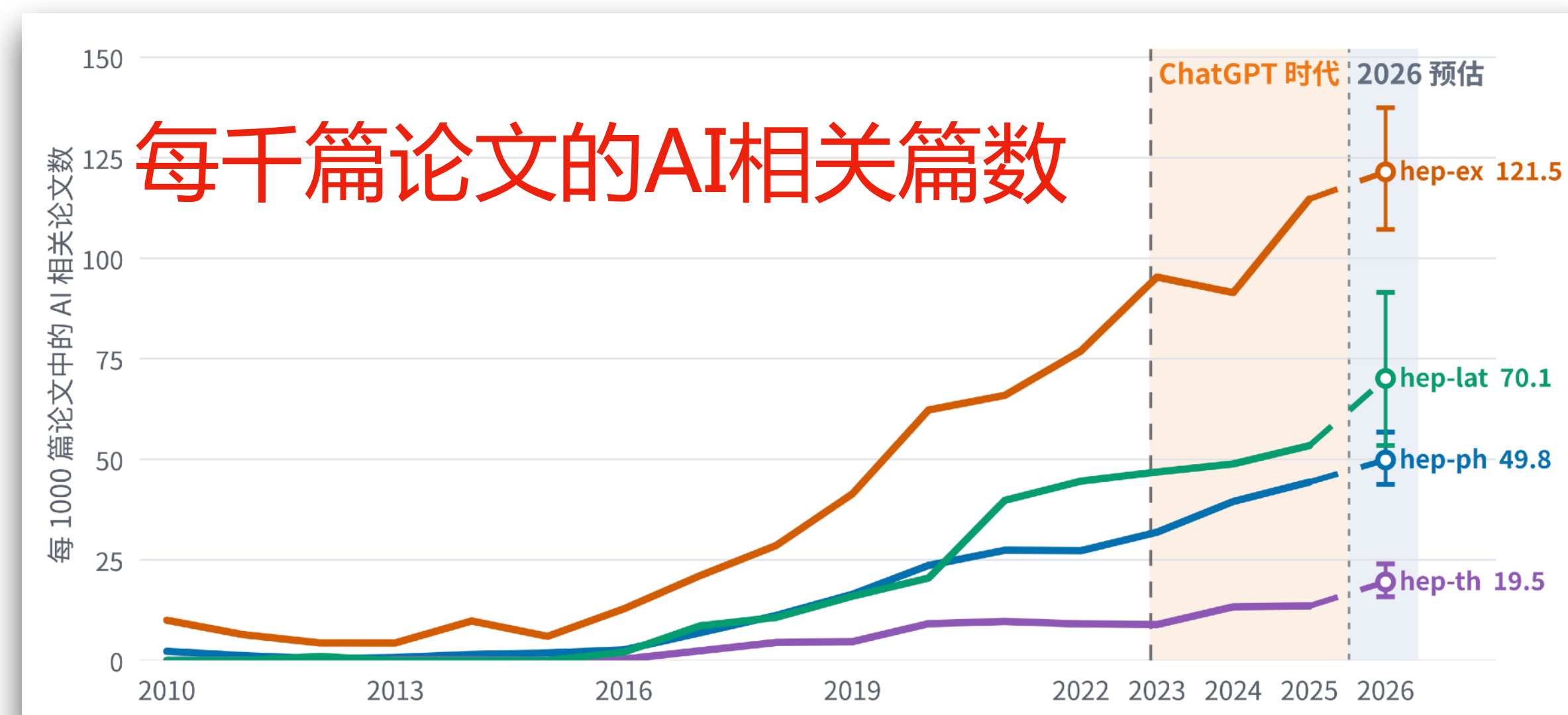
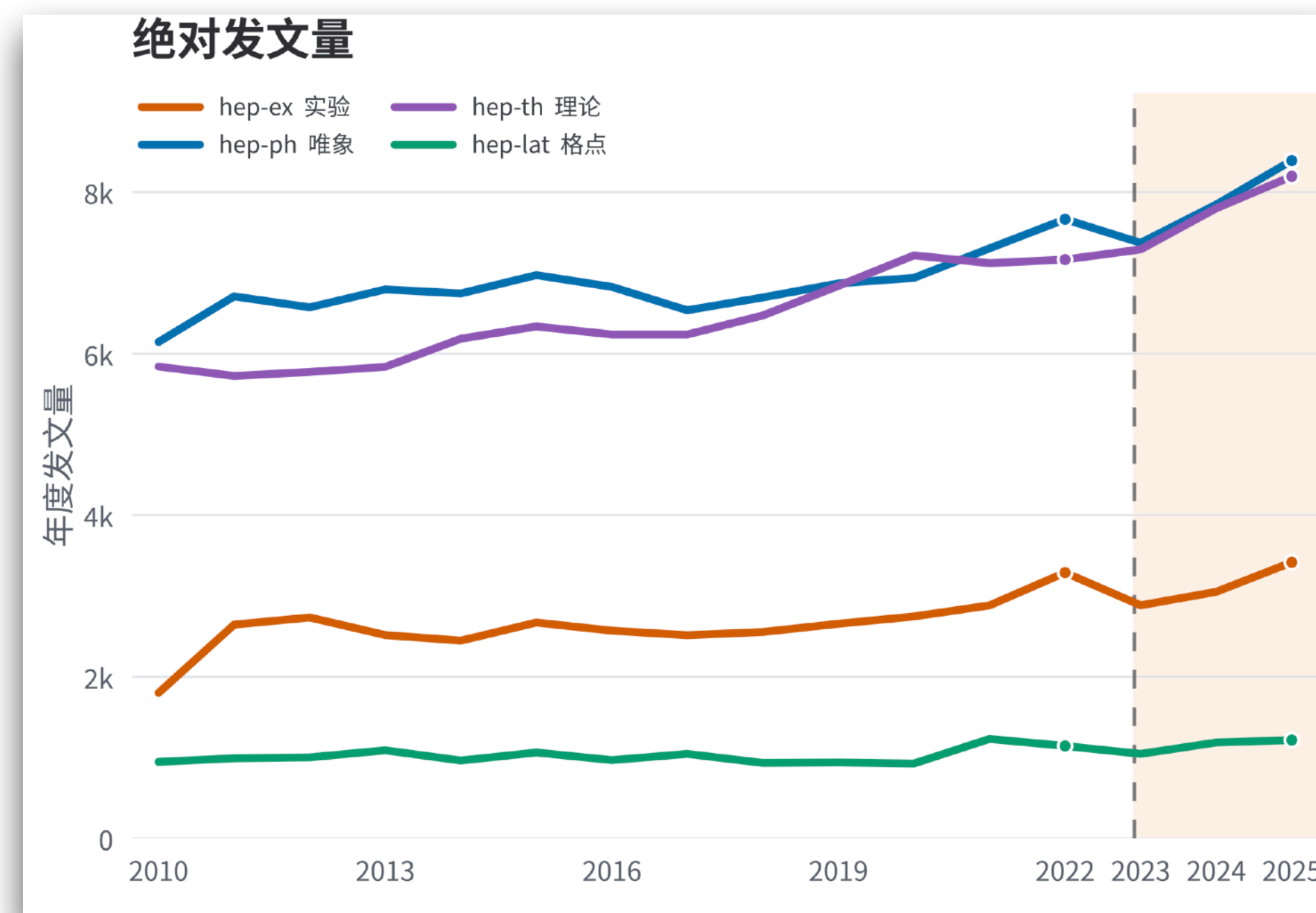


AI在高能物理中的应用

朱华星
北京大学

中国物理学会高能物理分会第十二届全国会员代表大会暨学术年会
2026年7月17日，华南师范大学



Outline

- AI for HEP
 - Specialized model
 - General model
- HEP for AI
 - Industrial impact
 - Physics of AI
 - Education impact
- Some final thoughts

Disclaimer: This is not a full review of all the progress made by our colleagues due to my limited capacity. Please consulted the excellent reviews below for more complete review including nuclear physics:

人工智能赋能粒子物理与核物理学： 从基础研究到重大应用

王贻芳^{1*} 何 源² 黄 耀³ 何万兵⁴ 焦 毅¹ 李聪乔⁵ 李 科¹ 刘北江¹ 马英起⁶
马余刚^{7*} 庞龙刚⁸ 齐法制¹ 谭思超⁹ 王春鹏¹⁰ 王 盟⁹ 徐小恒⁶ 徐 星² 赵振堂¹⁰
张英逊¹¹ 张正德¹ 赵红卫² 赵丽娜¹

国家科学技术思想库：人工智能赋能科学研究
National Think Tank in Science and Technology: AI Empowers Scientific Research

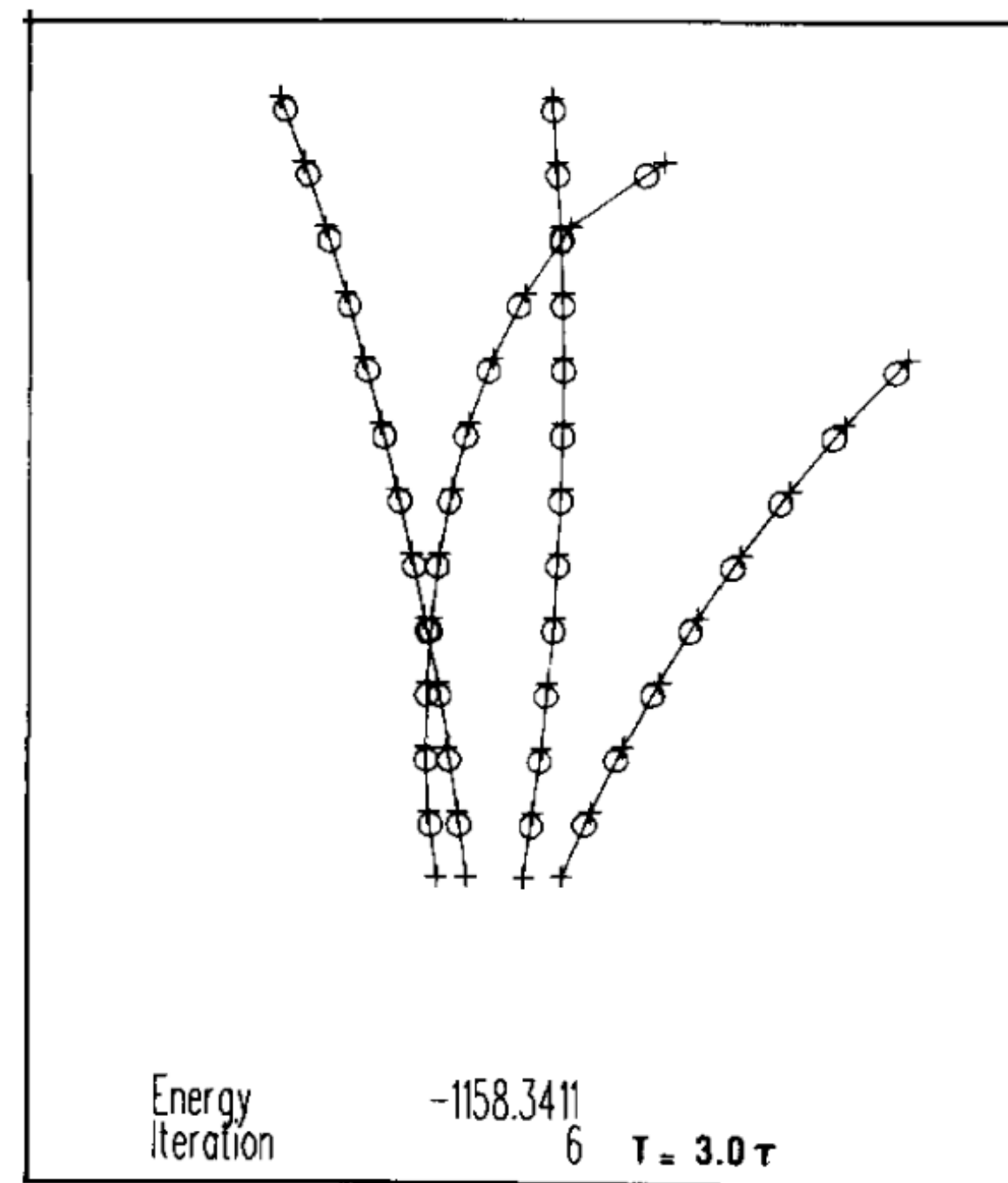
引用格式: 王贻芳, 何源, 黄耀, 等. 人工智能赋能粒子物理与核物理学: 从基础研究到重大应用. 中国科学院院刊, 2026, 41(6): 1089-1102, doi: 10.3724/j.issn.1000-3045.20260415007.
Wang Y F, He Y, Huang Y, et al. Artificial intelligence empowers particle physics and nuclear physics: From fundamental research to major applications. Bulletin of Chinese Academy of Sciences, 2026, 41(6): 1089-1102, doi: 10.3724/j.issn.1000-3045.20260415007. (in Chinese)

High-energy nuclear physics meets machine learning #2
Wan-Bing He (Fudan U., Shanghai and Fudan U.), Yu-Gang Ma (Fudan U., Shanghai and Fudan U.), Long-Gang Pang (Hua-Zhong Normal U., LQLP), Hui-Chao Song (Peking U.), Kai Zhou (Frankfurt U., FIAS) (Mar 12, 2023)
Published in: *Nucl.Sci.Tech.* 34 (2023) 6, 88 • e-Print: 2303.06752 [hep-ph]
pdf DOI cite claim reference search 99 citations

Colloquium: Machine learning in nuclear physics #2
Amber Boehnlein (Jefferson Lab), Markus Diefenthaler (Jefferson Lab), Nobuo Sato (Jefferson Lab), Malachi Schram (Jefferson Lab), Veronique Ziegler (Jefferson Lab) et al. (Dec 4, 2021)
Published in: *Rev.Mod.Phys.* 94 (2022) 3, 031003 • e-Print: 2112.02309 [nucl-th]
pdf links DOI cite claim reference search 240 citations

AI4HEP: Specialized model

Fundamental physics pursues nature's deepest questions, often pushing the boundaries of existing technology, or creating entirely new ones.



Tracking finding

B. Denby, 1987

SCHOONSHIP

A CDC 6600 program for symbolic evaluation of algebraic expressions.

M. Veltman

Abstract

A high speed program has been constructed capable of evaluating expressions of the form:

$$(A1 + B1 + \dots) * (A2 + B2 + \dots) * \dots + \dots$$

Symbolic manipulation

Schoonship, Veltman, 1963

Early history of AI in HEP

International Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High-energy and Nuclear Physics (AIHENP 1990)

19-24 March 1990. Lyon, France (C90-03-19)

Neural networks and high-energy physics

Carsten Peterson (Lund U.) (May, 1990)

Contribution to: AIHENP 1990, 465-480

Pattern recognition

Symbolic manipulations: Programs and physics applications

Andrey Kataev (Moscow, INR) (Mar, 1990)

Contribution to: AIHENP 1990, 647-656

Symbolic calc.

An integration of physics packages data modeling and human interface

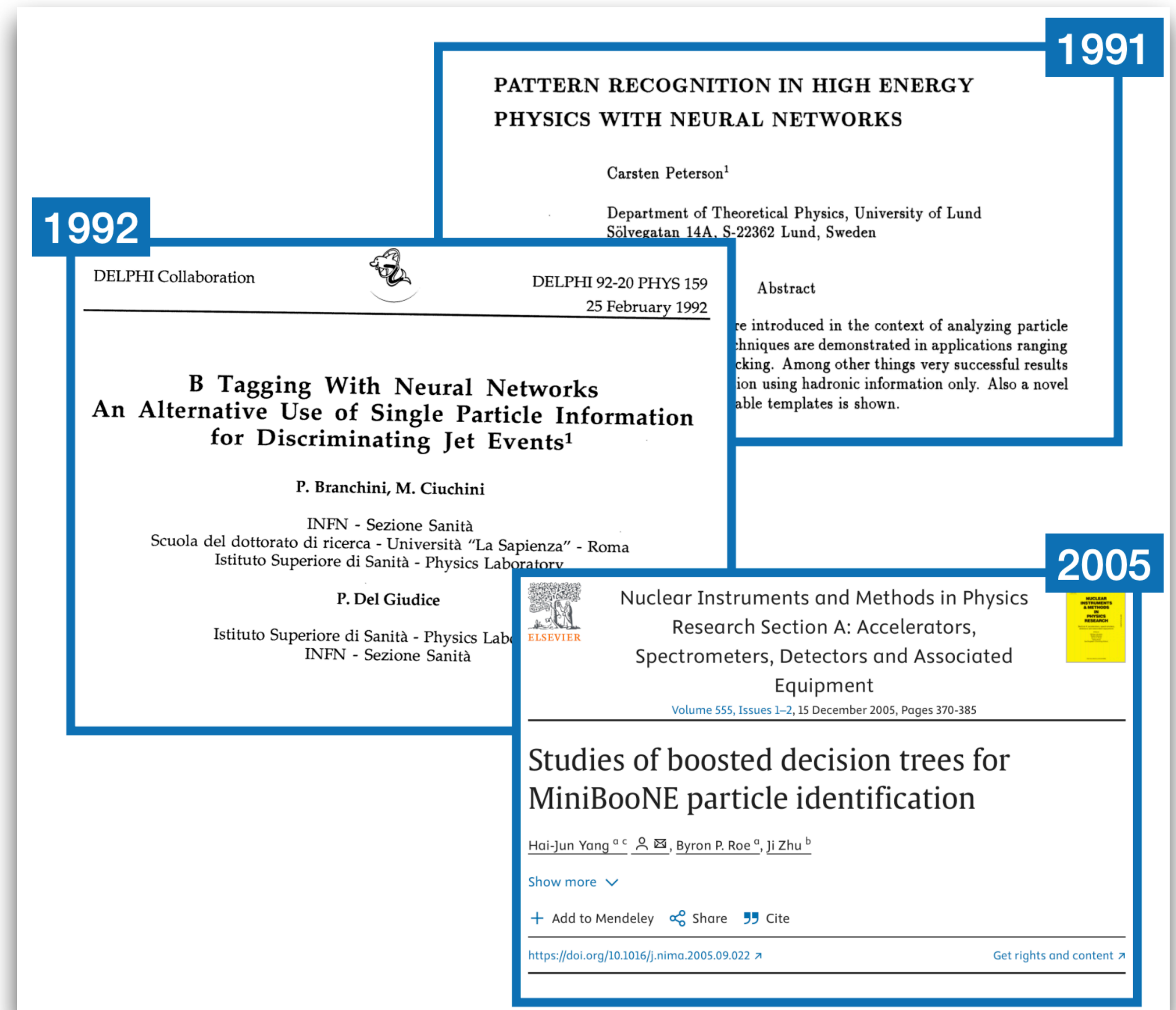
#13

J. Alberty (CERN), F. Anselmo (CERN), E. Barberio (CERN), O. Di Rosa (CERN), B. van Eijk (CERN) et al. (Mar, 1990)

Contribution to: AIHENP 1990, 117-123

Agentic system

<https://inspirehep.net/conferences/967194?ui-citation-summary=true>



Slide from M. Pierini, AI@LHC, Peking U., 2026

Jet tagging after 30 years

PHYSICAL REVIEW D **101**, 056019 (2020)

Jet tagging via particle clouds

Huilin Qu^{*}

Department of Physics, University of California, Santa Barbara, California 93106, USA

Loukas Gouskos[†]

CERN, CH-1211 Geneva 23, Switzerland

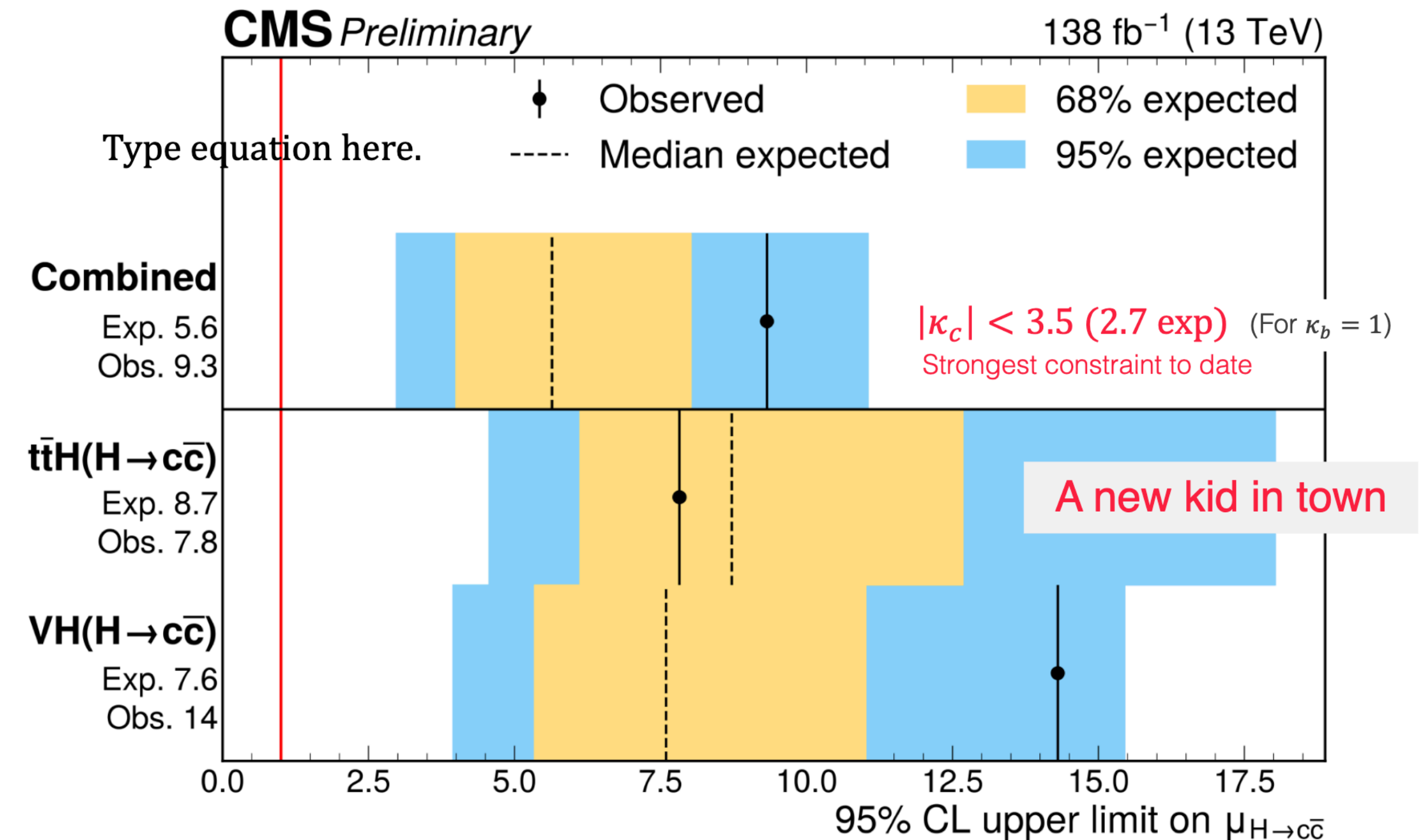
Inspirehep citation: 702 times

IUPAP EARLY CAREER SCIENTIST AWARDS
2026 IN PARTICLES AND FIELDS

For his crucial contributions and leadership in the development of advanced machine learning algorithms for identifying decays of boosted heavy particles and their application in the search for $H \rightarrow c\bar{c}$ decays at the Large Hadron Collider.

While deep neural network is a blackbox, physics insight can boost performance

- Symmetries: permutation, Lorentz
- Started a new era for jet tagging



Andreas Hoecker, EPS-HEP 2025 —Conference Highlights

Particle Transformer for Jet Tagging

Article | [Open access](#) | Published: 14 January 2026

Transforming jet flavour tagging at ATLAS

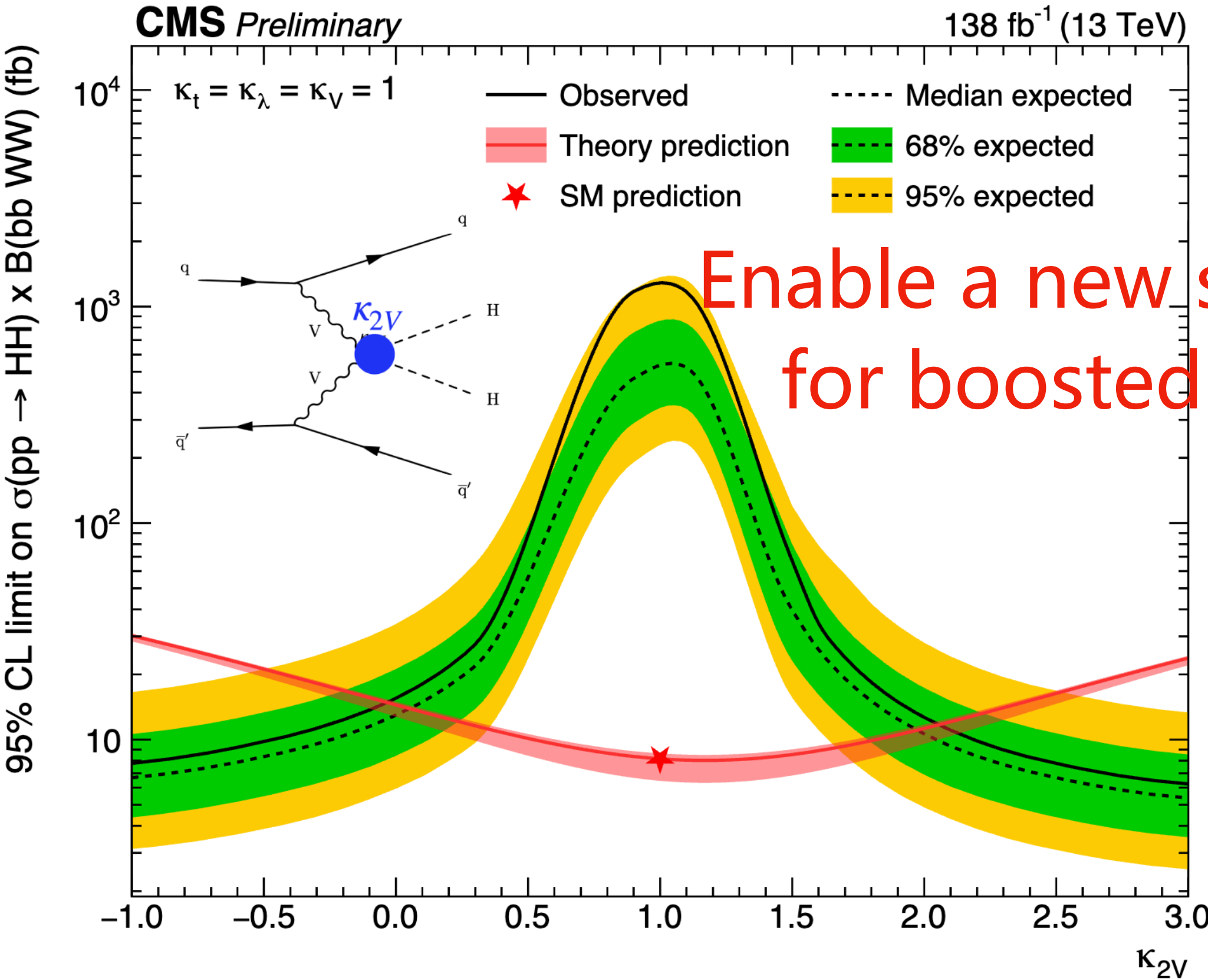
[The ATLAS Collaboration](#)

[Nature Communications](#) **17**, Article number: 541 (2026) | [Cite this article](#)

ICML 2022

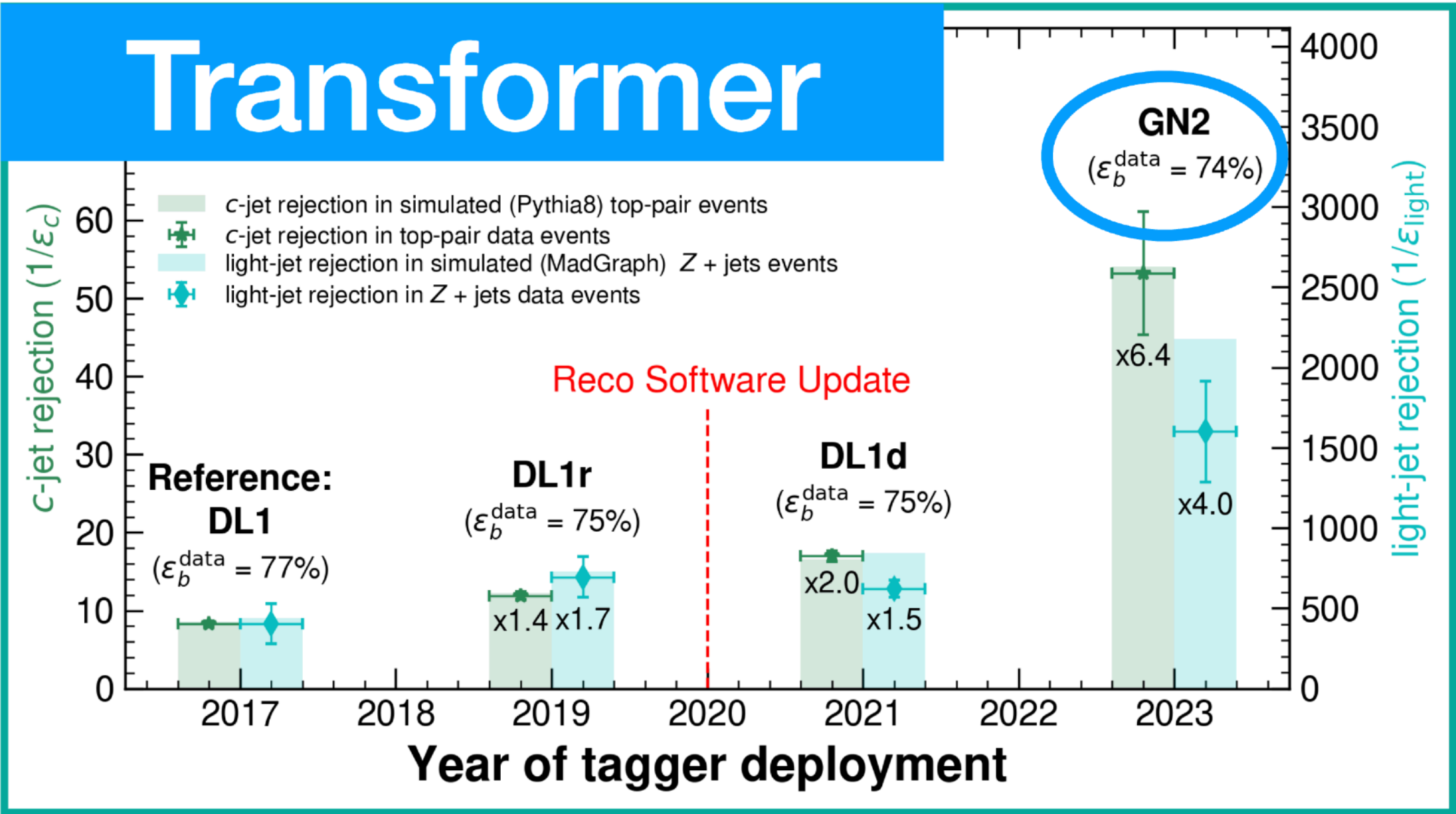
Huilin Qu¹ Congqiao Li² Sitian Qian²

CMS thesis award 2025



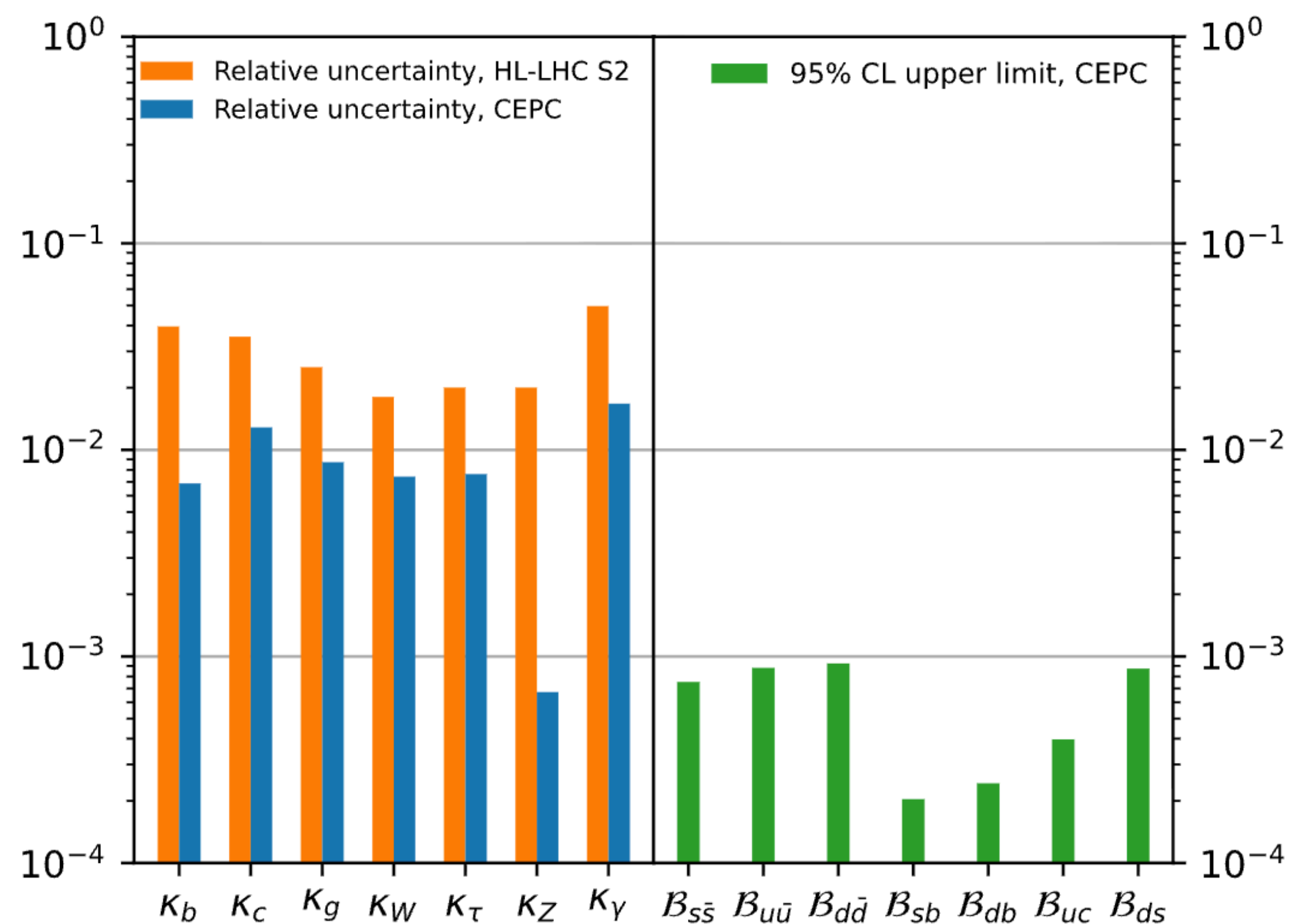
JAVIER DUARTE, ICHEP 2024

Next-generation flavor-tagging algorithm based on transformer



Bingxuan Liu, this conference

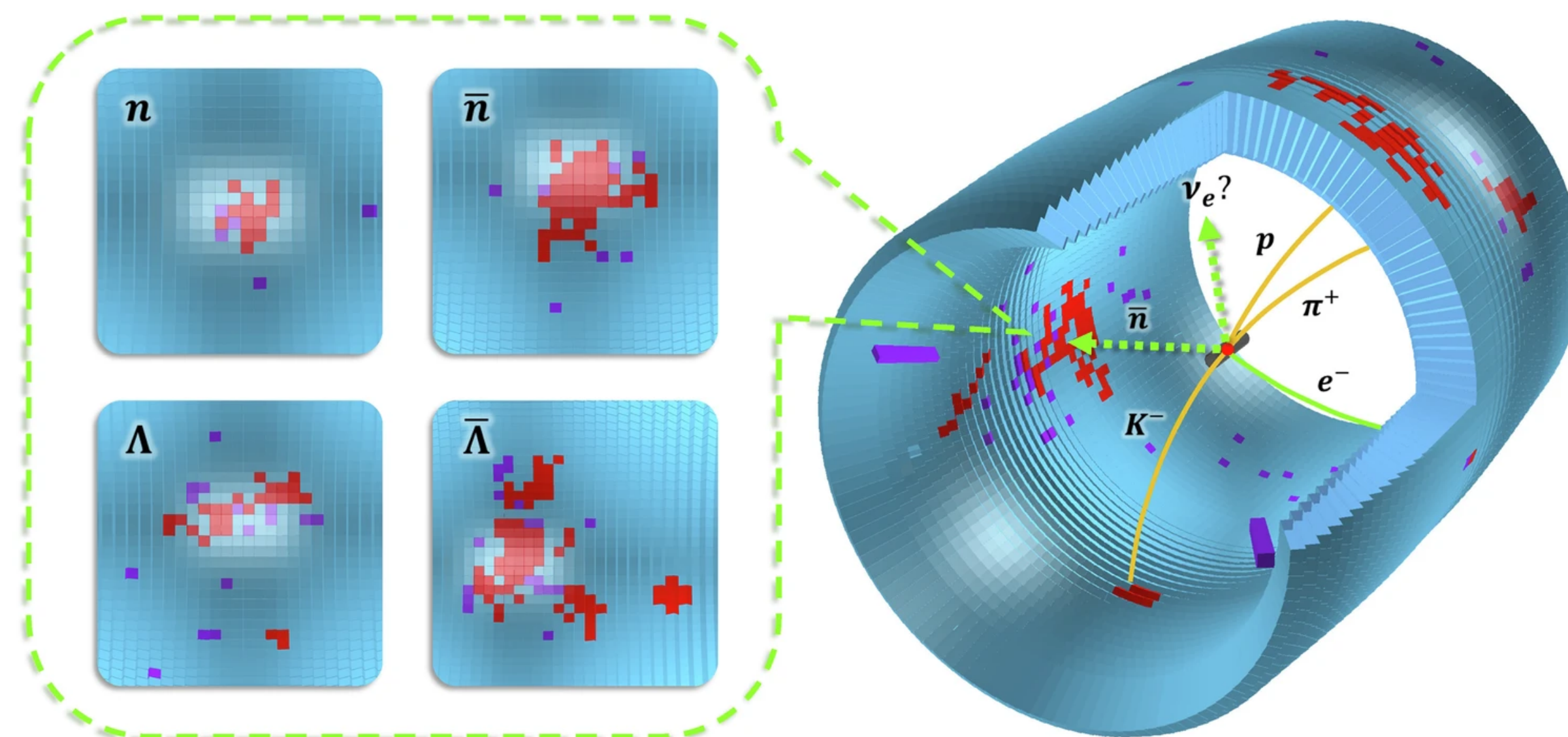
Improving effective Luminosity by 1 order of magnitude for jet origin identification



H. Liang et al., Phys.Rev.Lett. 132 (2024) 22, 221802

$$\Lambda_c^+ \rightarrow \mathbf{n} e^+ \nu_e \quad \text{V.S.} \quad \Lambda_c^+ \rightarrow \mathbf{\Lambda} (n \pi^0) e^+ \nu_e$$

Organize EMC showers as a point cloud
Achieve first observation of $\Lambda_c^+ \rightarrow n e^+ \nu_e$



Article | [Open access](#) | Published: 15 January 2025

Observation of a rare beta decay of the charmed baryon with a Graph Neural Network

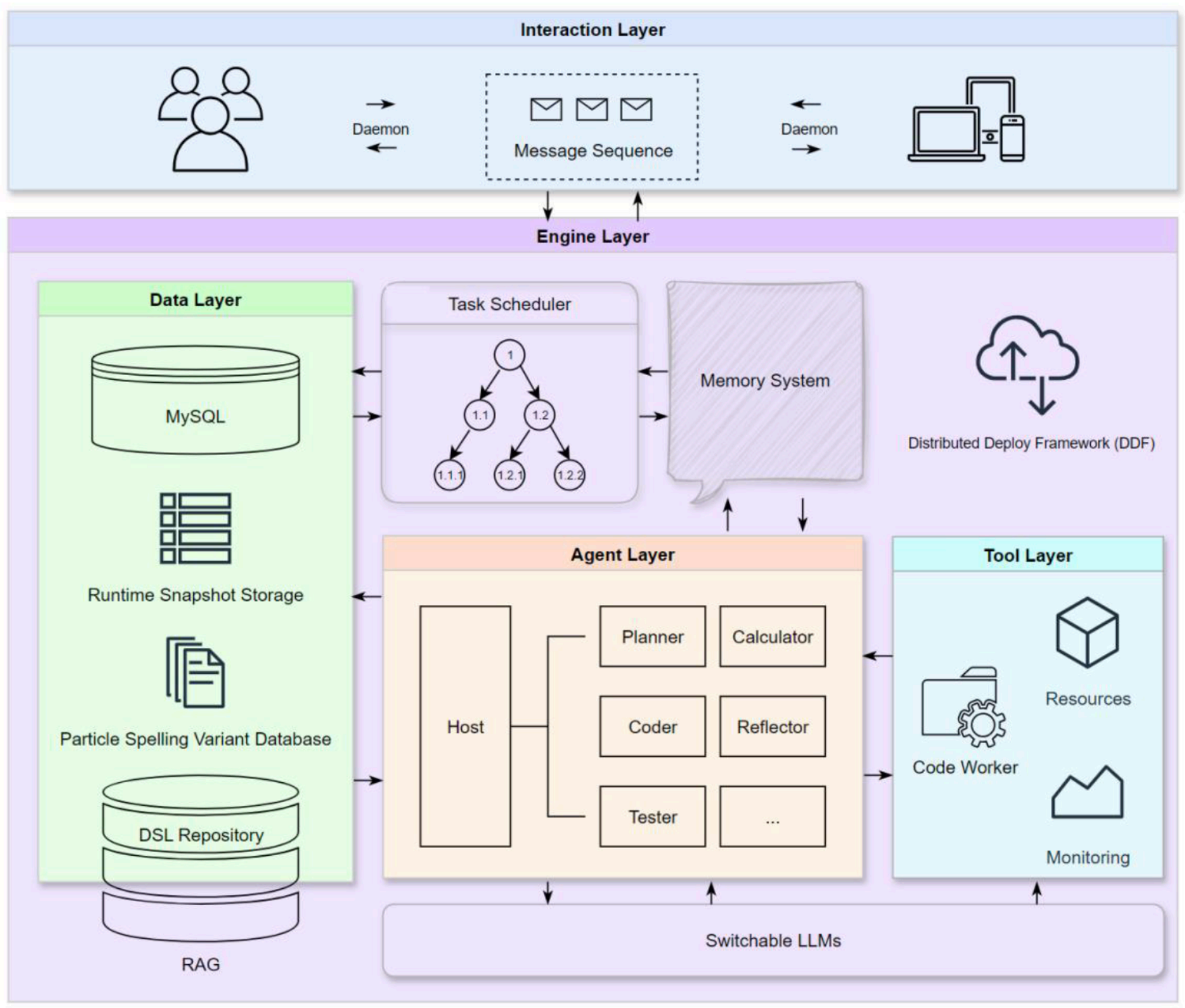
[The BESIII Collaboration](#)

[Nature Communications](#) **16**, Article number: 681 (2025) | [Cite this article](#)

AI4HEP: Agentic AI

General model as AI scientist

An agent-based system advancing HEP experiment
Reproduce the Zc(3900) discovery



M. He et al., 2604.22541

Aether: 专为科研人员定制的 AI助手

强大的编程和研究能力，科研全流程护航，兼容各种模型与提供商，支持微信、QQ、飞书一键接入，开箱即用，让科研触手可及。

两圈四点doublebox主积分在s=1,t=10展开计算

请帮我计算两圈四点doublebox拓扑的主积分在运动学量s=1, t=10下的结果，展开到维数正规化因子epsilon的正1阶。

已探索 4 次读取, 1 次搜索

Shell Checks AMFlow environment status

已探索 1 次搜索

我先检查了本机环境: Mathematica 和 AMFlow 都能找到, 但当前 Blade 的路径配置需要先修一下, 否则没法真正跑两圈积分。

我这里有一个关键歧义需要你确认, 我建议默认这样做:

- 按“无质量平面四点两圈 double box”标准族来算;
- 外腿满足 $p_1^2=p_2^2=p_3^2=p_4^2=0$, $(p_1+p_2)^2=s=1$, $(p_2+p_3)^2=t=10$;
- 结果采用 AMFlow 的归一化: 每圈 $1/(i\pi^{D/2})$;
- 并且我自动修复本机 Blade 配置后再计算。

如果你同意这个默认设定, 我就直接继续算到 ϵ^1 ; 如果不是这个族/归一化, 请直接给我:

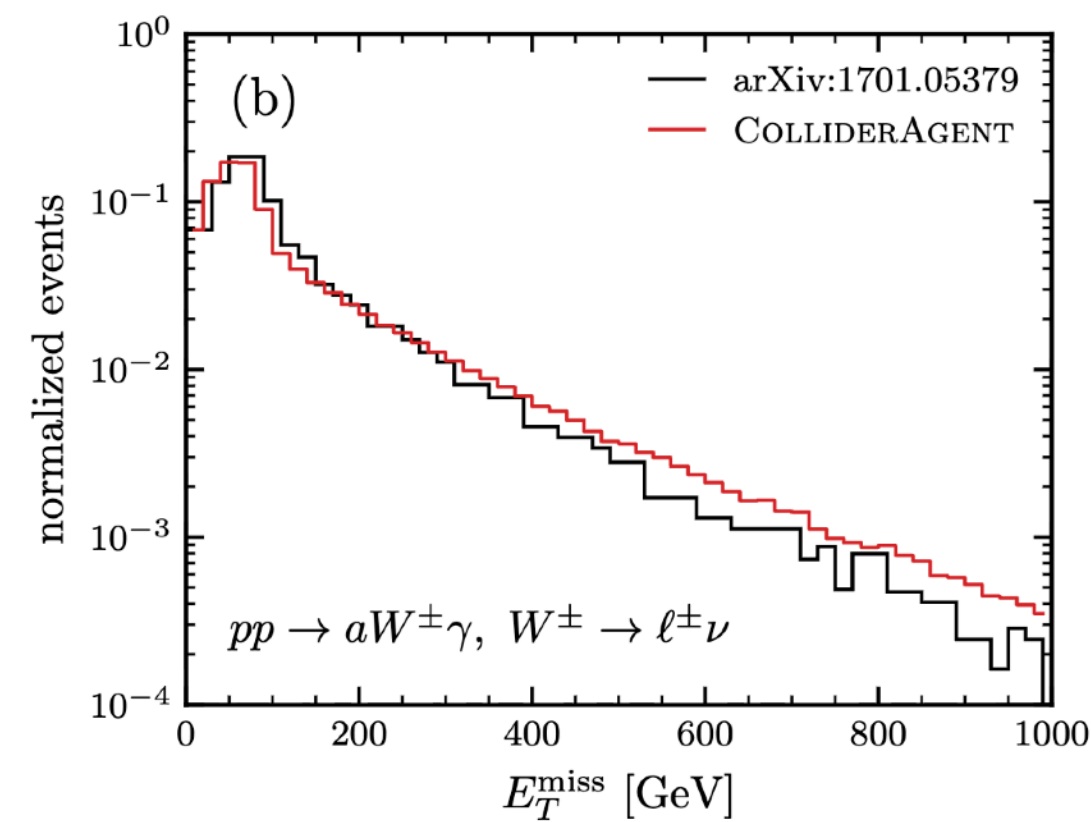
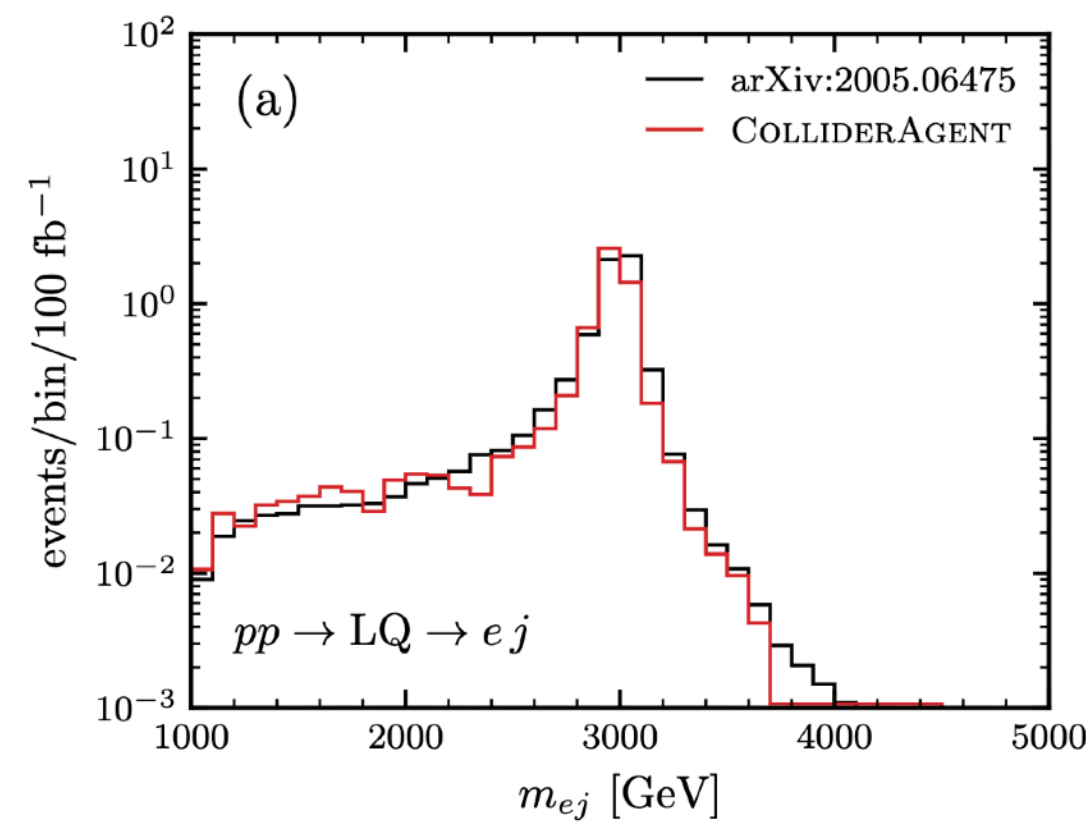
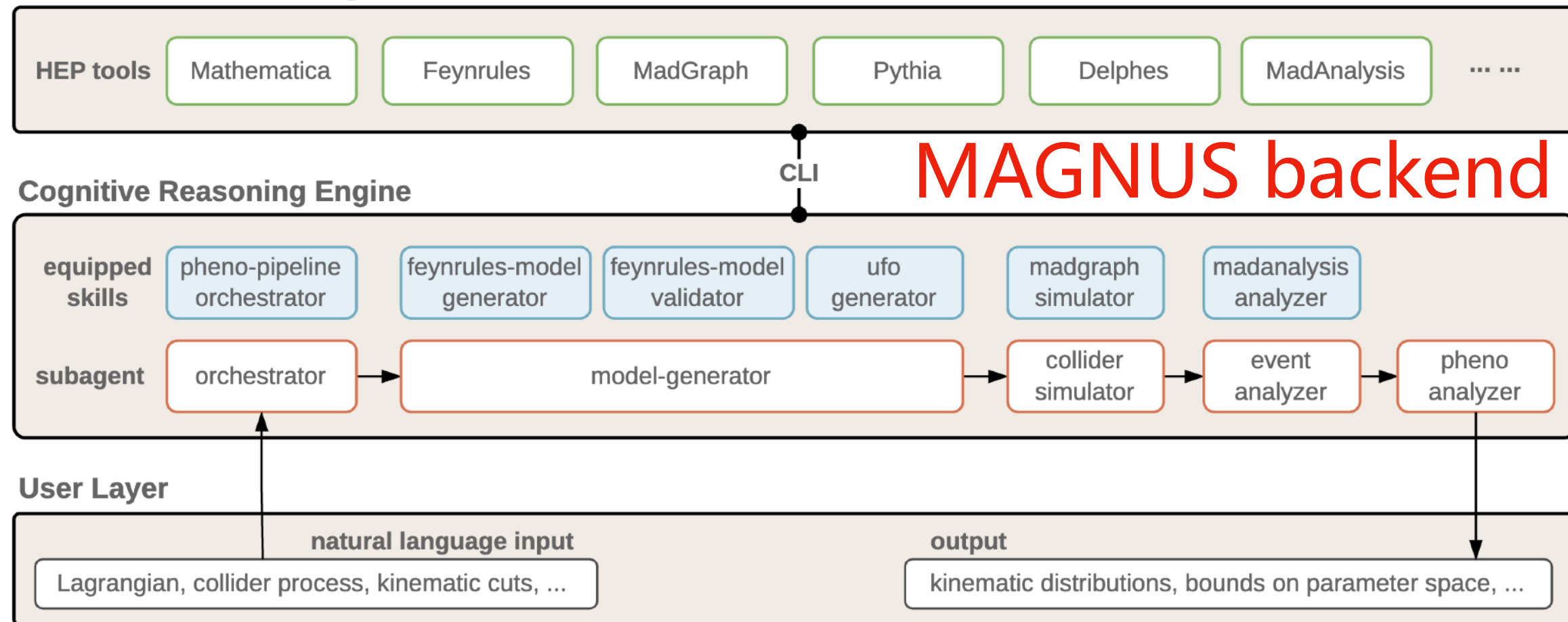
- 传播子列表;
- 你说的“主积分”具体是哪一组;
- 归一化约定。

<https://aether.aipphys.cn/>

Specialized agentic workflow

A framework incorporating major collider physics ecosystem

Execution Backend: Magnus

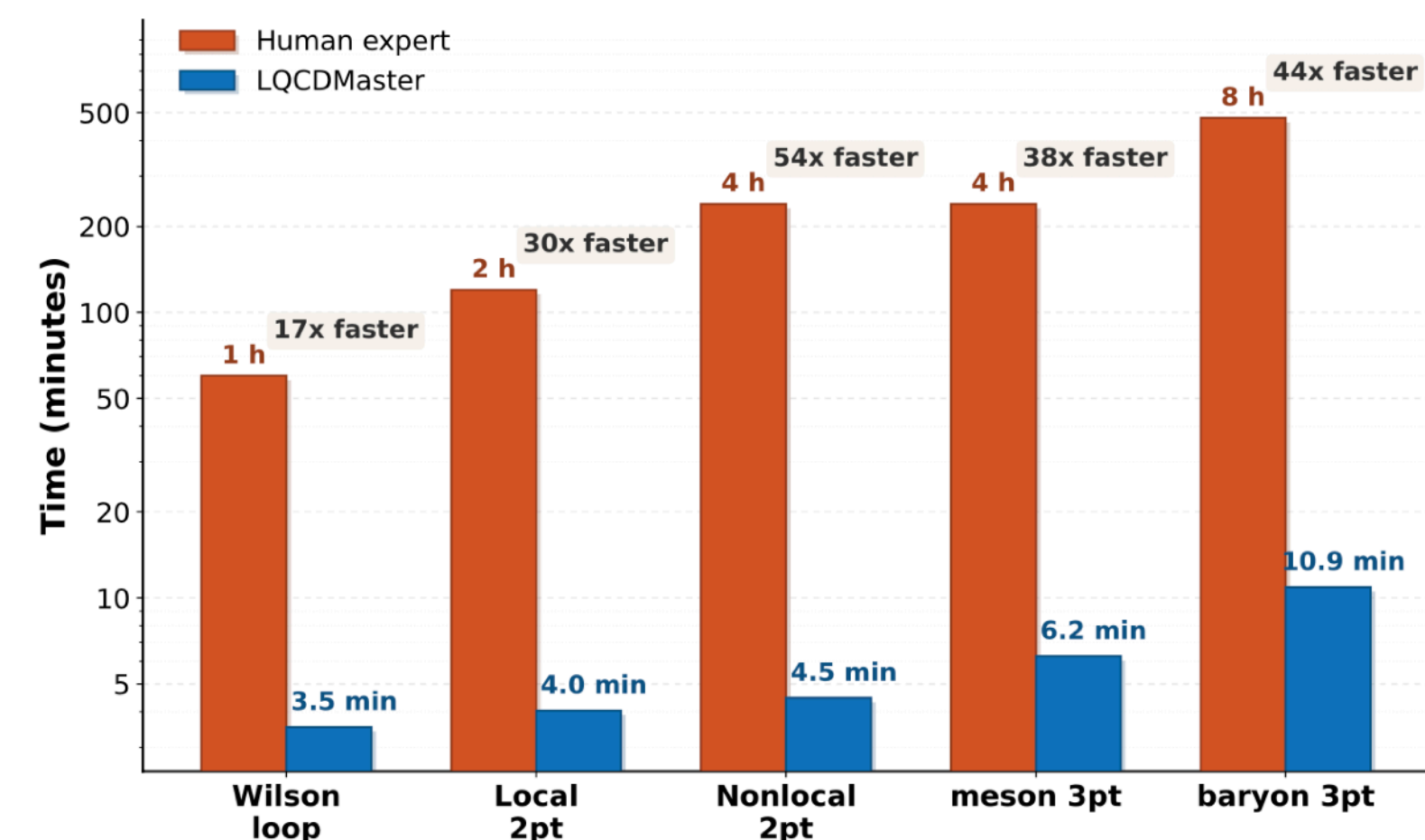
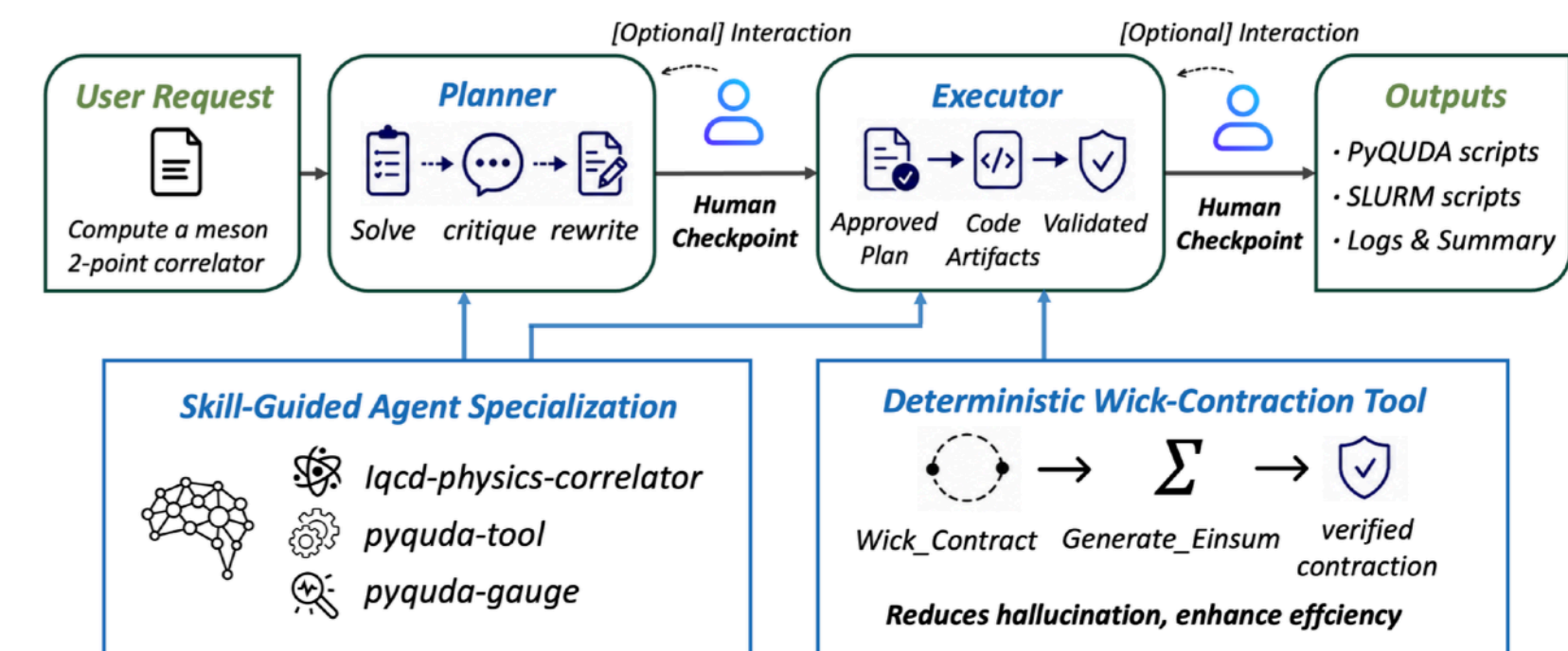


S. Qiu et al., 2603.14553

A lattice computation of light-cone distribution amplitudes with diagonal Wilson-line, a quantity accessible with standard methods **but never before computed**

LQCDMaster: Tool-Augmented, Skill-Guided LQCD Code Generation

From natural-language lattice-QCD research request to PyQUDA computing code

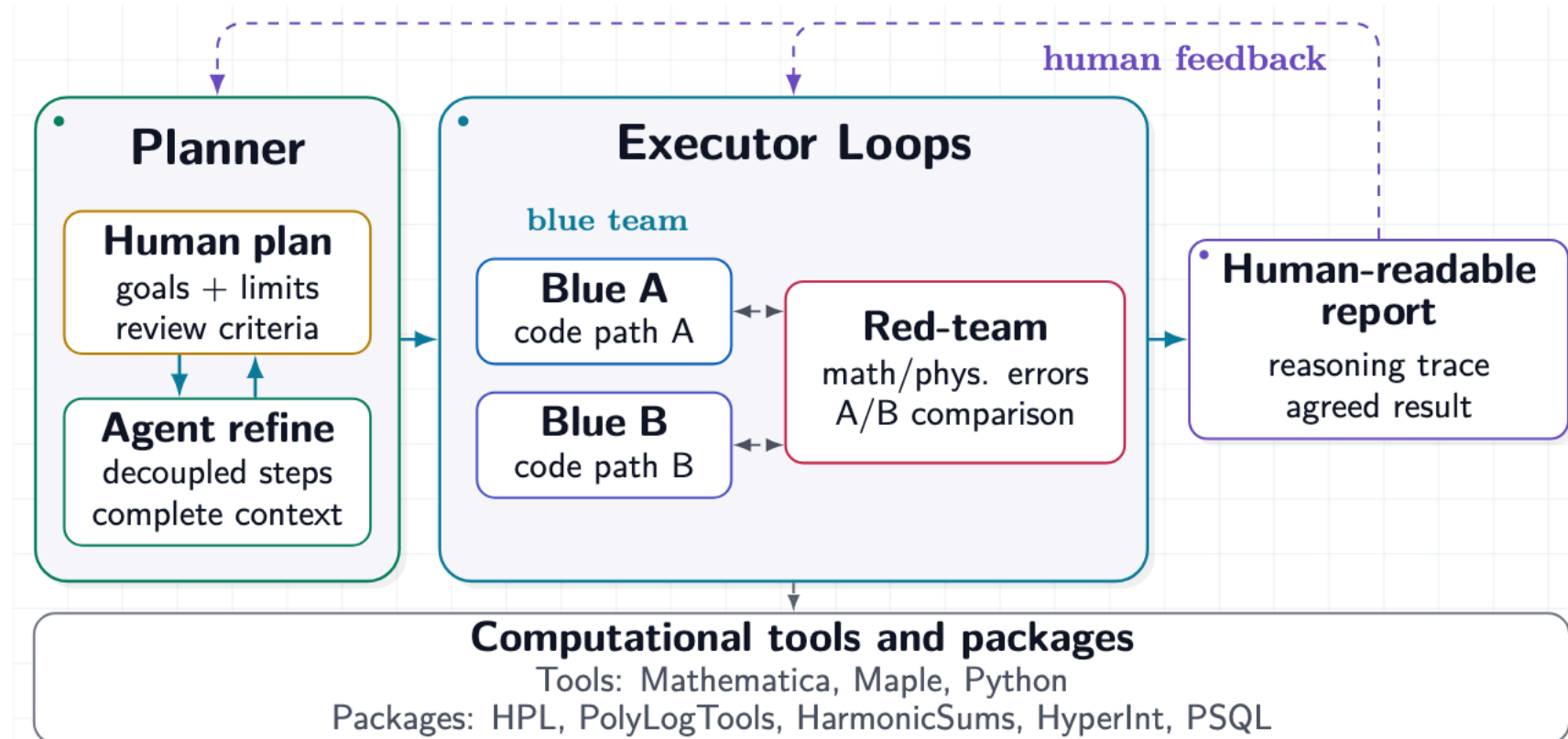


H.F. Gao, et al., 2607.15001

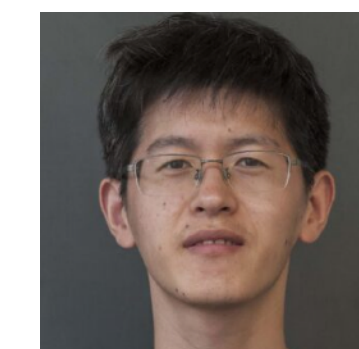
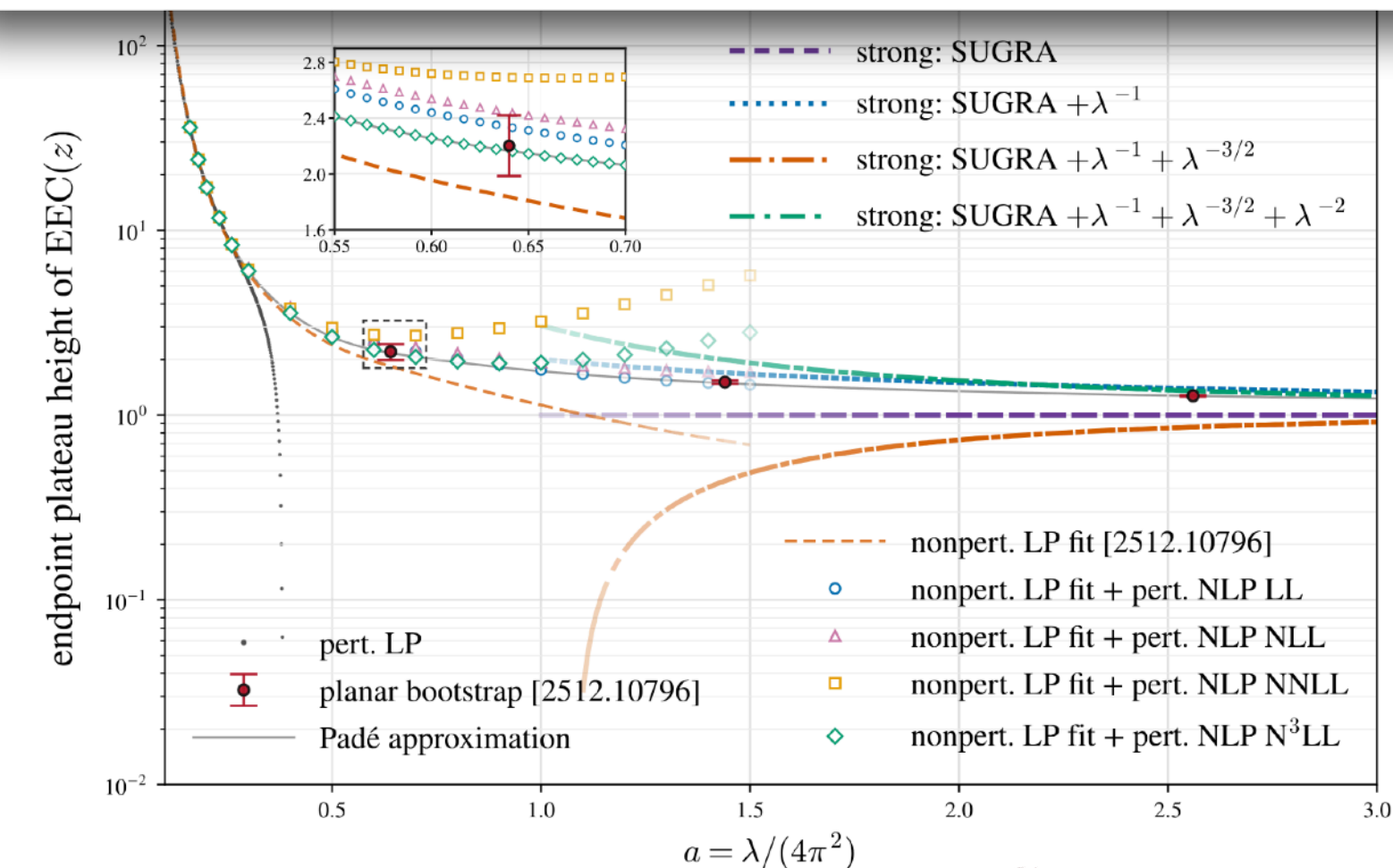
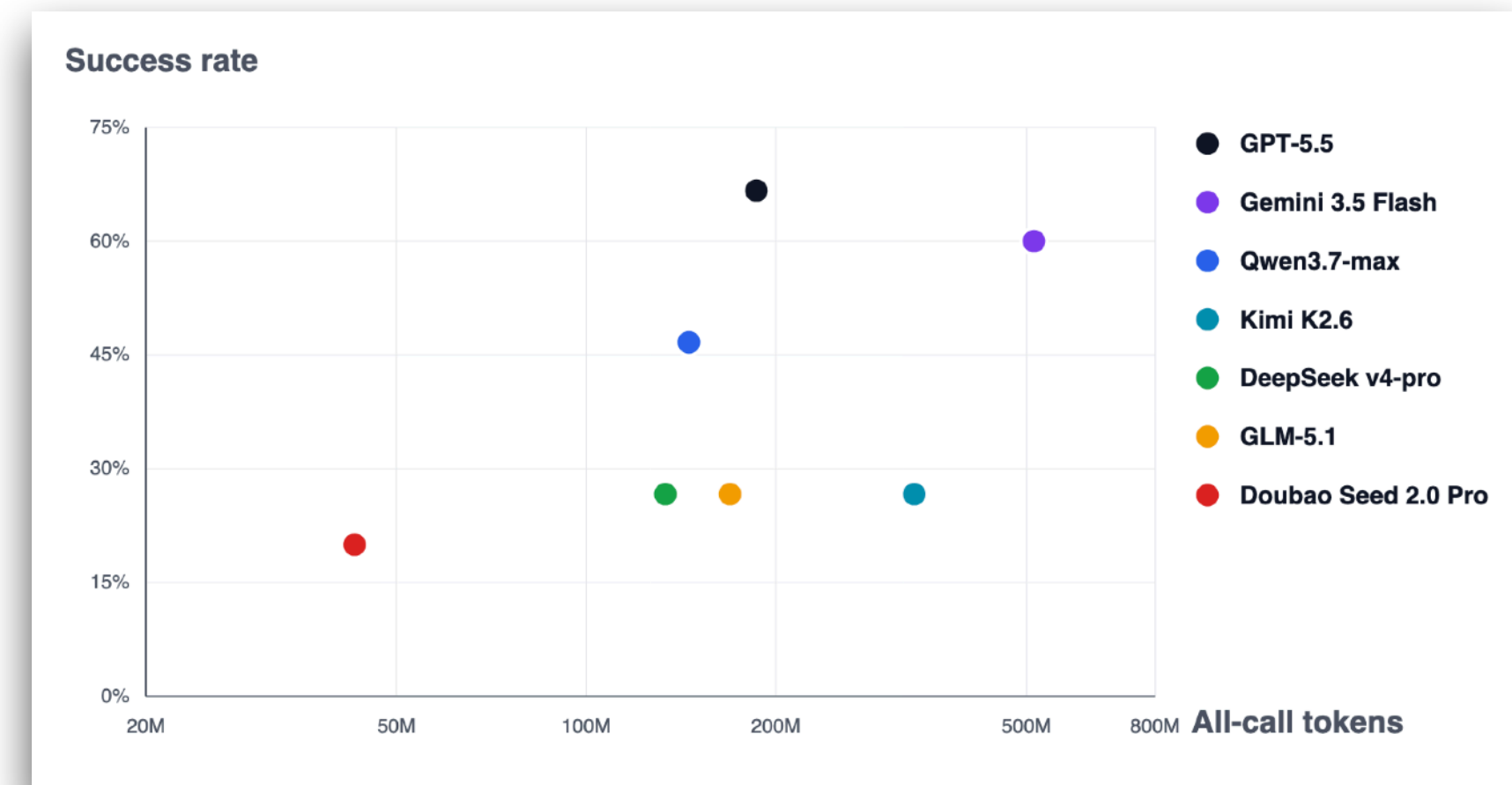
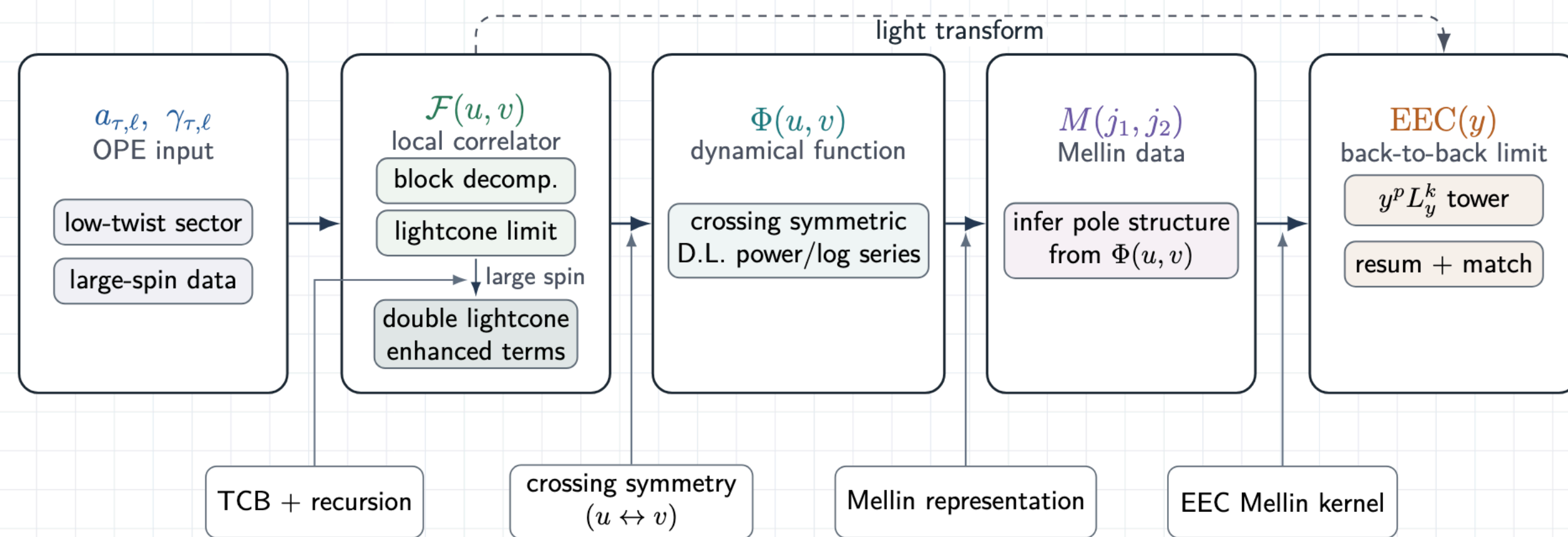
An agentic calculation of EEC@4 loops

A challenging problem for a new student

Requires substantial time even for expert,
main reason why this has not been done

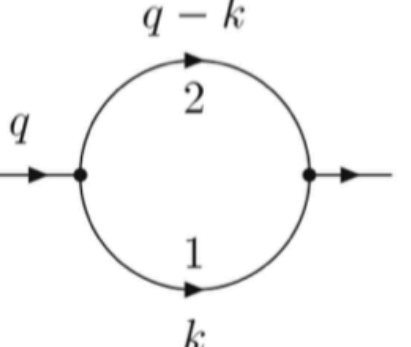


Back-to-Back Subleading-Power Resummation Roadmap



Hao Chen, Moult, Bo Wang, HXZ, in preparation

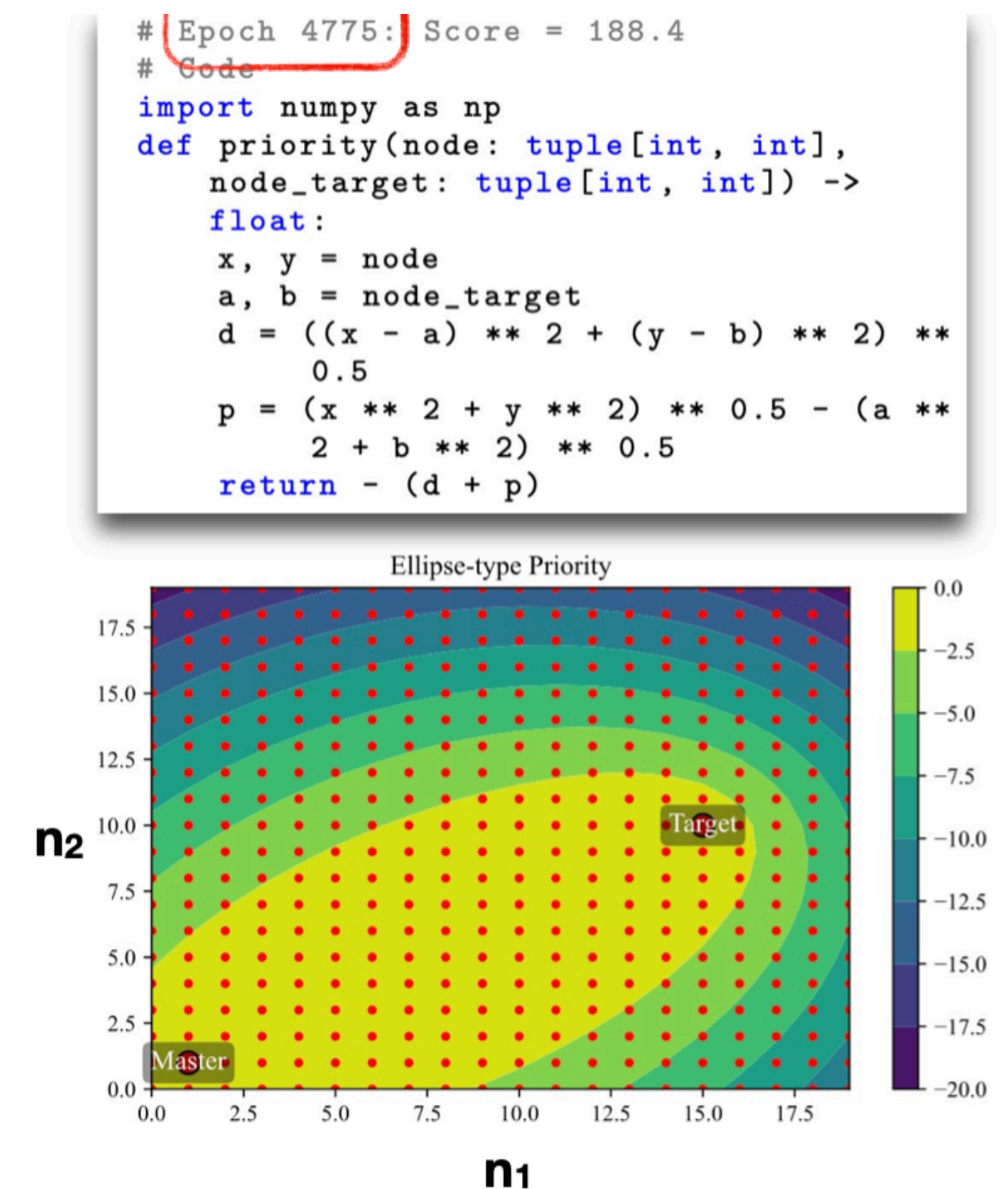
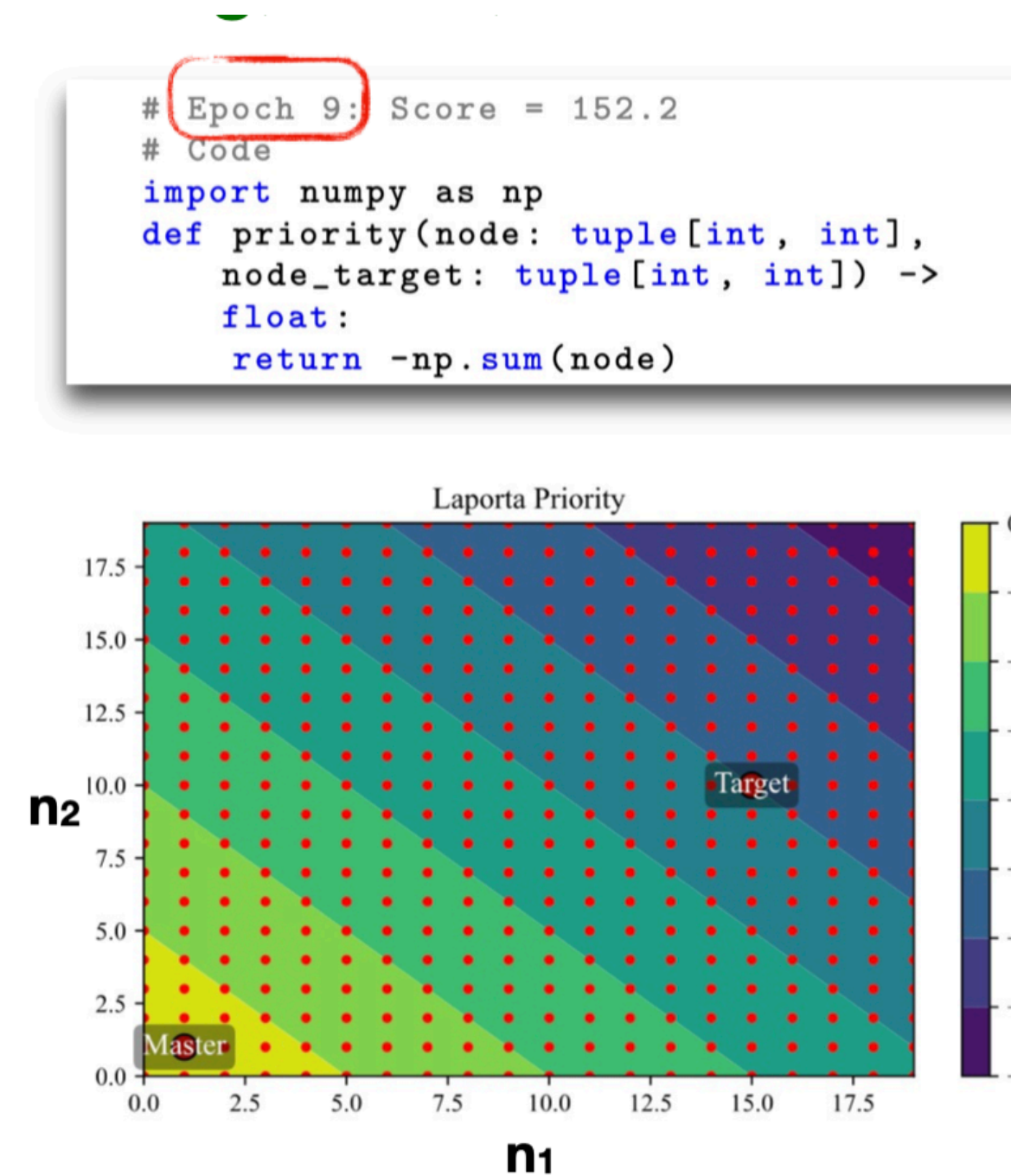
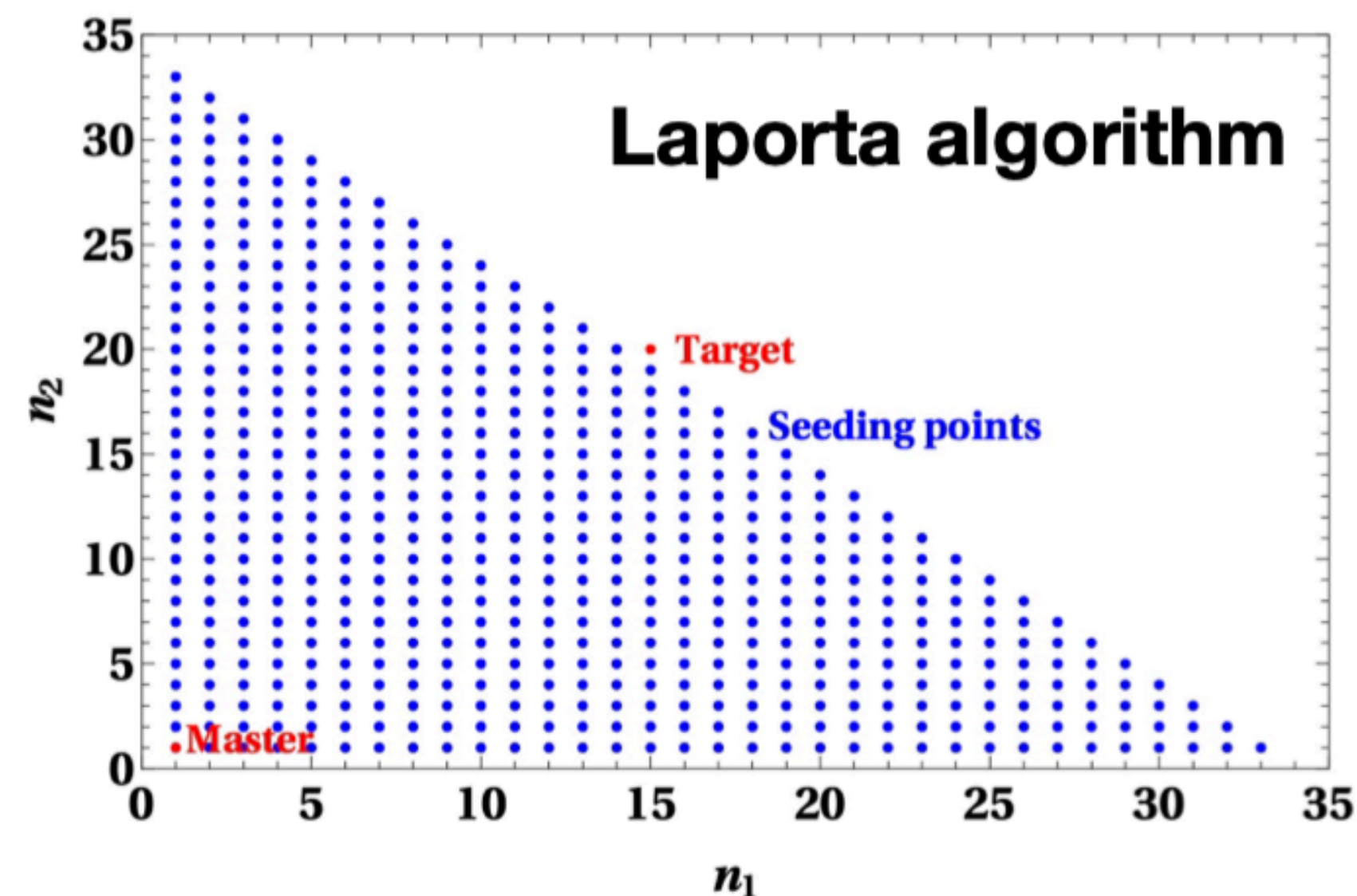
Specialized used of general model



$$I(n_1, n_2) = \int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2)^{n_1} ((q-k)^2)^{n_2}}$$

$$0 = (d - 2n_1 - n_2) I(n_1, n_2) - n_2 I(n_1 - 1, n_2 + 1) + n_2 I(n_1, n_2 + 1)$$

$$0 = n_1 I(n_1 + 1, n_2 - 1) - n_1 I(n_1 + 1, n_2) - n_2 I(n_1 - 1, n_2 + 1) + n_2 I(n_1, n_2 + 1) + (n_2 - n_1) I(n_1, n_2)$$



- Rediscover classic Laporta algorithm in 9 Epoch
- An Ellipse-type algorithm achieve 1000 times improvement

LLM as a mutation operator in evolution algorithm



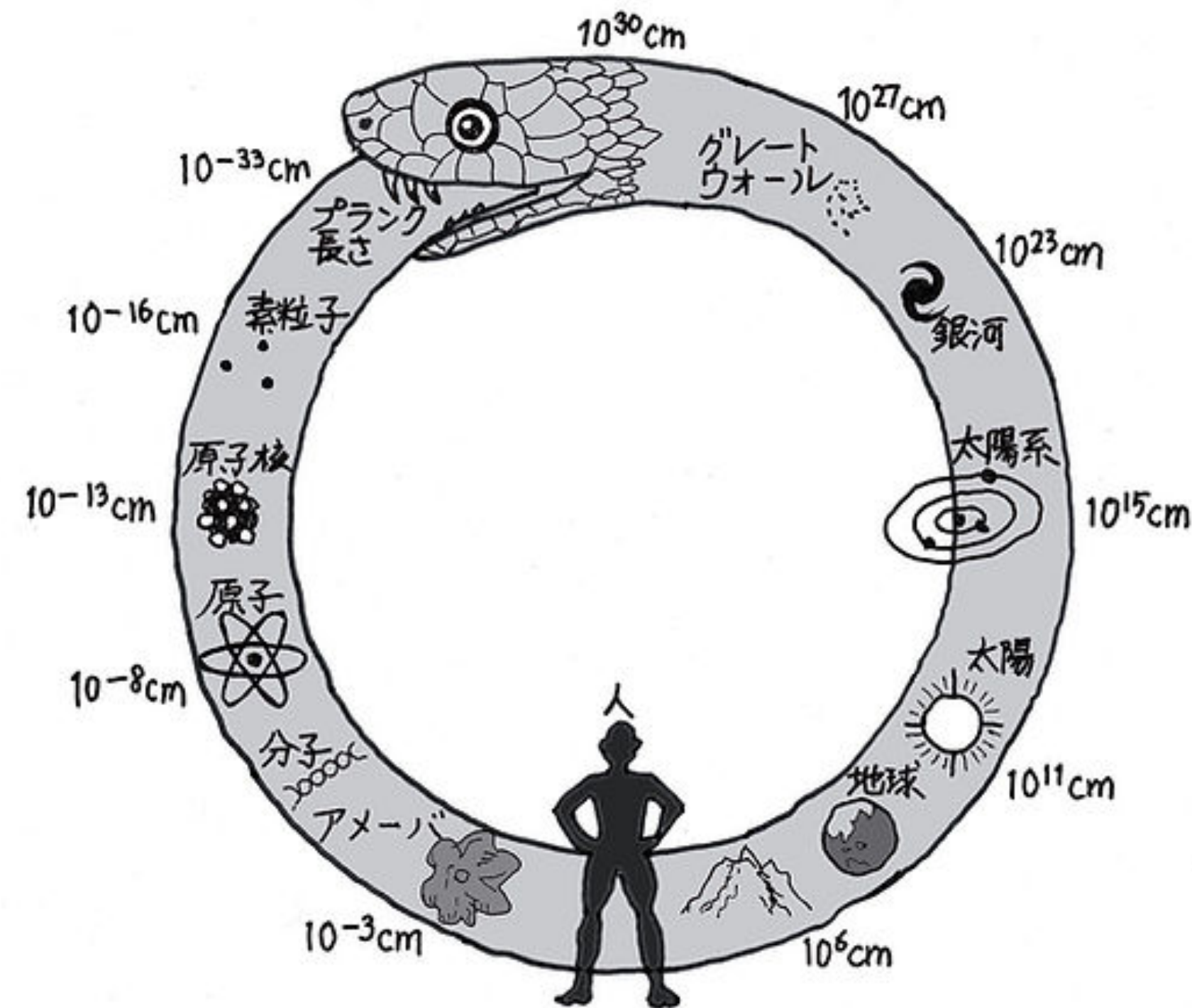
Zhuo-Yang Song, T.Z. Yang, Q.H. Cao, M.X. Luo, HXZ, JHEP 06 (2026) 225

HEP for AI

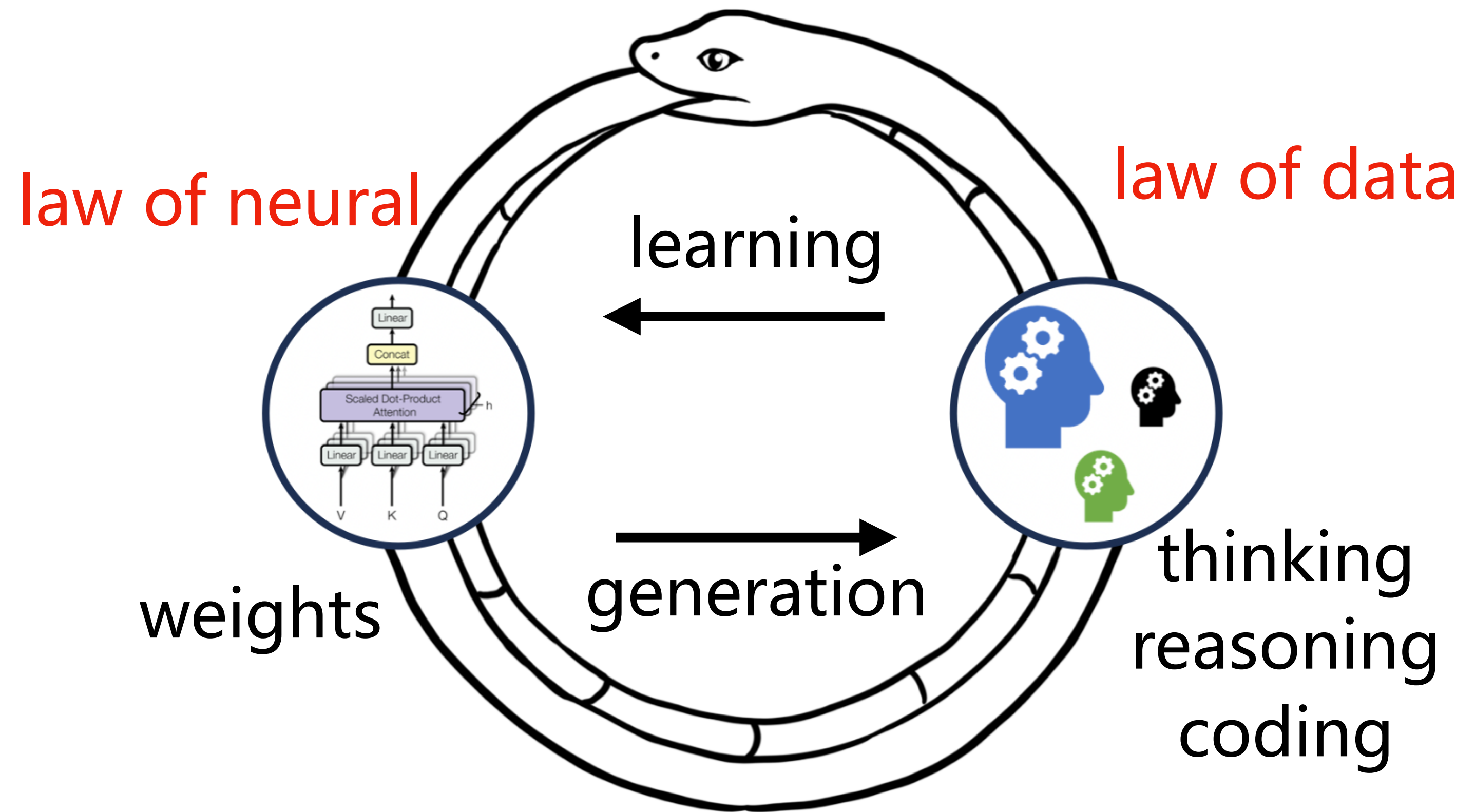
Unique position of High Energy Physics

- Look backward
 - The birth of WWW
 - High-energy experiments were among the first to adopt neural networks in scientific research
- High Energy Physics in the era of AI
 - Unique model of international collaboration through big experiment
 - International evaluation standard
 - Anticipate the challenges 10 years ahead

A particle physics' view



Universe exhibits diverse phenomena at different scales

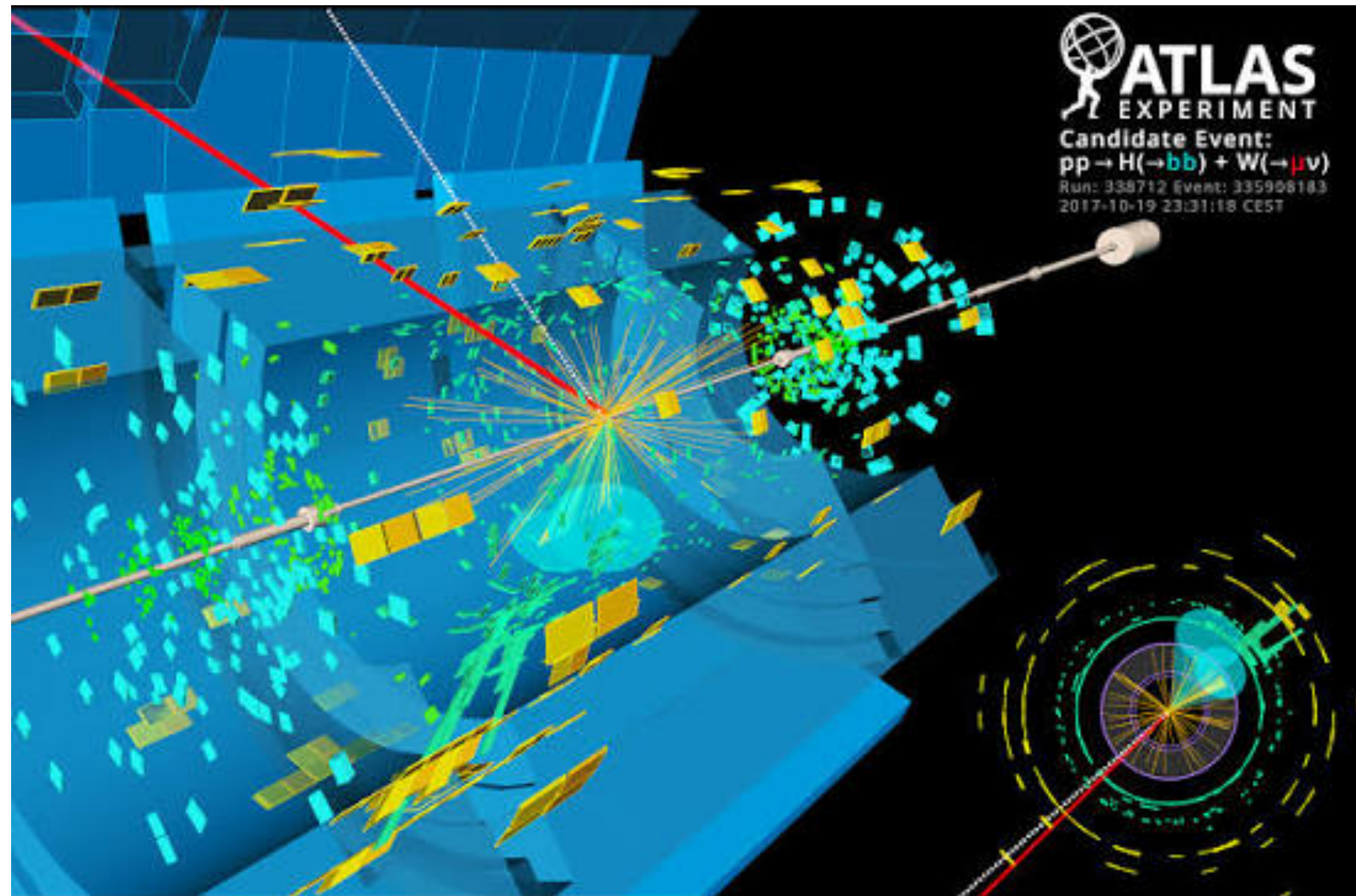


The two end also exhibits very different laws

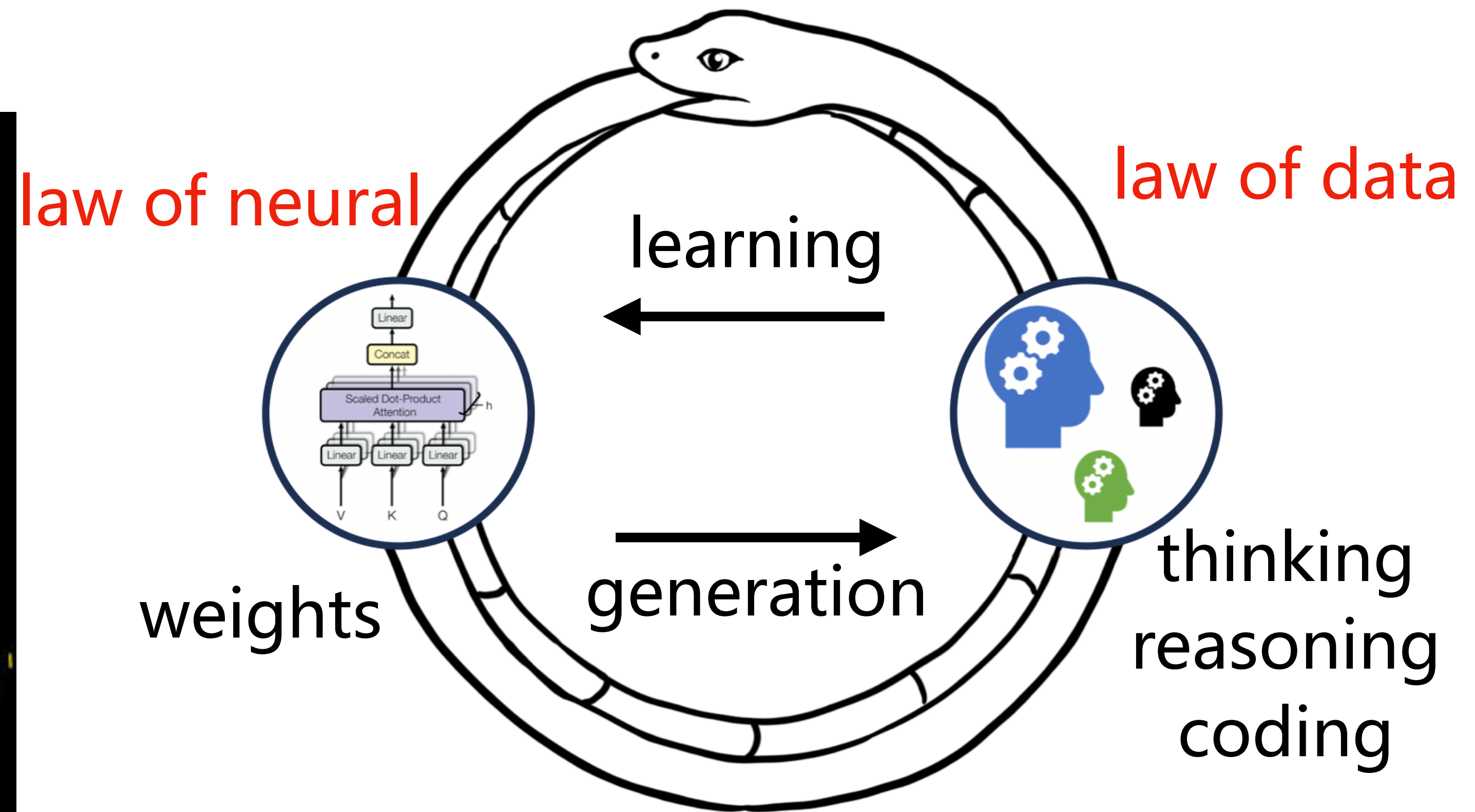
The laws of microscopic dynamics and macroscopic dynamics

HEP4AI: Physics of AI

A collider physics' view

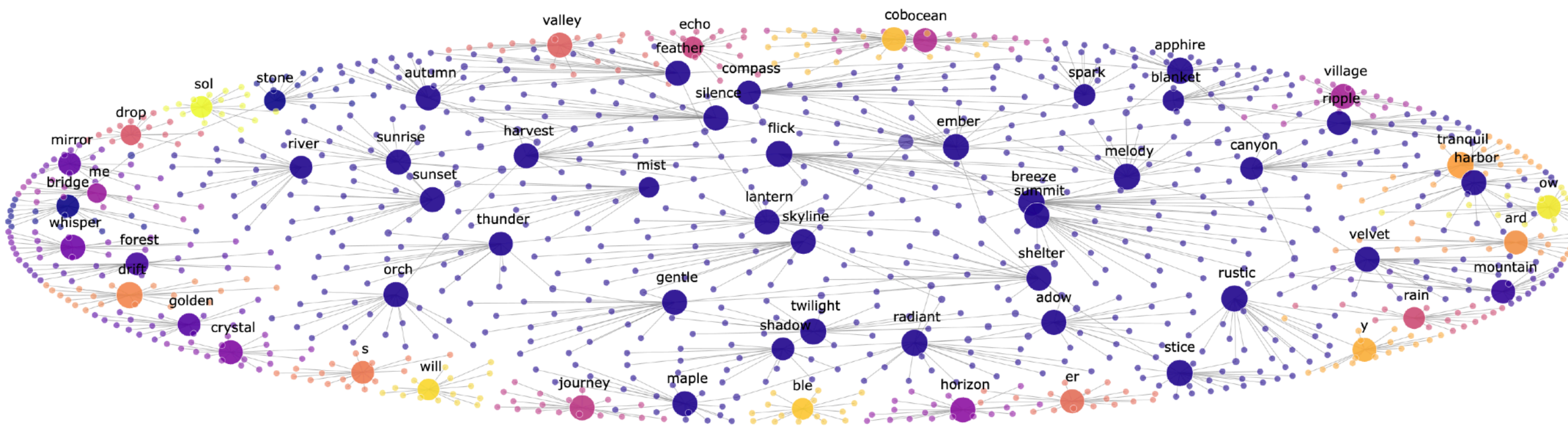


We cannot see the inner structure of proton.
We infer it from macroscopic distance and repeated experiments

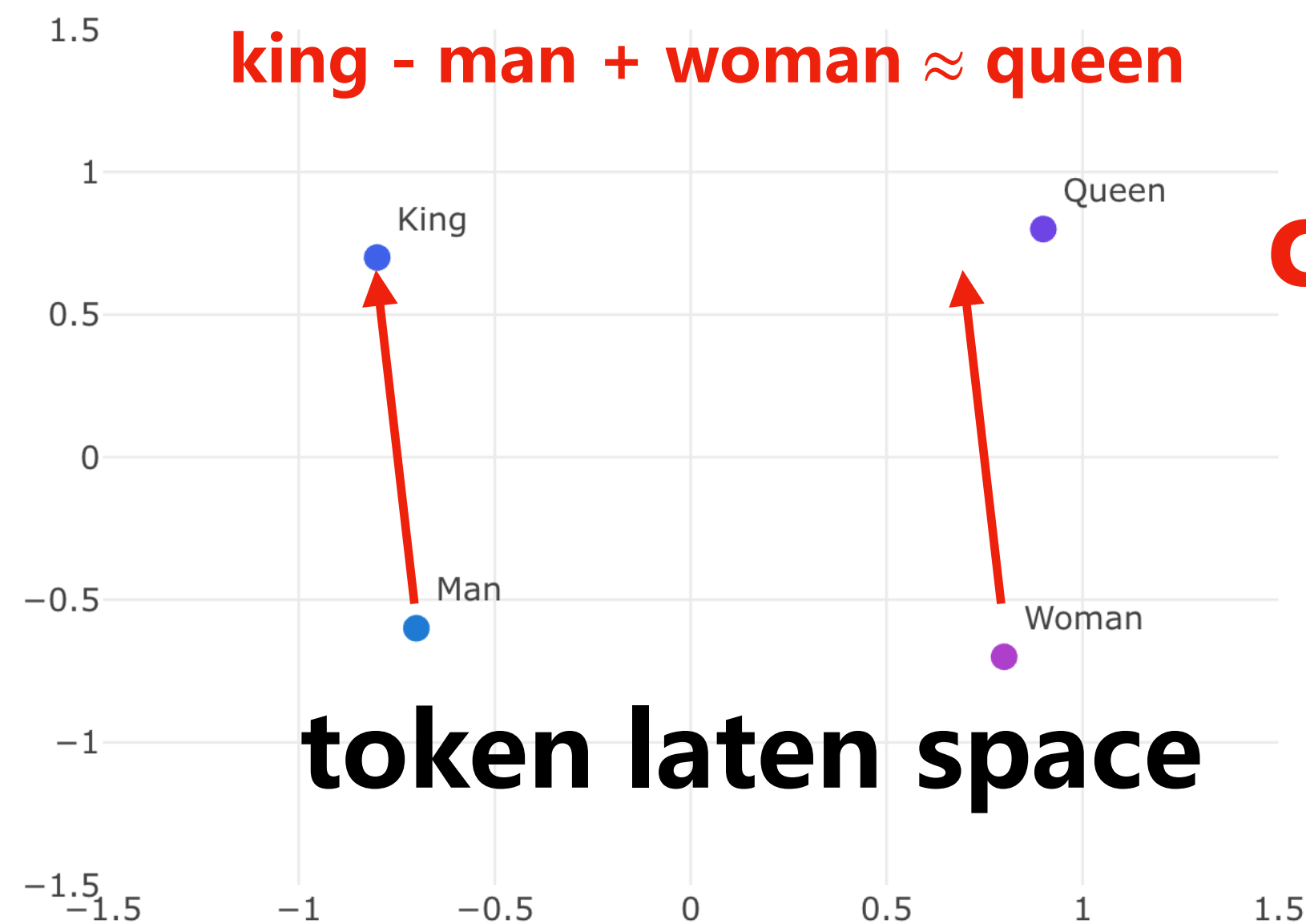


For genuine large model that exhibits frontier intelligence, very expensive/difficult to study the microscopic law. Instead, everyone can repeat the generation and study the law of data/generation

Space of exploration: from token to concept



Source: Huggingface

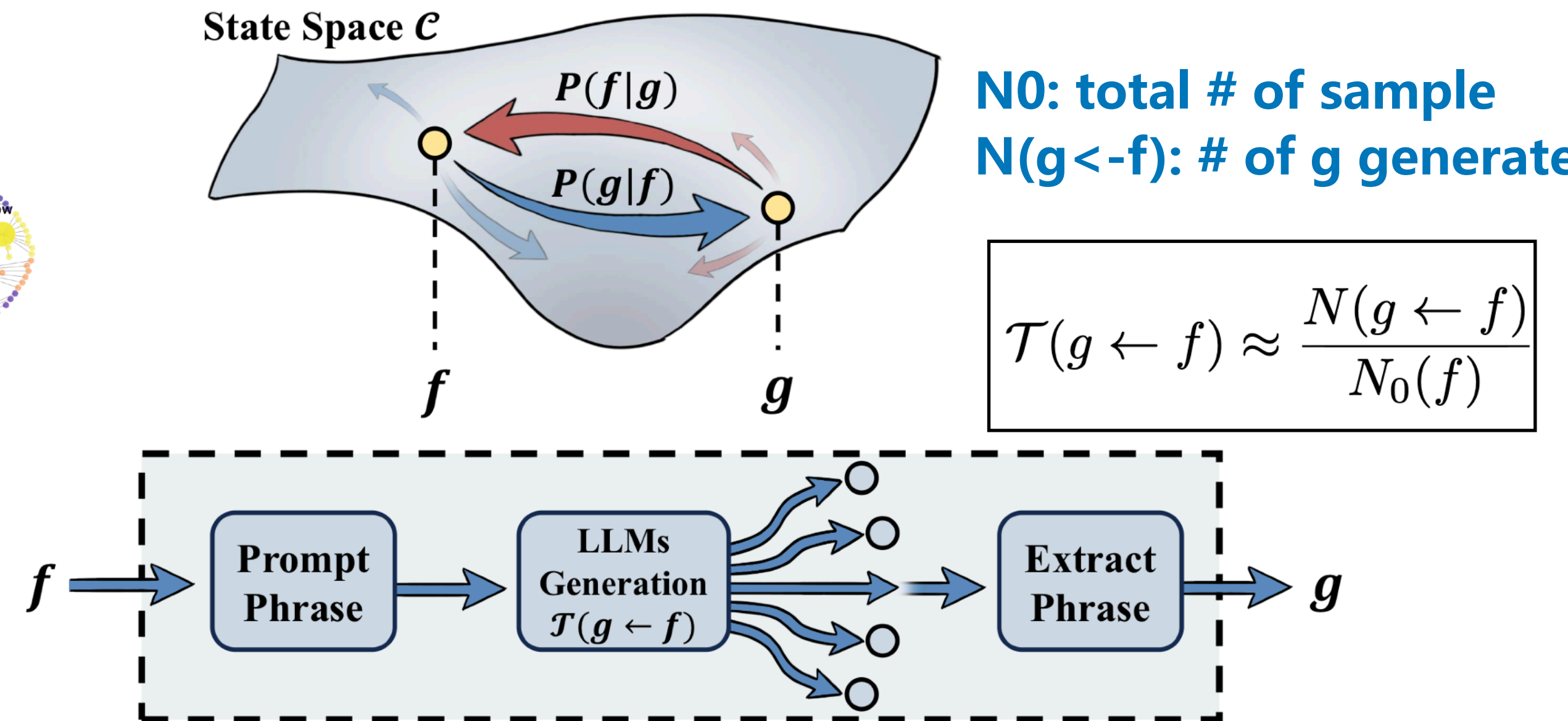


corse-graining

token $\rightarrow$ concept



Zhuo-Yang Song, Q.H. Cao, M.X. Luo, HXZ, 2512.10047



```
param1 * tanh(param2 * x + param3) + param4
param1 - (param2 / (x + param3))
param1 * x / (1 + param2 * log(x + 1))
param1 * tanh(param2 * x) + param3
param2 + param1 * (1 - exp(-x))
```

example from Symbolic Regression

Observation of evidence for detailed balance in LLM generation

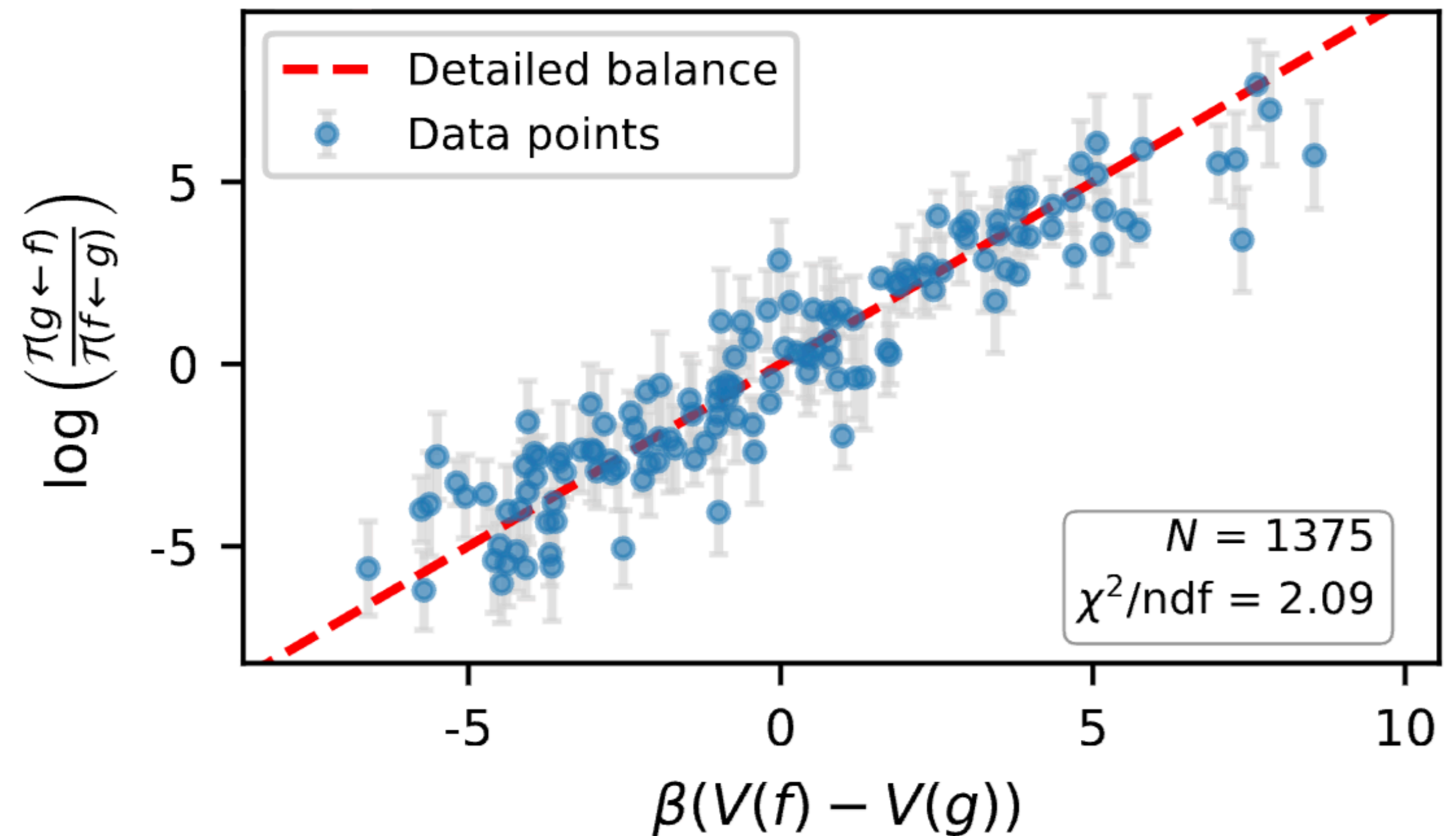
- The generation process exhibits a scaling law
- We interpret it as a consequence of **detailed balance**

$$\pi(g)\mathcal{T}(f \leftarrow g) = \pi(f)\mathcal{T}(g \leftarrow f)$$

$\pi(f)$ probability of staying in state f

$$\mathcal{T}(f \leftarrow g) = \mathcal{T}(g \leftarrow f)e^{-\beta(V_{\mathcal{T}}(f) - V_{\mathcal{T}}(g))}$$

- Can we apply more tools from equilibrium physics to study LLM agent?



Measured in the example of symbolic regression

```
param1 * tanh(param2 * x + param3) + param4  
param1 - (param2 / (x + param3))  
param1 * x / (1 + param2 * log(x + 1))  
param1 * tanh(param2 * x) + param3  
param2 + param1 * (1 - exp(-x))
```


HEP4AI: Industrial impact

Benchmark as a collaboration between science and industry community

Shaping the Future of Mathematics in the Age of AI

Johan Commelin* Mateja Jamnik† Rodrigo Ochigame‡ Lenny Taelman§
Akshay Venkatesh¶ 2603.24914

Another proposal being pursued by multiple groups is the creation of community-owned **benchmarks**. Good **benchmarks** can signal community values to those developing AI systems, while also helping mathematicians understand realistic use cases for, and limitations of, the technology. For example, **benchmarks** that focus on open-ended problems can counter overstated claims based on narrow task suites.

多个团队正在推进的另一项提议是创建社区共有的基准测试。优秀的基准测试不仅能向人工智能系统的开发者传递社区的价值取向，还能帮助数学家理解该技术的实际应用场景及其局限性。例如，聚焦于开放式问题的基准测试可以有效反驳那些基于狭隘任务套件而产生的夸大言论。

135,997

PHYBench: Holistic Evaluation of Physical Perception and Reasoning in Large Language Models

23 Apr 2025

reasoning chain-of-thought

PHYBench: Holistic Evaluation of Physical Perception and Reasoning in Large Language Models

Peking University Beijing Computational Science Research Center

StephenQS nondegeneracy +1

A comprehensive benchmark suite called PHYBench evaluates physical reasoning capabilities in large language models through 500 carefully curated physics problems and introduces a novel Expression Edit Distance (EED) Score metric, revealing significant performance gaps between current LLMs and human physics students across various domains of physics.

Problem Method Results Takeaways

Completed 1674

Science Olympiad track extends this format by employing international Olympiad medalists across a range of scientific subjects to craft problems specifically adversarial against state-of-the-art models. Complementary work such as PHYBench (Qiu et al., 2025), ChemBench Mirza et al. (2024), and

Cited by OPENAI benchmark: <https://openai.com/index/frontierscience/>

PHYBench

Peking University, CSRC

Physics

ACADEMIA

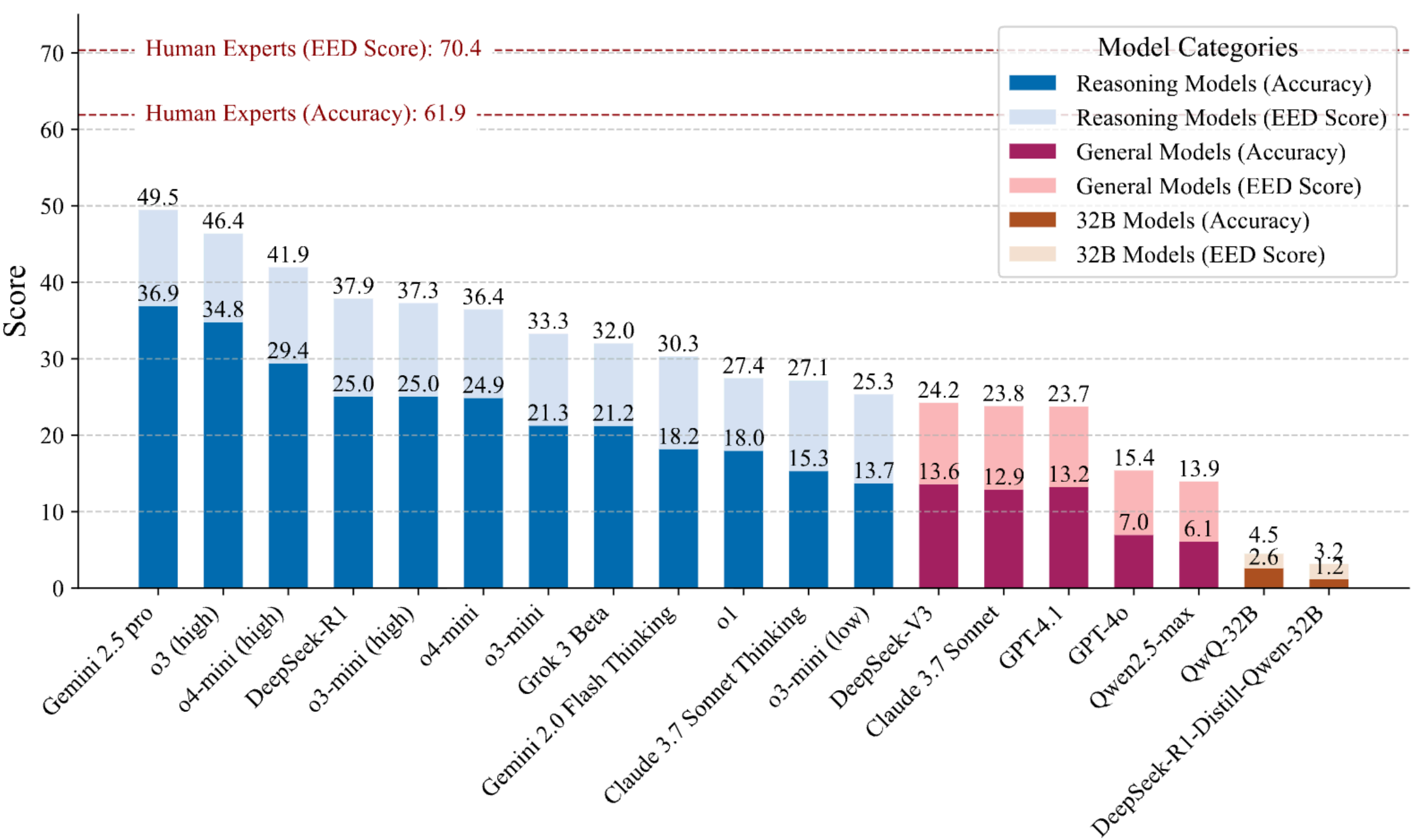
500 original physics problems.
Gemini 2.5 Pro: 36.9% vs. human experts: 61.9%.

<https://hai.stanford.edu/ai-index/2026-ai-index-report>



NIPS 2025
Shi Qiu, et al., , 2504.16074

Progress in model capability as measured from PHYbench



April 2025



Seed2.0 正式发布

Seed2.0 正式发布

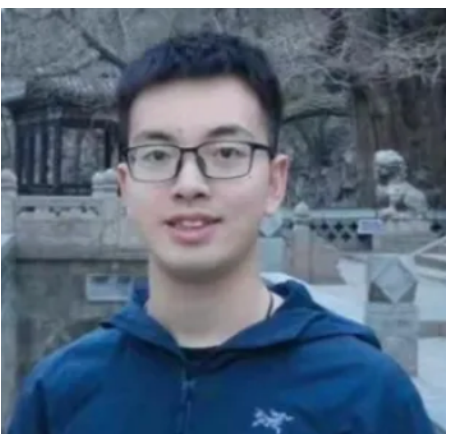
日期
2026-02-14

分类

模型发布

Capability	Benchmark	Seed2.0 pro	Seed2.0 lite	Seed2.0 mini	GPT-5.2 High	GPT-5 mini High	Claude Sonnet 4.5
STEM	GPQA Diamond	88.9	85.1	79	92.4	82.1	
	Superchem (text-only)	51.6	48	16.2	58	34.8	
	BABE	50	50.2	40.4	58.1	49.2	
	Phybench	74	73	56	74	60	
	FrontierSci-research	25	18.3	3.3	25	18.3	
	FrontierSci-olympiad	74	70	44	75	69	

Feb 2026



NIPS 2025

Shi Qiu, et al., , 2504.16074

PRBench: ultimate benchmark for AI intelligence

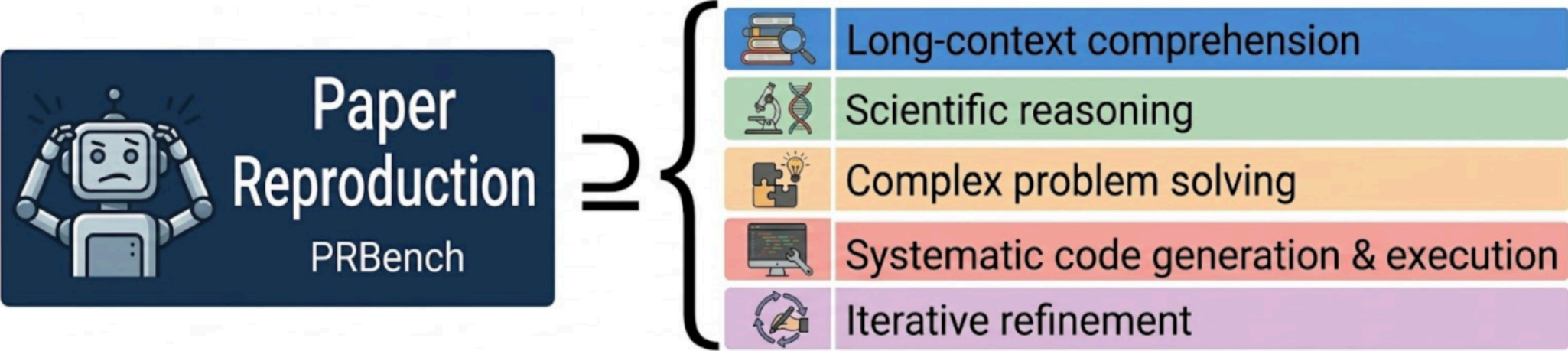
PRBench: End-to-end Paper Reproduction in Physics Research

Shi Qiu¹, Junyi Deng¹, Yiwei Deng¹, Haoran Dong¹, Jieyu Fu¹, Mao Li¹, Zeyu Li¹, Zhaolong Zhang¹, Huiwen Zheng¹, Leidong Bao¹, Anqi Lv¹, Zihan Mo¹, Yadi Niu¹, Yiyang Peng¹, Yu Tian¹, Yili Wang¹, Ziyu Wang¹, Zi-Yu Wang¹, Jiashen Wei¹, Liuheng Wu¹, Aoran Xue¹, Leyi Yang¹, Guanglu Yuan¹, Xiarui Zhan¹, Jingjun Zhang¹, Zifan Zheng¹, Pengfei Liu¹, Linrui Zhen¹, Kaiyang Li¹, Qichang Li¹, Ziheng Zhou¹, Guo-En Nian¹, Yunwei Xiao¹, Qing-Hong Cao¹, Linjie Dai¹, Xu Feng¹, Peng Gao¹, Ying Gu¹, Chang Liu¹, Jia Liu¹, Ming-xing Luo², Yan-Qing Ma¹, Liang-You Peng¹, Huichao Song¹, Shufeng Wang¹, Chenxu Wang¹, Tao Wang¹, Yi-Nan Wang¹, Chengyin Wu¹, Pengwei Zhao¹, and Hua Xing Zhu^{1,†}

¹School of Physics, Peking University, China

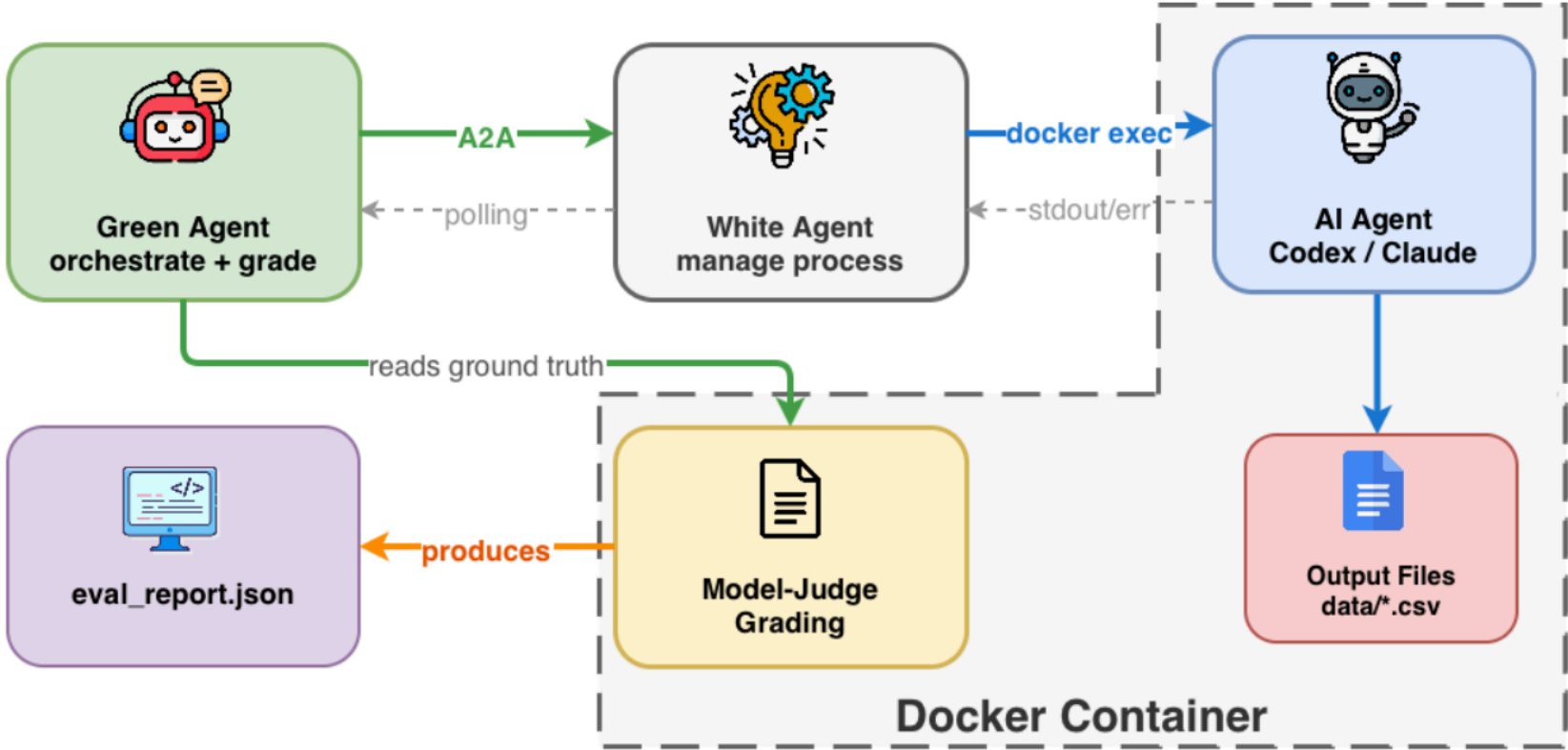
²Beijing Computational Science Research Center, China

March 31, 2026



A research level benchmark for scientific understanding, reasoning, scientific computing

Science is a endless frontier



Agent	Methodology	Code	Data Acc.	Instruction	Overall
<i>OpenAI Codex</i>					
GPT-5.3-Codex	78.00	43.00	21.00	92.00	34.00
<i>OpenCode-based Agents</i>					
GPT-5.3-Codex	72.00	36.00	16.00	90.00	28.50
Kimi K2.5 (1T)	61.90	22.00	11.40	80.60	20.57
DeepSeek V3.2 (671B)	63.20	19.30	8.90	84.20	18.50
Minimax 2.7 (230B)	55.60	16.30	10.00	86.00	17.97
GLM-5 (744B)	50.50	18.80	10.60	67.00	17.87
End-to-End Callback Rate			0.0		



Shi Qiu, et al., , 2603.27646

Education impact

3学分课程：物理学与人工智能

- 主讲教师：马滢青（负责人），陈基，冯旭，林晨，王力乐，宋慧超，张亿，朱华星

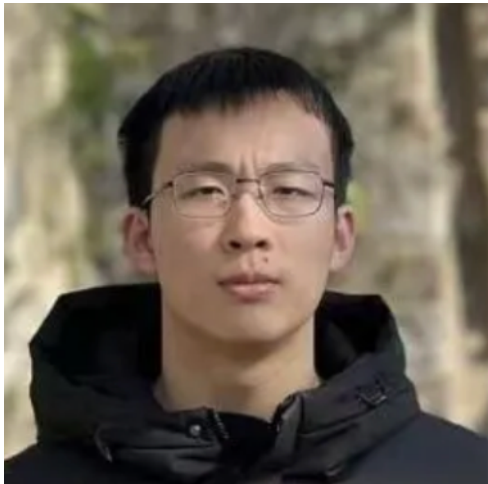
- 第一周：人工智能的历史和编程基础
- 第二周：大语言模型与智能体
- 第三周：提示词工程
- 第四周：机器学习的核心概念
- 第五周：经典算法和网络架构
- 第六周：Transformer等架构
- 第七周：人工智能中的搜索策略
- 第八周：凝聚态物理与人工智能
- 第九周：天体物理与人工智能
- 第十周：核物理与人工智能
- 第十一周：量子多体与人工智能
- 第十二周：扩散模型
- 第十三周：统计物理与人工智能
- 第十四周：粒子物理与人工智能

Participating in the forefront of the industry-academia collaboration



我在北京大学强基 | 上究星辰理，下探自然秘

Original 北大招办 北京大学招生办 2026年4月19日 18:51 北京



回来之后，我参加了物理学院的AI解题大赛，结果还不错，也因此有机会加入朱华星老师和曹庆宏老师的课题组继续做研究。这段时间里我陆续参与了两个项目——一个是评测大模型物理推理能力的基准PHYBench，另一个是用于高能物理对撞机唯象学研究的智能体系统ColliderAgent。做这些工作的过程里，我越来越清楚自己想去哪个方向，也把这些想

此外，我还积极拓展跨学科视野。大二时我参与了理论物理研究所牵头的PHYBench项目，与团队共同构建用于评测大语言模型科学推理能力的基准测试。经过一个多月的集中攻关，我们最终发表了一篇人工智能领域的会议论文，这段经历让我对物理与人工智能的交叉融合有了更深刻的认识。

Large Language Model training camp

Organize a team of high-school students to learn LLM from scratch using physics reasoning as target

Divided into 4 teams: AI infra, RL, data, agent

Complete a full cycle of model training

github.com/Rise-AGI/OpenFundus
github.com/Rise-AGI/magnus



Some final thoughts

Automation Without Understanding

Why the United States must preserve mathematical capacity in the age of AI

JUN-YONG PARK

<https://arxiv.org/pdf/2607.06377>

My father used to tell me: “When you cannot tell what is the right thing to do, do the most difficult thing well. That is the right thing to do.” Mathematical understanding is difficult. It is slow. It is unprofitable on any timescale a budget can see. It takes a kind of passion that no benchmark can measure. And it is part of the foundation on which everything else stands.

The choice facing the United States is not between human thought and machine intelligence. It is between automation with understanding and automation without it. A country that chooses the second path does not become more advanced. It becomes more dependent, more fragile, and less capable of directing its own future.

我父亲曾告诉我：“当你无法辨别什么是正确的事时，就把最难的事做好。那就是正确的事。”数学理解是困难的。它是缓慢的。在任何预算能够看到的任何时间尺度上，它都是无利可图的。它需要一种没有任何基准可以衡量的激情。而它是其他一切赖以立足的基础的一部分。

美国面临的选择不是人类思维与机器智能之间的选择。而是有理解的自动化与无理解的自动化之间的选择。选择第二条道路的国家不会变得更先进。它会变得更依赖、更脆弱，更无法掌控自己的未来。



take quantum mechanics, even if
you are a computer scientist,

If you're studying computing science or engineering of some kind, and you have a choice between a course that teaches you about a piece of technology that is currently fashionable like, I don't know, mobile app programming or LLM prompting or whatever it is, I am sure there are equivalent in various engineering disciplines, don't take those classes, **if you have a choice between mobile programming and quantum mechanics, take quantum mechanics, even if you're a computer scientist.**

MIT MGAIC Symposium: Foundations and Futures Keynote
Yann LeCun, Oct. 2025

