上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

AI scientist for LQCD: a new tool for flavor physics at STCF

高浩飞

2026.7

饮水思源 · 爱国荣校

目 录

01

简介

- 从LLM到AI agent与AI scientist
- 格点QCD计算

AI

The volume curves are staggering. Industry baselines establish that AI is now the primary author of modern software infrastructure: Microsoft reports that AI now generates 30% of its codebase; Salesforce estimates AI agents handle 30% to 50% of workloads; Meta has signaled a 50% target for late 2026; and GitHub Copilot's telemetry points to a 46% average code share across its entire user base. Most striking, Google says that 75% of its production code is AI-generated.

粒子与核物理研究所学术讲座
INPAC Special Particle Physics Seminar

Vibe Research 实战：四个月，十四篇直接反应论文是怎么做出来的

Abstract

过去四个月，我在大语言模型的日常协助下完成了十四篇直接核反应方向的理论论文。没有新机器、没有新理论、没有新学生，改变的只是研究者与工具之间的分工。这是一份完全基于个人实践的工作流报告。直接反应是一个理论框架几十年前就已封顶的子领域，真正卡住这个方向的从来不是想不出，而是做不完：角动量重耦、坐标变换、图表打磨、审稿回复，一个人一年能做完的物理远少于他能想到的物理。LLM恰好擅长这类结构清晰但繁琐的劳动。我会用真实案例拆解完整流程，说明 LLM 在哪些环节承担主力（文献综述、推导校验、代码原型、图像、初稿），哪些环节不可替代（选题、物理判断、数值合理性的直觉）。Vibe Research 不是让 AI替你做物理，而是把物理之外的一切工序交给 AI，再把省下的时间用来做更多的物理。但直接反应只是一个极端样本。无论是 ab initio结构、壳模型、核天体反应网络还是裂变聚变动力学，整个核物理都有同一个结构性处境：理论骨架早已搭好，真正的瓶颈是做不完。放射性束流实验的数据以近乎指数的速度增长，而能把数据翻译成物理的理论家几乎停滞。如果限制我们产出的从来不是想象力而是人力，那么当一个能够放大劳动力的工具第一次可用，我们面对的就不是多发几篇论文的问题，而是一次学科的重新定位：核物理是继续作为一门越做越精致的守成学科，还是在我们这代人手里重新成为一个对低能量子多体、对元素起源、对精密检验标准模型仍在主动设问的前沿。我这四个月的十四篇论文不是答案，只是一个早期的证据。

Prof. Lei Jin
Tongji University

Biography

金磊，同济大学物理科学与工程学院教授、博士生导师。金磊教授于2016年在西班牙塞维利亚大学获得博士学位，2016-2020年先后在俄亥俄大学及意大利 INFN-Pisa从事博士后研究，2020年入职同济大学，2021年入选国家级高层次青年人才计划。金磊教授从事量子少体理论、大爆炸核合成、原子核反应、及原子核的第一性原理计算等方面研究，在PRL、PLB等高水平期刊发表学术论50余篇。

Time: 10:00am, Monday, April 20th, 2026
Place: Science Building 5-203
https://meeting.tencent.com/dm/ZnnqGWUTxnCA
ID: 674736189 Password: 123456

Host: Prof. Yifei Niu
Contact: Chengcheng Zhang

Email: niuyf@sjtu.edu.cn
Email: Chengcheng_Zhang@sjtu.edu.cn

Vibe physics: The AI grad student

从大语言模型到AI Agents

上海交通大学 SHANGHAI JIAO TONG UNIVERSITY

- **大语言模型**（LLM，Large Language Model）：以自然语言为对象，在大规模文本数据上进行训练得到的，参数规模巨大的统计模型
- **核心能力：在给定上下文（context）中，基于条件概率预测下一个词元（token）**

  - **数据驱动学习**：不依赖复杂的人工规则，仅需大量的文本数据训练，自动学习语言模式和规律，适应多种语言和领域文本
  - **上下文感知**：存在前文语境信息，能依据上下文预测下一个词，生成的文本连贯性和逻辑性强

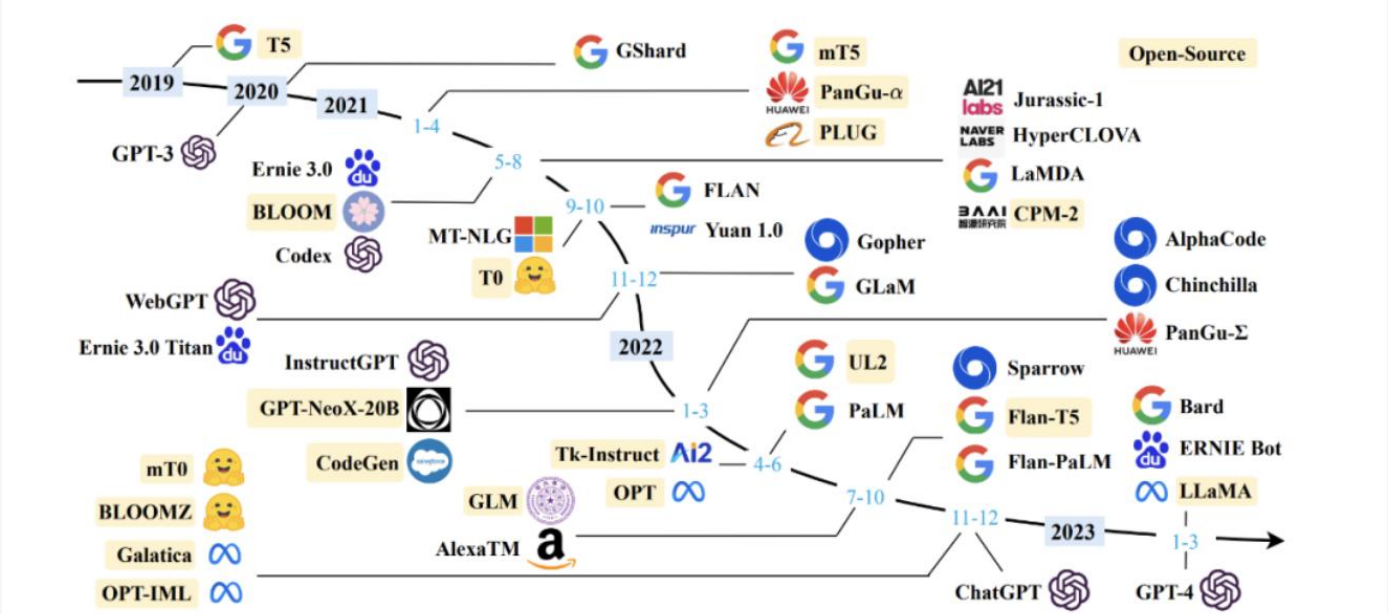


Fig. 1. A timeline of existing large language models (having a size larger than 10B) in recent years. We mark the open-source LLMs in yellow color.

从大语言模型到AI Agents

上海交通大学 SHANGHAI JIAO TONG UNIVERSITY

Agent（智能体）

=

LLM（大模型）+ Planning（规划）+Memory（记忆）+ Tools（工具）

主要局限

1 幻觉

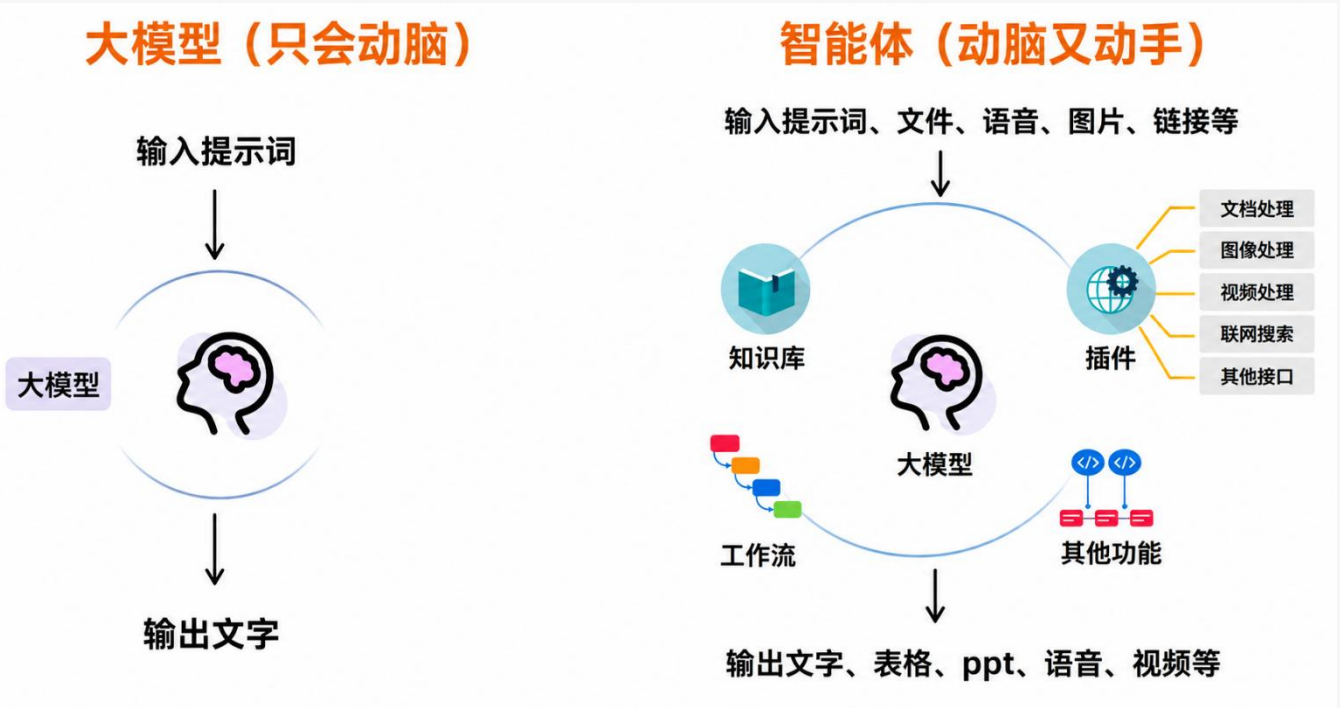
可能生成看似合理但并不真实的内容

2 逻辑和数学不稳定

复杂推理、严格逻辑与数值计算结果不稳定

3 上下文会缺失

长上下文中容易遗漏关键信息或遗忘前文约束



从 AI Agents 到 AI Innovators

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

OpenAI：通向通用人工智能（AGI）的五个阶段

Stages of Artificial Intelligence	
Level 1	Chatbots, AI with conversational language
Level 2	Reasoners, human-level problem solving
Level 3	Agents, systems that can take actions
Level 4	Innovators, AI that can aid in invention
Level 5	Organizations, AI that can do the work of an organization

[2022-2023]
GPT 4, Deepseek V3

[2024]
GPT o3, Deepseek R1

[2025-2026] AI Agent:
Openclaw, Manus ...

[2026-2027] ?

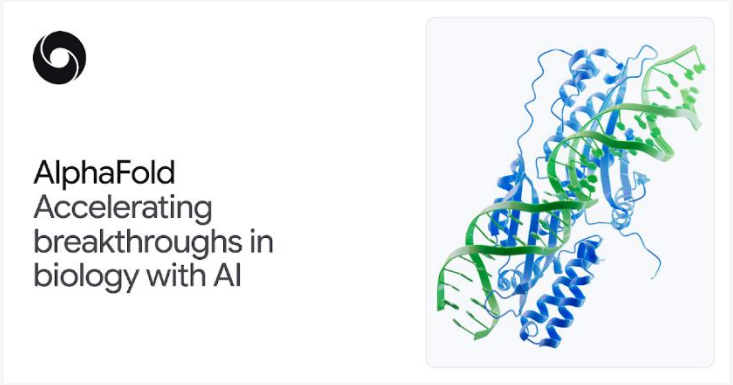
目前，AI能力的边界正在迅速拓展，
逐渐超越文本对话，成为任务的执行者

SJTU

从 AI scientific tools 到 AI Scientist

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

AI科研工具

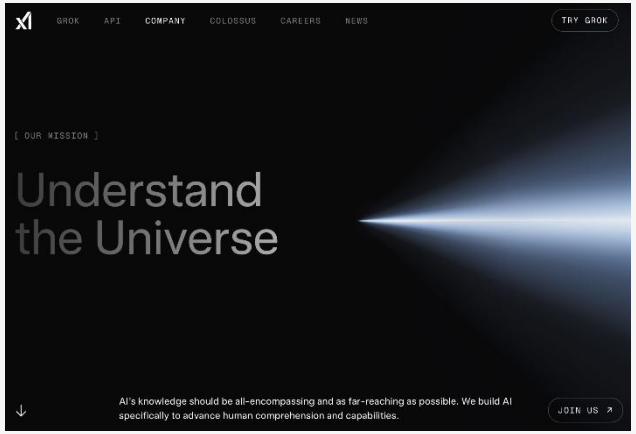


AlphaFold: 蛋白质结构预测

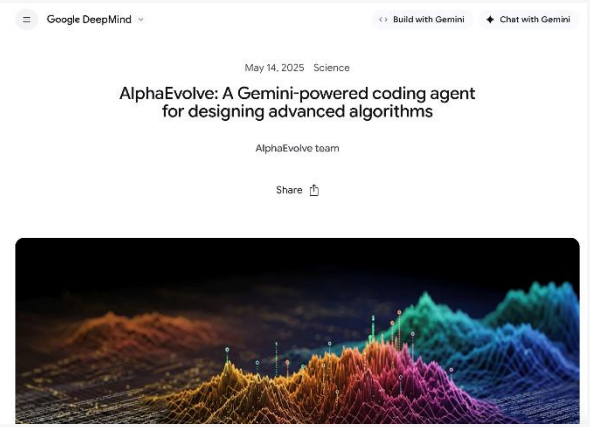
DeePMD-kit

DeePMD-kit：分子模拟

AI for Science领域的新探索：AI科学家



xAI



Google DeepMind

Microsoft Kosmos
AI Scientist for Autonomous Discovery

Microsoft

sakana.ai

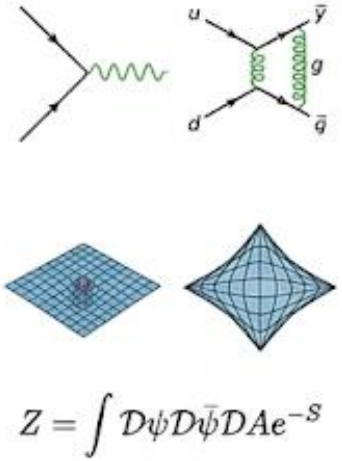
The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery

Sakana AI

where’s AI physicist ?

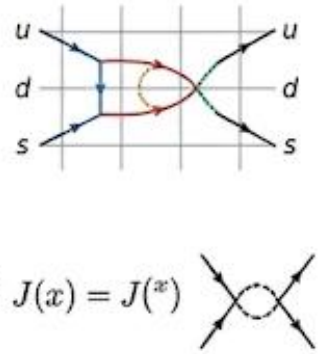
格点QCD计算全流程指南

步骤 1. 学习理论
(Theory Learning)



理解非微扰强相互作用。设定 m_q, g 等参数。

步骤 2. 构造关联函数
(Construct Correlation Functions)

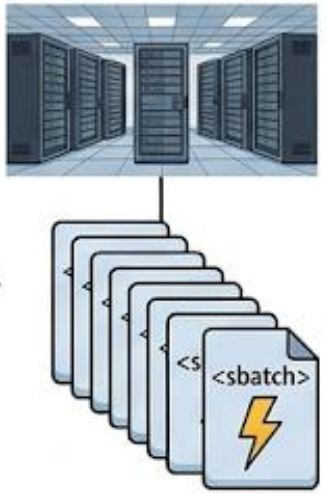


定义强子算符。构造关联函数。

步骤 3. 写代码
(Coding)

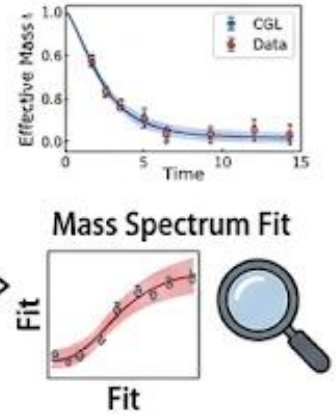
编写高效、可扩展的并行代码 (QUDA, MPI, OpenMP)。

步骤 4. 提交任务
(Submit Jobs)



在超算中心部署和管理并行任务。收集数据。

步骤 5. 分析与提取可观测量
(Analysis & Extract Observables)



实施统计分析。进行质量谱拟合、误差控制、连续极极化和手征外推。

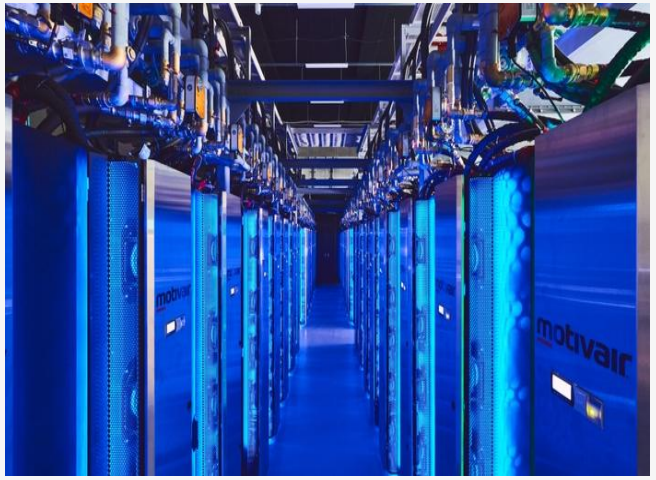
AI 大有可为

LQCD：科研范式的改变

Confinement of quarks*

Kenneth G. Wilson
Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14850
(Received 12 June 1974)

1974



~2000

AI physicist for Lattice QCD?

?

LQCD

计算：物理目标 → 计算程序

- 必要的物理知识：量子场论，算符构造，关联函数，…
- 专业的计算程序：Chroma，PyQUDA，…

LQCDMaster

分析：格点数据 → 物理结果

- 超长链条、海量数据、高维结构的复杂分析
- 耗时久，需至少两人检查

PhysMaster

2512.19799, PhysMaster

2603.22471, PhysMaster for LQCD

2604.15411, PRL-Bench

AI scientist实例

- LQCDMaster: 专业Lattice智能体

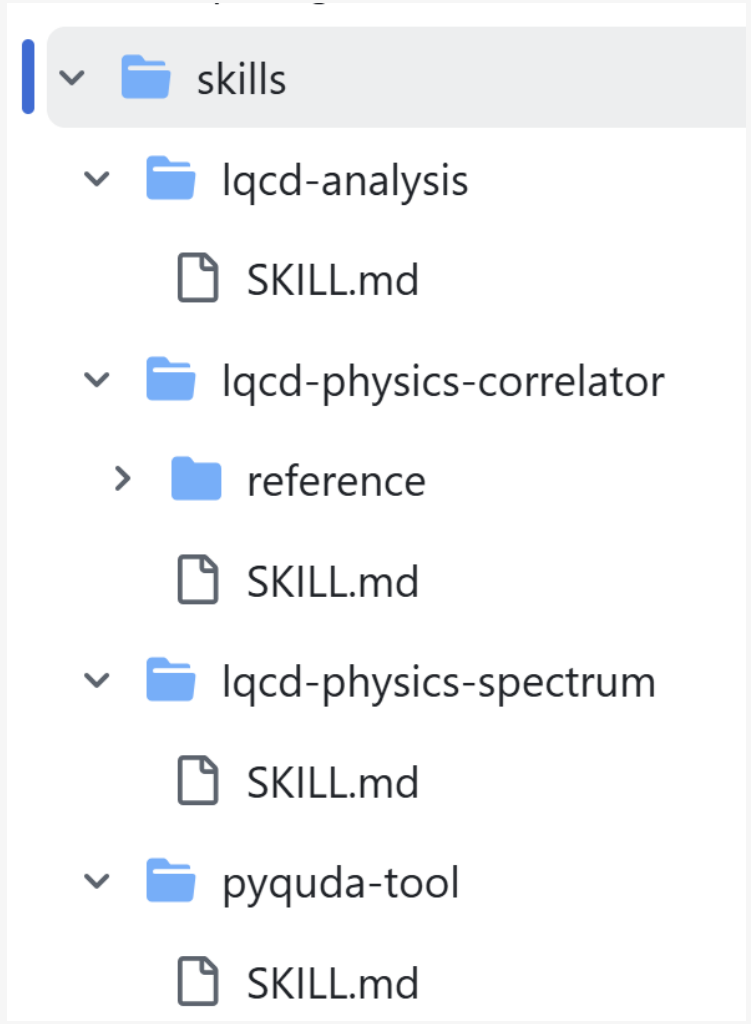
Skills：超级进化版提示词

元数据 (Metadata)： 对技能的简短描述。保存在全局上下文中，体积小，非常节省 **Tokens**（省钱又省心）。

行动指南 (Action Guide)： 真正的提示词，规定了 AI 每一步该怎么做。

资源文件 (Resources)： 可能包含 **Python 代码** 或其他执行程序，保证程序在调用 Skill 时能完成复杂的动作。

Skills：超级进化版提示词



name	lqcd-analysis
description	Lattice QCD analysis pipeline skill. Takes correlator data from lattice measurements and extracts physics results with statistical and systematic error control. Covers: source-time shifting, correlator folding, jackknife/bootstrap resampling, gvar/lsqfit integration, effective mass, multi-state correlated fits (hadron spectrum), matrix element extraction (ratio method, summation method, simultaneous C_2+C_3 two-state fit), dispersion relation and speed of light, fit diagnostics (χ^2/dof, Q-value, AIC, SVD cut), and scale conversion to physical units. Uses fit function templates from lqcd-physics-spectrum. Trigger on: "analyze correlators", "fit the data", "extract mass", "effective mass", "matrix element", "dispersion relation", "speed of light", "jackknife", "resampling", "lsqfit", or when contractions are done and results needed.

LQCD Analysis Pipeline

name	lqcd-physics-correlator
description	Lattice QCD physics reasoning skill. Derives the chain from a physics observable to a computable correlator: interpolating operator construction (mesons, baryons), correlator definition (two-point, three-point), Wick contraction with γ_5-hermiticity and flavor symmetry, propagator requirements, and einsum expressions for contractions. Hands spectral decomposition and fit templates off to lqcd-physics-spectrum. Uses DeGrand-Rossi gamma basis (PyQUDA convention). Trigger on: hadron masses, decay constants, form factors, matrix elements, operator construction, Wick contraction, disconnected diagrams, or "what correlators/propagators do I need".

LQCD Physics Reasoning

name	pyquda-tool
description	PyQUDA tool usage skill. Generates Python code that calls PyQUDA to solve quark propagators on lattice gauge configurations. Covers: configuration loading, quark parameter setup (Wilson/clover action, mass or kappa, clover coefficient, link smearing), multigrid solver configuration, source construction (point, Gaussian smearing with APE/HYP/stout links), propagator inversion, and residual verification. Reads ensemble parameters from ensemble_registry.yaml. Trigger on: "compute propagators", "solve propagator", "run inversions", "call PyQUDA", "solve Dirac equation", or when lqcd-physics-correlator has produced a propagator requirements list.

PyQUDA Tool Usage

LQCDMaster

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

Differences Among Agents

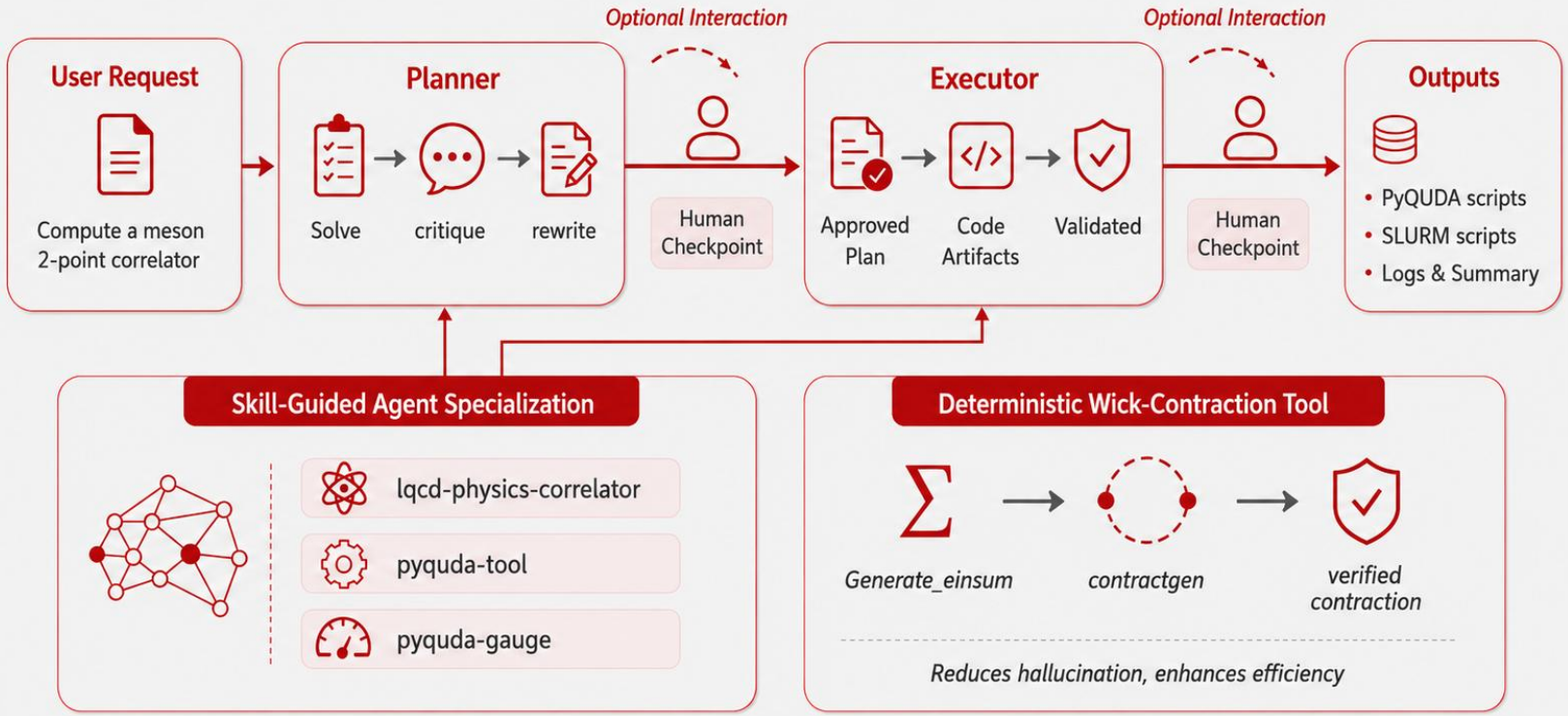
Compare agent systems by skills and tools

Agent	Skills	Tools
LQCDMaster	LQCD-specialized	PyQUDA, Wick, SLURM
Codex	general coding	editor, terminal
Deep Research	search & synthesis	web, retrieval

Agents differ mainly in specialization and tool access.

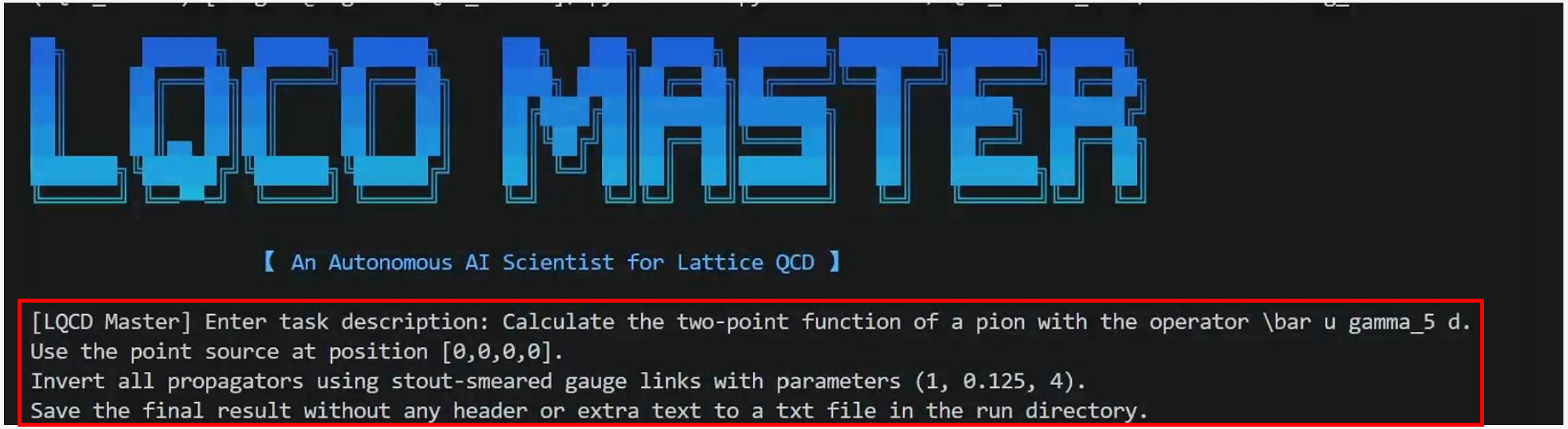
LQCDMaster: Tool-Augmented, Skill-Guided LQCD Code Generation

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



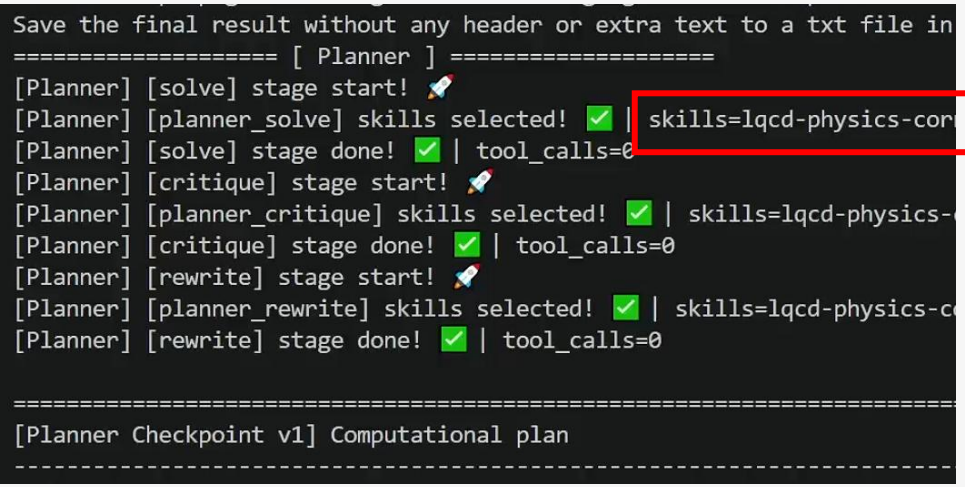
LQCDMaster

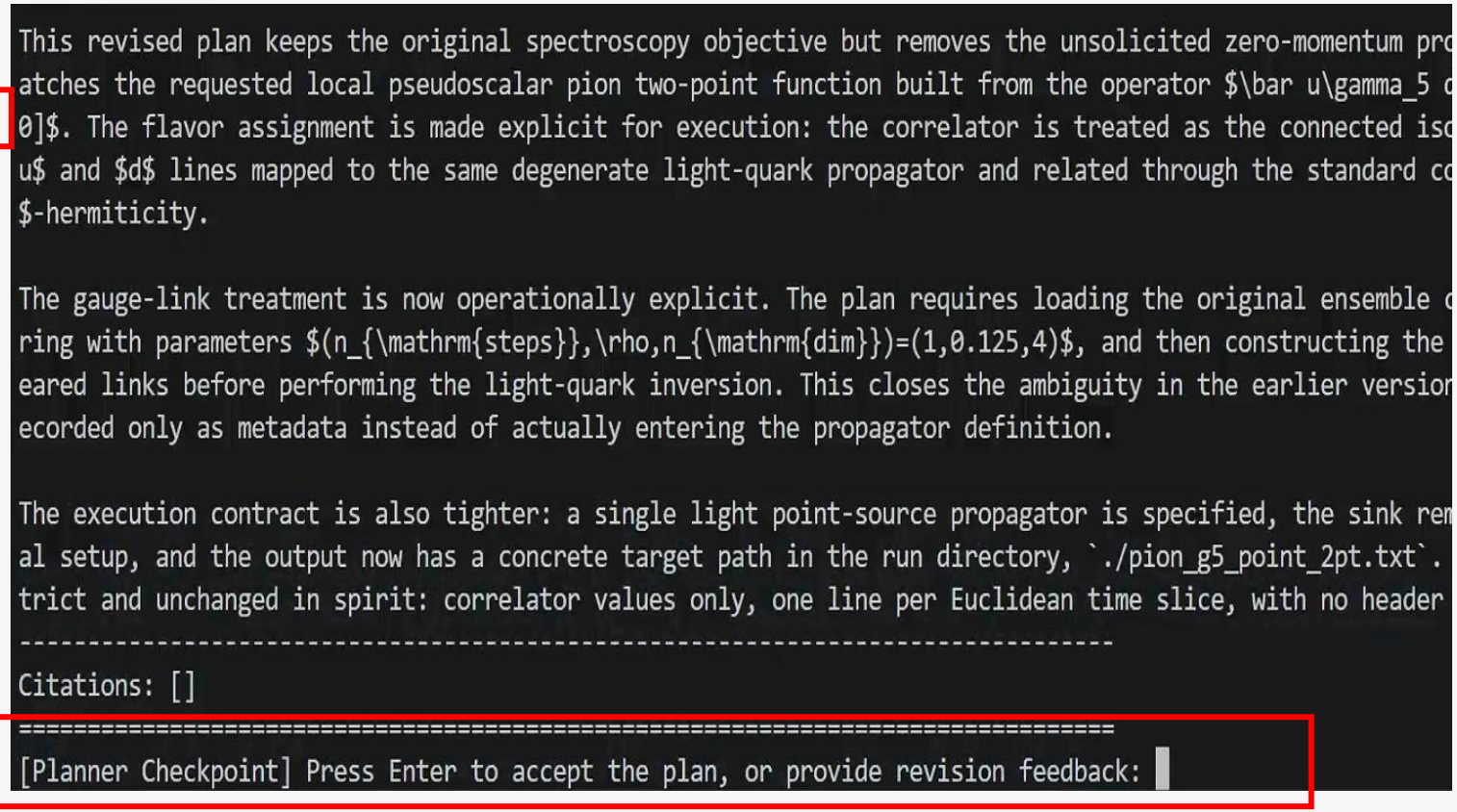
上海交通大學
SHANGHAI JIAO TONG UNIVERSITY



SJTU

LQCDMaster
1. Planner





LQCDMaster

2. Executor

```
================== [ Executor ] ====================
[Executor] [executor:v1] stage start! 🚀
[Executor] [executor_generate] skills selected! ✅ | skills=lqcd-physics-correlator, pyquda-to
[Executor] [executor_generate] tool call: generate_einsum 🔧
[Executor] [executor] stage done! ✅ | tool_calls=1
[Executor] [executor] LLM output keys: ['main_program', 'test_submit', 'full_submit', 'notes']
[Executor] [executor_static_analysis] stage start! 🚀
[Executor] [executor_static_analysis] stage done! ✅ | errors=0
[Executor] [executor_critique] skills selected! ✅ | skills=lqcd-physics-correlator, pyquda-to
[Executor] [executor_critique] stage done! ✅ | tool_calls=0
[Executor] [executor_critique:v1] found 2 error(s); trigger executor_rewriter static-check rou
[Executor] [executor_rewrite] skills selected! ✅ | skills=lqcd-physics-correlator, pyquda-tool
[Executor] [executor_rewriter] stage done! ✅ | tool_calls=0
[Executor] [executor_static_analysis] stage start! 🚀
[Executor] [executor_static_analysis] stage done! ✅ | errors=0
[Executor] [executor_critique] skills selected! ✅ | skills=lqcd-physics-correlator, pyquda-to
[Executor] [executor_critique] stage done! ✅ | tool_calls=0
[Executor] [executor_critique:v1] remaining_errors=1; pass_to_checkpoint_for_human_review
[Executor] [executor:v1] artifacts_saved=../LQCD_Master_runs/20260701meeting_test/executor_v1

================================================================================
[Executor Checkpoint v1] Scripts generated. Approval will trigger the test submission and queue
--------------------------------------------------------------------------------
```

```
--------------------------------------------------------------------------------
PLAN SUMMARY
Implementation is close, but it has one clear submission-output mismatch and one code-g
--------------------------------------------------------------------------------
WARNINGS
- Contraction code in `main.py` is not from the generate_einsum tool — this is a code-g
- The contraction `contract('wtzyxCBba, wtzyxCBba -> t', prop_1.data.conj(), prop_1.data
der the expected `gamma_5` cancellation pattern, but because it is handwritten rather t
index convention should be manually confirmed against the approved codegen scheme.
- The shell scripts are structurally acceptable, but `resource_path` is passed as `~/.c
 and cache persistence semantics should be confirmed in the actual runtime environment
--------------------------------------------------------------------------------
EXECUTOR CRITIQUE ERRORS
- Output path deviates from the approved plan: `main.py` writes `./pion_g5_point_2pt_cf
ectory file `./pion_g5_point_2pt.txt`.
================================================================================
[Executor Checkpoint] Press Enter to submit the test job, or provide revision feedback:
```

LQCDMaster

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

3. Sbatch & Revision

```
[Executor] [submit_test:v1] stage start! 🚀
[Executor] [submit_test] submit_script=/public/home/tangmen/work_pyquda/LQCD_Mas
[Executor] [submit_test:v1] stage done! ✅ | ok=True | job_id=116295877
[Queue] Waiting for test job output...

==================== [ Slurm Queue ] ====================
Current state: submitted, but Job 116295877 is not yet visible in this `squeue`

========== [ Workflow Completed ] ==========
[Run dir] Path: /public/home/tangmen/work_pyquda/LQCD_Master_runs/20260701meetin
[Status] Completed
[Trajectory] File saved to: /public/home/tangmen/work_pyquda/LQCD_Master_runs/20
```

PyQUDA script

```python
import os
import sys
import numpy as np
from opt_einsum import contract
from pyquda_comm import array
from pyquda_utils import core, io, source

# Run: mpirun -n 4 python3 main.py ~/.cache 10000

resource_path = sys.argv[1]
n_cfg = sys.argv[2]

latt_size = [24, 24, 24, 72]
grid_size = [1, 1, 1, 4]
backend = "cupy"

cfg_path_template = "/public/share/weiwang/clqcd/beta6.20_mu-0
20_mu-0.2770_ms-0.2400_L24x72_cfg_{n_cfg}.lime"
```

```
(venv-pyquda-ghf) (LQCD_master) [tangmen@login04 LQCD_Master]$ sq
             JOBID PARTITION     NAME             USER             ST       TIME  NODES NODELIST(REASON)
         116295877 kshdexclu16   pion2pt_test     tangmen           R       0:13      1 f15r4n00
```

LQCDMaster

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

Results

LQCDMaster

```
1.1632582004702274e+00 1.6298989353700609e-19
1.5899417140129657e-01 -3.1288345983559066e-20
6.6792772064338091e-02 -1.8309717069082567e-21
4.3159090846817828e-02 2.8240659410012015e-21
3.2299487133679798e-02 5.5189972692166903e-22
2.7268744090132607e-02 1.7910441711253317e-21
2.3438578907075451e-02 -1.2690798524872489e-21
1.8866329493605054e-02 2.5422652664104344e-23
1.4799854306247230e-02 5.8349250315531007e-22
1.1442865398272710e-02 -4.2147577271057789e-22
8.5568580471043123e-03 -6.8507581334097165e-22
6.7101015406195062e-03 3.9420353022519341e-23
5.4243075746006021e-03 2.5732692494828724e-22
4.4562053383694874e-03 9.2278744194189070e-24
3.5444740801028971e-03 -2.0748313980642531e-22
2.7404629726457790e-03 -9.0842046809605541e-23
2.2053743805644011e-03 4.7201464456982269e-23
1.8660592448400629e-03 5.6359714146602858e-23
1.5510799666623900e-03 -5.0935419996935300e-23
1.3171561952816122e-03 -2.5314429724148256e-23
1.1836375160887566e-03 -4.0472612544954405e-23
1.1198515815636760e-03 -5.4850223431593454e-23
9.7770928464498835e-04 -1.5486298429897916e-23
8.1756510049545378e-04 2.4344945854809999e-23
6.8547323463373294e-04 -6.2143334753871893e-24
6.0819821006694627e-04 1.1801971602605797e-24
5.3658265672072681e-04 4.6379231934554659e-24
4.8950185247911382e-04 -1.2682234867042764e-23
4.5012816904248991e-04 -1.0273755256948248e-23
4.0525277459911493e-04 1.6517936764238590e-23
3.6996232112879835e-04 1.2112544855428597e-24
3.6731253299061281e-04 -8.9428715025501767e-24
3.7249642400715219e-04 -2.1970684486900455e-24
3.3750675587723152e-04 1.9249275981950755e-23
2.8443344424865360e-04 -1.4010591970531877e-24
2.5479901715632627e-04 -5.0045233767632090e-24
2.3254081420827512e-04 -1.3779721037429684e-25
```

```
0 1.1632582004519192e+00 1.4005672981522223e-19
1 1.5899417138341790e-01 -5.6053213093206848e-20
2 6.6792772042648815e-02 -1.6306087298918203e-21
3 4.3159090818386550e-02 -3.0535495934667713e-21
4 3.2299487097856461e-02 -8.2890773268915713e-22
5 2.7268744041291315e-02 -5.9764568132516765e-22
6 2.3438578845013602e-02 -1.7744801281402388e-22
7 1.8866329426338885e-02 1.4769628744480308e-21
8 1.4799854242634027e-02 -6.2308908547881819e-22
9 1.1442865343009671e-02 -2.6357604143162727e-25
10 8.5568580034694162e-03 3.0638797931490269e-23
11 6.7101015067621657e-03 1.3823084809355477e-22
12 5.4243075475905751e-03 -1.3845458542208331e-23
13 4.4562053153985324e-03 8.1155461623207739e-23
14 3.5444740617283058e-03 -1.2063115173108889e-23
15 2.7404629587137051e-03 -2.1890303972571412e-23
16 2.2053743702848748e-03 -1.2985173449622994e-22
17 1.8660592366359829e-03 -2.9348689049514671e-23
18 1.5510799600792818e-03 5.8400030079355079e-23
19 1.3171561899026154e-03 4.1654010403819551e-24
20 1.1836375113576708e-03 -2.2888511529267745e-23
21 1.1198515771406975e-03 -3.5111069539927589e-23
22 9.7770928089549263e-04 -2.5345590744949977e-23
23 8.1756509742067268e-04 6.6513829717290804e-24
24 6.8547323207204011e-04 -3.4901069423407738e-25
25 6.0819820784988293e-04 -1.1956889020871173e-23
26 5.3658265455234274e-04 1.3441518938009812e-23
27 4.8950185025076122e-04 -2.3918889583702923e-23
28 4.5012816692392889e-04 4.2324027479362309e-24
29 4.0525277239996684e-04 1.5433621786282831e-23
30 3.6996231880816725e-04 -3.4282906364204203e-25
31 3.6731253053153923e-04 1.4321870230694023e-23
32 3.7249642157486881e-04 -3.7272863964274161e-24
33 3.3750675379376683e-04 -1.3489400828739602e-24
34 2.8443344255384330e-04 -1.6080625126320365e-23
35 2.5479901560404767e-04 4.5975510733493838e-24
36 2.3254081275916983e-04 -6.1408043964475191e-24
```

Manual

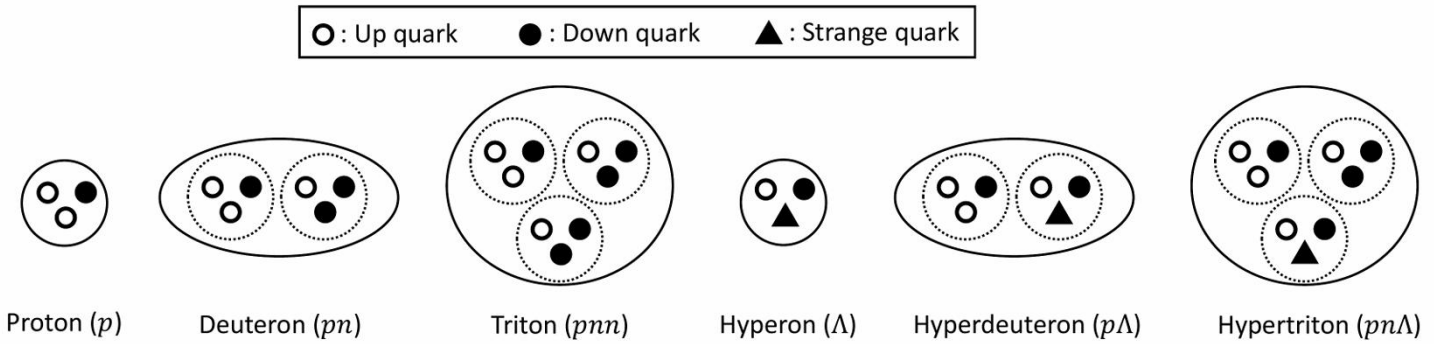
LQCDMaster

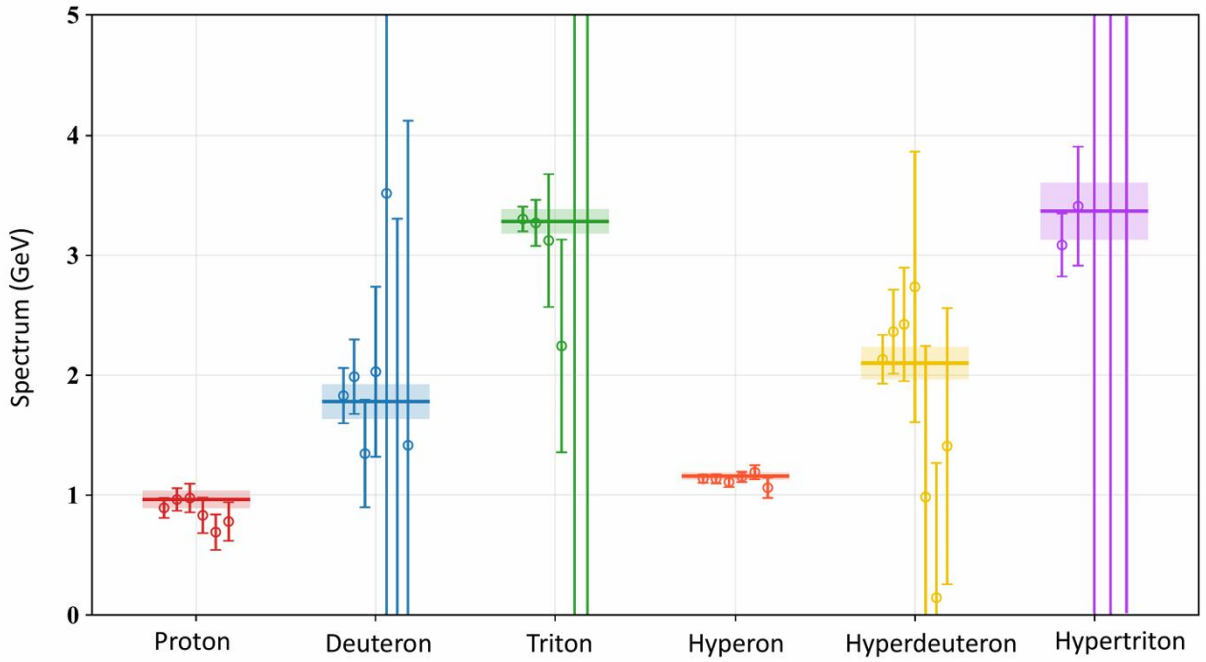
上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

Multi-hadron contractions

- 夸克数增加
- wick收缩的复杂度快速上升
- 程序难以实现

LQCDMaster 可以自动化完成这一步骤





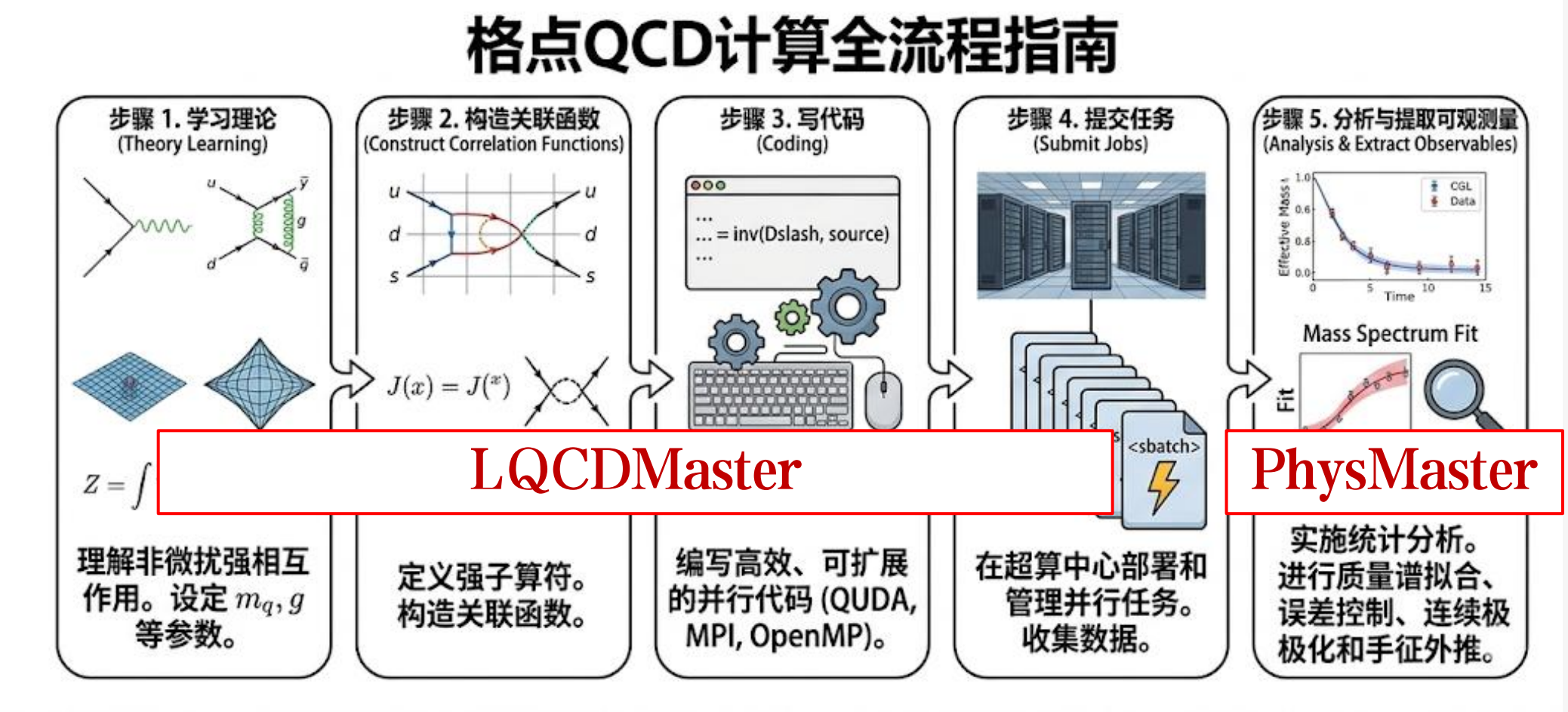
SJTU

03

展望与总结

迈向AI科学家之路

未来展望


格点QCD计算全流程指南
步骤 1. 学习理论
(Theory Learning)
理解非微扰强相互作用。设定 m_q, g 等参数。
步骤 2. 构造关联函数
(Construct Correlation Functions)
$J(x) = J(x)$
定义强子算符。构造关联函数。
步骤 3. 写代码
(Coding)
... = inv(Dslash, source)
编写高效、可扩展的并行代码 (QUDA, MPI, OpenMP)。
步骤 4. 提交任务
(Submit Jobs)
<sbatch>
在超算中心部署和管理并行任务。收集数据。
步骤 5. 分析与提取可观测量
(Analysis & Extract Observables)
CGL
Data
Effective Mass
Time
Mass Spectrum Fit
Fit
实施统计分析。进行质量谱拟合、误差控制、连续极极化和手征外推。
LQCDMaster
PhysMaster

未来展望

- **探索AI Agent解决新特定科学问题的方案：eg. 双智能体强化学习拓展设计空间；**
- **领域专精AI自主科学家：LQCDMaster** 超长程高效探索+稳健数值计算+可信知识库

我们希望，***PhysMaster / LQCDMaster / …*** 可以将智慧、思维活跃的物理学家从繁重的重复劳动中解放出来，加速灵感的生成与验证，创造物理研究新范式

2512.19799, PhysMaster

2603.22471, PhysMaster for LQCD

2604.15411, PRL-Bench