

高职级专家述职

未来院3室 许达
2025年7月15日入职

技术攻关：年度亮点工作及完成情况

入职时间：2025年7月15日

总体方向：深耕 AI for Quantum, AI for Science, AI for AI 交叉领域, 聚焦以下核心工作

软件工程：

- 量子噪声和人工智能量子纠错系统（合作，给量子计算机戴上“AI 降噪耳机”）
- 人工智能自动科研智能体系统（让 AI “像科学家一样思考”）
- 量子科研和量子代码大模型和多智能体（打造懂量子物理的“专用大脑”）

科学研究：

- 量子计算相关的物理和数学问题的研究（应用我们的科研智能体）
- 人工智能训练和推理算法研究（应用我们的科研智能体）

论文方面：已投稿到行业一流杂志论文 1 篇，已完成完成第 2 篇长论文正在投稿到行业一流杂志；多篇论文初稿在撰写中。

软件著作：已开发和正在开发的软件和计算工具6个，总代码超过13万行，其中2个软件著作权申请过程中。量子噪声和人工智能量子纠错系统超导部分完成 第一版，离子阱模块对接北京市科委量子项目；人工智能自动科研智能体系统第一版本已在本组及跨部门试用。这两个软件著作权申请过程中。量子科研和代码大模型及其软件正在研发过程中。另外有其他多个和研究有关的软件或计算工具也在开发中。

支撑价值：为“量子 × 大模型 × 智能体”的长期技术路线奠定关键基座, 形成可复用代码与规范。

技术攻关成果(1): 量子噪声模型和人工智能量子纠错译码

项目目标与角色 (给量子计算机戴上“AI 降噪耳机”)

项目目标: 面向超导 & 离子阱量子计算装置, 构建 AI 增强的量子纠错软件 QEC 平台。

担任角色: 技术负责人/架构师。主导模型架构设计与核心算法攻关。

- 超导模块: 完成 V1 版本的功能实现与仿真验证, 形成端到端可运行的纠错译码链路。
- 离子阱模块: 正在按北京市科委项目需求开发, 将于近期对接真实装置数据与实验流程。
- 模型创新: 引入多头隐含注意力 (Multi-Head Latent Attention) 架构, 加速训练和推理。
- 工程成果: 形成清晰的模块划分、单元测试与 Benchmark 脚本, 可复用性强。

创新思路:

传统纠错算法 (如 MWPM) 在处理复杂噪声时精度不够, 而普通深度学习模型又太慢, 跟不上量子比特的衰变速度。我们开发高效率的模型加快推理速度。

核心创新点:

架构革新: 引入了 **多头隐含注意力 (Multi-Head Latent Attention)** 机制。传统模型像是在“死记硬背”所有的噪声数据, 而我们的模型学会了“抓重点”(Latent), 只关注关键的错误特征, 从而大幅提升了推理速度和准确率。

通用适配: 构建了模块化架构, 一套代码可以同时适配 **超导** 和 **离子阱** 两种完全不同的物理平台。

预期成效 (Impact):

近期: 已完成 V1 版本, 实现端到端的纠错仿真链路。

远期: 软件直接部署在“无极一号”等真实装置上, 实现实时的在线纠错。

技术攻关成果(2):人工智能自动科研多智能体

项目目标与角色 (AI “像科学家一样思考”)

项目目标: 让智能体贯穿科研全链路:选题 → 检索 → 代码实验 → 数据分析 → 论文写作。

担任角色: 项目发起人/核心开发者。主导智能体工作流设计与 Prompt 工程优化。

- 已实现能力: 自动论文检索与阅读辅助、论文思路头脑风暴、实验代码生成、结果记录整理、初稿撰写与润色。
- 应用成效: 已辅助本组及跨部门同事完成多篇论文草稿/初稿, 显著缩短科研迭代周期。
- 论文进展: 在 AI 赋能量子与量子纠错方向已投稿 1 篇, 高水平论文 1 篇将于近期投稿。
- 工程贡献: 积累开发智能体的工程经验。

创新思路:

目前的普通大模型, 通常是“直觉式”回答, 缺乏深思熟虑。科研需要严谨的逻辑推理。

思路转变: 我们不仅训练模型, 更侧重于**推理时的计算 (Test-Time Compute)**, 强迫 AI 在回答前“多想一会儿”, 而且优先思考最“有戏”的方向。

核心创新点:

预算审慎推理 (Budgeted Deliberation): 我们提出了一个统一框架, 允许 AI 根据问题的难易程度, 动态调整“思考时间”。

风险敏感搜索 (Risk-Sensitive MCTT): 结合了控制理论, 让 AI 在探索科研路径时, 既能大胆假设, 又能像下棋一样预判风险, 避免在错误的方向上浪费算力。

预期成效:

效率提升: 显著缩短科研迭代周期, 目前已辅助团队完成了多篇论文初稿。

行业卡位: 在“AI for Science”领域建立专业的模型, 智能体, 以及工程化标准。(谷歌和openai有类似软件但不公开)

技术攻关成果(3):量子科研和量子代码大模型/多智能体

项目目标与角色 (打造懂量子物理的“专用大脑”)

项目目标: 面向量子科学研究与量子编程场景的专业代码大模型。

担任角色: 技术骨干。负责数据管线构建与微调策略, 推理时算法等的研发。

- 代码侧进展: 已完成数据管线与训练脚本开发, 本周启动第一阶段微调训练
- 训练策略: 采用参数高效微调(如 LoRA 等)与并行加速策略, 提高算力利用率
- 能力规划: 重点提升量子算法设计、量子电路编程与科研分析脚本的生成质量
- 推理时算法: 探索推理时算法增强模型的科学推理与自检能力
- 继续研究最先进的训练, 推理算法, 以提高表现和计算效率

创新思路:

有些通用模型写量子 电路代码(如 Qiskit)经常出错, 有些模型价格昂贵。

思路转变: 使用高质量的量子专用数据管线进行微调, 并结合参数高效微调(LoRA)技术降低训练成本。

核心创新点:

强化学习推理增强: 利用我们自己的推理时算法, 产生高水平的代码与论文, 然后让模型在生成的数据上训练。

预期成效:

工具落地: 发布 我们自己的 模型, 使其在量子 编程任务上达到可用水平。

填补空白: 解决量子领域缺乏 专用代码大模型的行业痛点

技术影响力：

行业技术卡位

1. 围绕“AI 增强量子纠错”和“量子专用代码大模型”布局，聚焦目前国内外空白/稀缺方向
2. 支撑北京市科委量子项目，助力公司深度融入“北京量子高地”建设
3. 人工智能量子纠错、自动科研智能体、量子科研和代码大模型都属于稀缺软件。AI量子纠错软件有望在真机上实验
4. 通过实际 Demo 和内部分享，帮助更多同事尝试使用智能体，提高科研生产力
5. 通过科研智能体，辅助攻关重大科学问题

专业队伍培养：

角色定位

- 作为 4-2 级技术专家，重点在“量子 ×AI 软件工程”方向发挥带头作用。
- 在软件架构设计、代码质量控制和工程工具选型等方面提供指导。
- 与同事共同建立规范化的代码管理与开发流程(Git 规范、Code Review、自动化测试)。
- 搭建并维护团队自有算力环境, 沉淀 One-Click 训练与评测脚本, 降低入门门槛。
- 通过技术分享会和 Pair-Programming 形式，帮助同事熟悉软件开发流程。

知识沉淀

- 撰写人工智能量子纠错架构设计、AI-Scientist 工作流说明等内部技术文档多份。
- 撰写各类内部技术调研报告约 20 篇，为团队技术决策提供有力支撑。
- 分享人工智能领域最新进展

问题与改进：

当前短板

- 对公司内部业务场景和产业化需求理解仍需加深，科研与业务结合度有提升空间。

科研推进难点

- 外部的商业模型用于大规模的科学研究极为昂贵，我们急需开发和部署自己的科研模型，解决卡脖子问题。
- 大模型训练算力资源紧张，影响迭代速度和多方案对比实验。

针对性改进措施

- 将量子与 AI 技术更紧密地对接具体业务问题。
- 深入研究能够提高大模型训练效率，降低训练费用的方法；深入研究推理时算法，提高大模型的能力

未来规划：

未来规划

2026 年科研与产品化目标

- 人工智能量子纠错系统：完成离子阱 V1 在我院“无极一号”等装置上的在线部署与场景验证, 形成 v1.0 文档与 Benchmark。
- 量子大模型：发布 v0.x → v1.0, 在典型量子编程任务上达到可用水平, 产出 2 - 3 篇算法相关论文。
- 自动科研智能体系统：接入自研科研推理模型, 形成研究院标杆级自动化科研流水线案例。
- 团队建设重点：重点培养“量子 +AI”复合型人才, 建立常态化技术分享机制, 提升团队整体工程与算法能力。
- 软件工程：2025 年已完成代码总行数约 13 万行(含三大项目核心代码), 2026持续保持高质量工程标准。
- 论文目标：2026 年内完成多篇重要问题的投稿。
- 资源诉求：当前仅有约 5 块 Ascend 910B 算力卡, 可满足基础实验但难以支撑多模型并行与大规模实验。

核心贡献：

提出了“预算审慎推理” (Budgeted Deliberation) 的统一形式化框架, 并设计了10 种超越线性 CoT 的测试时计算算法 (如 IGD, RS-MCTT 等)。

核心技术点：

- Budget-Performance Frontier (BPF): 定义并量化了测试时计算投入与模型性能的边界曲线。
- Index-Guided Deliberation (IGD): 引入 Gittins 索引机制, 实现多推理线程的动态最优调度。
- Risk-Sensitive MCTT: 结合风险敏感控制理论, 改进了蒙特卡洛思维树搜索的鲁棒性。

当前进展：

- 代码实现: 已完成核心算法代码开发, 包括 IGD 调度器与 RS-MCTT 搜索树。
- 测试验证: 正在进行合成数据与基准测试验证, 初步验证了 BPF 曲线的收益递减规律。

附录 2：黑洞的量子纠错理论

核心贡献：

这一工作将黑洞蒸发刻画为一个具有有限记忆的量子过程，把传统的几何 - 熵叙述彻底翻译到量子纠错语言中。核心贡献之一是证明：所谓“岛屿”实际上等价于算符代数量子纠错中的一个纠错子系统，黑洞内部自由度可视为编码在辐射中的量子码。Page 曲线的“转折点”被重新解释为量子纠错意义下“能否从辐射中高保真恢复内部信息”的阈值条件。进一步地，这一工作基于过程张量给出具体的恢复通道构造和误差界，展示如何在晚期辐射上实现对内部算符的近完美重建。总体上，它把黑洞信息悖论系统地重写为一个关于量子纠错码结构和可恢复性的物理问题。

创新点：

黑洞信息即量子编码--我们证明了黑洞内部的信息并没有消失，而是像**量子纠错码**一样被编码在辐射中

附录 3: 量子算法模拟黑洞

核心贡献：

我们这个工作是用量子算法模拟黑洞。我们把具有有限记忆的“地平线记忆梳”（HMC）模型，具体实现为一套可在门模型量子计算机上运行的量子电路与仿真流程。做法是将系统划分为“内区 - 记忆 - 辐射”三个寄存器，并给出相应的离散化映射；在每一步演化中，先用随机快搅拌电路（single-ary/qTPE）把内区与记忆中的信息快速打散，再通过由 RXX、RYY、RZZ 组合成的部分 SWAP 发射门，把一部分信息以么正方式转移到辐射寄存器，从而在有限资源下重现带有 Page 转折的辐射熵曲线。与此同时，我们在 Qiskit 中实现了可直接调用的 BlackHole Comb 仿真代码，包含熵的计算、资源开销估计和噪声缓解方案，使黑洞的非马尔可夫量子动力学变成一个可以在现有量子平台上实际运行和检验的量子算法问题。

附录 4： 时空彭罗斯不等式的谱几何证明(量子纠错视角)

核心贡献：

提出“谱质量泛函”与“谱共形流”新框架，**在一定的条件下**，严格证明了广义相对论核心猜想之一时空彭罗斯不等式。此不等式是重大公开问题。

附录 5:二次增长图上的谱 Zeta 函数渐近律

核心贡献:

证明了在具有二次体积增长的图(如二维晶格、无序介质)上,谱 Zeta 函数遵循精确的渐近律。与量子科学的联系:

- 量子场论正则化:谱 Zeta 函数是量子场论中计算单圈有效作用量(One-loop Effective Action)和卡西米尔效应(Casimir Effect)的核心工具(Zeta Function Regularization)。
- 无序量子系统:该理论适用于随机电导模型(Random Conductance Model)和渗流簇(Percolation Clusters),为研究无序量子材料中的安德森局域化(Anderson Localization)与量子输运提供了几何谱学基础。
- 量子引力离散化:在二维量子引力(如 Liouville Quantum Gravity)的离散模型中,谱渐近行为反映了时空的有效维度与共形异常。

本工作的副产品:发现了新的圆周率的渐进表达式