

从拉格朗日量到排除线：当AI开始学会“做”对撞机唯象学

仇 是¹ 蔡则宇¹ 卫家燊¹ 李泽宇¹ 殷艺轩¹ 曹庆宏¹
刘 畅¹ 罗民兴² 袁兴博³ 朱华星¹

(1. 北京大学 100091; 2. 北京计算科学研究中心 100193; 3. 华中师范大学 430079)

粒子物理学的核心问题之一,是探究自然界在最基本层次上的构成与相互作用规律。自20世纪中叶以来,这一探究主要依托大型粒子对撞机来推进:将质子或电子加速至接近光速,使其发生高能碰撞,在碰撞产生的末态粒子中寻找新物理存在的迹象。欧洲核子研究中心(CERN)的大型强子对撞机(LHC)是目前世界上能量最高的粒子对撞机,自2008年运行以来,已在其中发现了希格斯玻色子^[1,2],并对大量超出标准模型的新物理(BSM)假说给出了严格的实验约束^[3]。

然而,在理论预言与实验检验之间,存在一条往往被低估的技术鸿沟。研究者从一个理论模型出发,需要经历一套高度专业化的计算与模拟流程,才能最终得到可与实验数据直接比较的物理结论。这一流程通常依次涉及以下环节:基于拉格朗日量写出费曼规则并导出可以被后续程序读取的模型文件^[4-6],利用矩阵元生成器计算散射截面并产生初级事例^[7],通过部分子簇射(parton shower)与强子化程序处理末态粒子演化^[8],借助快速探测器模拟工具将末态粒子转化为探测器上的信号^[9],再经由统计分析程序完成判选、分布计算与置信区间提取,最终形成排除线或发现显著性^[10]。整条流程涉及FeynRules、MadGraph、Pythia、Delphes、MadAnalysis等多个专业软件,各工具之间存在复杂的数据格式依赖与操作约定,且大量关键经验以隐性形式存储于研究者的个人积累之中,难以被系统化传承。

这一现实决定了,对撞机唯象学研究中除了要

思考物理问题本身,还要把大量的时间与精力消耗在“让各软件正确配合”的流程性执行。一个理论想法也许半天就能完整表述,但将其转化为一张可发表的排除线图,往往需要数天乃至数周,原因往往不是物理难度,而是中间存在大量必须手动打通的技术环节。这一瓶颈在需要大规模参数扫描的唯象学研究中尤为突出,严重制约了研究效率与探索空间。

与此同时,人工智能领域的快速发展正在为上述困境提供新的解决思路。特别是2025年,大语言模型(LLM)在代码生成、文档理解与多步骤推理方面的能力提升显著。而AI智能体(AI agent)技术的出现,进一步使其从“回答问题”转向“执行任务”:智能体不仅能够理解和规划复杂任务,还能自主调用外部工具、读写文件、运行程序,并在出现错误时进行自我修正^[12-16]。这一能力特征,与对撞机唯象学工作流的结构高度契合——后者本质上就是一个由多个阶段、多种工具和明确数据依赖构成的有结构执行链条^[17]。

本文介绍的ColliderAgent系统,正是在上述背景下提出的。该系统以多智能体架构为核心,结合同面向高能物理软件生态定制的以结构化领域知识为基础的智能体技能(Skills)与双层解耦的执行架构,尝试实现从自然语言描述的物理目标出发、经由完整唯象学计算流程、直至产出物理结论的端到端自动化。其目标并非以AI取代物理学家,而是提出一种新的科研协作范式:将复杂的软件调用、脚

本编排、错误修复与重复性迭代交由智能体承担,使物理学家得以将精力集中于模型假设的提出、分析策略的设计与物理结论的判断——即那些在现阶段真正不可替代的专业职能。本文以下各节将依次介绍对撞机唯象学的研究范畴与流程性挑战、AI智能体的基本概念与多智能体架构、Collider-Agent的系统设计、评测基准,以及这一工作所代表的范式意义与未来展望。

一、对撞机唯象学的研究范畴

对撞机唯象学(collider phenomenology)是连接理论物理与实验物理的核心桥梁学科。理论物理学家以拉格朗日量、场算符、耦合常数及对称性破缺机制为基本语言,描述基本粒子的相互作用;而实验物理学家在探测器中直接测量的,却是末态粒子(如电子、 μ 子、光子)、横动量等信息,以及从中间接提取的各类分布与统计显著性。两者之间,隔着一整套高度专业化的计算与模拟流程^[3]。

唯象学的核心任务之一,正是建立并执行这套流程。其基本步骤如下:从一个理论模型出发,明确所引入的新粒子内容与相互作用形式;将模型转化为费曼规则,并写成事件生成器可读取的模型文件;通过矩阵元计算程序,模拟质子碰撞中的部分子级散射过程,生成初级事例;经过部分子簇射(parton shower)与强子化(hadronization)处理,得到接近真实末态的强子事例;再借助快速探测器模拟,将强子事例转化为探测器上直接观测到的重建事例;然后再对这些事例施加筛选条件(analysis cuts)、计算截面、绘制分布、开展统计检验,最终得到排除线或发现前景。

这一流程高度依赖一套成熟的软件工具链,主要包括:用于模型定义的FeynRules^[4-5]、负责矩阵元与事件生成的MadGraph5_aMC@NLO(以下简称MadGraph)^[7]、用于部分子簇射与强子化的Pythia^[8]、实现快速探测器模拟的Delphes^[9],以及用于末态分析的MadAnalysis^[10]。这些软件各司其职、接口相互

衔接,共同构成现代高能物理(HEP)计算研究的重要基础设施。

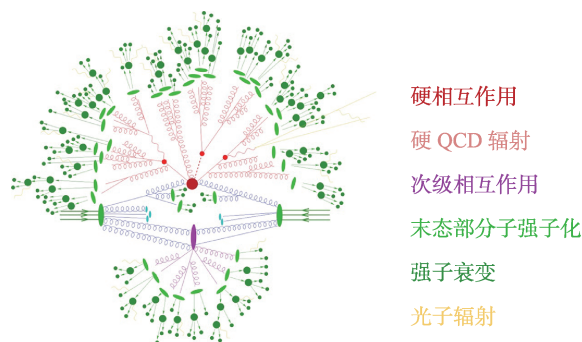


图1 强子对撞机上希格斯玻色子伴随正反顶夸克对产生过程的示意图,包括其衰变及末态强子化等各阶段^[11]。硬相互作用(大红色团块)发生后,两个顶夸克与希格斯玻色子相继衰变(小红色团块);与此同时,伴随有额外的硬QCD辐射(红色)以及次级相互作用(紫色团块);随后,末态部分子发生强子化(浅绿色团块),所产生的强子进一步衰变(深绿色团块)

二、唯象学研究中的流程性知识

在对撞机唯象学的研究中,真正耗费研究者精力的,除了理论推导本身外,就是流程性的执行工作。而且后者往往是更加耗费时间的。

所谓流程性知识,是指那些嵌入在软件操作、模型约定和计算实践中的隐性专业知识。例如:某类超出标准模型(BSM)理论模型中的粒子应如何命名,才能与FeynRules的导出逻辑保持兼容^[4-5];MadGraph的参数卡(parameter card)中的某一宽度(width)项是否应当固定,如何固定,会对截面计算产生何种影响^[7];软件执行的运算和操作与实际的物理公式有哪些隐含的区别;各软件输出格式之间存在哪些隐含的约定差异;哪些看似合理的操作会引入系统偏差,却没有在文档中明确说明^[8-10]。

这类知识往往不出现在正式教材或软件的使用手册中,而是以口口相传的方式在研究组内部流通。它决定了一项计算能否顺利进行,也决定了一个新加入者需要多长时间才能真正具备独立开展研究的能力。

因此,唯象学研究效率低下的根本原因,往往

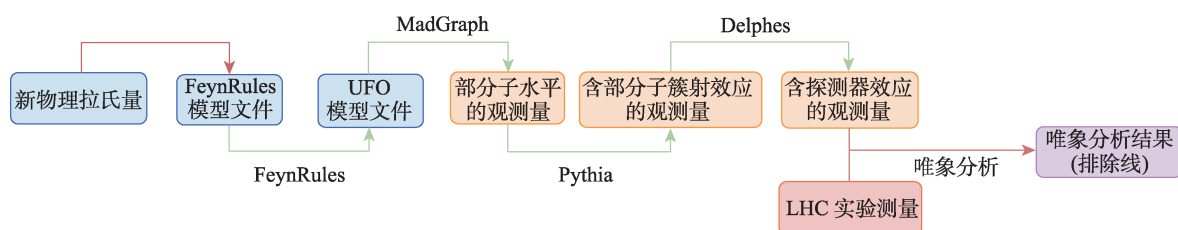


图2 对撞机物理研究的一般工作流程示意。图中每个方框代表研究过程中的一个阶段性成果；红色箭头表示需要较多人工参与(如理论推导或代码编写)的环节，绿色箭头则表示可借助现有程序包半自动化完成的环节

不在于物理问题本身的难度，而在于理论想法与可发表结果之间存在大量必须手动打通的操作环节。一个模型的物理图像也许半天就能阐明，但将其转化为一张包含完整统计分析的结果图，却常常需要数天乃至数周。而一名研究生可能需要几个月的时间才能熟悉相关软件的使用。

三、AI智能体的基本概念

近年来，人工智能领域的一个重要进展在于，大语言模型(LLM)的应用模式正从被动回答转向主动执行。这一变化催生了“AI智能体”(AI agent)的概念^[12-13]。

所谓AI智能体，是指一类不仅能够理解和规划任务，还能够自主调用外部工具、读写文件系统、执行程序代码、检查中间结果，并在必要时进行自我修正的AI执行系统。与传统的聊天式语言模型不同，智能体不以生成回答为终点，而以完成任务为目标^[12-16]。

从技术架构来看，现代AI智能体通常包含以下几个核心组件：负责任务规划与推理的语言模型主干；封装外部能力的工具集合(如代码执行环境、文件操作接口、网络搜索模块)；维护任务状态与上下文信息的记忆机制；以及组织多步骤工作流的调度逻辑。这种架构使得智能体能够处理需要多个步骤、多种工具协同配合才能完成的复杂任务^[13-17]。

值得注意的是，真实科研任务的复杂性远超单次提示所能处理的范围。一个完整的唯象学计算流程，本质上就是一个由多个阶段、多个工具和多类中间产物构成的有结构工作流。这一特点使得高能物

理成为AI智能体技术的一个极具潜力的应用场景：其软件生态成熟、接口规则明确、专业约定丰富，既对智能体的领域知识提出了较高要求，又为其执行能力的发挥提供了清晰的结构框架。

当任务的复杂程度超出单一智能体的合理处理范围时，多智能体系统(multi-agent system)是一种自然的架构延伸。其基本思路是：将一个复杂任务分解为若干子任务，分配给多个功能专一的子智能体分别执行，彼此通过结构化消息传递进行协调。这种分工方式的核心优势在于，每个子智能体只需维护与自身任务直接相关的上下文，从而避免单一智能体在处理长链条任务时因上下文积累而产生的推理质量下降。多智能体架构的具体设计，包括角色划分、协调机制与任务调度。具体内容将在第四节结合 ColliderAgent 的系统实现进一步展开^[12,14,17]。

四、ColliderAgent 的核心设计：skills、多智能体与双层解耦

ColliderAgent 的系统设计围绕三个核心概念展开：结构化的领域知识单元(skills)、面向分工协作的多智能体架构(multi-agent architecture)，以及将科学逻辑与计算执行分离的双层解耦设计。三者相互配合，共同构成从物理模型到实验结论的自动化执行链条。下文依次介绍各设计要素的原理基础与系统意义。

4.1 Skills: 结构化的领域知识单元

在智能体系统的能力扩展范式中，skills代表了一种比提示词工程和工具调用更高层次的抽象。

提示词是临时性的、难以模块化复用的;工具调用使智能体能够访问外部函数或 API,但每次调用仅执行并返回结果,不携带对任务的领域理解。Skills 则在此基础上引入了一种结构更完整的知识单元:它是一个可按需动态加载的模块化知识包,内部通常包含操作指令、工作流引导、可执行脚本、参考文档与元数据,在智能体判断当前任务与该 skills 相关时被检索并加载,从而以一致、可复现的方式完成对应的专业化操作^[18]。

在 ColliderAgent 中,skills 的具体形态与上述定义高度对应,但其内容由高能物理领域专家定制,面向唯象学计算流程的各个阶段。一个针对模型实现阶段的 skills,通常包含 FeynRules 模型文件的书写规范与字段说明、粒子命名约定与参数卡格式要求、典型 BSM 模型样例,以及导出过程中的常见问题与规避方法。智能体在执行相应任务时,不依赖对模型的泛化猜测,而是检索并调用相关 skills,按照其中的规范完成操作^[5-6,18]。

这一设计可以用实验室中传授操作经验的方式加以类比。导师带新入组的研究生,不会只给一句“把这个模型跑出来”,而是会提供说明文档、历史样例、标准模板与常见报错的处理建议。这些材料合在一起,构成了一种可传递的操作经验。Skills 机制所做的,正是将这种非正式的隐性经验,转化为可存储、可版本管理、可跨项目共享、可持续改进的结构化知识资产。在传统研究组中,大量专业知识以隐性形式存储在个人记忆中。这类知识难以传承、容易随人员流动而流失。Skills 机制为这种经验的沉淀与复用提供了一种系统性的解决路径。

4.2 多智能体架构:面向长链条科研任务的分工协作

对撞机唯象学的完整计算流程涵盖模型实现、事件生成、探测器模拟与末态统计分析等多个阶段。每个阶段的关注点、所调用的工具以及所需的推理方式均存在显著差异。如果由单一智能体全程承担,在实践中容易导致两类问题^[4,7-10]。

其一是上下文积累引发的干扰。模型实现阶段产生的大量技术细节会持续占据智能体的上下文窗口,在进行统计分析时形成无关信息的干扰,降低推理准确性。其二是角色混淆导致的推理风格不稳定。同一个智能体在不同阶段需要分别扮演理论建模者、软件工程师与物理结果分析师的角色,这种频繁切换会使其输出质量趋于不稳定。

ColliderAgent 采用多智能体分工架构来应对上述挑战。系统由若干功能专一的子智能体(sub-agent)构成,各自承担特定阶段的任务,彼此通过结构化消息传递进行协调。具体而言:协调智能体编排器(orchestrator)负责理解用户意图并完成任务分解;模型智能体负责 FeynRules 模型文件的生成与验证;模拟智能体负责 MadGraph 脚本编写、事件生成以及 Pythia 与 Delphes 的参数配置;分析智能体负责判选流程(cutflow)与分布输出;唯象解释智能体负责参数扫描、排除线计算与物理结论提炼^[12-14]。

这种分工方式的优势在于,每个子智能体只需维护与自身任务直接相关的上下文信息,信息负担显著降低,推理稳定性相应提升。这一设计与实验室中小型研究组的协作模式高度类似:每位成员专注于自己手中的那部分工作,通过明确的接口与其他成员交换信息,而非要求每个人同时掌握项目的全部细节。

4.3 双层解耦:科学逻辑与计算执行的结构性分离

ColliderAgent 在系统层面引入了另一项关键设计:将科学逻辑层(认知层)与计算执行层(基础设施层)明确分离,即双层解耦架构。

这一设计可以用大型实验项目的组织结构加以类比。实验项目通常存在两个运作层次:研究决策层关注的是科学目标的确立与任务的逻辑组织——研究哪个模型、分析哪个末态过程、选取何种统计策略;而计算基础设施层关注的则是任务的实际执行——作业提交至哪个节点、软件环境如何配置、运行状态如何监控、失败作业如何重试、输出文件

表 1 ColliderAgent 中主要 subagent 与 skill 的对应关系

subagent	核心 skill	主要职责
orchestrator	pheno-pipeline-orchestrator	分解自然语言任务,组织端到端 workflow
model-generator	feynrules-model-generator; feynrules-validator; ufo-generator	从拉格朗日量生成 .fr 与 UFO 模型,并做一致性验证
collider-simulator	madgraph-simulator	生成 MadGraph 脚本与 cards,并完成事件生成、Pythia 和 Delphes 流水线
event-analyzer	madanalysis-analyzer	执行事例筛选、cutflow 与直方图提取
pheno-analyzer	基于底层 LLM 推理	完成 profile likelihood、排除线提取、参数扫描与可视化
shared utility	magnus	作业提交、监控、重跑、结果回收与错误恢复

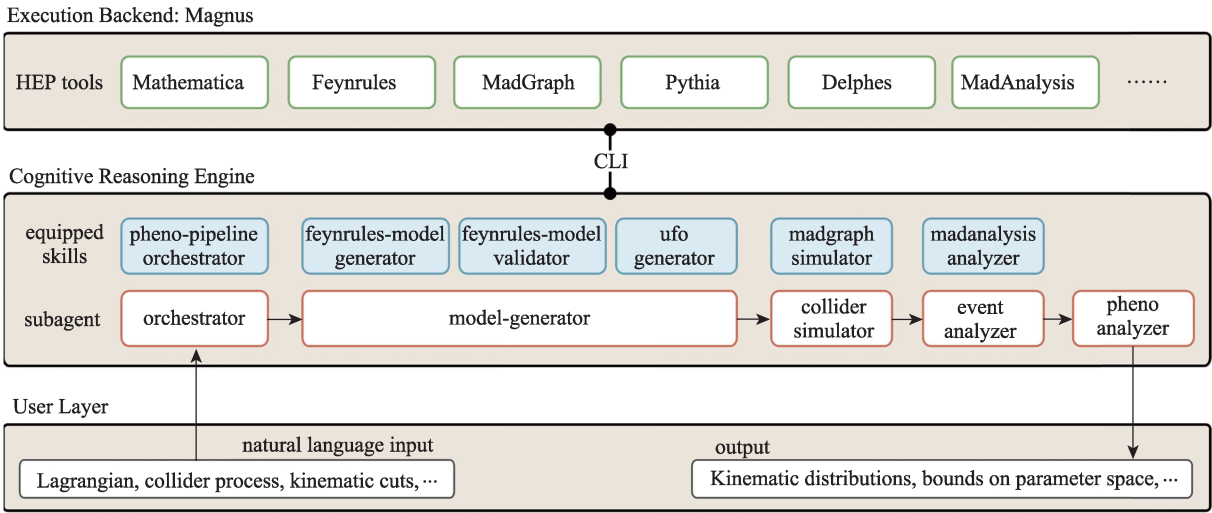


图 3

存储于何处。这两个层次虽然紧密协作,但在关注点与专业要求上存在本质区别。

在 ColliderAgent 中,上层由多智能体系统及对应的 skills 承担,负责科学任务的理解、分解、规划与结果判断;下层则由统一的计算执行后端 Magnus 负责,处理作业的实际提交、运行监控与资源管理。两层之间通过清晰定义的接口通信,互不渗透^[19]。

这种结构性分离具有两方面的系统意义。其一,避免了科学逻辑与工程细节的耦合。若两层混杂,研究者往往会将大量精力消耗在配置环境、调试作业等非物理问题上,而非聚焦于实际的科学判断。其二,提升了系统的可移植性与可扩展性。上层的任务逻辑与具体的计算平台解耦后,替换底层基础设施(如从本地集群迁移至云计算平台)无须改动上层的物理推理流程;反之,当支持的工具链从 FeynRules、MadGraph、Delphes 扩展至其他软件时,

也无须重新设计计算调度机制。

从更长远的视角来看,这种双层解耦设计为系统的持续演化预留了结构性空间。只要上下两层之间的接口保持稳定,ColliderAgent 的应用范围便可随软件生态的扩展而自然延伸,乃至推广至其他计算物理领域的类似 workflow 场景。

4.4 Magnus: 执行层的具体实现

在 ColliderAgent 的双层解耦架构中承担执行层职能的,是一个名为 Magnus 的开源计算平台。可以把 Magnus 理解为一套统一的“作业调度中心”:它将底层的高性能计算集群包装成一个标准化的执行环境,使研究者与 AI 智能体能够以完全相同的方式提交计算任务、调用各类物理软件工具,并将经过验证的计算流程保存下来,供日后反复使用^[19]。

Magnus 引入的核心概念是“蓝图”(Blueprint)。

顾名思义,蓝图是一份标准化的操作说明:它规定了完成某项计算任务需要哪些输入(例如模型文件的路径、粒子质量等参数),以及这项任务应当如何在集群上运行(例如调用哪个软件环境、分配多少计算资源)。研究者或智能体只需按照蓝图填入参数,平台便会自动完成后续的一切执行细节。这一设计的意义在于,将原本分散在个人经验中的操作知识固化为可重复调用的标准程序。以FeynRules模型验证为例:过去,研究者需要亲自知道如何正确启动验证程序、如何设置各类参数、如何处理常见报错,这些经验往往只存在于个人笔记或口头传授中,随着人员流动极易流失。有了对应的蓝图之后,任何人只需填写必要的输入信息,即可执行同一套经过反复验证的标准流程,无需重新摸索。在ColliderAgent中,五个蓝图分别覆盖了从FeynRules模型验证、UFO格式导出、MadGraph事件生成,到Pythia强子化与MadAnalysis末态分析的各个阶段,串联起从模型文件到物理结论的完整计算链条。

蓝图机制还带来了一个重要特性:人机对称。Magnus平台本身并不区分操作的发起者是人类还是智能体——研究者在网页界面上点击“提交”,与智能体通过程序接口调用同一个蓝图,触发的是完全相同的执行流程。这意味着,智能体的每一步操作对研究者而言都不是黑盒:任何时候都可以用相同的参数手动重新运行某个蓝图,独立核验智能体产出的中间结果。对于将AI引入科研流程这件事而言,这种透明性与可介入性至关重要。它让研究者始终保持对计算过程的掌控与信任。

在更底层的运行机制上,Magnus通过容器化技术保证每个计算任务在严格隔离的环境中运行,避免不同任务之间相互干扰或意外修改公共文件系统。软件环境由平台统一管理,首次使用时自动下载配置,之后可直接复用,研究者无需手动维护复杂的软件依赖。当多个用户或多个智能体同时提交作业时,调度器会根据优先级合理分配计算资源,保障整体运行效率。

最后值得指出的是,Magnus本身并不包含任

何高能物理特有的逻辑。它的全部功能(蓝图定义、作业调度、环境隔离、文件管理)都是通用的科学计算基础设施。ColliderAgent是Magnus上的第一个大规模应用,但同样的执行架构可以自然适用于其他具有类似 workflow 结构的计算科学场景,例如凝聚态物理中的第一性原理计算、计算化学中的分子动力学模拟,乃至任何需要多阶段、多工具协同运行的科学计算任务。这种通用性,正是4.3节所述双层解耦设计在基础设施层面的直接体现。

五、系统评测:端到端论文复现作为基准任务

衡量一个面向科研的AI系统是否真正有效,最直接的标准是:它能否在给定自然语言描述的情况下,独立完成一项具有实际物理意义的完整计算任务。为此,本工作构建了一套以已发表的对撞机唯象学论文为基础的端到端复现任务集,作为系统能力的评测基准^[20,25]。

任务集的覆盖范围经过刻意设计,以确保对典型唯象学工作流的全面检验。具体而言,既包含部分子级(parton-level)分析,也包含探测器级(detector-level)分析;既涵盖具体BSM模型,也涉及有效场论(EFT)框架;评测指标从单一运动学分布的复现延伸至二维参数扫描与排除线提取;对撞机实验设定则同时考虑LHC当前运行条件与未来对撞机方案。所有任务均以自然语言prompt形式给出,系统须自主完成从模型实现到物理结论的全部中间步骤,不提供任何预先编写的可执行代码。

以下几个代表性案例说明了任务集所覆盖的典型难度层次。

图4(a)单标量leptoquark共振产生。Leptoquark是一类连接轻子与夸克的假想粒子,在多种新物理模型中频繁出现。该任务的关键难点不在于物理概念本身,而在于过程依赖质子内部的轻子成分,需使用包含轻子PDF的特殊初态设定,且后续模拟工具对入射粒子类型存在格式限制,须进行

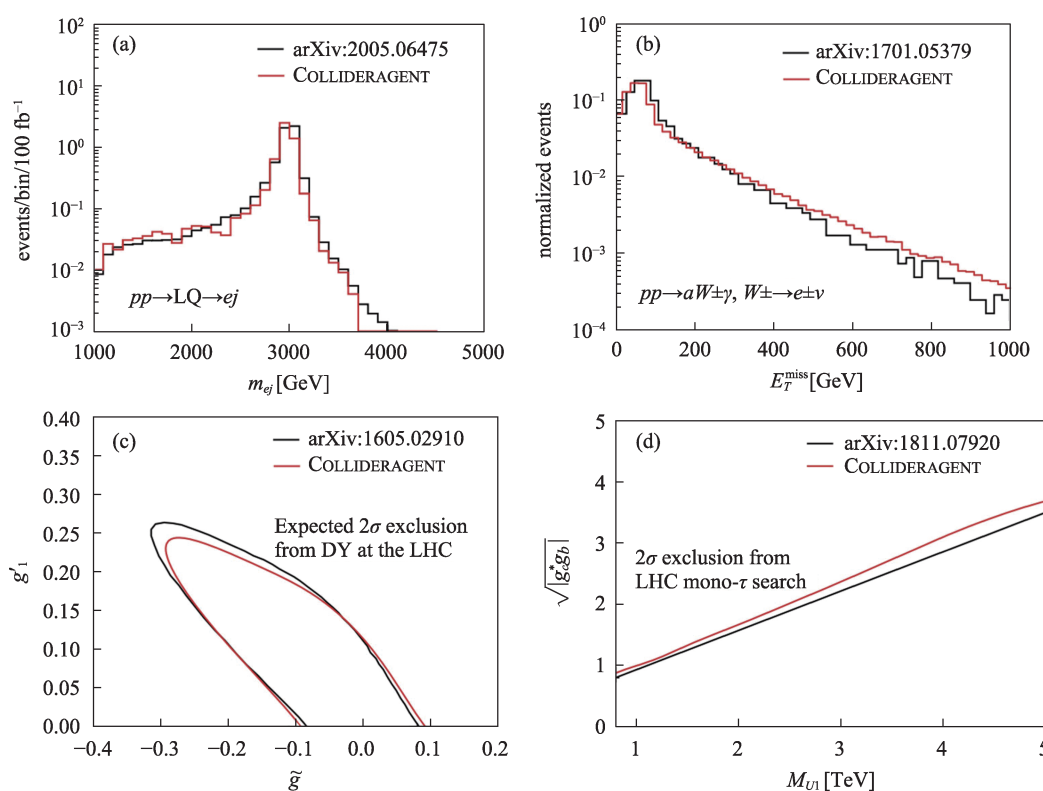


图4

额外的兼容性处理。这类细节是真实唯象学工作中最常见的时间消耗点,对系统的流程适应能力构成实质性检验^[21-22]。

图4(b)ALP有效场论算符与缺失横能分布。类轴子粒子(axion-like particle, ALP)的EFT基准任务要求系统从高维算符出发,经过顶角结构的正确实现,复现含不可见粒子末态的归一化缺失横动量分布。该任务验证的是系统在“理论输入—过程生成—可观测输出”这条完整逻辑链上的贯通能力,而非单纯的软件调用^[23-24]。

图4(c)U(1)扩展模型的二维参数扫描。对含重规范玻色子Z'的U(1)'模型进行二维参数扫描,并提取不同Z'质量下的排除区域。此类任务的主要难度在于工程层面:脚本组织、批量作业提交、失败重试、结果汇总与统一绘图,在参数空间较大时会迅速演变为繁重的重复性劳动。该基准评测的正是系统对此类高度重复计算流程的自动化处理能力^[25,27]。

图4(d)U1向量leptoquark的mono-τ探测器级分

析。这是任务集中流程最完整的案例,要求系统从信号模型模拟出发,经Delphes快探测器模拟、实验判选条件施加,直至基于profile likelihood的统计解释,最终给出排除线。这一任务代表了当前系统所能处理的最完整意义上的碰撞物理研究——其复杂度已接近一篇典型唯象学论文的核心计算部分^[28-30]。

评测结果表明,ColliderAgent在上述任务上的多次独立复现中总体成功率较高,且复现得到的物理结果与论文基本吻合。这一结论表明,将自然语言描述的物理问题转化为完整计算结果,这条技术路径已具备实质可行性,并在若干代表性任务上得到了系统性验证^[20-21,23,25,28]。

六、ColliderAgent所代表的范式创新

ColliderAgent的核心价值,并不在于“让AI学会了又一套物理软件”,而在于它揭示了一种新的科研分工方式。

在传统的对撞机唯象学研究中,大量能力被集中于单个研究者身上:既需要理解理论模型,又需要熟悉软件工具链,还需要编写脚本、调试报错、监控作业、处理格式兼容问题。一项本应以物理判断为核心的工作,常常被大量执行性劳动所裹挟。ColliderAgent提出的思路,是将这两类性质不同的能力明确拆分。其一是执行性能力,包括代码生成、软件调用、参数配置、批量作业提交、错误修复与输出整理;其二是专业性能力,包括判断模型是否物理合理、选择敏感的信号区域、确定具有物理意义的判选条件、对统计结果作出正确解释,以及将结论置于领域整体背景下加以评价。在这种新分工下,前者可以越来越多地交由AI智能体承担,后者则仍然由物理学家主导。

这一分工的意义在于重新凸显物理学家最不可替代的价值,而非削弱其作用。当前,智能体再擅长执行计算流程,也无法自主判断一个模型是否值得研究;再能够生成精确的排除线图,也无法天然理解其背后的物理含义;再能够顺利跑通全部代码,也不能替代人类作出“数值结果虽然出来了,但物理上站不住脚”这类判断。因此,这篇工作所提出的,本质上是一种“AI负责执行、专家负责判断;AI负责搬运、专家负责建模;AI负责迭代、专家负责定义问题”的新型科研范式。

七、未来展望

从更长远的角度审视,ColliderAgent的意义也许不止于对撞机唯象学本身,而在于它所指向的一种更普遍的变化:领域知识开始被系统地沉淀为可调用、可复用、可版本化的操作结构^[18-19]。

Skill机制是这种结构的具体体现。它使原本散落在论文附录、软件文档、课题组口耳相传与个人工作笔记中的隐性经验,开始有机会被组织成真正可执行的知识资产。这类资产今天服务于一个智能体系统,明天同样可以服务于新入组的学生、跨课题组的合作者,乃至其他学科中具有类似 workflow 结构的研究场景。从这个意义上说,Collider-

Agent所展示的,不只是“AI完成了一项物理任务”,而是“物理学的程序性知识,第一次开始被较系统地工程化”。当知识可以被工程化,它便更容易被共享、被累积、被改进,也更有可能演变为真正意义上的学科基础设施。

因此,未来科研效率的实质性提升,可能并不主要来自新一代基础模型的能力跃迁,而是来自这样一种结构性转变:强大的智能体系统承担基础代码撰写、软件调用、流程迭代、批量计算与错误修复;物理学家则专注于提供真正稀缺的资源——领域知识、模型理解、分析策略与物理判断,并将这些内容进一步具体化为 skills、模板与文档。AI正在成为高能物理研究的“执行层基础设施”,而专业研究者继续扮演“定义问题、保证正确、赋予物理意义”的角色。从拉格朗日量到排除线,这中间曾经横亘着一条漫长而耗人的技术链。ColliderAgent所做的,是第一次较为完整地展示:这条链,正在被重新组织。

参考文献

- [1] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC[J]. Physics Letters B, 2012, 716: 1-29.
- [2] CMS Collaboration. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC[J]. Physics Letters B, 2012, 716: 30-61.
- [3] Boyle P, Pedro K, Qiang J. CompF2: Theoretical Calculations and Simulation Topical Group Report[R]. 2022.
- [4] Christensen N D, Duhr C. FeynRules-Feynman rules made easy[J]. Computer Physics Communications, 2009, 180: 1614-1641.
- [5] Alloul A, Christensen N D, Degrande C, et al. FeynRules 2.0 - A complete toolbox for tree-level phenomenology[J]. Computer Physics Communications, 2014, 185: 2250-2300.
- [6] Degrande C, Duhr C, Fuks B, et al. UFO-The Universal FeynRules Output[J]. Computer Physics Communications, 2012, 183: 1201-1214.
- [7] Alwall J, Frederix R, Frixione S, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations[J]. JHEP, 2014, 07: 079.
- [8] Bierlich C, et al. A comprehensive guide to the physics and usage of PYTHIA 8.3[J]. SciPost Physics Codebases, 2022: 8.
- [9] de Favereau J, Delaere C, Demin P, et al. DELPHES 3, A modular

- framework for fast simulation of a generic collider experiment[J]. JHEP, 2014, 02: 057.
- [10] Conte E, Fuks B, Serret G. MadAnalysis 5, A User-Friendly Framework for Collider Phenomenology[J]. Computer Physics Communications, 2013, 184: 222-256.
- [11] Gleisberg T, Hoeche S, Krauss F, et al. Event generation with SHERPA 1.1, JHEP, 2009, 02:007.
- [12] Yao S, Zhao J, Yu D, et al. ReAct: Synergizing Reasoning and Acting in Language Models[C]//ICLR. 2023.
- [13] Yang J, Jimenez C E, Wettig A, et al. SWE-agent: Agent-Computer Interfaces Enable Automated Software Engineering [EB/OL]. arXiv:2405.15793, 2024.
- [14] Wang X, et al. OpenDevin: An Open Platform for AI Software Developers as Generalist Agents[EB/OL]. arXiv:2407.16741, 2024.
- [15] Bran A M, Cox S, Schilter O, et al. ChemCrow: Augmenting large-language models with chemistry tools[J]. Nature Machine Intelligence, 2024, 6: 525-535.
- [16] Boiko D A, MacKnight R, Kline B, et al. Autonomous chemical research with large language models[J]. Nature, 2023, 624: 570-578.
- [17] Zhou S, Xu F F, Zhu H, et al. WebArena: A Realistic Web Environment for Building Autonomous Agents[C]//ICLR. 2024.
- [18] Anthropic. Equipping agents for the real world with Agent Skills [EB/OL]. 2024.
- [19] Cai Z Y, Wei J S. Magnus: An agentic infrastructure automating scientific discoveries[EB/OL]. 2026.
- [20] Badea A, Chen Y, Lee Y-J. Agentic AI-Physicist Collaboration in Experimental Particle Physics: A Proof-of-Concept Measurement with LEP Open Data[EB/OL]. arXiv:2603.05735, 2026.
- [21] Buonocore L, Haisch U, Nason P, et al. Lepton-Quark Collisions at the Large Hadron Collider[J]. Physical Review Letters, 2020, 125: 231804.
- [22] Buonocore L, Nason P, Tramontano F, et al. Leptons in the proton [J]. JHEP, 2020, 08: 019.
- [23] Brivio I, Gavela M B, Merlo L, et al. ALPs Effective Field Theory and Collider Signatures[J]. European Physical Journal C, 2017, 77: 572.
- [24] Georgi H, Kaplan D B, Randall L. Manifesting the Invisible Axion at Low-energies[J]. Physics Letters B, 1986, 169: 73-78.
- [25] Accomando E, Coriano C, Delle Rose L, et al. Z-prime, Higgses and heavy neutrinos in U(1)-prime models: from the LHC to the GUT scale[J]. JHEP, 2016, 07: 086.
- [26] Basso L, Belyaev A, Moretti S, et al. Phenomenology of the minimal B-L extension of the Standard model: Z-prime and neutrinos [J]. Physical Review D, 2009, 80: 055030.
- [27] CMS Collaboration. Search for resonant and nonresonant new phenomena in high-mass dilepton final states at $\sqrt{s}=13$ TeV [J]. JHEP, 2021, 07: 208.
- [28] Greljo A, Martin Camalich J, Ruiz-Alvarez J D. Mono-tau Signatures at the LHC Constrain Explanations of B-decay Anomalies [J]. Physical Review Letters, 2019, 122: 131803.
- [29] ATLAS Collaboration. Search for High-Mass Resonances Decaying to tau nu in pp Collisions at $\sqrt{s}=13$ TeV with the ATLAS Detector[J]. Physical Review Letters, 2018, 120: 161802.
- [30] CMS Collaboration. Search for a W-prime boson decaying to a tau lepton and a neutrino in proton-proton collisions at $\sqrt{s}=13$ TeV[J]. Physics Letters B, 2019, 792: 107-131.

科苑快讯

科学家们对地核深处的突然逆转感到困惑

太平洋下铁水流动的惊人逆转,让科学家们对地球磁场的演变有了清晰的认识。

地球磁场被认为主要是由在我们脚下3000千米处组成地球外核的过热、旋转的液态铁海洋产生的。它就像自行车发动机中旋转的导体一样,产生电流,从而产生不断变化的电磁场。

科学家以前认为,外核是一个运动相对稳定的系统,然而现在其循环变化可能比预期要突然得多。这项研究为创造地球磁场的湍流过程提供了新的线索,并可能指出外核的运动与地球内部更深处发生的变化之间的联系。

科学家现在需要弄清楚,太平洋下大规模的水流逆转,是代表了一种短暂的波动,是重复振荡的一部分,还是核心循环的一种新的稳定平衡。

研究人员研究了欧空局(ESA)的Swarm卫星观测之后表明,这种逆转可能是暂时振荡或岩心动力学中较长自然周期的一部分。这些发现还可以帮助科学家对研究地球外核、内核核下地幔边界有了更多的了解,这是地球深部动力学的关键区域。

(高凌云编译自2026年7月6日SciTechDaily网站)