小和图划分之间的权衡。给定一个更大的片上内存，马恺声研究组可以缓冲更多的中间数据来融合更多的层。然而，较大的片上内存将导致较大的芯片面积，如果仔细搜索该方案，适当数量的融合层可能就足够了。为此，马恺声研究组需要对图划分和内存大小选择进行协同设计。

这项工作的大纲包括一个通用层融合流水线生成器和一个基于遗传算法的定制优化器。流水线生成器以存储容量和特定图划分作为输入，生成具有相应代价的流水线方案；而基于遗传算法的优化器则利用内存开销的反馈来更新和探索容量和图划分。

实验表明，与之前的贪心算法和动态规划相比，Cocco 获得了更低的外部内存访问次数、更低的带宽需求和更稳定的图划分优化。与其他典型方法相比，通过协同探索，Cocco 还将成本降低了 1.89% 至 50.33%。



该成果研究论文：Zhanhong Tan, Zijian Zhu, Kaisheng Ma, “Cocco: Hardware-Mapping Co-Exploration towards Memory Capacity-Communication Optimization”, ASPLOS 2024.

大规模深度学习芯粒加速器架构和映射的联合探索

随着深度神经网络（DNN）解决越来越复杂的问题，其规模和复杂性迅速增长，导致计算和存储需求增加。尽管采用先进技术和扩大单芯片尺寸已推动了许多拥有数百亿晶体管的大规模单片加速器的发展，但摩尔定律的终结和光掩模尺寸的限制对进一步集成提出了挑战。

Chiplet 技术利用先进封装技术组合小功能芯片，提供了克服这些限制的解决方案。基于 Chiplet 的 DNN 推理加速器（如包含 36 个芯片的 Simba）已出现，但也带来了新的架构设计和 DNN 映射挑战，具体如下：

架构设计的主要挑战是确定最佳 Chiplet 粒度。虽然 Chiplet 技术改善了面积限制和产量，但也引入了更高的封装费用和 D2D 互连成本。D2D 互连在能量和面积上更密集，带宽低于片上线路。这些不利影响统称为“Chiplet 成本”。因此，需要在使用更多小 Chiplet 以提高产量和使用更少大 Chiplet 以降低成本之间找到平衡，这仍是未解决的挑战。

DNN 映射的主要挑战来自于 Chiplet 技术允许的更大规模和昂贵的 D2D 链路。随着加速器规模的增加，保持高利用率和能量效率变得更难。LP 映射（在空间上映射多个层）广泛用于大规模加速器以应对这一挑战。LP 映射的核心是空间映射（SPM），决定了哪个层的哪部分分配到哪个核心，对性能和能量效率有显著影响。然而，大多数现有策略仍是启发式的，未能全面定义、探索或理解 LP SPM 的问题和优化空间，限制了优化机会的利用。第二个挑战在于 D2D 链路的高能耗和低带宽，因此，设计能够自动减少 D2D 通信的空间映射策略对于提升 Chiplet 加速器性能和效率至关重要，但现有文献中明显缺乏这类方法。

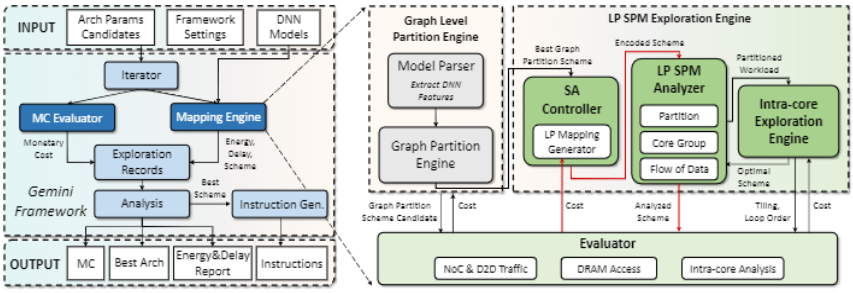


图 1. Gemini 联合探索框架

马恺声研究组针对这一挑战而提出了一种面向层的编码方法，用于表示多核 Chiplet DNN 推理加速器中的 LP SPM 方案。利用这种编码方法，该研究组明确了 LP 映射的优化空间，计算出其巨大规模，远超现有的启发式策略，并分析了隐藏在这一空间中的潜在优化机会。在后摩尔时代的 Chiplet 背景下，这些机会尤为重要。据研究人员所知，这是第一个明确而系统地定义 DNN 推理加速器 LP SPM 优化空间的工作。

基于上述编码和高度可配置的通用硬件模板，该研究组开发了 Gemini，一个用于大规模 DNN Chiplet 加速器的映射和架构共同探索框架。Gemini 包括两个主要引擎：映射引擎和货币成本评估器。在映射引擎中，研究人员开发了一种具有五个专门设计操作符的模拟退火（SA）算法，以探索由该研究组编码方法定义的广阔空间，并自动最小化昂贵的 D2D 通信。货币成本评估器则评估了具有不同架构参数的加速器的 MC。据研究人员所知，Gemini 是第一个联合探索大规模 DNN Chiplet 加速器映射和架构优化空间的框架，不仅考虑能耗和性能，还考虑了 MC。

与采用 Tangram SPM 的 SOTA Simba 架构相比，Gemini 的共同优化架构和映射在各种 DNN 和批量大小下，平均实现了 1.98 倍的性能和 1.41 倍的能效提升，而 MC 仅增加了 14.3%。

该成果研究论文：Jingwei Cai, Zuotong Wu, Sen Peng, Yuchen Wei, Zhanhong Tan, Guiming Shi, Mingyu Gao, and Kaisheng Ma, "Mapping and Architecture Co-exploration for Large-scale DNN Chiplet Accelerators", HPCA 2024.

二、数据库系统

主要完成人：张焕晨研究组

使内存中的学习型索引在磁盘上高效

近年来，学习型索引已成为数据库管理系统的研究热点。设计高效的学习型索引能够显著提升系统的检索速度、空间利用效率、智能化水平和自适应性。然而，目前最先进的学习型索引主要针对内存场景设计，与基于磁盘的数据库管理系统并不完全匹配。然而，最新研究显示，将现有的内存中的学习型索引直接扩展到磁盘时，其性能无法超越 B+ 树。为让内存中的学习型索引在磁盘上也能高效运行，而不需要为磁盘场景专门设计基于磁盘的学习型索引结构，张焕晨研究组提出了一套通用的转换和优化准则（图 1），并将其应用于现有的内存型学习型索引，以充分利用磁盘的特性来提高索引的性能。

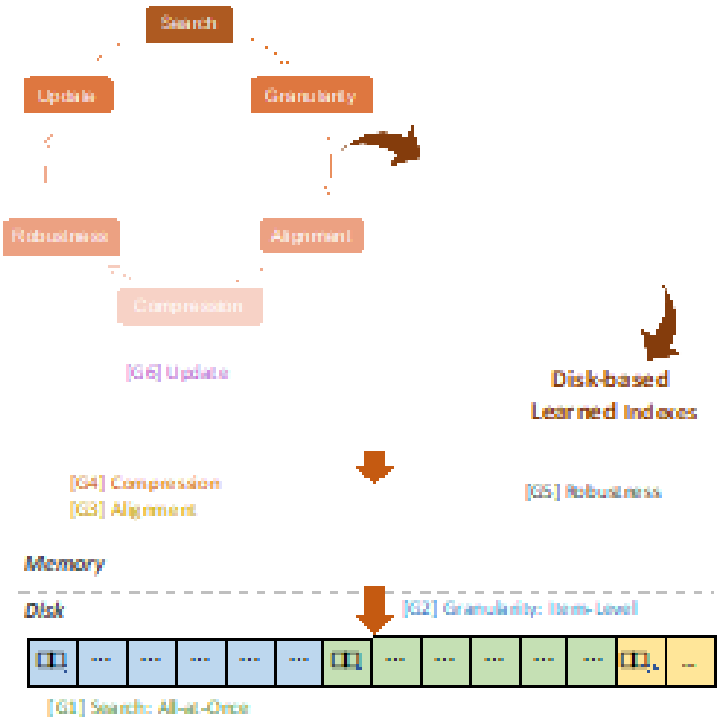


图 1 通用的转换和优化框架

这一通用转换方法能够有效将内存中的学习型索引转化为适用于磁盘场景的索引，并在提升索引结构的速度、空间压缩率、鲁棒性、智能化水平等方面具有显著优势，同时减轻了研究人员设计磁盘场景下学习型索引的负担。具体包括:（1）结合磁盘特性和学习型索引的特点，提出了内存和磁盘交互、模型预测的粒度两方面的优化策略，以加速基于磁盘的学习型索引的响应速度。（2）总结了多种学习型索引的构建算法，并提出通用的优化，可以显著减少索引需要的模型的数量，同时保持高效的查询性能。此外，还为模型们的各个组成部分分别提出相应的压缩技术，可以进一步压缩索引结构整理的空间大小。（3）从功能性和鲁棒性的角度分别提供相应的优化策略，有助于研究人员更好地理解和高效使用学习型索引。（4）这一通用转换方法可应用于现有的内存中的学习型索引，使研究人员无需为磁盘场景重新设计索引结构，从而有效提高了开发效率。

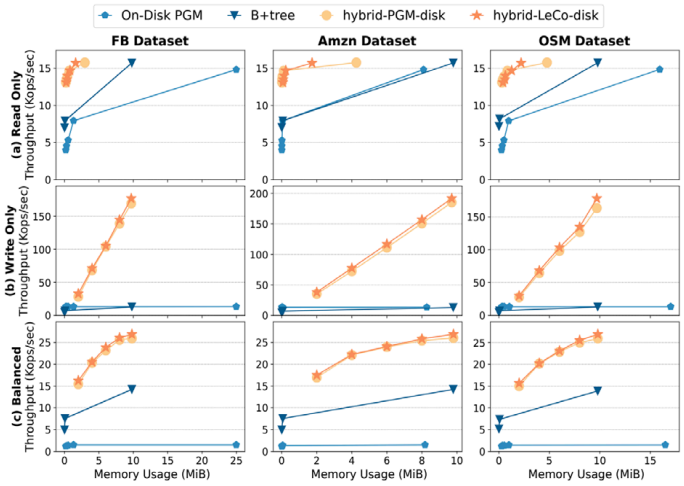


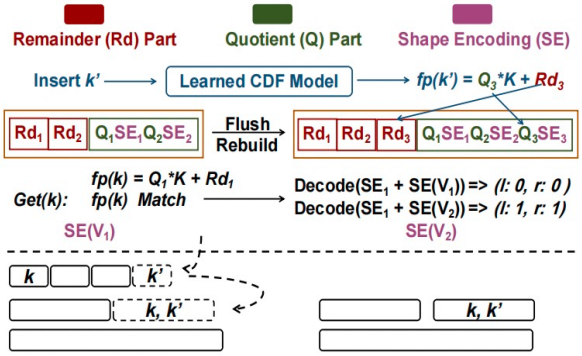
图 2 张焕晨研究组的索引结构（hybrid-PGM-disk 和 hybrid-LeCo-disk）和其他的索引们在不同的工作负载和数据集上的吞吐量和空间用量

张焕晨研究组的实验表明，与传统的 B+ 树和之前基于磁盘的学习型索引的实现相比，应用张焕晨研究组的指南开发的索引在吞吐量和空间效率方面实现了帕累托改进。该工作为基于磁盘的学习型索引的设计提供了全新的视角和方法，进一步推动了将人工智能与数据库融合的学习型索引的发展。

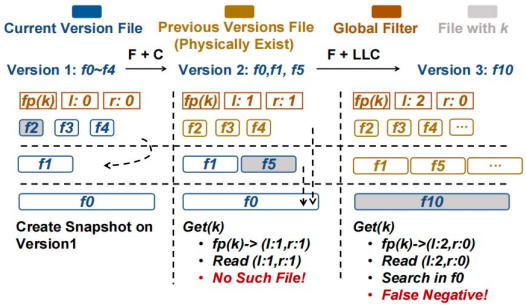
该成果研究论文：Jiaoyi Zhang, Kai Su, and Huanchen Zhang, "Making In-Memory Learned Indexes Efficient on Disk", SIGMOD 2024.

通过形状编码构建 LSM-Tree 全局范围过滤器

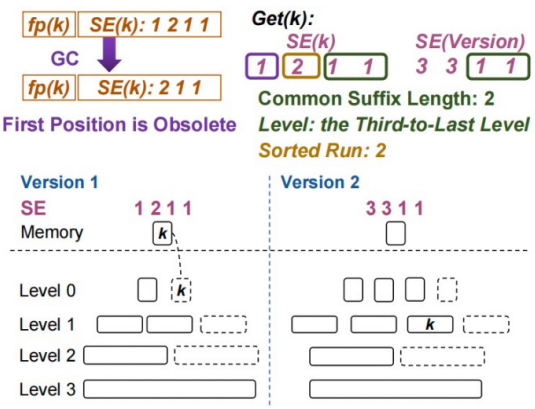
LSM-Tree 是一种非常常见的键值对数据存储引擎。LSM-Tree 在磁盘上维护多个排序数组，牺牲了部分查询性能来换取高效的写入性能。为了减少无用的磁盘读写，LSM-Tree 通常会在内存中为每个排序数组维护一个过滤器来加速查询。随着数据量的扩大以及排序数组个数的增加，查询这些过滤器的时间逐渐变成了性能瓶颈。 该文提出了一种全局范围过滤器结构来减少查询过滤器的时间开销。在全局范围过滤器当中，张焕晨研究组通过形状编码来保证查询结果在并发场景下的正确性。



过滤器数据结构在 LSM-Tree 键值对数据存储引擎当中是一个常见的数据结构，随着磁盘速度的提升，查询过滤器的开销逐渐成为了性能瓶颈。过往工作通过建立全局过滤器，减少过滤器查询次数的方式来降低这一瓶颈。但是过往的全局过滤器往往无法加速范围查询，同时在并发查询的场景下会产生正确性的问题。



张焕晨研究组基于形状编码提出一种新的全局范围过滤器，在不产生额外内存开销的情况下将全局过滤器支持了范围查询且可以保证结果的正确性。基于形状编码的全局范围过滤器具有显著的优点，具体包括：1. 可以同时加速点查询和范围查询。2. 解决了全局过滤器在并发读写场景下查询结果不正确的问题。3、该数据结构的维护代价很低，在LSM-Tree 合并多个排序数组的时候不需要额外处理，由此提高了 LSM-Tree 的写性能。



该成果研究论文：Hengrui Wang, Te Guo, Junzhao Yang, Huanchen Zhang, "GRF: A Global Range Filter for LSM-Trees with Shape Encoding" , SIGMOD 2024.

三、密码学

主要完成人：陈一镭研究组、宋一凡研究组

通过密码学方法来证明值域规避问题和远点问题的困难度

最近在显式构造问题的研究领域引入了一种系统性的方法，通过使用元问题（meta problems）来探索显式构造问题的复杂性，即值域规避问题（缩写为 Avoid）和远点问题（缩写为 RPP）。这些元问题的上限和下限为之前独立研究的特定显式构造问题的复杂性提供了统一的视角。以前的工作很大程度上未解决的一个有趣问题是：Avoid 和 RPP 对于简单电路（例如低深度电路）是否困难。

在该文中，陈一镭研究组证明，在合理的密码学假设下，即使输入电路与常数深度电路一样简单，范围回避问题和远程点问题也无法通过非确定性搜索算法有效解决。这扩展了由 Ilango、Li 和 Williams (STOC’23) 针对采用 NP 见证加密（Witness encryption）的确定性算法建立的困难度结果，其中值域规避问题的输入是通用布尔电路。

该研究组的主要技术贡献是一种新颖的见证加密结构，其灵感来自于 NP 中某些不太可能是 NP 完全的承诺语言（promise language）的公钥加密。研究人员引入了一种通用方法，将具有特定属性的公钥加密方案转换为与初始公钥加密方案相关的语言的见证加密方案。通过这种通用方法，该研究组提出基于标准格或基于编码的 PKE 方案的变体，他们在合理的假设下获得了 NP\ coNP/poly 中某些语言的可证明安全的见证加密方案。此外，他们证明，在 Rudich 的超级比特（RANDOM'97）的广义安全概念下，他们的见证加密结构对于非确定性敌手来说也可能是安全的，这对于证明 Avoid 和 RPP 针对非确定性算法的难度至关重要。

该成果研究论文：Yilei Chen (Tsinghua University, Shanghai Qi Zhi Institute), Jiatu Li (MIT), "Hardness of Range Avoidance and Remote Point for Restricted Circuits via Cryptography", STOC 2024.

Problem	Algorithms	Hardness	Capability
Avoid	$\mathsf{FP}^{\mathsf{NP}}$-reducible to $\mathsf{Hard}_{\varepsilon}$ [Kor21]	Not likely in FP [ILW23] Not likely in SearchNP (*our results*)	Most explicit constructions [Kor21; GLW22]
NC^1-Avoid	FP-reducible to NC^0-Avoid [RSW22]	Not likely in SearchNP (*our results*)	Good Linear Code, Rigid Matrices, etc. [GLW22]
ACC^0-Avoid & ACC^0-RPP	$\mathsf{FP}^{\mathsf{NP}}$ (weak parameters) [RSW22; CHLR23]	Not likely in SearchNP (*our results*)	Best known lower bounds against ACC^0 [CHLR23]
NC^0-Avoid	FP (weak parameters) [GLW22; GGNS23]	Not likely in SearchNP (*our results* + [RSW22])	NC^0-Avoid with strong parameters simulate NC^1-Avoid [RSW22]
XOR-RPP	$\mathsf{FP}^{\mathsf{NP}}$ (weak parameters) [CHLR23] FP (very weak parameters) [APY09]	-	Imply “help function lower bounds” [AS10]

图中列举来 Avoid 和 RPP 的各个子问题的已知算法，困难度（该研究组的主要贡献），以及证明这些问题的困难性的意义（Capability）

可容忍泄露的隐私电路

密码学中的一个理想目标是设计一个可以对隐私数据做一般计算任务的电路 / 硬件使得它可以天然的防范任意的侧信道攻击，这一目标被抽象为隐私电路。隐私电路一般可分为两类，第一类是可抵御泄露的隐私电路，对于某一类泄露函数 L，它要求任意关于电路内部的 L 范围内的信息泄露与隐私数据无关，这类隐私电路天然的需要对隐私数据进行初始的加工；第二类是可容忍泄露的隐私电路，对于某一类泄露函数 L，它要求任意关于电路内部的 L 范围内的信息泄露可以规约到对于输入和输出的 L 范围内的信息泄露。在这篇工作中，宋一凡研究组主要关注后者。注意到后者不需要对输入输出进行任何加工，这意味着当侧信道攻击局限于 L 中的泄露函数时，可容忍泄露的隐私电路可以被视为一个理想的安全硬件（即其内部对外不可见）。

该文中，宋一凡研究组给出了首个针对全局泄露函数类的可容忍泄露隐私电路的构造，具体而言，宋一凡研究组的构造可以防范深度 1 的泄露函数类（逻辑与、或、异或）。同时宋一凡研究组给出了从无状态可容忍泄露的隐私电路到有状态可抵御泄露的隐私电路的一般构造，对于深度 1 的泄露函数类，首次实现有状态可抵御泄露的隐私电路大小与泄露数据量成次平方关系。

该成果研究论文：Ishai, Y., Song, Y., "Leakage-Tolerant Circuits. Advances in Cryptology", EUROCRYPT 2024.

基于图的针对完美安全半诚实多方安全计算协议的自动化验证框架

证明多方安全计算协议的安全性一直是一个困难的任务。在目前主流的基于模拟的多方安全计算定义中，安全性要求构造一个模拟器来证明攻击者可获得的信息可以由模拟器生成。然而模拟器的构造通常与具体协议相关且需要人力实现，这对验证某一给定的多方安全计算协议的安全性造成障碍。同时，多方安全计算协议的安全性也会因为实现过程的疏忽而丧失，并且这类疏忽在现实中很难被察觉。

在这篇工作中，宋一凡研究组提出了一个针对完美安全半诚实多方安全计算协议安全性的自动化验证框架。宋一凡研究组的框架是完全可靠的：任何在宋一凡研究组框架下被证明安全的协议都满足基于模拟的多方安全计算定义。对于完备性，宋一凡研究组的框架可以在多项式时间对任意给定的恶意参与方集合验证任意基于 BGW 协议的安全性。与之前的工作不同，宋一凡研究组的框架有望对任意的多方安全协议进行验证。实验方面，宋一凡研究组测试了两类半诚实安全协议：BGW 协议与二元域到算术域转化协议，宋一凡研究组的的框架可以在合理的时间内完成协议的安全性验证。

该成果研究论文：X. Xie, et al., "GAuV: A Graph-Based Automated Verification Framework for Perfect Semi-Honest Security of Multiparty Computation Protocols", IEEE Symposium on Security and Privacy 2024.

在异步网络下构建具有线性复杂度的多方安全计算协议

多方安全计算允许互不信任的若干个参与方利用各自的隐私数据完成一个共同的计算任务，Ben-Or、Canetti、Goldreich[STOC’ 93] 和 Ben-Or、Kelmer、Rabin[PODC’ 94] 最早奠基了多方安全计算协议在异步网络下的可行性。在这之后，尽管有很多工作在提升协议的通讯效率，目前在无条件安全以及最优恶意参与方数量的情况下取得的最好结果对于每个电路门的通讯量仍与参与方数量的四次方成正比，与之相对的，在同步网络模型下，最好结果对于每个电路门的通讯量仅与参与方数量呈线性关系。

这篇工作中，宋一凡研究组在异步网络下实现线性复杂度这个问题上取得一定进展。宋一凡研究组的多方安全计算协议对于每个电路门的开销为与参与方数量成线性关系的通讯量以及进行常数次秘密分享。在此之前，最好的工作[IEEE Trans. Inf. Theory’ 17] 需要参与方数量平方关系的通讯量以及进行参与方数量成线性关系次的秘密分享。利用目前最优的秘密分享方案 [J.Crypto’ 23]，宋一凡研究组的协议实现总复杂度与参与方数量的三次方成正比。宋一凡研究组注意到一个同期工作给出了首个线性复杂度的秘密分享方案，当应用他们的秘密分享方案时，宋一凡研究组的协议在异步网络下首次实现总复杂度与参与方数量成线性关系。

该成果研究论文：Goyal, V., Liu, C. Song, Y., "FTowards Achieving Asynchronous MPC with Linear Communication and Optimal Resilience", CRYPTO 2024.

在异步网络下具有线性复杂度的秘密分享

异步网络下的秘密分享机制（ACSS）允许一个参与方将自己的信息以 Shamir 秘密分享方案分享给其他参与方并保证所有参与方可以最终获得他们的信息。ACSS 是异步网络下多方安全计算协议构造的重要组成部分，已知的最好结果 [J. Crytol’23] 的通讯量与参与方数量成立方关系，另一方面，同步网络下的秘密分享机制仅需要线性通讯复杂度。

这篇工作中，宋一凡研究组在异步网络下给出首个线性复杂度的秘密分享机制，利用 Choudhury 与 Patra[IEEE Trans. Inf. Theory’ 17] 的方案，宋一凡研究组可以得到通讯复杂度与参与方数量成平方关系的异步网络下的多方安全计算协议。宋一凡研究组注意到一个同期工作优化了 Choudhury 与 Patra[IEEE Trans. Inf. Theory’17] 的方案，当与他们的方案结合时，宋一凡研究组得到了首个在异步网络下通讯复杂度与参与方数量成线性关系的多方安全计算协议。

该成果研究论文：Ji, X., Li, J. Song, Y., "Linear-Communication Asynchronous Complete Secret Sharing with Optimal Resilience", CRYPTO 2024.

四、计算机网络

主要完成人：房智轩研究组

具有收敛动力学的演化系统最优控制

许多复杂系统（如物联网、区块链、分布式机器学习等）对决策者的动作具有即时反馈，但其反馈需要较长时间才能收敛到稳态。这给没有先验知识的决策者带来了很大的困难，因为决策者需要精细地控制对系统收敛过程的探索与等待时长，从而保证学习精度与效率的平衡。这种动态系统的决策控制问题在许多现实领域中都可以观察到，如无线网络中的动态调度控制，多阶段的多智能体博弈等。

房智轩研究组研究了具有收敛动力学的演化系统最优控制，使决策者在缺乏系统稳态反馈先验知识的情况下，能通过使用设计的“乐观 - 悲观收敛和置信界限算法 Optimistic-Pessimistic Convergence and Confidence Bounds（OP-C2B）”，当等待某个动作接近稳态不值得时就迅速切换行动，来缩减学习与等待的时间。该算法通过利用“收敛界限”来确定系统离稳定状态有多远，并通过在维持对可行的动作集的悲观评估的同时在该集合内进行乐观动作选择来实现的。

该研究组证明 OP-C2B 算法能够在有限动作集的情况下，保证亚线性的遗憾值和约束违背。特别地，当系统的收敛速率是线性或超线性时，OP-C2B 可实现对数级的遗憾值和约束违背。此外，团队还将该算法推广到无限的决策者动作集，并证明了在移动众包和资源分配等重要的博弈控制场景中展示了算法的优越性。

该成果研究论文：Qingsong Liu and Zhixuan Fang, “Learning the Optimal Control for Evolving Systems with Converging Dynamics”，ACM SIGMETRICS 2024.

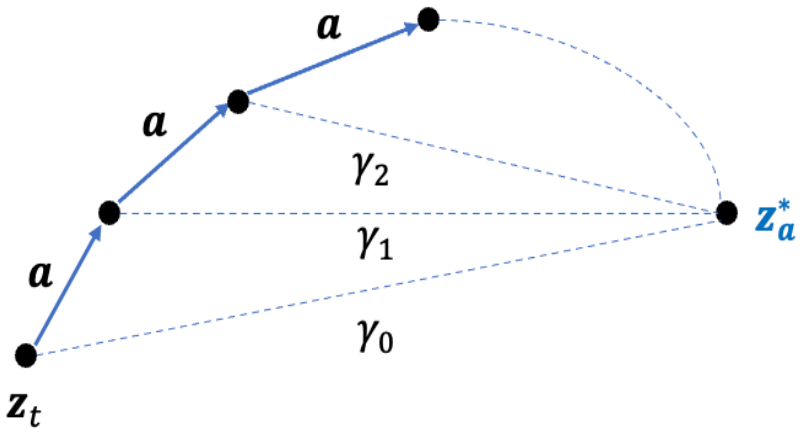


图 1： 收敛动力学示意图

量子信息科学

一、离子阱量子计算与模拟

主要完成人：段路明研究组、吴宇恺研究组

首次实现基于数百离子量子比特的量子模拟计算

离子阱系统被认为是最有希望实现大规模量子模拟和量子计算的物理系统之一。多个实验验证了离子量子比特的高精密相干操控，该系统的规模化被认为是主要挑战。此前研究人员在 Paul Trap（保罗型离子阱）中实现了最多 61 个离子一维阵列的量子模拟。虽然基于 Penning Trap（彭宁型离子阱）可实现更大规模约两百离子的量子模拟，但因缺乏单比特分辨探测能力而难以提取量子比特空间关联等重要信息，无法用于量子计算和精密的量子模拟。

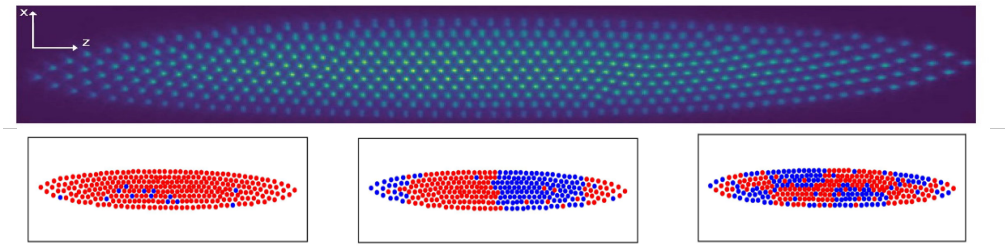


图 1 实验获得 512 离子二维阵列图像与典型 300 离子单点分辨测量结果

段路明研究组利用低温一体化离子阱技术和二维离子阵列方案，大规模扩展离子量子比特数并提高离子阵列稳定性，首次实现 512 离子的稳定囚禁和边带冷却，并首次对 300 离子实现可单比特分辨的量子态测量。

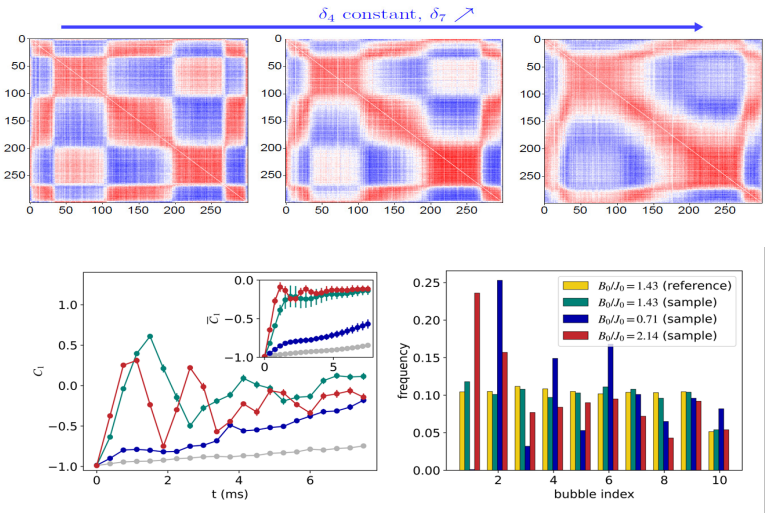


图 2 典型 300 离子长程横场伊辛模型量子模拟计算结果

研究人员进而利用 300 个离子量子比特实现可调耦合的长程横场伊辛模型的量子模拟计算。一方面，研究人员通过准绝热演化制备阻挫伊辛模型的基态，测量其量子比特空间关联，从而获取离子的集体振动模式信息，并与理论结果对比验证；另一方面，研究人员对该模型的动力学演化进行量子模拟计算，并对末态分布进行量子采样，通过粗粒化分析验证其给出非平庸的概率分布，超越经典计算机的直接模拟能力。该实验系统为进一步研究多体非平衡态量子动力学这一重要难题提供了强大的工具。

该成果研究论文：S.-A. Guo, Y.-K. Wu, J. Ye, L. Zhang, W.-Q. Lian, R. Yao, Y. Wang, R.-Y. Yan, Y.-J. Yi, Y.-L. Xu, B.-W. Li, Y.-H. Hou, Y.-Z. Xu, W.-X. Guo, C. Zhang, B.-X. Qi, Z.-C. Zhou, L. He & L.-M. Duan, "A Site-Resolved 2D Quantum Simulator with Hundreds of Trapped Ions", Nature 2024.

首次在离子阱系统实现合成维度

合成维度是量子模拟研究拓扑物态的重要手段，可用于研究实验室难以直接获得的高维材料性质，也可通过调节格点之间的跃迁振幅实现可控的人工磁场，模拟实验中难以直接达到的磁场强度。此前，在中性原子、光子学等系统中已通过内部能级和离散的频率、动量、角动量等模式实现过合成维度。近期，段路明研究组、吴宇恺研究组首次在离子阱系统，利用离子在空间振动的无穷维福克态（Fock state）实现合成维度，扩展了合成维度的实现方式和在离子阱系统的应用。

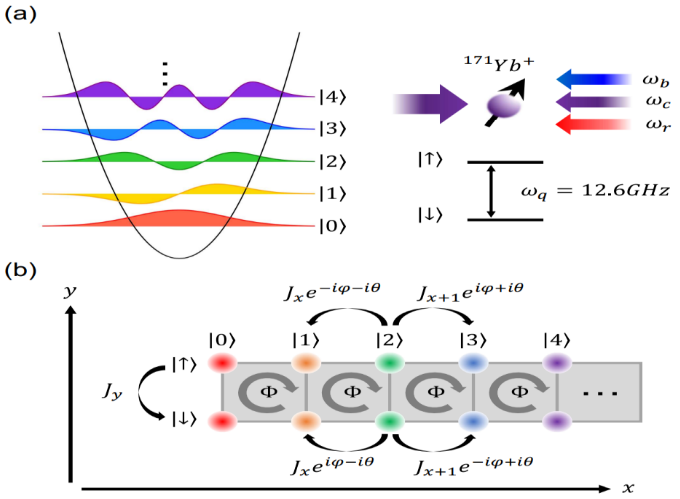


图 1 实验方案示意图

为获得合成维度，研究人员设计方案利用多频激光实现离子内部能级和空间振动自由度之间的耦合，进而将两种自由度映射为双脚梯（two-leg ladder）格点，从而产生格点之间相位可调的跃迁，以实现可控的人工磁场（图 1）。研究人员首先在不同实验参数下验证了所获得的各向异性 Harper-Hofstadter 模型量子模拟结果的正确性，进而演示了该模型中拓扑手性边缘态的运动。研究人员设计准绝热演化路径制备该模型的基态，并测量其手性电流随着人工磁场的变化，演示了该模型的量子相变特性（图 2）。该工作有望推广到更高的合成维度，拓展了离子阱量子模拟平台在拓扑量子物态研究的应用，同时也将 Harper-Hofstadter 模型的量子相变与自旋 - 玻色子模型的量子相变关联起来，加深对其物理机制的理解。

该成果研究论文：Y. Wang, Y.-K. Wu, Y. Jiang, M.-L. Cai, B.-W. Li, Q.-X. Mei, B.-X. Qi, Z.-C. Zhou, and L.-M. Duan, "Realizing Synthetic Dimensions and Artificial Magnetic Flux in a Trapped-Ion Quantum Simulator", Phys. Rev. Lett. 132, 130601 (2024).

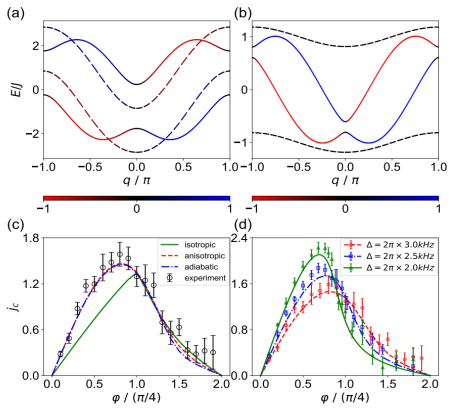


图 2 实验测量各向异性 Harper-Hofstadter 模型量子相变

离子阱系统中实现自旋 - 玻色子模型的量子模拟实验

在物理学领域，开放量子系统是指与外部量子系统相互作用的量子力学系统。环境和系统之间的相互作用通常会导致系统的量子耗散，最初编码在系统中的信息会丢失，量子耗散是一个重要而困难的统计力学问题。为解决开放量子系统问题，研究人员们提出了一系列模型，其中自旋 - 玻色子耦合模型是描述物质 - 光相互作用的基本物理模型。它有许多应用，从研究量子系统中的强耦合和量子热动力学，到研究马尔可夫和非马尔可夫条件下量子系统中的信息流动。

段路明研究组利用离子自旋二能级作为信息编码，将一系列外部声子振动模式作为环境实现了自旋 - 玻色子模型的量子模拟，首先是理论上通过环境谱密度描述了系统与环境的耦合强度在不同离子上的情况（图 1 右）。

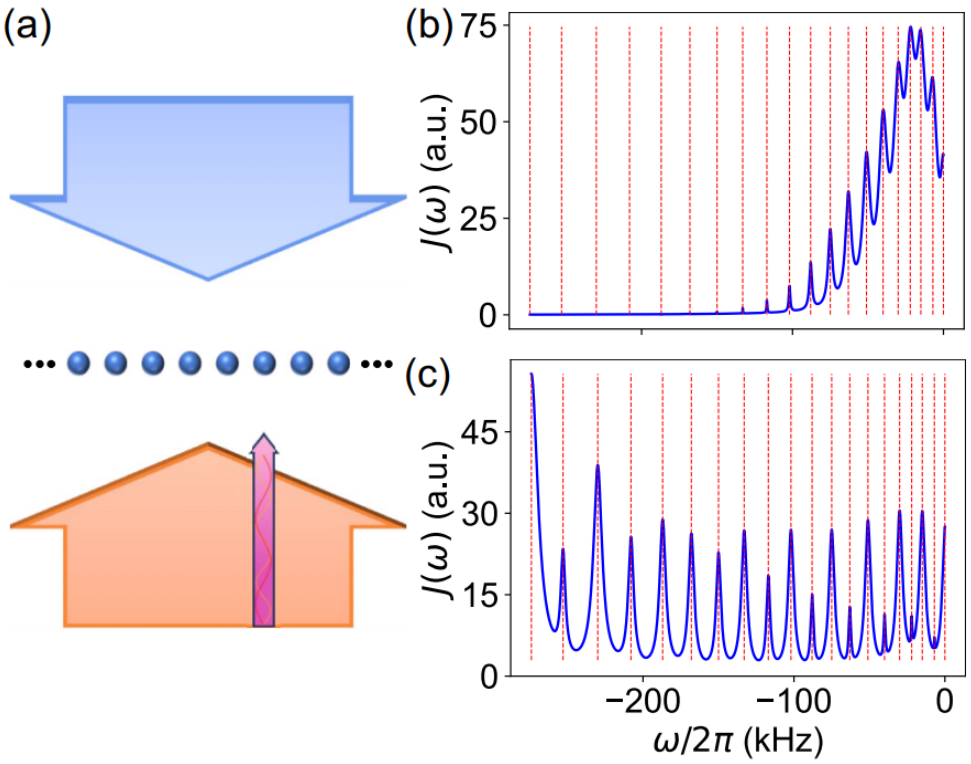
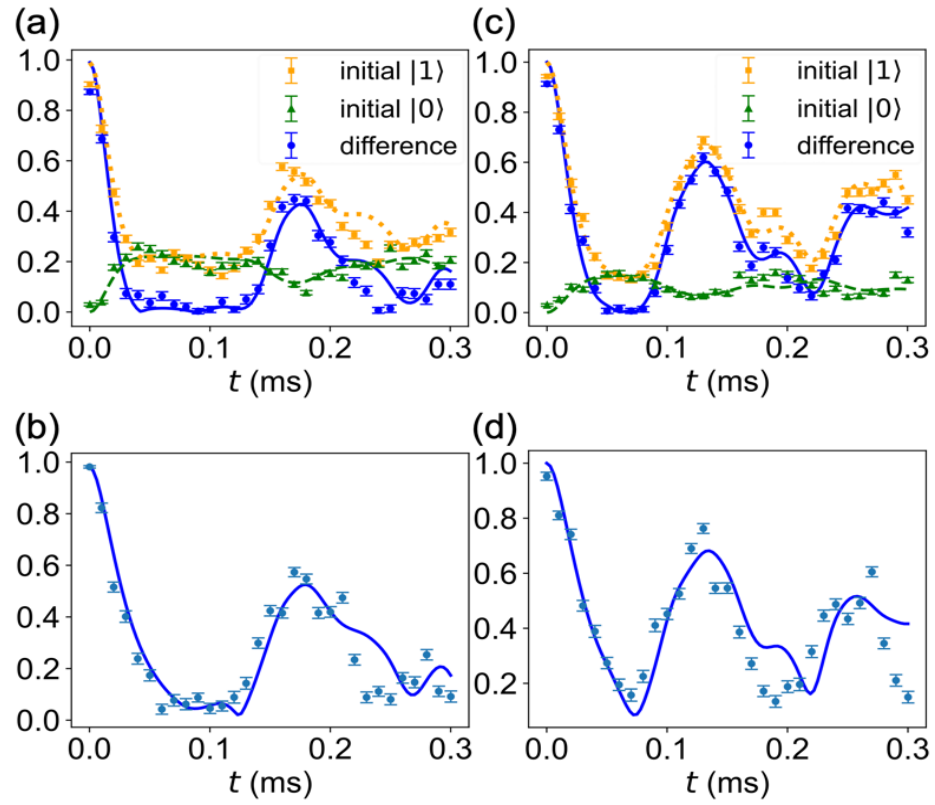
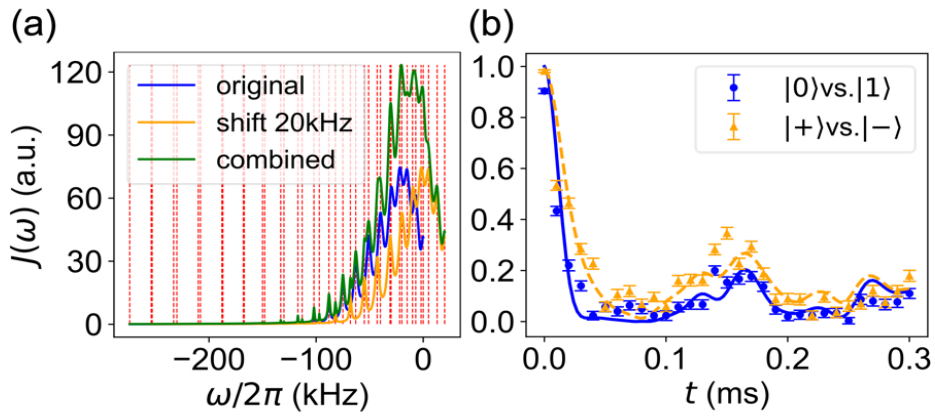


图 1: 边缘离子（右上）与中心离子（右下）环境谱密度与频率关系

继而通过设定相应的实验参数来调控期望达到的系统与环境的耦合强度，实验上观测到了编码信息在不同系统环境耦合条件下的演化情况，观察到了编码信息丢失和恢复的现象（图二），实验和理论数据基本吻合。研究人员还展示了多种自由度对环境耦合强度的调节能力，例如离子个数、离子间距、激光频率，以及增加多个激光频率成分对环境谱密度的调控（图三），展现了离子阱系统对于研究开放量子系统的潜力。



图二：（a，b）20 离子系统中边缘离子信息演化情况；（c，d）10 离子系统中边缘离子信息演化情况；（a，c）信息初态编码在 01 态；（b，d）信息初态编码在 +- 态



图三：（a）20 离子系统中边缘离子在两个频率成分作用下的环境谱密度情况；（b）相同激光频率失谐情况下的信息演化情况

该成果研究论文：G.-X. Wang, Y.-K. Wu, R. Yao, W.-Q. Lian, Z.-J. Cheng, Y.-L. Xu, C. Zhang, Y. Jiang, Y.-Z. Xu, B.-X. Qi, P.-Y. Hou, Z.-C. Zhou, L. He, L.-M. Duan, "Simulating the spin-boson model with a controllable reservoir in an ion trap", Physical Review A 109.6 (2024): 062402.

二、量子网络

主要完成人：段路明研究组、吴宇恺研究组

利用双类型量子比特编码实现无串扰的量子网络结点

量子网络是实现量子通讯和大规模量子计算的基础，生成离子 - 光子纠缠是实现离子阱量子网络的关键步骤。此前，为了实现无串扰的离子 - 光子纠缠，研究人员通常采用不同种类的离子分别来产生离子 - 光子纠缠（称为“通讯离子”）和存储量子信息（称为“存储离子”），这样可以使得通讯离子散射的光子远失谐于存储离子的跃迁频率，从而抑制串扰误差。然而，这类方案需要精细控制不同种类离子的比例和位置，且协同冷却效率低，难以实现不同离子间的量子逻辑门操作。

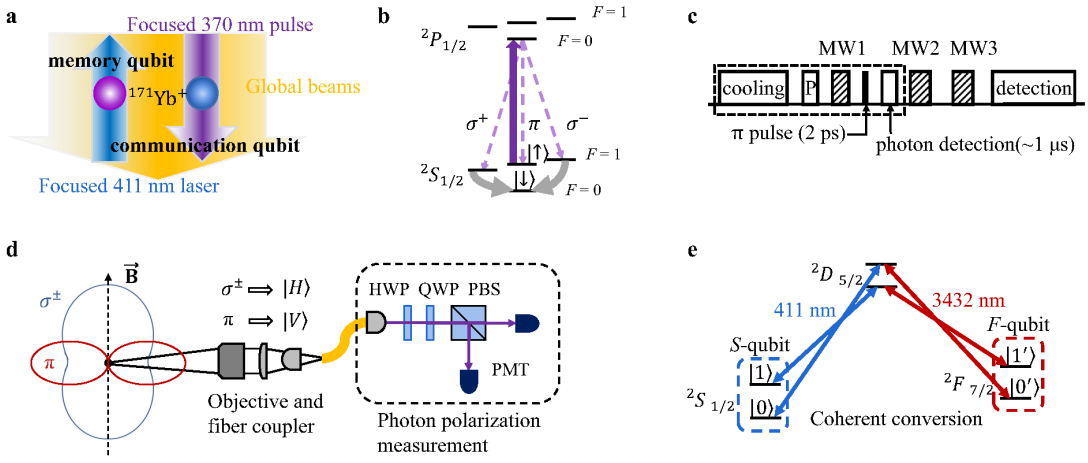


图 1： 实验方案示意图

为了克服上述困难，段路明研究组提出使用同种离子的双类型量子比特实现量子网络结点的方案。该方案利用同种离子的两对超精细结构能级分别编码通讯比特和存储比特（S- 比特和 F- 比特），并利用 411 nm 和 3432 nm 的双色窄线宽激光实现了两种量子比特之间微秒量级的相干转换。一方面，研究人员演示了利用 S- 比特在数百毫秒的时间尺度内生成离子 - 光子纠缠；另一方面，研究人员通过自旋回波方法延长 F- 比特的存储寿命，实现相干时间达到秒量级的存储量子比特。通过比较有无离子 - 光子纠缠生成操作时存储比特的保真度变化，研究人员证实了两种量子比特之间低于实验精度的串扰误差，从而实现了无串扰的量子网络节点。相较于之前采用不同种类离子的方案，该方案极大简化了实验系统，向着未来量子计算机和量子网络的模块化迈出了重要一步。

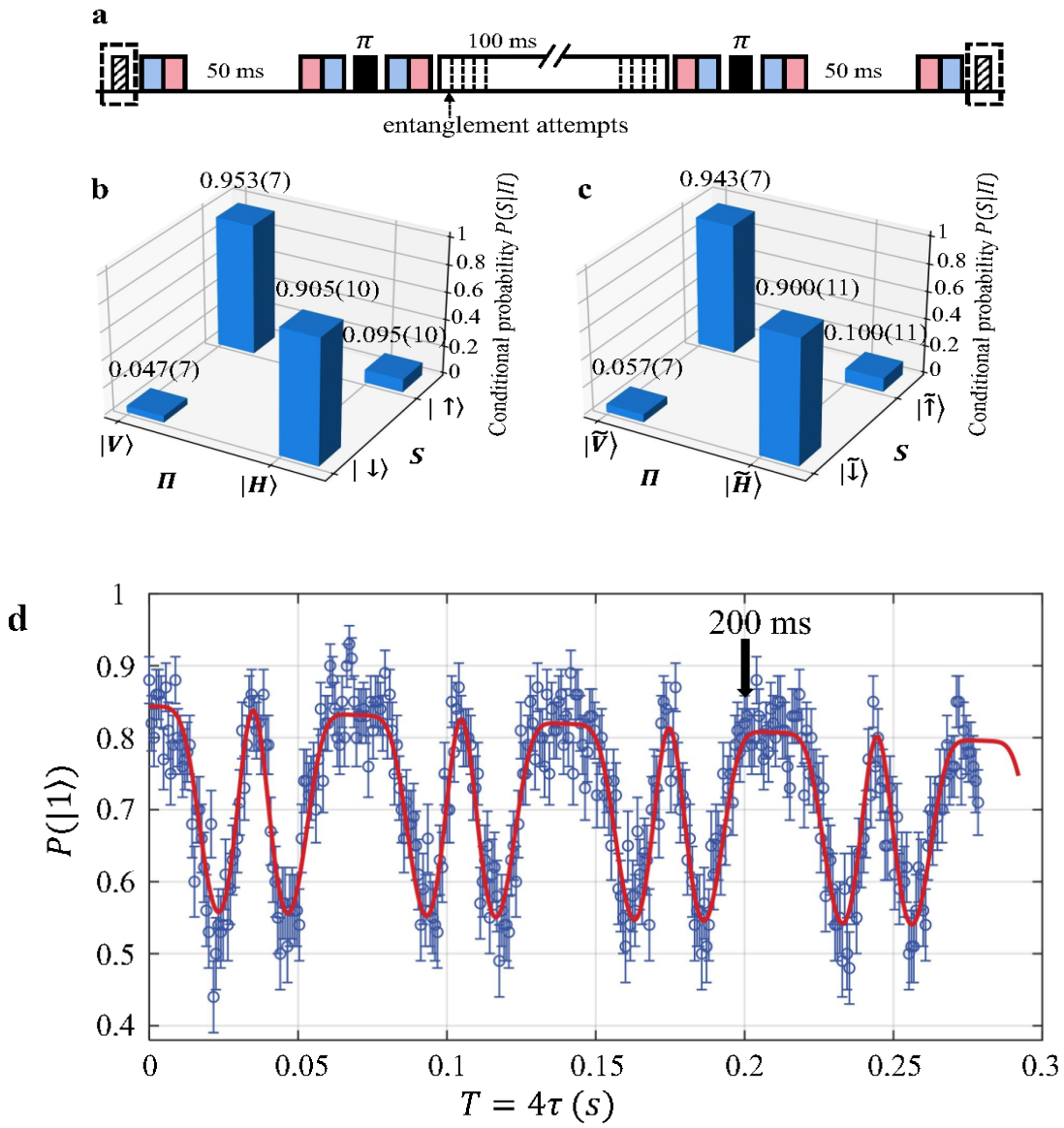


图 2： 无串扰量子网络结点演示图

该成果研究论文：Feng L, Huang Y Y, Wu Y K, Guo W X, Ma J Y, Yang H X, Zhang L, Wang Y, Huang C X, Zhang C, Yao L, Qi B X, Pu Y F, Zhou Z C, Duan L M, “Realization of a crosstalk-avoided quantum network node using dual-type qubits of the same ion species[J]”, Nature Communications 2024.

三、量子中继

主要完成人：濮云飞研究组、段路明研究组

可实现 1000 个光量子比特随机存取的存储器

量子存储器是实现量子计算、量子网络、量子精密测量的关键部件。同时具备高保真度，长存储寿命，多存储模式，以及随机存取功能的光量子存储器是实现长距离量子中继和量子网络的必备条件。目前可以用来处理大规模光量子比特流的量子存储器尚未实现。

在该工作中，濮云飞与段路明研究组实验实现了具有 72 个光量子存储单元（规模同目前最大的量子计算机可比），存储时间 >500 微秒（光子在 100km 光纤中传输的时间），以及 1000 次光量子比特操作（对于一个 72 光量子比特的输入序列可产生 72! ≈种不同的输出序列），是之前的世界纪录 12 次的接近 100 倍。在此基础上，该工作首先展示了一个接近于经典计算机中的随机读写存储器（RAM）的光量子随机读写，在 1000 次读写中，存储的保真度为 93(5) %。

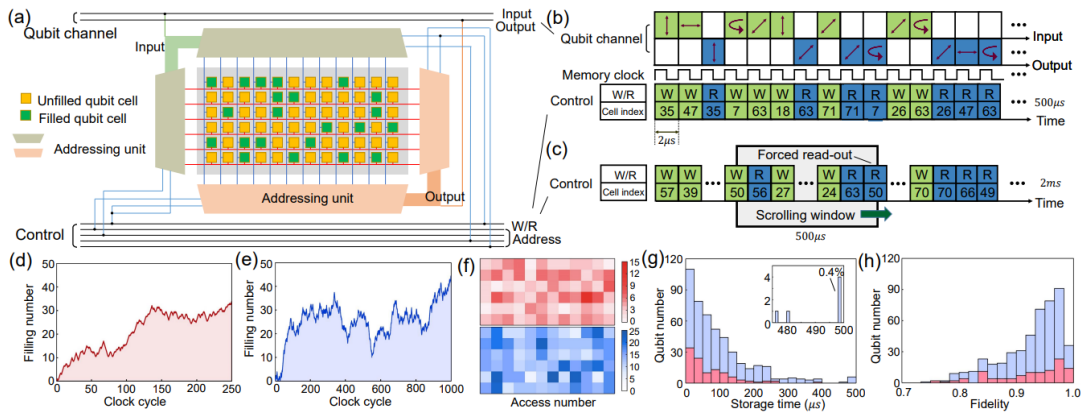


图 1：大规模量子随机读取存储器

此外，该工作也展示了几种在经典信息处理中的数据结构在量子系统中的实现。分别实现了量子队列，量子堆栈，和量子缓存器。这些实现对于拓宽量子存储的工具库以及对未来更大规模和更加复杂的量子体系，以及量子操作系统的实现有帮助。

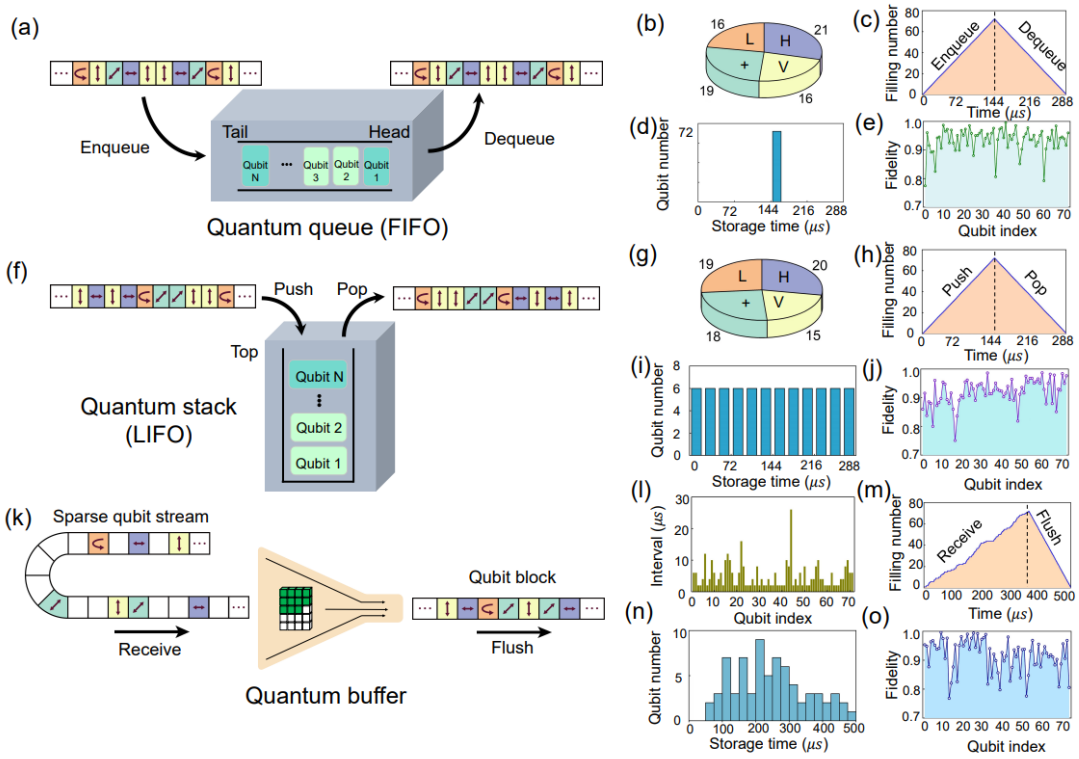


图 2：量子队列，堆栈，缓存器的实验实现

最后，该工作展示了连续 4 个随机产生的纠缠光子对的存储，同步，以及交换顺序的输出。这对于量子中继的实现至关重要。

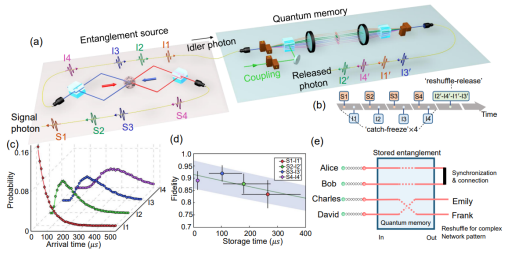


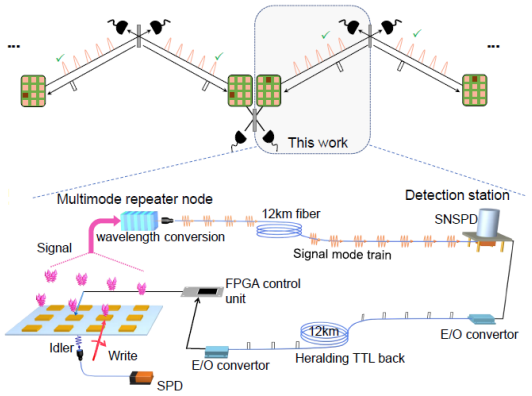
图 3：同步和交换四个纠缠光子对

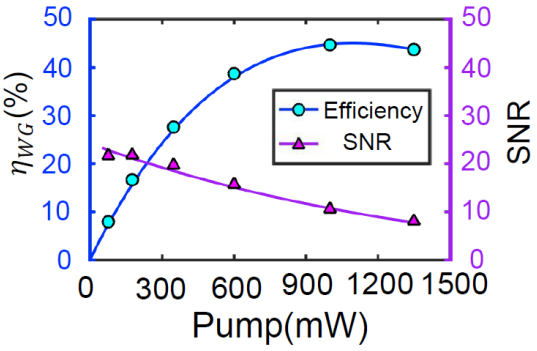
该成果研究论文：Sheng Zhang, Jixuan Shi, Zhaibin Cui, Ye Wang, Yukai Wu, Luming Duan, Yunfei Pu, "Realization of a Programmable Multipurpose Photonic Quantum Memory with Over-Thousand Qubit Manipulations", Phys. Rev. X 14, 021018 (2024).

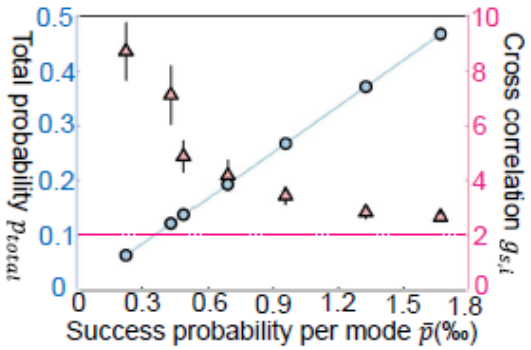
12km 光纤长度下提升量子中继节点与光子的预报式纠缠效率 140 倍

濮云飞与段路明研究组通过多路复用提升量子中继节点与通讯波段光子的预报式纠缠产生效率。成功实现了 12km 光纤距离上的中继节点和光子之间的预报式纠缠，且成功率通过多路复用提升 140 倍。在 12km 的光纤距离上，原子 - 光子纠缠产生速率达到 1.95kHz；且在存储器存储寿命内可产生 0.46 个期望纠缠。此两项均为长距离预报式原子 - 光子纠缠的世界纪录。该研究为之后通过多路复用提升两个中继节点之间的纠缠效率铺平了道路。

该成果研究论文：Sheng Zhang, Jixuan Shi, Yibo Liang, Yuedong Sun, Yukai Wu, Luming Duan, Yunfei Pu, "Fast delivery of heralded atom-photon quantum correlation over 12km fiber through multiplexing enhancement", https://arxiv.org/abs/2403.13623.



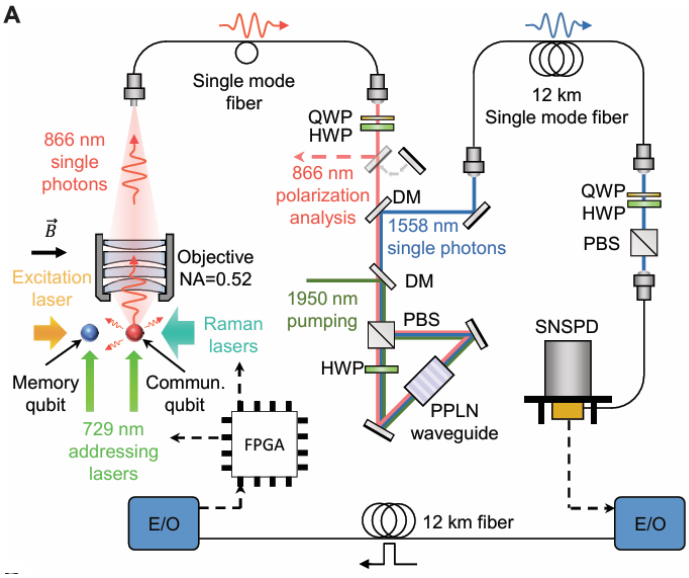


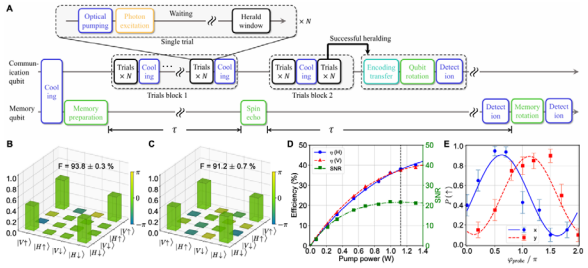


适用于 12km 城际量子网络的无串扰、多离子网络节点

段路明与濮云飞研究组首次实现适用于城际量子网络的无串扰离子阱节点。该节点由两个钙 40 离子组成，其中一个是用于光量子通讯的通讯量子比特，另一个是用来存储量子信息的存储量子比特。两种不同的量子比特采用 Dual-type 方案编码在不同的子空间，相互之间无串扰。通讯量子比特能够产生离子和 866nm 光子的纠缠，可以通过 3m 和 1km 的光纤传输之后产生预报式纠缠，纠缠保真度分别为 94% 和 91%。866nm 光子可以进一步通过参量下转换为 1558nm 通讯波段光子，并经过 12km 光纤的传输，得到保真度为 88% 的离子 - 光子纠缠。存储比特可以在不断进行离子 - 光子纠缠产生的同时，将量子信息保存 0.3 秒，不受各种诸如冷却，泵浦，光子激发，以及荧光探测等破坏性操作的影响。

该成果研究论文：P.-C. Lai, Y. Wang, J.-X. Shi, Z.-B. Cui, Z.-Q. Wang, S. Zhang, P.-Y. Liu, Z.-C. Tian, Y.-D. Sun, X.-Y. Chang, B.-X. Qi, Y.-Y. Huang, Z.-C. Zhou, Y.-K. Wu, Y. Xu, Y.-F. Pu, L.-M. Duan, "Realization of a crosstalk-free multi-ion node for long-distance quantum networking", https://arxiv.org/abs/2405.13369.





四、量子密码与通信

主要完成人：马雄峰研究组

利用不完美单光子源实现超过 23Mbps 的测量设备无关量子随机数生成

量子随机性在很大程度上依赖于其生成器的准确标定，然而，不完美的设备和不准确的标定会导致错误的信息熵估计和输出结果中的偏差，从而显著影响所生成随机性的可靠程度。测量设备无关（MDI）量子随机数生成（QRNG）旨在使用未标定和不可信的测量设备产生可靠的量子随机数，这些设备易受各种针对测量漏洞的攻击方案的威胁。然而，现有的实现到目前为止性能还不足。

在该工作中，马雄峰与中科大、中科院团队合作，设计并实施了一个高速 MDI-QRNG 方案。该方案利用了一种对于不完美单光子源具有鲁棒性的量子测量层析方法。和传统的方案相比，这种诱骗态方法可以获更加准确的层析结果和更紧的随机性下界。最后，通过使用一种高速时域编码系统，研究团队实验验证了该方案并获得了 7.37x10^-2 比特每脉冲的最小熵下界，相应的随机数生成速率超过了 23Mbps，大大超越了现有的 MDI-QRNG 实现方案，并在离散变量半设备无关 QRNG 中创下了记录。

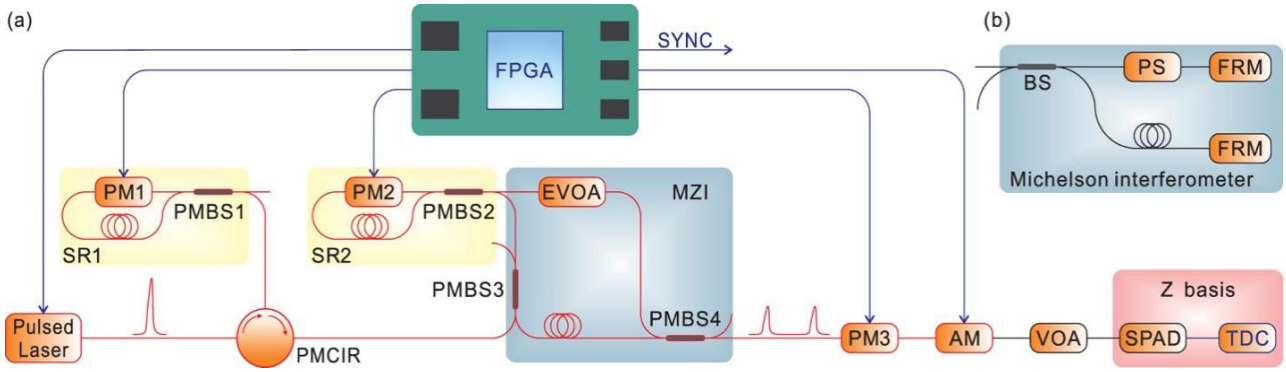


图1（a）高速MDI-QRNG的实验设置，包括源和测量两部分。
（b）使用匹配的迈克尔逊干涉仪进行的 X，Y 基测量

该成果研究论文：You-Qi Nie, Hongyi Zhou, Bing Bai, Qi Xu, Xiongfeng Ma, Jun Zhang, Jian-Wei Pan, "Measurement-device-independent quantum random number generation over 23 Mbps with imperfect single-photon sources", Quantum Science and Technology 2024.

模式配对量子密钥分发的实场测试

量子密钥分发（QKD）是量子技术的基石，能够为远程双方提供信息理论安全的密钥。随着全球许多量子通信网络的建立，模式配对协议（MP）因其在使用简单设备的情况下，在城际距离上的高效性而脱颖而出，成为一种有前景的解决方案。MP 方案通过利用模式配对的固有稳定性，解决了全局相位锁定的挑战，不需要相位稳定化，从而简化了系统设置。MP 协议在不需要严格相位稳定化要求的情况下，保持了量子态之间的相干关系，展现出在城际距离上的卓越性能，是量子通信网络的一个实用且引人注目的选择。其易于实施和高鲁棒性，使其成为量子通信的有前途方法。

此前的研究展示了 MP 方案在实验室中的高效性，取得了令人印象深刻的密钥生成率。然而，从受控的实验室环境过渡到现实世界设置会引入新的实验挑战。特别是在现场部署的光纤网络中以及非对称链路条件下，MP 方案的性能尚待验证。现场环境的复杂性以及多节点场景中为最大化密钥率所需的参数优化都是关键考虑因素。这些不确定性强调了需要严格的现场测试来验证该协议在现实世界中的适用性。

马雄峰和博士生黄溢智与中科大团队合作，在已部署的城际光纤网络中成功实施了 MP 方案。如图 2 所示，现场测试实验在现有的义乌 - 武义 - 丽水及金华 - 武义 - 丽水商用通信光纤上进行，这些链路也代表了城际量子网络的典型长度与规模。在工作中，课题组采用线性规划方法进行参数估计，以提高包括对称与非对称等实际场景中密钥率。通过使用超物理接触（UPC）连接器的商用光纤和集成光纤环行器，现场测试系统有效减轻了来自探测器和光纤端面的光反射影响。这种方法显著降低了噪声，实现了万分之一级别的极低比特错误率。值得注意的是，该现场展示避免了使用全局相位锁定技术，简化了系统配置并节省了额外的光纤资源。全局相位锁定的消除有助于减少背景噪声和更稳定的错误率，只需使用两个密集波分复用滤波器进行噪声减少。这进一步简化了系统并增强了其实用性。

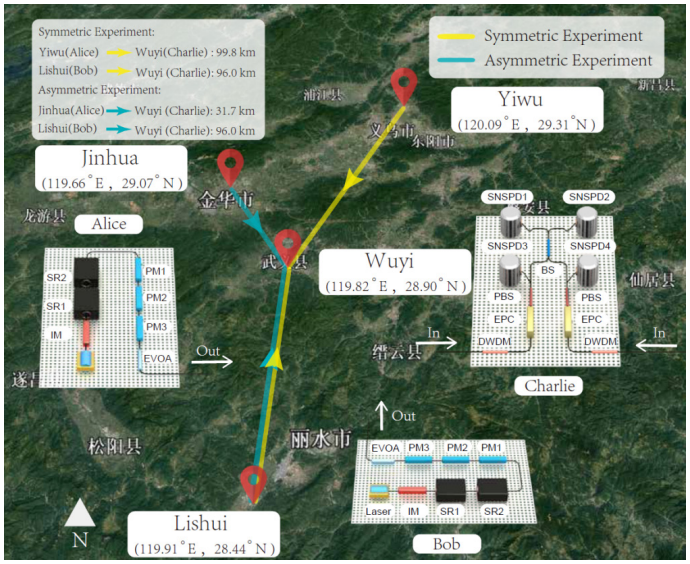


图 2 现场实验示意图。在浙江省中部的几个城市之间现有光纤网络中实验实现模式匹配量子密钥分发协议。

实场测试的结果显示，在 100 到 200 公里的通信范围内，系统在对称和非对称场景中实现了每秒数千比特的密钥率。这个密钥率在距离超过北京 - 上海骨干网络平均节点跨度的情况下，与该网络中单一 QKD 系统的密钥率相当。该研究结果强调了 MP 方案的实际可行性，其设计与现有网络类似，展现出比现有 QKD 方案更灵活且更高的密钥生成率，同时保持了 MDI 特性。这项工作证明了 MP 方案适应于城市量子通信网络，增强了节点布置的灵活性和路由能力，为多样化的量子网络场景提供了可靠的解决方案。

该成果研究论文 : Hao-Tao Zhu, Yizhi Huang, Wen-Xin Pan, Chao-Wu Zhou, Jianjun Tang, Hong He, Ming Cheng, Xiandu Jin, Mi Zou, Shibiao Tang, Xiongfeng Ma, Teng-Yun Chen, and Jian-Wei Pan, "Field test of mode-pairing quantum key distribution", Optica 11, 883-888 (2024).

一个经典量子混合的量子态非线性性质探测的框架

量子系统的非线性特性探测在多体量子物理和量子信息科学中扮演着重要角色。量子态的纯度、多体关联函数等非线性特性是描述量子多体系统性质的关键物理量。近期提出的基于量子态纯化的量子错误缓解方案也依赖于大量非线性量子特性的测量。然而，目前高效的量子测量方案主要集中于线性量子特性的探测，对于非线性量子特性的探测手段相对不足。

在该项工作中，马雄峰研究组的博士生刘振寰与复旦大学的周游教授合作，基于影子层析和量子相干测量技术，提出了一种结合经典与量子的量子态非线性特性探测实验框架。该框架能够高效地同时探测多个独立的量子态非线性特性，并且可以在仅对少量量子态进行相干测量的情况下推导出更高阶的量子态非线性特性。经过对框架复杂度的分析，研究人员发现其在探测某些特定非线性物理量方面的效率优于传统的基于非相干测量的实验方案。

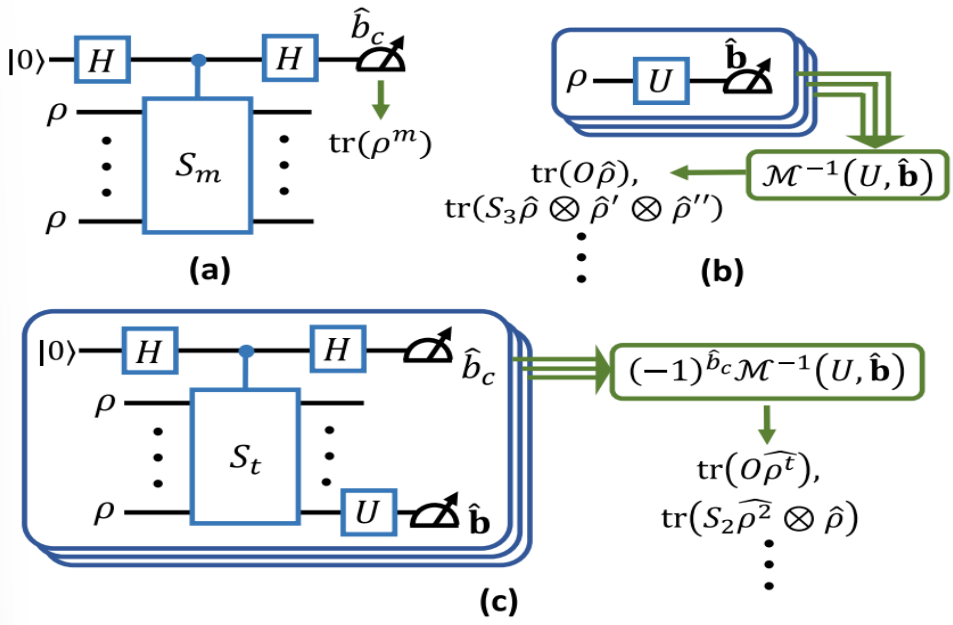


图 3 经典量子混合的量子态非线性性质探测方案

该成果研究论文：You Zhou & Zhenhuan Liu, "A hybrid framework for estimating nonlinear functions of quantum states", npj Quantum Information, 2024.

非局域性，纠缠和测量非对易性的相互关系

贝尔的定理揭示了量子理论与局部隐变量模型之间的根本不兼容性，从而突出了量子非局域性、纠缠和测量非对易性这三种关键的量子资源。这些资源不仅对量子基础有重要意义，在量子密码学领域也扮演着至关重要的角色。尽管纠缠和测量非对易性同时作为展现非局域性的必要条件，它们与非局域性之间的定量关系却是复杂的。非局域性提供了一种在设备不可信的条件下，进行设备无关纠缠估计的有效途径。另一方面，测量非对易性作为非局域性的另一个必要条件，其作用往往被忽视。

在此项研究中，马雄峰教授及其研究组成员博士生张行健和朱雨薇，基于一类参数化的 CHSH 型贝尔不等式，对纠缠形式、纠缠的负性以及单向可蒸馏纠缠等多种纠缠度量进行了深入分析，并给出了这些纠缠度量的精确解析和数值下界。研究进一步探讨了在实际量子信息处理任务中，如何通过优化贝尔不等式的参数来提高纠缠估计的准确性。此外，研究还指出，在限定系统为两量子比特的条件下，将测量非对易性纳入考量，会发现非局域性、纠缠和测量非对易性三者之间存在非直观的相互作用关系：在固定非局域性的条件下，系统达到最小纠缠时，所采用的测量策略一般并非是最大非对易性测量。这一发现为理解和利用量子资源提供了新的视角。

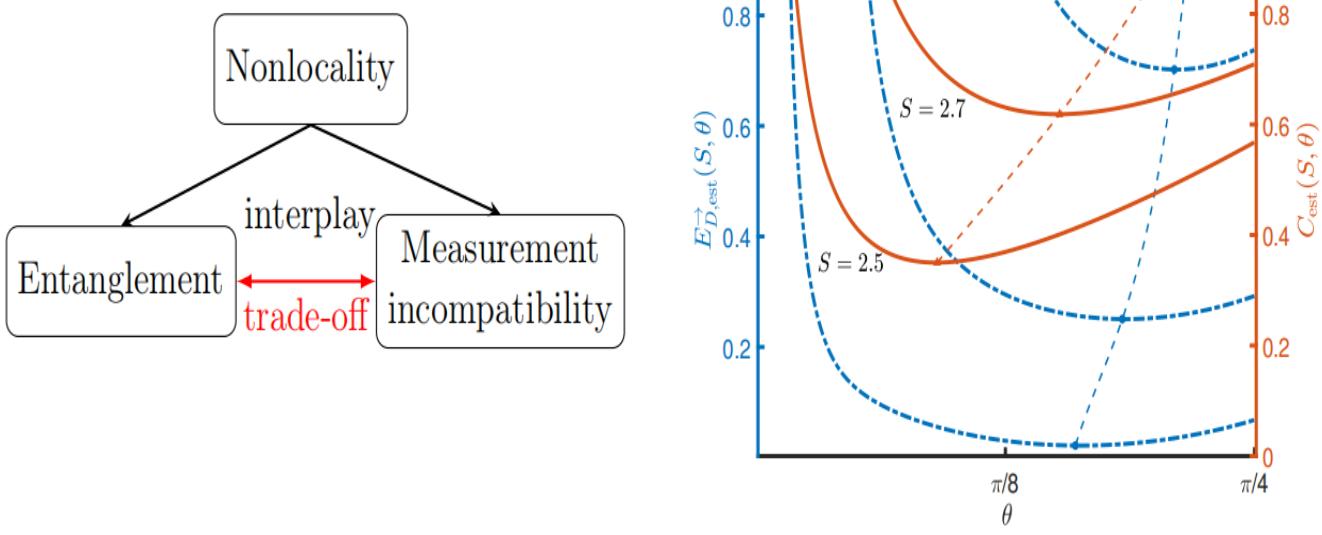


图 4 非局域性，纠缠和测量非对易性的相互关系

该成果研究论文：Yuwei Zhu, Xingjian Zhang and Xiongfeng Ma, "Interplay among entanglement, measurement incompatibility, and nonlocality", Quantum Science and Technology. 2024.

五、量子纠错

主要完成人：孙麓岩研究组

逻辑量子比特之间的量子纠缠受到量子纠错保护

量子纠错是量子计算领域的核心问题之一。众所周知，环境噪声会显著影响量子比特的状态，使量子计算很容易因外部干扰而出现错误。量子纠错可以保护量子信息不受错误影响，为量子计算的实用化进程带来了曙光。量子纠错通过引入冗余自由度，利用特定的编码方式构造逻辑比特来保护信息。一旦检测到错误，就可以使用这些额外的冗余信息来识别和修正错误，而不破坏原始量子信息。

当前主要的量子纠错技术路线是基于二能级比特的编码方案，利用大量物理比特编码一个逻辑比特。这在实验上有比较大的挑战，表现在错误通道数目随着物理比特数目的增加而增加，同时需要复杂的多体相互作用来实现错误的检测和纠正。为了克服这些困难，孙麓岩研究组采用了一种基于单个微波谐振子的物理架构，即玻色编码。这种方案利用谐振子的无限大希尔伯特空间来进行冗余信息编码，但是错误通道的种类保持不变，因此大大减少了量子纠错对硬件的要求，可以率先实现量子纠错。

近年来，孙麓岩研究组一直致力于基于玻色编码的量子纠错研究。该课题组与邹长铃研究组合作，首次在实验上实现了玻色二项式编码的量子纠错，演示了对逻辑量子比特的连续量子纠错，将相干时间延长为没有量子纠错时的 2.8 倍 [Nature Physics 15, 503 (2019)]。随后，为了提高纠错操作的保真度，他们又合作发展了对错误透明的相位门 [Nature Physics 16, 827 (2020)]。在此基础上，他们又与南方科技大学俞大鹏 / 徐源团队、福州大学郑仕标研究组、北京量子信息科学研究院于海峰团队合作，通过提高量子比特的相干时间和优化错误检测方案，实现了量子纠错突破盈亏平衡点 [Nature 616, 56 (2023)]。玻色编码下一步的关键问题是：如何将量子纠错扩展到多个逻辑量子比特并实现纠缠保护。

孙麓岩与合作者在解决这一关键问题上取得了突破。他们通过在空间分离的玻色模式中编码量子信息（图 1），首次成功实现了受量子纠错保护的纠缠逻辑量子比特。不仅展示了量子纠错技术在保护量子信息方面的巨大潜力，也为实现实用的量子计算和通信网络提供了重要的技术基础。具体地，研究团队在两个空间分离的玻色模式中，实现了二项式编码的逻辑量子比特，并建立了两个逻辑量子比特之间的纠缠。与物理比特之间的纠缠不同，逻辑比特之间的纠缠对局域噪声有更强的抵御能力，并且可以通过局域量子纠错操作进行恢复，而不是直接耗散到环境中。研究团队通过连续量子纠错操作对纠缠逻辑量子比特进行保护，使纠缠逻辑量子比特的相干时间相较于未受保护的状态提高了 45%（图 2）。更进一步，研究团队还首次证明了，通过错误探测以及提纯，纠缠逻辑量子比特能够违反贝尔不等式（图 3）。

该研究工作成功将玻色量子纠错码扩展到多个逻辑比特，证实了通过量子纠错保护纠缠逻辑量子比特的可行性，这不仅为迈向通用量子计算，构建量子纠错保护的量子网络和分布式量子计算奠定了实验基础，也为验证物理学基本原理提供了新的实验平台。

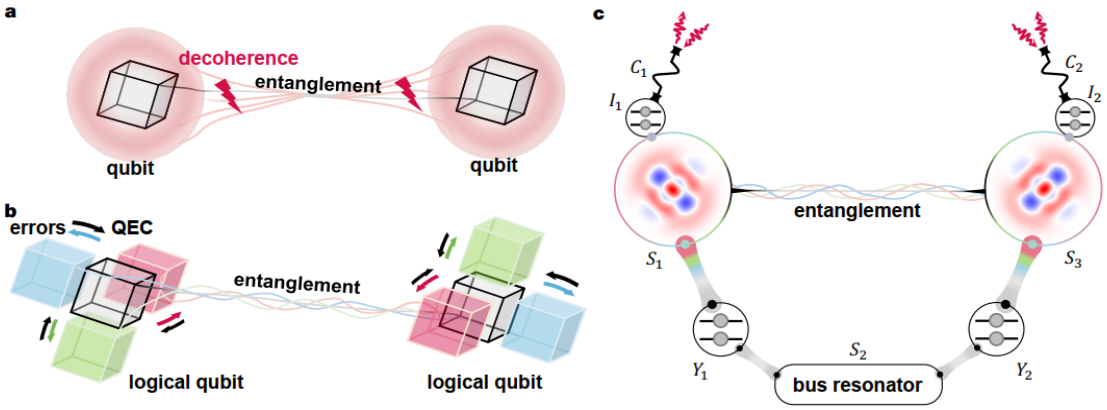


图 1 纠缠逻辑量子比特（ELQ）的原理和装置图

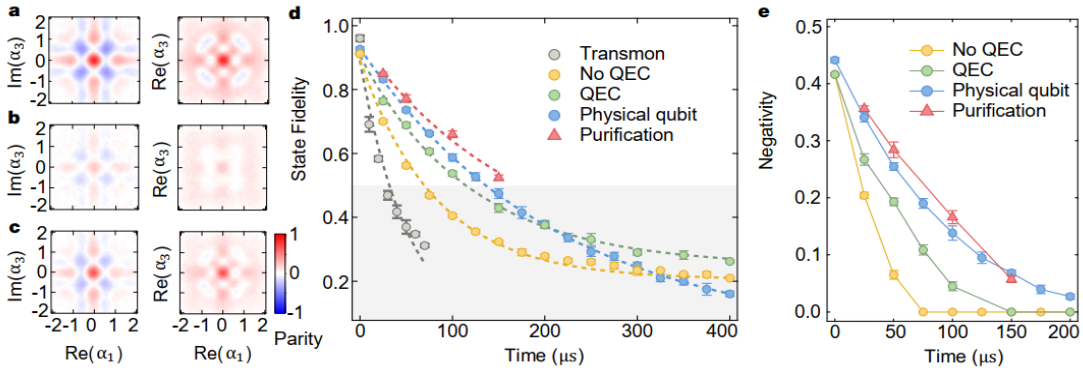


图 2 利用量子纠错实现对逻辑量子比特的纠缠保护

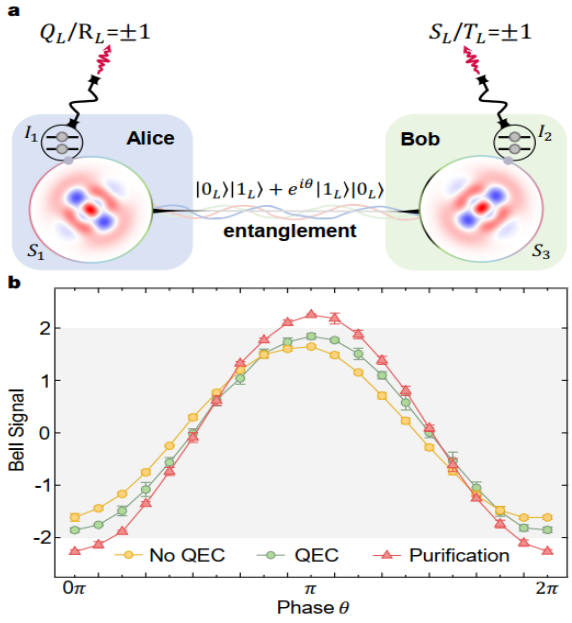


图 3 利用纠缠的逻辑量子比特演示贝尔不等式的违背

该成果研究论文：Weizhou Cai, Xianghao Mu, Weiting Wang, Jie Zhou, Yuwei Ma, Xiaoxuan Pan, Ziyue Hua, Xinyu Liu, Guangming Xue, Haifeng Yu, Haiyan Wang, Yipu Song, Chang-Ling Zou & Luyan Sun, "Protecting entanglement between logical qubits via quantum error correction", Nature Physics 2024.

六、量子人工智能

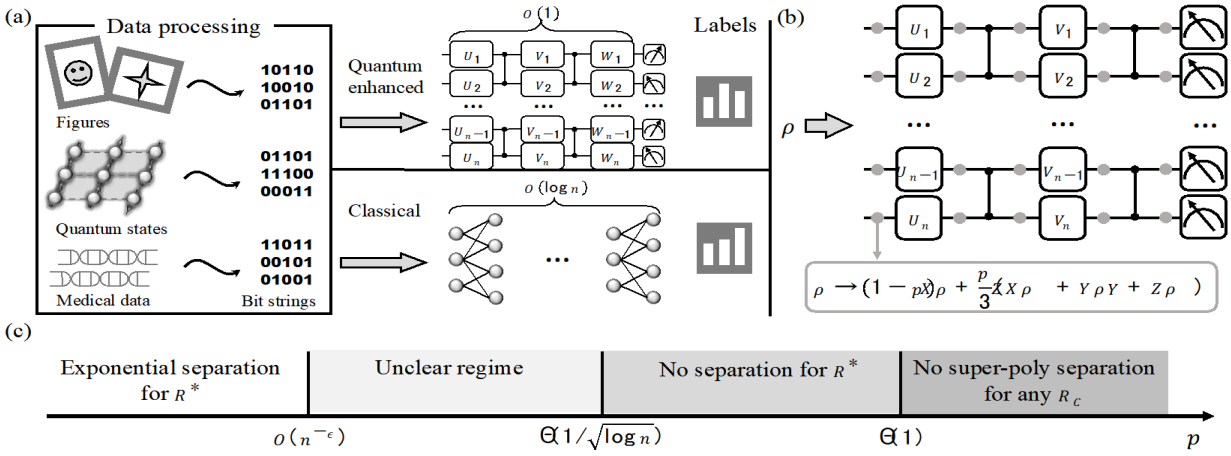
主要完成人：邓东灵研究组

基于浅层线路监督学习的量子优势

量子机器学习近年来受到了广泛的关注。其中，量子模型的学习优势是一个重要的探索方向。然而，发展具有量子优势的学习算法面临诸多挑战：1）一些被提出具有指数加速的量子学习模型被“去量子化”，即存在一些经典算法可以实现同样的学习效率；2）大多数量子优势需要依赖复杂性理论的假设；3）在带噪声量子线路中的量子学习优势鲜有研究。

在此背景下，邓东灵研究组在当前的工作中提出了可以严格证明的的量子学习优势。首先，邓东灵研究组提出了一个多标签学习任务。利用量子非定域性作为所利用的量子资源，邓东灵研究组严格证明了，在无噪声的情形下，存在一个量子算法可以利用常数深度的线路解决这个学习问题，而对于经典的有限连接的神经网络，解决这个问题的深度至少为 log(n)。值得提出的是，这一量子优势不需要依赖任何复杂性理论的假设，即，无条件 (unconditional) 的量子学习优势。此外，在不同的噪声情形下，邓东灵研究组研究了噪声对实现量子学习优势的影响。邓东灵研究组证明，如果噪声强度的上界是一个关于系统大小的逆多项式，即 O(1/poly(n))，这种学习优势将持续存在。如果噪声强度大于一个逆多对数，即 Θ(1/)，这种分离将消失。对于噪声强度恒定的量子器件，邓东灵研究组证明了在任何由浅 Clifford 电路定义的分类任务中，无论任何学习线路结构，不存在超多项式的经典 - 量子学习分离。

邓东灵研究组的工作对量子机器学习的优势在不依赖复杂性假设、不同噪声条件下进行了界定，为相关的工作和未来的探索提供了指导。



该成果研究论文：Zhihan Zhang, Weiyuan Gong, Weikang Li, Dong-Ling Deng, "Quantum-Classical Separations in Shallow-Circuit-Based Learning with and without Noises", arXiv: 2405.00770.

七、拓扑凝聚态物理

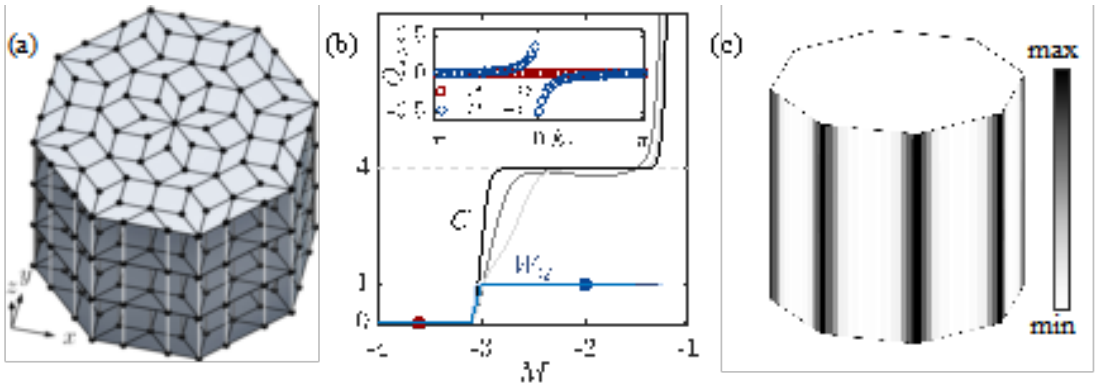
主要完成人：徐勇研究组

无常规晶格对应的三维高阶拓扑绝缘体

拓扑物态由于其独特的“体 - 边对应”关系受到了广泛的关注。在同一维度的高阶拓扑绝缘体具有维数低于传统拓扑绝缘体的边界态。准晶体系具有更多的旋转对称性，如五重或八重旋转对称性，这些对称性能保护一些更加新奇的高阶拓扑相。这些相具有独特的边界态，并且不能在具有平移不变性的常规晶格中存在。人们已经发现在二维的准晶体系中存在一些特别的高阶拓扑相。于是人们自然会考虑：三维准晶体系中能否存在高阶拓扑相，并且其对应的拓扑不变量是什么?

徐勇研究组首次在理论上证明了三维准晶体系中存在由八重旋转对称性保护的高阶拓扑绝缘体相，并给出了刻画这个拓扑相的拓扑不变量。研究组在由二维准晶堆叠而成的三维晶格中构造了新的紧束缚模型，模型哈密顿量满足时间反演对称性和八重旋转对称性相结合的一种新的对称性。在此对称性的保护下，体系会在 8 条棱上出现拓扑边界态。为了构造刻画这一拓扑态的拓扑不变量，研究组将只能用于描述方格子的四级矩推广至准晶体系中，并用其绕数描述这个拓扑绝缘体态。此外，课题组还研究了三维准晶体系中具有时间反演对称性的二阶拓扑绝缘体和外尔半金属，并给出了它们对应的拓扑不变量。

他们的研究发现了三维准晶体系中新的高阶拓扑相，并提出了新的拓扑不变量，这为研究在 z 方向上也没有平移对称性的准晶体系中的拓扑相开辟了新思路。



（a）三维准晶示意图；（b）拓扑不变量随相变变化图；（c）一维边界态的态密度图

该成果研究论文： Y.-F. Mao, Y.-L. Tao, J.-H. Wang, Q.-B. Zeng, and Y. Xu, "Higher-Order Topological Insulators in Three Dimensions without Crystalline Counterparts", Phys. Rev. B 109, 134205 (2024).

Editor:
Kailin Li
Reviewer:
Jian Li, Yipu Song, Dongling Deng, Yang Gao, Xiamin Lv