

丘成桐中学科学奖 · 物理奖赛道候选课题 40 则

本方案面向 2027 赛季（2026 赛季提交截止为 2026 年 9 月 15 日，从零启动已来不及）。倒排时间表以 2026 年 9 月为第 0 周、2027 年 9 月 15 日提交为终点，共约 52 周。

学生画像与资源假设：普通重点高中生，有一定编程与数学基础；计算资源为个人笔记本、纯 CPU、无 GPU 无集群；实验资源不超过中学常规实验室；全部使用公开数据；由中学教师即可指导。凡超出这一假设的路线，都在标签里标明了资源门槛。

四条必须先知道的赛制事实（取自官方公告，2020 – 2026 全部公示文章）：指导老师不得来自培训机构或任何谋利机构，违者取消资格；总决赛全程英文（研究报告、PPT、答辩）；入围总决赛的研究报告会在 11 月全网公示；提交材料含论文查重报告与实验视频。规则每年会改，报名开放后须重新核对当年《参赛规则》与本学科《参赛指南》。

这个赛道的评委在奖励什么

统计 2020 – 2025 共 65 项物理奖获奖论文，结论非常干脆：这是六个赛道里唯一一个由桌面实验主导的赛道。

"厨房物理"是这个赛道的正统。历届金奖依次是：绳子自发打结的物理机制、葡萄等离子体（燃烧还是放电）、软物质弹性模量的改进测量法、听声辨球数（声学特征 + 机器学习）、旋转诱导的自发对称破缺与滞回环。银铜奖里是黏性液滴变形、环在碗中的混沌运动、意大利面为什么断成三截、纸张交错堆叠的摩擦放大、液体晃荡的非线性动力学。共同特征：现象来自日常生活，装置自制，但分析深度做到了量纲分析 + 标度律 + 理论模型 + 实验验证的完整闭环。

流体与软物质是最大的单一门类，约占三分之一：液滴撞击、水跃、射流、沉降、渗流、毛细。声学与振动其次（管道共鸣、声悬浮、音乐盒、金属板啸叫）。这两族的器材成本最低，恰好匹配普通高中生的资源条件。

天体物理与数据分析类存在，但天花板明显偏低。TESS 假阳性识别管线、M82 星系中心物理条件、暗扇区相互作用解哈勃张力、系外行星大气反演——全部止步于铜奖或优胜奖，六年里没有一篇进入金银。这不代表不该做，而是意味着：如果走数据分析路线，必须自建独立管线并做出可被检验的物理判断，不能只是"用公开数据跑了一遍现成方法"。2025 的银奖是自研物端棱镜光谱仪 + 配套 AI 分类模型——注意它赢在自己造了仪器。

反复获奖的学校：重庆育才、南京外国语、人大附、北师大实验、北京十一学校。这些学校有稳定的物理实验传统与器材积累。普通学校要竞争，最可行的路径是选一个器材极简、但现象本身尚未被系统定量的日常现象。本方案 P11 – P20 全部按"总预算 ≤ 3000 元、无高校仪器依赖"设计。

本赛道 20 条路线的分族逻辑

- P01 – P10 计算、理论与公开观测数据族：适合动手条件差、但编程与数学更强的学生。结构性优势是没有器材风险与实验重复性风险；结构性劣势是要在评委的口味逆风里靠分析深度取胜，因此每条路线都强制要求自建管线并做方法学校验。
- P11 – P20 桌面实验与低成本测量族：正面迎合这个赛道的传统。核心竞争力不是现象多新奇，而是测量不确定度控制得多好——每条路线都写死了误差源、重复次数与标度拟合的稳健方法。丘奖要求提交实验视频，这一族的记录规范也写进了方法路径。

两族独立排序：P01/P11 最稳，P10/P20 风险与上限最高。

P01 · TESS 扩展任务新扇区对热木星 WASP-12b 类轨道周期衰减的自建凌星计时管线独立检验

推荐优先级 ★★★★★ · 天体物理公开数据 · 纯 CPU 轻量 · 时序拟合 / MCMC · 结论正负均有效

1 · 研究问题

用自建的 lightcurve + batman + emcee 凌星计时管线，从 TESS 扩展任务（Extended Mission）2023 – 2026 年新扇区光变曲线中重建 WASP-12b 等热木星的中凌时刻（mid-transit time），拟合出的轨道周期衰减率 dP/dt 与已发表值（ $32.53 \pm 1.62 \text{ ms/yr}$ ）是否在 95% 置信区间内一致？对照组（无衰减报道的热木星）的 dP/dt 是否与零一致，从而证明管线无系统偏置？

2 · 研究背景与空白

热木星（hot Jupiter）因潮汐耗散会向宿主恒星缓慢旋进，轨道周期衰减（orbital decay）表现为中凌时刻相对线性星历（ephemeris）的二次偏离。测量它只需要长期、均匀的凌星计时——这是公开测光数据能直接支撑的少数正在发生的天体演化"测量"之一。

已有工作：Patra et al. (2017) 与 Yee et al. (2020) 确立 WASP-12b 为轨道衰减最强证据；Turner et al. (2021, AJ) 用 TESS 主任务数据确认衰减率 $32.53 \pm 1.62 \text{ ms/yr}$ ；Leonardi et al. (2024, A&A, TASTE V) 用地面数据延长基线；2025 年 arXiv:2510.06589 结合新凌星时刻重新分析轨道动力学。方法（凌星模型拟合 + 星历模型比较）完全公开。

空白在时间覆盖：TESS 扩展任务持续产生新扇区，任何"截止到 2026 – 2027 年最新数据"的联合计时分析都是天然的新样本独立检验；且加入一组"应为零信号"的对照热木星做管线零点检验的做法在学生可及层面并不常见。方法门槛低、数据全公开、衰减"看到/没看到"都构成有效结论，非常适合一年期课题。

3 · 可检验假设

- **H1:** 2023 – 2026 新扇区中 WASP-12b 的中凌时刻相对常周期线性星历的偏离 $> 3\sigma$ ，且二次星历拟合给出的 dP/dt 与 Turner et al. (2021) 的 $32.53 \pm 1.62 \text{ ms/yr}$ 在 95% 置信区间内一致。
- **H2（管线零点假设）：**5 – 10 颗文献未报道衰减的对照热木星，用同一管线拟合的 dP/dt 与 0 在 2σ 内一致；若不一致，说明管线存在系统偏置，H1 的结论无效。

4 · 量化验收标准

1. **方法学校验（硬门槛）：**用自建管线对 10 颗已确认凌星行星重跑单凌星拟合， R_p/R_* 、 a/R_* 、凌星中点与 NASA Exoplanet Archive 公布值偏差 $\leq 5\%$ （中凌时刻偏差 ≤ 1 分钟）。此步不过关，后续全部结论无效。
2. 每个中凌时刻附 MCMC 后验分布与收敛诊断（链长 ≥ 50 倍积分自相关时间，或 Gelman – Rubin $\hat{R} < 1.01$ ）。
3. 星历模型比较预先写死：线性 vs 二次模型用 $\Delta\chi^2$ 与 BIC 双口径报告； dP/dt 置信区间用自助抽样（bootstrap）给出，不单独依赖单次最小二乘协方差。
4. 对照样本 ≥ 5 颗零结果检验全部报告（不得只报 WASP-12b）。
5. 全部脚本、凌星时刻表与后处理开源，第三方可一键重跑。

5 · 数据与工具

用途	来源 / 工具
TESS 光变曲线	MAST 归档, SPOC 2 分钟 PDCSAP 光变; 用 <code>lightkurve</code> 按目标、按扇区检索下载 (单目标单扇区约数十 MB, 全项目 < 5 GB)。覆盖 2018 至今全天分区, 每扇区约 27 天
凌星物理模型	<code>batman</code> (内置 Mandel–Agol 凌星模型与多种临边昏暗律; TTV 星历拟合脚本需自行实现)
后验采样	<code>emcee</code> (内置仿射不变采样器; 收敛诊断需自行调用 <code>autocorr</code> 模块)
历史凌星时刻	文献汇编表 (Yee 2020 / Turner 2021 / Leonardi 2024 附表) —— 仅作联合拟合的历史锚点与对照, 不计入本项目的数据贡献
行星参数先验	NASA Exoplanet Archive—— 仅作校验与先验, 不计入贡献

算力: `batman` 单次模型评估毫秒级; 单凌星 MCMC ($32 \text{ walkers} \times 2 \times 10^4$ 步) 数分钟; 全项目约数百次拟合, 纯 CPU 笔记本一至两天内完成。**降规模版本:** 只做 WASP-12b + 3 颗对照; H1 主结论完整保留, 仅损失对照组统计力。

6 · 方法路径

1. 装环境, 跑通 `lightkurve`/`batman`/`emcee` 官方示例, 完成第 4 块硬门槛校验。
2. 核实 WASP-12 天区 2023 – 2026 年是否有新 TESS 扇区 (Web TESS Viewing Tool), 并确定对照热木星名单 (明确、可复现的入选判据: $V < 13$ 、周期 < 3 天、扇区数 ≥ 3)。
3. 实现逐凌星切窗、去趋势与单凌星 MCMC 拟合, 输出中凌时刻及不确定度。
4. 与文献历史时刻联合, 做线性/二次星历模型比较, `bootstrap` 给 dP/dt 置信区间。
5. 对照组重复同一流程, 做零点检验。
6. 交叉校验: 用与主管线独立的梯形模型 + 最小二乘重估中凌时刻, 比较两法差异。

7 · 新颖性边界

- 本课题不声称发现轨道衰减——WASP-12b 衰减已由 Patra/Yee/Turner 等发表; 也不提出新计时方法。
- 已有工作: Turner et al. (2021) 已用 TESS 主任务扇区确认衰减; Leonardi et al. (2024) 与 2025 年重分析已延长地面基线。
- 本项目贡献 (主结论): 自建管线 + 截至 2026/27 年最新 TESS 扇区的扩展基线独立检验 + 对照组零点检验。若 dP/dt 一致, 普适性获得一次独立支持; 若偏移, 给出偏移量与可能的采样/系统因素分析——两种结果都有效。
- 与历届获奖作的关系: 2021 年铜奖"TESS TOI 假阳性自动识别管线"做的是分类问题, 与本题 (精密计时) 无重叠。
- 价值: 轨道衰减率数值随基线增长而收紧, 任何新的独立时刻表对该领域都有实际增量。

8 · 决策门槛 (go / no-go)

- 第 6 周末: 硬门槛校验须通过 (10 颗行星、偏差 $\leq 5\%$)。不过关 $\rightarrow$ 第 8 周前只保留官方 SPOC 参数复核, 课题降级为"多行星 TESS 星历维护与 TTV 搜索" (同一拟合框架, 目标从 dP/dt 改为线性星历更新),

主结论框架（计时 + 模型比较）不变。

- 第 10 周：核实清楚 WASP-12 天区新扇区可用性（不要边做边发现）。若 2023 年后无新扇区 → 降级路径 A：用全部既有扇区做联合重分析 + 对照组零点检验（仍是独立检验）；降级路径 B：主目标换为 Kepler-1658b 等其他衰减候选（判据：已有 ≥ 2 个 TESS 扇区）。
 - 第 30 周：冻结样本与统计口径，进入写作与复跑验证；预算裁剪顺序：对照组数量 → 交叉校验方法数 → 主目标不可裁。
 - 适合学生：会 Python、愿意学 MCMC 的任何队伍；本题是十题中失败模式最少的一条。
-

P02 · 二维稀释伊辛模型临界行为的 Wolff 集团蒙特卡洛检验：弱普适性与对数修正两种图像的定量区分

推荐优先级 ★★★★★ · 统计物理 / 蒙特卡洛 · 纯 CPU (需 numba 加速) · 有限尺寸标度分析 · 无数据下载风险

1 · 研究问题

位稀释 (site-diluted) 二维伊辛模型的表现观临界指数随稀释度漂移，应解读为"弱普适性" (ν 连续变化、 γ/ν 不变) 还是"纯伊辛指数 + 对数修正"? 在笔记本可达的系统尺寸 $L \leq 256$ 上，有限尺寸标度 (finite-size scaling) 能否定量区分这两种图像，若不能，区分所需的最小 L 是多少?

2 · 研究背景与空白

二维伊辛模型有 Onsager 精确解 ($T_c = 2/\ln(1+\sqrt{2})$, $\nu = 1$, $\gamma/\nu = 7/4$)，是检验一切自旋模拟管线的黄金标尺。加入无序 (随机置空格点) 后，Harris 判据在 $\alpha = 0$ 处处于边界情形，理论预言临界行为极其微妙：比热从 $\ln L$ 变为 $\ln \ln L$ 发散。

已有工作分成两派：Ballesteros et al. (cond-mat/9707179) 的大规模蒙特卡洛支持"纯伊辛指数 + 对数修正"；而基于 Wang – Landau 算法的研究 (arXiv:0807.0075, Physica A 2008) 报告 $\nu > 1$ 、 γ/ν 保持 $7/4$ ，支持"弱普适性"解读。两派使用的系统尺寸多在 $L \leq 128$ 量级，无序样本数 $10^2 - 10^3$ 。

空白在问题方向： 争论的焦点其实是"同一批数据用两种模型都能拟合"——很少有工作在同一套现代统计口径 (bootstrap 置信区间、预注册拟合区间、无序涨落显式报告) 下对两种图像做正面模型比较。这正适合学生：物理答案已知候选、模拟代码百行量级、算力标尺文献明确、"区分不了"本身也是定量结论。

3 · 可检验假设

- H1: 对稀释度 $p \in \{1.0, 0.9, 0.8, 0.7\}$ (p 为格点保留率)，磁化率比值 γ/ν 与纯伊辛值 $7/4$ 在 95% 置信区间内一致 (弱普适性与对数修正两派在此点预测相同，作为管线一致性检验)。
- H2: 若把关联长度数据强行按纯幂律拟合，得到的有效 $\nu(p)$ 随稀释单调上漂；该漂移的幅度与形状可被" $\nu = 1 + \text{对数修正}$ "模型在等自由度下拟合得不差于" ν 连续变化"模型 (用 AIC/交叉验证判定)。

4 · 量化验收标准

1. 方法学校验 (硬门槛)：纯二维伊辛 ($p = 1$)、 $L = 32 - 256$ ：Binder 累积量交叉定出的 T_c 与精确值偏差 $\leq 0.1\%$ ， ν 、 γ 、 β 与精确值偏差 $\leq 2\%$ 。此步不过关，后续全部结论无效。
2. 每个 (p, L) 点无序平均 ≥ 500 个独立样本，显式报告样本间涨落与热化判据 (热化步数 ≥ 20 倍积分自相关时间)。
3. 拟合口径预先写死：双对数最小二乘之外，加非线性最小二乘与含修正标度项拟合交叉验证；全部指数误差用自助抽样给置信区间。
4. Wolff 集团算法与 Metropolis 在 $L = 16$ 小系统上互检：能量、磁化统计量一致性 ≤ 0.5 个合并标准差。
5. 模型比较用双成本口径：等自由度与等参数个数下的 AIC/BIC 都报告。
6. 全部代码与原始测量数据开源，第三方可一键重跑。

5 · 数据与工具

用途	来源 / 工具
模拟引擎	自写 Python + numba (Wolff 集团翻转与 Metropolis 均需自行实现, 约 200 行; 无现成框架依赖是本题优点)
精确解对照	Onsager 解析结果 (教科书级) —— 仅作参考, 不计入贡献
文献基准	cond-mat/9707179 与 arXiv:0807.0075 的表列指数—— 仅作参考, 不计入贡献
分析	numpy/scipy/matplotlib; bootstrap 自写 (约 30 行)

算力 (基于文献标尺): 1997 年 Ballesteros 等在当年硬件上完成 $L \leq 128$ 、 10^3 无序样本的研究; 现代单核 Wolff 实现典型速度 $\sim 10^7$ 自旋更新/秒 (numba 实测速度需第 5 周核实)。估算: $L = 128$ 、每样本 10^4 测量扫略、500 样本 $\approx 8 \times 10^{13}$ 自旋更新 $\approx$ 单核 2-3 天后台; 4 个 p 值 $\times$ 5 个 L 值总计约 3-4 周后台运行, 可行。降规模版本: $L \leq 96$ 、样本 300 $\rightarrow$ H1 (γ/ν 检验) 完整保留; 损失的是对数修正与弱普适性的区分能力, 此时主结论转为"给出区分两图像所需最小 L 的定量判据"。

6 · 方法路径

1. 实现 Metropolis 单自旋翻转, 在 $L = 16$ 纯伊辛上跑通并对照精确内能曲线。
2. 实现 Wolff 集团算法, 完成两算法互检与第 4 块硬门槛校验。
3. 核实 numba 加速后的更新速度 (阈值见第 8 块), 据此冻结 (p, L) 网格。
4. 对每个 (p, L) 采集能量、磁化、磁化率、Binder 累积量, 无序平均并记录涨落。
5. 做有限尺寸标度: Binder 交叉定 $T_c(p)$, 数据塌缩 (data collapse) 与幂律拟合并行。
6. 对" ν 漂移"与"对数修正"两模型做预注册的正面比较 (AIC/BIC + 交叉验证)。
7. 独立交叉校验: 用比热的 $\ln L$ vs $\ln \ln L$ 发散形状做与指数拟合无关的第二判据。

7 · 新颖性边界

- 本课题不声称发现新普适类, 也不声称解决弱普适性争论——该争论已有 25 年文献史 (点名上述两派)。
- 已有工作: 两派各自给出了指数数值与解读, 但采用的误差口径、拟合区间选择互不相同, 直接比较困难。
- 本项目贡献 (主结论): **同一管线、同一套预注册统计口径下对两种图像的正面定量比较与独立复现**, 并给出笔记本尺度下"可区分/不可区分"的明确边界。这属于方法层面的可复现性贡献 + 独立检验, 在计算物理中是被承认的贡献类型, 但定位必须在论文中讲清楚, 否则会被误读为重复工作。
- "在 $L \leq 256$ 内两模型不可区分"同样是有效结论, 但必须给出误差棒证明有能力分辨设定的差异幅度。
- 历届获奖图谱中无自旋模型蒙特卡洛先例, 题材上避开了实验流体类的拥挤赛道。

8 · 决策门槛 (go / no-go)

- 第 4 周末: 硬门槛校验 (纯伊辛指数 $\leq 2\%$) 通过; 不过关 $\rightarrow$ 排查实现而非继续加规模。
- 第 5 周末: 核实 numba 实现速度。若 $< 5 \times 10^6$ 自旋更新/秒 $\rightarrow$ 具名降级: 其一, 改用 C 扩展重写核心循环 (一周预算); 其二, 直接采用降规模版本 ($L \leq 96$), 主结论框架 (两模型比较) 不变。

- 第 20 周末：若在已达尺寸内两模型 AIC 差 < 2 （不可区分） $\rightarrow$ 把主结论正式切换为"最小可区分 L 的外推判据"，补充一组 $L = 320$ 的单点长跑（后台 1-2 周）作为外推锚点。
 - 预算裁剪顺序：p 值个数（4 $\rightarrow$ 3） $\rightarrow$ 交叉校验判据数 $\rightarrow$ L 网格密度；硬门槛与无序样本下限不可裁。
 - 适合学生：对编程与统计有耐心、能接受"负结果也要写清楚"的队伍；无任何外部数据依赖，是十题中不可抗力风险最低的一条。
-

P03 · Gaia DR3 疏散星团成员的自建聚类判定管线与质量分层-年龄关系的统一口径独立检验

推荐优先级 ★★★★★ · 天体物理公开数据 · ADQL + 无监督聚类 · 纯 CPU 极轻量 · 统计取向

1 · 研究问题

用自建的五维天测（位置 + 自行 + 视差）HDBSCAN/GMM 聚类管线对约 20 个疏散星团（open cluster）重建成员表后，在统一的质量分层指标 Λ_{MSR} 口径下，“动力学年龄越老分层越强”的已发表趋势是否在一个独立选取的星团样本上成立？

2 · 研究背景与空白

疏散星团中的大质量星通过两体弛豫向中心沉降（质量分层，mass segregation），分层程度随星团年龄与弛豫时间（relaxation time）之比增长，是恒星动力学的教科书预言。Gaia DR3 的微角秒级天测使得从公开数据里“从零”重建星团成员表成为高中生可完成的操作。

已有工作：Cantat-Gaudin et al. (2020) 给出千余星团的成员星表；A&A 2023 (arXiv:2303.08474) 发布了基于深度网络的 DR3 成员判定工具；MNRAS 2023 (arXiv:2305.17728) 用 DBSCAN+GMM 方法在样本中找到 10 个星团的明确分层证据，并报告老星团分层速率更高。指标端，Allison et al. (2009) 的最小生成树比值 Λ_{MSR} 是标准分层度量。

空白在评价维度与样本：已发表的分层-年龄结论来自不同作者、不同成员判定口径、不同质量代理的异质汇编；用同一条管线、同一个 Λ_{MSR} 定义、同一质量代理贯穿一个独立星团样本的均匀检验，正是“换样本独立检验”这一最稳差异化轴。数据量小、算法现成、正负结论都有效。

3 · 可检验假设

- H1：以 $\log(\text{age}/t_{\text{relax}})$ 为横轴， Λ_{MSR} （前 10 大质量成员）显著 > 1 的星团比例随之上升；Spearman 秩相关 $\rho > 0$ 且 bootstrap $p < 0.05$ 。
- H2：存在 $\text{age} < t_{\text{relax}}$ 却 $\Lambda_{\text{MSR}} > 1$ 的年轻星团个例（原初分层候选）；若一个都没有，“分层皆为动力学演化产物”在本样本获得支持——两种结果都有效。

4 · 量化验收标准

1. 方法学校验（硬门槛）：对 5 个高置信近邻星团（如 Pleiades、Praesepe、NGC 2516 等）重建成员表，与 Cantat-Gaudin (2020) 概率 > 0.7 成员的交并比 (IoU) $\geq 80\%$ 。此步不过关，后续全部结论无效。
2. Λ_{MSR} 严格按 Allison et al. (2009) 定义实现，误差棒用 ≥ 500 次随机参照集 + bootstrap；“显著分层”判据 ($\Lambda_{\text{MSR}} - 1 > 2\sigma$) 预先写死。
3. 质量代理双方案对照：G 星等排序法与 PARSEC 等龄线插值法都报告，两法给出的 H1 结论必须分别陈述。
4. 场星污染用邻近控制天区估计并显式报告每星团污染率上限。
5. 全部 ADQL 查询语句、管线与成员表开源。

用途	来源 / 工具
天测与测光	Gaia Archive (gea.esac.esa.int) ADQL 锥形检索, 逐星团按需拉取 (每星团 < 10 ⁵ 行, 全项目 < 1 GB); 覆盖全天, DR3 天测五参数 + G/BP/RP 测光
聚类	<code>scikit-learn</code> (GMM 内置)、 <code>hdbscan</code> 包 (内置层级密度聚类; 成员概率的自举稳定性评估需自行实现)
等龄线	PARSEC (CMD 3.7 网页接口按需生成) ——用于质量代理
成员表对照	Cantat-Gaudin (2020) via CDS/VizieR—— 仅作校验与对比, 不计入本项目的数据贡献
星团年龄	文献汇编 (Dias 等星表) —— 仅作标签来源, 不计入贡献

算力: HDBSCAN 对 $10^5 \times 5$ 维数据秒级; 全项目任何笔记本轻松完成, 无降规模必要 (但仍给: 星团数 20 $\rightarrow$ 10, H1 统计力下降、框架不变)。

6 · 方法路径

1. 装环境, 写通 ADQL 锥形检索, 对 Pleiades 拉数据并可视化自行 - 视差空间。
2. 实现 GMM 与 HDBSCAN 双聚类 + 概率判据 (判据须明确、可复现), 完成硬门槛 IoU 校验。
3. 按预注册判据 (距离 < 2 kpc、|b| 覆盖、年龄跨 3 个量级) 选定约 20 个星团样本。
4. 实现 Λ_{MSR} (最小生成树自写或用 `scipy.sparse.csgraph`), 对每星团计算并给误差棒。
5. 用双质量代理分别做 $\Lambda_{\text{MSR}} - \log(\text{age}/t_{\text{relax}})$ 相关分析 (t_{relax} 按标准公式由成员数与半质量半径估计)。
6. 场星污染控制区检验 + 把 IoU 硬门槛在样本中随机抽 3 星团复检。
7. 独立交叉校验: 用与 MST 无关的第二指标 (累积径向分布的 KS 检验分质量段比较) 复核分层判定。

7 · 新颖性边界

- 本课题不声称发现质量分层现象 (教科书结论), 不声称提出新的成员判定方法 (点名 arXiv:2303.08474 与 arXiv:2305.17728 已有更强的 ML 工具)。
- 已有工作: MNRAS 2023 在其选取的星团集上完成了"判定 + 分层"一体化分析并报告年龄趋势。
- 本项目贡献 (主结论): 在独立选取的星团样本上、以统一管线与统一统计口径重新检验分层 - 年龄关系, 并量化质量代理选择对结论的影响 (评价维度补充)。趋势复现或不复现都是有效结论; 不复现时须用误差棒证明有能力分辨文献报告的趋势幅度。
- 与历届获奖作的关系: 2024 年铜奖为 M82 星系中心条件分析 (河外、光谱), 与本题 (银河系内、天测统计) 无重叠。

8 · 决策门槛 (go / no-go)

- 第 6 周末: 硬门槛 IoU $\geq 80\%$ 通过。不过关 $\rightarrow$ 第 8 周前降级: 改用"以 Cantat-Gaudin 概率为初值的半自动复核"模式 (此时成员判定不再计入贡献, 贡献收窄为统一口径的 Λ_{MSR} 均匀测量), 主结论框架 (分层 - 年龄关系) 保留。

- 第 14 周末：核实低质量端完备性（DR3 在 $G \approx 18$ 后天测质量下降,不要边做边发现）；若目标星团完备极限差异过大 $\rightarrow$ 把 Δ_{MSR} 计算统一截断到共同完备星等，重跑全样本。
 - 第 26 周末：冻结样本与口径；预算裁剪顺序：星团数 $\rightarrow$ 第二分层指标 $\rightarrow$ 双质量代理不可裁（它是差异化轴的一半）。
 - 适合学生：喜欢天文、统计感好、编程中等即可；本题动手门槛是十题中最低的之一，风险集中在"结论可能平淡"，而非"做不出来"。
-

P04 · 格点高分子自回避行走的 pivot 算法自建实现：Flory 指数高精度复现与 θ 点跨越标度的独立测量

推荐优先级 ★★★★★ · 软物质理论建模 / 蒙特卡洛 · 纯 CPU (需 numba) · 算法实现取向 · 可复现性贡献

1 · 研究问题

自建的 pivot + 局部移动混合蒙特卡洛能否在二维/三维格点上把自回避行走 (self-avoiding walk, SAW) 的 Flory 指数 ν 复现到 1% 以内, 并进一步在相互作用自回避行走 (ISAW) 上独立测得 θ 点 (theta point) 附近 ν 从膨胀值向 θ 值的跨越行为?

2 · 研究背景与空白

SAW 是高分子排除体积效应的最小模型, 其端距标度 $R \sim N^\nu$ 的指数是普适量: 二维精确值 $\nu = 3/4$ (Nienhuis), 三维最精值 $\nu = 0.587597(7)$ 来自 Clisby (2010) 对 pivot 算法的快速实现 (行走长度达 2.68×10^8 步)。加入近邻吸引后, ISAW 在 θ 温度发生塌缩转变: 二维 θ 点的 $\nu_\theta = 4/7$ 有精确预言。文献同时明确指出: pivot 算法的效率在 θ 点附近显著退化、塌缩相中更甚 (Sokal 综述, hep-lat/9509032; 多重马尔可夫链方法即为此而生)。

空白在可复现性维度: 高精度数值全部来自少数几个高度优化的专用实现; 一个结构透明、百行级、可被第三方审查的实现能把 ν 做到多准、把 θ 点跨越测到多清楚, 本身是有教学与方法价值的定量问题。物理结论已知使得每一步都可校验, 特别适合把"算法效率退化"从障碍转写成研究内容。

3 · 可检验假设

- H1: 三维 SAW ($N \leq 10^4$, 含修正标度项外推) 测得 ν 与 0.587597 偏差 $\leq 1\%$; 二维与 $3/4$ 偏差 $\leq 1\%$ 。
- H2: 二维 ISAW 在文献 θ 温度处, 回转半径标度指数与 $4/7$ 在 95% 置信区间内一致; 且 pivot 接受率随温度降至 θ 点呈文献描述的显著退化, 退化曲线可定量表征。

4 · 量化验收标准

1. 方法学校验 (硬门槛): 无相互作用 SAW 在 $N = 10^2 - 10^3$ 区间, 二维 ν 偏差 $\leq 2\%$ 、平均接受率与 Madras – Sokal 文献报告的 $O(N^{-D})$ 行为定性一致; 短链 ($N \leq 12$) 与精确枚举计数完全一致。此步不过关, 后续全部结论无效。
2. 每个 N 的有效独立样本数 $\geq 10^3$ (按积分自相关时间折算并显式报告)。
3. 指数拟合口径预先写死: 双对数最小二乘、含首阶修正标度项 $(1 + b \cdot N^{-\Delta})$ 的非线性拟合、逐步剔除小 N 的稳定性分析三者并行, 误差用 bootstrap 置信区间。
4. θ 点采样用 Metropolis 局部移动 + 多重马尔可夫链 (multiple Markov chain) 双方案, 两法在重叠温度点统计量一致 $\leq 1\sigma$ 。
5. 全部代码开源, 附精确枚举校验脚本。

用途	来源 / 工具
模拟引擎	自写 Python + numba : pivot (哈希表查自交, 单次尝试 $O(N)$)、局部移动 (BFACF 或爬虫移动) 均需自行实现; 无外部数据依赖
精确对照	二维 $v = 3/4$ 精确值; 短链精确枚举 (自写深度优先, $N \leq 12$ 秒级) —— 仅作校验, 不计入贡献
文献基准	Clisby (2010) $v = 0.587597(7)$; Sokal 综述 hep-lat/9509032; θ 点文献值—— 仅作对照, 不计入贡献
分析	numpy/scipy; bootstrap 自写

算力 (基于文献标尺): Clisby 的专用数据结构达 $N \sim 10^8$ ——**明确超出本项目范围**。简单哈希实现单次 pivot 尝试 $O(N)$: $N = 10^4$ 、每链 10^6 次尝试 $\approx 10^{10}$ 基本操作, numba 下约小时级/链长点; 全部 8-10 个 N 点约数天后台, 可行。 θ 点附近因自相关暴涨, 预算再乘 5-10 倍, 限制在 $N \leq 3 \times 10^3$ 。降规模版本: $N \leq 3 \times 10^3$ 全线 $\rightarrow v$ 复现精度目标放宽到 2%, θ 点跨越定性结论保留, 修正标度项拟合的杠杆臂缩短。

6 · 方法路径

1. 写精确枚举器 ($N \leq 12$) , 建立绝对正确的对照锚点。
2. 实现 pivot 算法与哈希自交检测, 完成硬门槛校验 (枚举一致 + 接受率行为) 。
3. 测 2D/3D SAW 的 $R^2(N)$, 做三口径指数拟合与外推, 对照精确/文献值。
4. 加入近邻吸引能, 实现 Metropolis 局部移动与多重马尔可夫链, 跑通高温端。
5. 核实 θ 点附近积分自相关时间量级 (第 8 块门槛项) , 冻结温度网格。
6. 测量跨越区 $v_{\text{eff}}(T, N)$, 给出 θ 点标度检验与接受率退化曲线。
7. 独立交叉校验: 用回转半径与端距两种尺寸度量分别拟合 v , 检验一致性。

7 · 新颖性边界

- 本课题不声称改进 v 的世界纪录 (点名 Clisby 2010 的 0.587597(7) 远超本项目可达精度) , 不声称发现 θ 点物理。
- 已有工作: pivot 算法及其快速实现、 θ 点的多重马尔可夫链研究均已发表 (Sokal 综述、Springer 多重马尔可夫链章节) 。
- 本项目贡献 (主结论) : **结构透明实现所能达到的精度边界的系统定标 + θ 点跨越与算法效率退化的独立定量测量**。这是可复现性贡献, 定位必须讲清楚; "简单实现在 $N \leq 10^4$ 内即可达 1%"或"达不到"都是有信息量的结论。
- 历届获奖图谱中软物质方向有 2022 金奖 (软物质弹性模量测量, 实验) , 与本题 (格点模拟) 无重叠。

8 · 决策门槛 (go / no-go)

- 第 6 周末: 硬门槛 (枚举一致 + $2D\ v \leq 2\%$) 通过; 不过关 $\rightarrow$ 排查实现, 不加规模。
- 第 10 周末: 核实 θ 点附近自相关时间。若多重马尔可夫链仍无法在两周预算内产出 10^3 有效样本 $\rightarrow$ 具名降级: 放弃 θ 点, 课题收窄为"SAW 修正标度指数 Δ_1 的三口径测量与简单实现精度定标", 标度分析框架完整保留。

- 第 24 周末：冻结全部拟合口径；预算裁剪顺序：3D 大 N 端 $\rightarrow$ θ 点温度网格密度 $\rightarrow$ 2D 主线不可裁。
 - 适合学生：编程较强、对算法本身有兴趣的队伍；风险集中在 θ 点采样效率，已配置明确退路。
-

P05 · 地表气温长程相关性 DFA 标度指数的新时间窗独立检验与"站点 vs ERA5 再分析"同点偏置量化

推荐优先级 ★★★★★ · 地球物理与大气公开数据 · 时间序列统计 · 纯 CPU 极轻量 · 双数据源对照

1 · 研究问题

用 1991 – 2025 年新 35 年时间窗与自选全球站网重新估计日气温距平的去趋势涨落分析 (detrended fluctuation analysis, DFA) 标度指数 α , "内陆站 $\alpha \approx 0.65$ 、海岸/岛屿更高"的经典分区结论是否依然成立? 在同一地点上, ERA5 再分析格点序列的 α 是否相对实测站存在系统偏移?

2 · 研究背景与空白

气温距平的长程相关性表现为 DFA 涨落函数 $F(s) \sim s^\alpha$ 且 $\alpha > 0.5$, 被视为气候系统慢分量 (海洋、土壤记忆) 的统计指纹。Koscielny-Bunde et al. (1998, PRL) 提出内陆站 $\alpha \approx 0.65$ 的"普适律"; Fraedrich & Blender 及后续工作 (如中国 557 站研究、澳大利亚地理依赖研究、希腊 ERA-Interim 的 MF-DFA) 确认了海陆差异与纬度依赖, 也对普适性提出修正。方法与数据全公开。

空白在时间覆盖与评价维度: 其一, 经典结论的数据多止于 2010 年前后, 1991 – 2025 这一增暖加速期的独立时间窗检验 (趋势去除方法对 α 的影响也需重估) 是标准的换样本轴; 其二, "同一位置上站点观测与 ERA5 的 α 是否系统不同"直接量化再分析同化平滑对标度统计的影响——本次检索未找到完全相同的系统同点对照工作 (未找到不等于不存在, 第 4 周文献复查后冻结表述)。计算量几乎为零, 正负结论都有效。

3 · 可检验假设

- H1: 内陆站 α 分布中位数落在 0.60 – 0.70, 海岸/岛屿站显著更高 (Mann – Whitney U 检验 $p < 0.01$) 。
- H2: ERA5 同点序列的 α 系统性高于配对站点值 (再分析平滑抬升相关性), 配对差中位数 > 0.02 且 Wilcoxon 符号秩 $p < 0.01$; 若无系统差, 则"再分析可替代站点做标度分析"获得一次定量支持——两种结果都有效。

4 · 量化验收标准

1. 方法学校验 (硬门槛): 自写 DFA 对合成分数高斯噪声 ($H = 0.6/0.7/0.8$ 各 100 条、长度与真实序列相同) 恢复 α 平均偏差 ≤ 0.02 ; 并对 3 个文献明确给出 α 值的站点复现至偏差 ≤ 0.05 。此步不过关, 后续全部结论无效。
2. 以 DFA2 为主口径, DFA1/DFA3 做敏感性; 拟合区间预先写死为 10 天 – 4 年并报告区间端点敏感性。
3. 站点质控口径预先写死: 记录长度 ≥ 30 年、缺测 $< 10\%$ 、GHCN 质量标志过滤规则全部写入脚本。
4. 季节循环去除双方案对照 (逐日气候态扣除 vs 前 4 阶傅里叶谐波), 两方案结论分别报告。
5. 所有 α 的误差棒用滑动窗 + bootstrap 给出; 站网选择判据 (分层随机抽样按大洲/海陆/纬度带) 预注册。
6. 全部脚本与站点清单开源。

5 · 数据与工具

用途	来源 / 工具
站点日气温	GHCN-Daily (NOAA NCEI, ncei.noaa.gov 按站号直接下载 CSV, 单站约 1–5 MB; 覆盖全球十万余站、逐日、1880s 至今), 目标站网 100–200 站
再分析	ERA5 单层 2 m 气温 (Copernicus CDS, cds.climate.copernicus.eu, 需免费注册; 用 CDS API 按点/按年切片下载, 单点 35 年小时序列 ~ 数十 MB, 先日均化; 0.25° 分辨率、1940 至今)。CDS 新版接口与配额需第 6 周核实
合成校验信号	自写分数高斯噪声生成器 (numpy FFT 法) —— 仅作校验
文献 α 值	Koscielny-Bunde 1998 等表列值—— 仅作参考, 不计入本项目的数据贡献
分析	numpy/scipy; DFA 自写 (约 60 行, 不依赖黑箱包, 便于审查)

算力: 全项目任何笔记本数小时内完成, 无算力风险。降规模版本: 站网 200 $\rightarrow$ 60 站、ERA5 只取配对子集 $\rightarrow$ H1/H2 框架完整保留, 仅统计力下降。

6 · 方法路径

1. 自写 DFA 与 fGn 生成器, 完成硬门槛校验 (合成信号 + 文献站点复现)。
2. 按预注册判据抽取全球站网, 写死质控脚本并输出站点清单。
3. 计算全部站点 α (DFA1/2/3、双去季节方案), 做海陆/纬度分组统计检验 (H1)。
4. 核实并跑通 CDS API, 按站点坐标下载 ERA5 同点序列并日均化。
5. 做配对 α 差分析 (H2), 并按纬度带/地形分层看偏置结构。
6. 敏感性总表: 拟合区间、DFA 阶数、去季节方案三因素对主结论的影响矩阵。
7. 独立交叉校验: 对 20% 抽样站点用功率谱指数 β ($\alpha = (\beta+1)/2$ 关系) 复核 DFA 结果。

7 · 新颖性边界

- 本课题不声称发现气温长程相关性 (1998 年已发表) 也不声称提出新统计方法。
- 已有工作: Koscielny-Bunde et al. (1998) 与后续区域研究 (中国、澳大利亚、希腊等) 已给出 α 的地理分布图景; MF-DFA 应用于再分析数据也已存在。
- 本项目贡献 (主结论): 1991 – 2025 新时间窗的全球分层站网独立检验 + 站点/ERA5 同点配对偏置的系统量化。前者是最稳的换样本轴; 后者是评价维度补充, 对所有用再分析数据做标度分析的研究是直接有用的定标信息。
- 若分区差异不显著, "经典分区在新窗口不显著" 同样是有效结论, 需用误差棒证明可分辨 0.05 量级的 α 差。
- 历届获奖图谱中无大气统计类先例; 2022 有一项微塑料 – 沉积物与气候相关的提名奖, 主题无重叠。

8 · 决策门槛 (go / no-go)

- 第 4 周末: 硬门槛 (合成信号 ≤ 0.02) 通过; 不过关 $\rightarrow$ 排查 DFA 实现。
- 第 6 周末: 核实 CDS API 注册、配额与下载速度 (不要边做边发现)。若 ERA5 获取受阻 $\rightarrow$ 具名降级: H2 改为 " α 对海拔与离海距离的回归结构" (纯 GHCN 即可), H1 与全部统计框架保留。

- 第 16 周末：站网 α 全表完成。若海陆差异效应量 $<$ 预注册最小可检测差 $\rightarrow$ 扩站网至 300 站（一周预算）或正式接受"不显著"结论并转入敏感性归因。
 - 预算裁剪顺序：ERA5 配对范围 $\rightarrow$ DFA 阶数敏感性 $\rightarrow$ 站网规模下限 60 站不可再裁。
 - 适合学生：统计兴趣强、编程入门即可的队伍；本题是十题中"每周都有产出"最有保障的一条。
-

P06 · α -FPUT 非线性晶格能量均分时间随能量密度的标度指数：辛积分器自建实现的独立测量

推荐优先级 ★★★☆☆ · 经典力学数值 / 非线性动力学 · 纯 CPU 较重（需 numba/C） · 长时间积分 · 标度律检验

1 · 研究问题

用自建辛积分器测量 α -FPUT (Fermi – Pasta – Ulam – Tsingou) 链的能量均分时间 T_{eq} 随能量密度 ϵ 的标度指数，在 $N = 32 - 256$ 、笔记本可达的 ϵ 区间内，测得的指数与近期文献报告的普适标度律是否一致？有限尺寸（小 N ）对该标度的偏离从何处开始？

2 · 研究背景与空白

FPUT 问题（1955）是计算物理的开端：弱非线性晶格不但不快速热化，反而出现能量回归（recurrence）与超长的亚稳态。均分时间 T_{eq} 随 ϵ 的标度是刻画“通往热化之路”的核心量：波湍流理论对 α 链给出高阶共振主导的幂律预测，数值上 Onorato 等与 Benettin – Ponno 系列工作给出了不同参数区间的指数；2025 年 PRE (arXiv:2504.04725) 研究了预热化（prethermalization）阶段，2026 年综述 (arXiv:2603.23347) 明确指出：历史文献中表现互异的热化指数主要源于非线性微扰强度的参数化不一致，而非机制不同。

空白在独立检验与参数边界：该综述给出的统一图像是新近论断，正需要独立实现、独立参数化核对的检验；且小 N （32 – 128）下标度律的适用边界对文献常用的大 N 结论是有价值的补充。物理全已发表、判据可量化、算力标尺文献明确。

3 · 可检验假设

- H1：在统一采用该综述的微扰参数化后， $T_{eq}(\epsilon)$ 在可达区间内呈幂律，指数与目标文献值在 bootstrap 95% 置信区间内一致（目标值在第 3 周精读后预注册； α 链文献量级为陡幂律，具体数值以精读为准，不预先照抄）。
- H2： N 从 256 降到 32 时，幂律指数的漂移超出统计误差（有限尺寸效应可检出）；若检不出，则“小 N 即达渐近标度”成立——两种结果都有效。

4 · 量化验收标准

1. 方法学校验（硬门槛）：复现 $N = 32$ α -FPUT 经典能量回归（FPUT recurrence），回归时间与文献值偏差 $\leq 5\%$ ；辛积分全程能量漂移 $|\Delta E/E| \leq 10^{-6}$ 。此步不过关，后续全部结论无效。
2. 均分判据预先写死：谱熵 $\eta(t)$ (spectral entropy) 降至阈值法，双阈值 (0.1/0.05) 敏感性并行报告；不得事后挑阈值。
3. ϵ 扫描 ≥ 6 个点、每点 ≥ 8 个随机初始相位实现；指数拟合用双对数最小二乘 + 稳健回归 (Theil – Sen) 交叉验证，bootstrap 给置信区间。
4. 时间步收敛扫描：dt 减半后 T_{eq} 变化 $< 2\%$ 方可采信；守恒量漂移随 dt 的标度须与积分器阶数一致（作为实现正确性的第二证据）。
5. 全部代码与原始时序数据开源。

5 · 数据与工具

用途	来源 / 工具
积分引擎	自写 velocity-Verlet / 四阶辛积分器 + numba （需自行实现；无外部数据依赖）
谱分析	numpy FFT（正则模投影自写，约 20 行）
文献基准	FPUT 原始报告回归时间、arXiv:2504.04725、arXiv:2603.23347 的标度结果—— 仅作参考，不计入本项目的数据贡献
可视化	matplotlib

算力（基于文献标尺）：文献中低 ϵ 端 T_{eq} 可达 $10^8 - 10^9$ 自然时间单位——该区间明确超出笔记本范围。可达区间估算： $dt = 0.05$ 、单实现积分 10^7 时间单位 = 2×10^8 步； $N = 64$ 时约 1.3×10^{10} 粒子更新，numba 单核按 $\sim 10^7 - 10^8$ 更新/秒计约 0.5 - 4 小时/实现（实测速度第 5 周核实）；6 个 $\epsilon \times 8$ 实现 $\times 2$ 个 $N \approx 2 - 4$ 周后台，可行。降规模版本： $N \in \{32, 64\}$ 、 ϵ 只取高端 4 点 $\rightarrow$ 指数仍可估，置信区间变宽、杠杆臂缩短， H_2 的大 N 端结论放弃。

6 · 方法路径

- 实现辛积分器，用谐振链 ($\alpha = 0$) 解析解验证频率与能量守恒。
- 复现 $N = 32$ FPUT 回归，完成硬门槛校验。
- 实测 numba 更新速度，据此冻结 (N, ϵ) 网格与总预算表。
- 实现正则模能量与谱熵在线计算（避免存全时序），跑通单点全流程。
- 扫描 ϵ 与 N ，每点多实现取均值与散布，估 T_{eq} 及其误差。
- 做双口径指数拟合与阈值敏感性矩阵，检验 H_1/H_2 。
- 独立交叉校验：对 2 个 ϵ 点改用不同阶积分器（二阶 vs 四阶）复算 T_{eq} ，一致性 $\leq 2\%$ 。

7 · 新颖性边界

- 本课题不声称发现新热化机制，不声称裁决波湍流理论——标度预测与大规模数值均已发表（点名 Onorato 系、Benettin - Ponno 系与 2026 综述）。
- 已有工作：2026 综述已提出"参数化不一致解释指数分歧"的统一论断，并配有其作者群的数值支持。
- 本项目贡献（主结论）：**对该新论断的第三方独立检验（独立实现、独立参数化核对）+ 小 N 有限尺寸边界的系统测量**。综述式论断恰恰最需要独立复现；一致或偏离都有效，偏离时须用收敛与守恒证据排除数值伪影后才可报告。
- 历届获奖图谱中非线性动力学多为实验摆/流体（2023 随机共振铜奖等），无晶格热化数值先例。

8 · 决策门槛 (go / no-go)

- 第 5 周末：硬门槛 + 速度基准。若 $numba < 10^7$ 粒子更新/秒 $\rightarrow$ 具名降级：其一，核心循环改 C（一周预算）；其二，直接采用降规模网格。两者均保留主结论框架。
- 第 20 周末：检查最低 ϵ 点是否能在预算内达到均分。达不到 $\rightarrow$ 截断 ϵ 区间、把"该点未热化的下限 $T_{eq} >$ " 作为数据点上限报告（文献同样处理方式），拟合框架不变。

- 第 30 周末：冻结判据与网格；预算裁剪顺序：N 网格 $\rightarrow$ 实现数 ($8 \rightarrow 5$) $\rightarrow \epsilon$ 点数下限 5 不可再裁。
 - 适合学生：动力学/编程兴趣强、能接受"大部分时间在等后台跑"的队伍；风险集中在算力预算而非方法。
-

P07 · GWOSC 公开应变数据上双黑洞并合的自建匹配滤波重探测与数据条件化选择的参数偏置量化

推荐优先级 ★★★☆☆ · 引力波公开数据 · 信号处理 · 纯 CPU 中等 · 注入-恢复实验

1 · 研究问题

自建的 PyCBC 匹配滤波 (matched filtering) 管线能否重现 GWTC-3 星表中 ≥ 10 个双黑洞事件的信噪比与啁啾质量 (chirp mass) ? 功率谱密度 (PSD) 估计方式 (Welch 段长、窗函数、中位数 vs 均值合并) 与高通截止频率的选择, 会给恢复的信噪比和参数带来多大的系统偏置?

2 · 研究背景与空白

匹配滤波是引力波探测的核心: 以理论波形为模板与应变数据做噪声加权互相关, 信噪比峰值即候选事件。LIGO/Virgo 的全部事件应变数据经 GWOSC 完全公开, PyCBC 官方文档给出了在笔记本上数分钟内重现 GW150914 (信噪比 $\sim 19-20$) 的完整教程; GW Open Data Workshop 每年发布成套练习。GWTC-3 给出了事件参数的官方值。

空白在评价维度: "重现教程"人人可做, 但把数据条件化 (data conditioning) 选择 $\rightarrow$ 恢复参数偏置做成系统的注入-恢复 (injection-recovery) 扫描, 在学生可及文献层面未找到相同的系统工作 (专业管线内部有各自的定标, 但未见以此为主题的独立量化; 未找到不等于不存在, 第4周复查)。该问题判据清晰、正负都有效, 且天然给出"学生自己动手"的强证据。

3 · 可检验假设

- H1: 对 ≥ 10 个 GWTC-3 双黑洞事件, 自建管线恢复的信噪比与公布值偏差 $\leq 10\%$ 、啁啾质量偏差 $\leq 5\%$ 。
- H2: 在含毛刺 (glitch) 数据段上, Welch 均值 PSD 相对中位数 PSD 会引起可检出的信噪比系统差 ($|\Delta \text{SNR}|/\text{SNR} > 5\%$), 而在干净段上两者一致 $\leq 2\%$ ——即 PSD 稳健化的收益可被注入实验定量测出。

4 · 量化验收标准

1. 方法学校验 (硬门槛): 重现 GW150914 在 H1 探测器的匹配滤波信噪比, 与 PyCBC 文档/文献值偏差 $\leq 5\%$; 峰值时刻偏差 $\leq 10 \text{ ms}$ 。此步不过关, 后续全部结论无效。
2. 注入-恢复 ≥ 200 个模拟双黑洞信号 (参数从预注册网格抽取) 注入真实 O3 噪声, 报告各条件化配置下的参数恢复偏置与散布; 背景显著性用时间滑动 (time slides) 估计。
3. 检测性能不报 accuracy (极不平衡问题, 几乎全是噪声样本, accuracy 恒接近 1 而无信息): 报告恢复率-误报率曲线与给定误报率下的召回。
4. 条件化参数网格 (段长 $\{4, 16, 64\} \text{ s} \times$ 窗 $\{\text{Hann, Tukey}\} \times$ 合并 $\{\text{均值, 中位数}\} \times$ 高通 $\{15, 20, 30\} \text{ Hz}$) 预注册, 不得事后增删。
5. 全部脚本开源; 每个事件附条件化配置-结果对照表。

5 · 数据与工具

用途	来源 / 工具
应变数据	GWOSC (gwosc.org)：按事件下载 32–4096 s、4/16 kHz 应变段（单段数十 MB，全项目 < 10 GB；支持按事件、按时间段的切片下载），覆盖 O1–O4 已发布数据
滤波与波形	PyCBC（内置：matched filter、IMRPhenom 系列波形近似、Welch/中位数 PSD 估计、注入工具；需自行实现：条件化网格扫描与偏置统计脚本）。文档明示 GW150914 示例可在笔记本分钟级运行
时频检查	gwpyp（内置 Q 变换谱图，用于毛刺段挑选）
事件参数	GWTC–3 星表（GWOSC/Zenodo）—— 仅作对照与标签来源，不计入本项目的数据贡献

算力（基于文献标尺）：单事件单模板匹配滤波即 PyCBC 教程量级（分钟）；本项目总量 $\approx 10 \text{ 事件} \times 18 \text{ 配置} + 200 \text{ 注入} \times 18 \text{ 配置} \approx 4 \times 10^3$ 次滤波，每次 $\sim 10 \text{ s}$ 量级 $\rightarrow$ 十余小时后台，可行。全模板库盲搜 ($10^5 - 10^6$ 模板) 明确超出范围，本项目只做目标化滤波。降规模版本：事件 $10 \rightarrow 5$ 、注入 $200 \rightarrow 50$ 、网格 $18 \rightarrow 8$ 配置 $\rightarrow$ H1/H2 框架保留，偏置曲面分辨率下降。

6 · 方法路径

1. 装 PyCBC/gwpyp (conda-forge)，跑通官方 GW150914 教程，完成硬门槛校验。
2. 选定 ≥ 10 个 O1 – O3 双黑洞事件（预注册判据：网络信噪比 > 10 、单事件公开段完整），逐事件重现信噪比与啁啾质量（模板网格细扫啁啾质量）。
3. 用 gwpyp Q 变换按明确判据挑选"干净段"与"含毛刺段"各 ≥ 10 段（判据须可复现）。
4. 实现条件化网格扫描器，对真实事件跑全配置矩阵。
5. 做注入-恢复实验：注入参数、恢复偏置、时间滑动背景全流程。
6. 汇总偏置曲面，检验 H1/H2，给出"推荐条件化配置 + 定量理由"。
7. 独立交叉校验：对 2 个事件用 gwpyp + 自写频域相关（不经 PyCBC 高层接口）复算信噪比，偏差 $\leq 5\%$ 。

7 · 新颖性边界

- 本课题不声称发现引力波事件（全部事件已由 LVK 合作组发表），不声称改进探测管线——专业管线（PyCBC Live、GstLAL）远超本项目。
- 已有工作：GWOSC 教程与开放数据工作坊已把"重现 GW150914"标准化；GWTC 系列给出官方参数。
- 本项目贡献（主结论）：**条件化选择对恢复信噪比/参数的偏置的系统注入-恢复量化**，即评价维度的转换：教程展示"能测到"，本项目回答"测得的数字对分析选择有多敏感"。这对所有基于开放数据的教学与独立复核工作是直接可用的定标结果。
- 与历届获奖作的关系：丘奖物理图谱中尚无引力波数据分析先例（2020 有 Hubble 张力理论铜奖，主题无重叠），题材差异化明显。

8 · 决策门槛 (go / no-go)

- 第 6 周末：硬门槛 ($\text{GW150914} \leq 5\%$) 通过。

- 第 8 周末：核实 PyCBC 在学生机器（macOS/Windows + conda）安装与运行稳定性（依赖链较重，不要边做边发现）。失败 → 具名降级：改用 `gwpv` + 自写频域匹配滤波（教学级实现,反而强化"自己实现"的方法学证据），事件数收窄到 5，H1/H2 框架保留。
 - 第 22 周末：注入-恢复主矩阵完成度检查。若进度 $< 50\%$ → 启动降规模版本。
 - 预算裁剪顺序：配置网格密度 → 注入数 → 事件数下限 5 不可再裁；硬门槛与交叉校验不可裁。
 - 适合学生：信号处理/编程中上、对天体物理有热情的队伍；风险集中在环境安装与数据段选择的工作量。
-

P08 · TESS 20 秒短节奏 M 矮星耀斑的自建检测 + 注入-恢复管线与耀斑频率分布幂律指数的稳健独立拟合

推荐优先级 ★★★☆☆ · 天体物理公开数据 · 检测管线 + 稳健幂律统计 · 纯 CPU 中等 · 完备性改正

1 · 研究问题

自建的耀斑检测与注入-恢复管线在 TESS 20 秒短节奏 (20-s cadence) 光变曲线上, 对约 30 颗活跃 M 矮星测得的耀斑频率分布 (flare frequency distribution, FFD) 幂律指数 α , 经完备性改正后是否落在文献报告的 TESS 波段区间 (约 1.5 – 1.7)? 不做完备性改正会把 α 低估多少?

2 · 研究背景与空白

恒星耀斑能量服从幂律分布 $dN/dE \propto E^{-\alpha}$, α 是磁重联能量释放统计的核心参数, 且 α 是否 > 2 决定"微耀斑加热日冕"类论证能否成立。TESS 自 2020 年起提供 20 秒节奏光变, 显著改善小耀斑的时间分辨。

已有工作: Günther et al. (2020) 用 TESS 2 分钟数据系统测量 M 矮星耀斑; Pietras et al. (2022) 报告 TESS 波段 $\alpha \approx 1.5 - 1.7$ (依赖能量估计方法); 2025 年 A&A 有年轻运动星群 M 矮星耀斑频率研究 (arXiv:2506.04465), 2026 年 A&A 用 CHEOPS 延伸 TESS 的 FFD 并正面比较幂律与对数正态模型 (arXiv:2510.15453); 20 秒节奏的准周期脉动研究 (arXiv:2506.22131) 确认了短节奏数据的可用性。注入-恢复被上述文献确认为可靠检测小耀斑的关键手段。

空白在样本与评价维度: 以 20 秒节奏为主、统一管线贯穿、并把"完备性改正前后 α 的偏移量"作为显式测量目标的独立样本分析, 是标准的换样本 + 评价维度轴; 幂律 vs 对数正态的模型比较在 20 秒数据上也刚由 2026 年论文打开, 独立检验正当其时。

3 · 可检验假设

- H1: 完备性改正后的 α (MLE 口径) 落在 [1.4, 2.0], 且与 Pietras et al. (2022) 区间在 95% 置信区间内相容。
- H2: 不做完备性改正、直接对全部检出耀斑拟合, 会使 α 被系统低估 ≥ 0.2 (低能端漏检削平分布); 该偏移量可由注入-恢复曲线定量预测。

4 · 量化验收标准

1. 方法学校验 (硬门槛): 对 3 – 5 颗已有公开耀斑表的恒星重跑检测, 能量排名前 20 的耀斑匹配率 $\geq 80\%$ 、峰值流量偏差 $\leq 30\%$ (公开耀斑表仅作对照标签, 不计入本项目的数据贡献)。此步不过关, 后续全部结论无效。
2. 幂律拟合口径预先写死: 以 Clauset – Shalizi – Newman 最大似然法 (含 $E_{\min}$ 选择的 KS 判据) 为主, 双对数最小二乘仅作对照展示并说明其偏差来源; 不确定度用 bootstrap 置信区间; 同时报告幂律 vs 对数正态的似然比检验。
3. 注入-恢复: 每星 $\geq 10^3$ 个人工耀斑 (经验模板、幅度-时长网格预注册), 给出检测完备性曲线; 检测判据 (连续 $\geq N$ 点超过 $k\sigma$) 预注册且全样本统一。
4. 检测性能不报 accuracy (耀斑点占比 $\ll 1\%$, accuracy 无信息): 报告注入样本上的召回-精确率与 PR 曲线。
5. 能量估计双方案 (相对幅度积分 $\times$ 静态光度 vs 黑体近似), 两方案 α 分别报告。

6. 全部脚本与耀斑表开源。

5 · 数据与工具

用途	来源 / 工具
光变曲线	MAST 归档 TESS SPOC 20 秒节奏目标光变（lightkurve 按 TIC 编号检索下载，单星单扇区约 10–50 MB；20 秒产品自 Sector 27 起覆盖部分目标，目标样本的 20 秒可用颗数需第 6 周核实）
目标选择	TIC/已发表活跃 M 矮星清单—— 仅作标签来源，不计入贡献
检测与注入	自写去趋势（滑动中位数）+ 阈值检测 + 注入器；altaipony 包（内置耀斑检测与注入-恢复， 其能力边界需核实 ；本项目以自写为主、该包为交叉校验）
幂律拟合	powerlaw 包（内置 CSN 方法与分布比较）+ 自写 bootstrap
恒星参数	TIC 星表（半径、温度，用于能量定标）—— 仅作输入参数

算力：检测 + 注入 $30 \text{ 星} \times 10^3 \text{ 注入} \rightarrow$ 数小时至 1 天，笔记本可行。降规模版本： $30 \rightarrow 10 \text{ 星}$ 、注入 $10^3 \rightarrow 300/\text{星} \rightarrow \text{H1/H2 框架保留}$ ， α 的样本方差增大。

6 · 方法路径

1. 装环境，用 lightkurve 拉 2–3 颗著名耀斑星（如 Proxima Cen、AU Mic 类）目视检查数据质量。
2. 实现去趋势 + 阈值检测，完成硬门槛校验（与公开耀斑表匹配）。
3. 核实 20 秒样本规模，按预注册判据（光谱型 M0–M6、 ≥ 2 个 20 秒扇区、污染标志干净）冻结样本。
4. 实现注入-恢复，产出每星完备性曲线。
5. 计算耀斑能量（双方案），构建改正前/后 FFD，做 MLE 幂律拟合与模型比较。
6. 汇总 α 分布，检验 H1/H2，并按光谱型/自转周期分组看趋势。
7. 独立交叉校验：用 altaipony 对 5 颗星复跑全流程， α 偏差 $\leq 1\sigma$ 。

7 · 新颖性边界

- 本课题不声称发现耀斑幂律行为（已发表 30 年），不声称建立新的检测算法。
- 已有工作：Günther 2020（2 分钟节奏系统样本）、Pietras 2022（ α 区间）、2026 CHEOPS+TESS 论文（模型比较）已覆盖主要图景。
- 本项目贡献（主结论）：**20 秒节奏独立样本上、统一注入-恢复口径下的 α 测量 + "完备性改正对 α 偏移量"的显式量化 + 幂律/对数正态比较的独立检验**。H2 的偏移量对所有不做注入-恢复的 FFD 文献是直接的定标警示； α 相容与否两种结果都有效。
- 与历届获奖作的关系：2025 银奖为自制物端棱镜光谱仪 + AI 分类（仪器 + 分类），与本题（公开测光时序统计）无重叠。

8 · 决策门槛（go / no-go）

- **第 6 周末：硬门槛（匹配率 $\geq 80\%$ ）+ 20 秒样本规模核实（ ≥ 15 颗合格星）**。匹配率不足 $\rightarrow$ 阈值扫描并把"检测判据敏感性"升格为主结论组成部分；样本不足 $\rightarrow$ **具名降级**：退回 2 分钟节奏（样本充足），20 秒

仅保留个例对照，FFD 框架不变。

- 第 18 周末：注入-恢复完备性曲线全样本完成度 $\geq 50\%$ ，否则启动降规模版本。
 - 第 28 周末：冻结口径；预算裁剪顺序：分组分析 $\rightarrow$ 双能量方案（保留主方案 + 附录） $\rightarrow$ 注入数下限 300/星 不可再裁。
 - 适合学生：细致、能忍受数据清洗琐碎工作的队伍；风险集中在 20 秒样本可用性，退路明确。
-

P09 · ZTF 类星体光变阻尼随机游走参数的高斯过程重估与基线长度偏置的仿真定量检验

推荐优先级 ★★☆☆☆ · 天体物理公开数据 / 统计建模 · 高斯过程 + 蒙特卡洛仿真 · 纯 CPU 中等 · 偏置研究

1 · 研究问题

用 celerite 类高斯过程 (Gaussian process) 对数百条 ZTF 类星体光变拟合阻尼随机游走 (damped random walk, DRW) 参数后, 以 ZTF 真实采样模式做的注入仿真能否定量重现 Kozłowski (2017) 的基线偏置判据 (特征时标 $\tau \leq 10\%$ 基线才可靠)? 在该判据筛选后的"合格"子样本上, τ 与光度的标度关系与 MacLeod et al. (2010) 是否一致?

2 · 研究背景与空白

类星体紫外/光学光变可用 DRW (阻尼指数协方差的高斯过程) 近似描述, 特征时标 τ 被认为联系吸积盘热时标, τ -光度-黑洞质量标度关系 (MacLeod et al. 2010, SDSS Stripe 82) 是该领域基准。已知的方法学陷阱: Kozłowski (2017) 证明当 τ 接近光变基线长度时拟合系统偏置; 后续工作 (如 2024 年 arXiv:2409.09637) 进一步指出即使 $\tau < 10\%$ 基线仍可能有可观偏置; Suberlak et al. (2021) 用 SDSS+PS1 延长基线改进参数; 2025 年 arXiv:2508.12076 用 SDSS+PS1+ZTF 共 22 年基线研究超出 DRW 的行为。

空白在问题方向: 已有工作要么用超长联合基线回避偏置, 要么在附录里讨论偏置。把"单一巡天 (ZTF, 约 6-7 年基线) 究竟能可靠回收什么"作为主问题、用真实采样窗口的注入仿真给出定量答案、再在合格子样本上独立检验 MacLeod 标度——这是评价维度转换 + 换样本的组合。风险高于前面各题: 数据清洗重、且"合格子样本"可能偏小。

3 · 可检验假设

- H1: ZTF 真实采样注入仿真中, τ 恢复偏置的绝对值随 τ /基线单调上升, 在 $\tau/T = 0.1$ 处偏置绝对值 $> 20\%$ (方向与幅度以仿真为准并与 Kozłowski 2017 的解析预期对照)。
- H2: 仅保留仿真判定"可回收"的子样本后, τ -绝对星等关系斜率与 MacLeod et al. (2010) 在 95% 置信区间内一致; 若不一致, 偏离量与波段/红移分布差异的关系可被仿真解释。

4 · 量化验收标准

1. 方法学校验 (硬门槛): 对 200 条合成 DRW 光变 (已知 τ 、 σ , 先规则采样后 ZTF 真实采样窗口) 做 GP 拟合, 规则采样情形恢复参数中位偏差 $\leq 10\%$ 。此步不过关, 后续全部结论无效。
2. 光变质控口径预先写死: 观测点数 ≥ 100 、时基 ≥ 4 年、剔除坏点标志; 样本选择判据 (红移、星等范围) 预注册。
3. τ 不确定度双口径: MCMC 后验宽度与 bootstrap 重采样都报告; 样本级标度关系拟合用含测量误差的回归 (如 LINMIX 类方法) 而非普通最小二乘。
4. 全部注入网格 ($\tau/T \in [0.01, 1]$, 10 点对数均匀 $\times$ 每点 ≥ 100 实现) 预注册。
5. 真实样本 ≥ 300 条 (SDSS DR16Q 类星体表仅作目标选择标签, 不计入本项目的数据贡献)。
6. 全部脚本与拟合表开源。

5 · 数据与工具

用途	来源 / 工具
光变曲线	ZTF 公开数据释放 (IRSA, irsa.ipac.caltech.edu: 光变曲线 API 按坐标/编号查询, 单条 KB—MB 级; 覆盖 2018 至今、g/r 双波段、北天, 节奏约 1—3 天)。API 批量查询限速需第 8 周核实
目标选择	SDSS DR16Q 类星体星表 (SDSS 官网/CDS) —— 仅作参考, 不计入贡献
GP 拟合	<code>celerite2</code> (内置 O(N) DRW/SHO 核与似然; MCMC 用 <code>emcee</code> ; 含误差回归需自行实现或用 <code>linmix</code> 移植, 需核实)
仿真	自写 DRW 生成器 (精确离散递推, 约 20 行) + ZTF 真实时间戳复用
对照关系	MacLeod et al. (2010) 标度系数 —— 仅作参考, 不计入贡献

算力: `celerite2` 单条拟合 (含 MCMC) 秒到分钟级; 300 条真实 + 10^3 条仿真 $\approx$ 1–2 天后台, 可行。降规模版本: 真实样本 300 $\rightarrow$ 100、仿真每点 100 $\rightarrow$ 50 实现 $\rightarrow$ H1 (仿真偏置曲线) 完整保留, H2 统计力下降。

6 · 方法路径

1. 写 DRW 生成器并用解析自协方差验证; 装 `celerite2/emcee` 跑通文档示例。
2. 完成硬门槛校验 (规则采样恢复 $\leq 10\%$) 。
3. 用 ZTF 真实时间戳做全网格注入仿真, 产出偏置曲面与"可回收"判据 (H1) 。
4. 从 DR16Q 选样、经 IRSA 拉取光变、按预注册质控过滤。
5. 对真实样本做 GP 拟合, 应用可回收判据分割样本。
6. 在合格子样本上做含误差回归, 检验 H2 并与 MacLeod 2010 对照。
7. 独立交叉校验: 对 30 条光变用结构函数 (structure function) 法独立估 τ , 与 GP 结果秩相关 $\rho > 0.7$ 方视为一致。

7 · 新颖性边界

- 本课题不声称发现 DRW 行为或标度关系 (MacLeod 2010 已发表), 不声称提出新的偏置理论 (Kozłowski 2017 已给出) 。
- 已有工作: Kozłowski 2017 (偏置判据)、Suberlak 2021 (SDSS+PS1 改进)、arXiv:2409.09637 (阻尼检验)、arXiv:2508.12076 (22 年联合基线、超越 DRW) 已覆盖长基线路径。
- 本项目贡献 (主结论): 以 ZTF 单巡天为对象、真实采样窗口下偏置的系统仿真定标 + 合格子样本上 MacLeod 关系的独立检验。"单巡天能回收什么"的定量答案对即将到来的大巡天时代 (LSST 早期数据同样是短基线) 有直接参考价值; H2 一致与否都有效。
- 历届获奖图谱无类星体光变先例。

8 · 决策门槛 (go / no-go)

- 第 8 周末: 硬门槛通过 + IRSA 批量获取速率核实 (≥ 500 条/周可达)。获取受阻 $\rightarrow$ 具名降级: 改用 ZTF 按天区打包的公开光变文件本地切片 (数据量升至数十 GB 但一次性), 或样本收窄到 100 条。

- 第 20 周末：统计合格子样本规模。若 < 100 条 $\rightarrow$ 具名降级：课题重心正式移到 H1（纯仿真偏置定标为主结论,真实数据仅作个例演示），仿真框架与统计口径完整保留。
 - 第 30 周末：冻结样本；预算裁剪顺序：波段数 ($g+r \rightarrow r$) $\rightarrow$ 结构函数交叉校验样本数 $\rightarrow$ 仿真网格不可裁（它是主结论底座）。
 - 适合学生：统计成熟度较高、能读英文方法论文的队伍；风险在数据获取与子样本规模，两条退路都已具名。
-

P10 · Gaia DR3 宽双星低加速度区速度差统计的自建重建：筛选口径能否翻转"引力反常"结论的敏感性曲面

推荐优先级 ★☆☆☆☆ (高风险高回报) · 天体物理公开数据 / 引力检验 · ADQL + 前向蒙特卡洛 · 统计重 · 仅限强学生

1 · 研究问题

用自建的 Gaia DR3 宽双星 (wide binary) 筛选与相对速度统计管线重建样本后, Chae (2023 – 2025) 报告的低加速度 ($\sim 10^{-10} \text{ m/s}^2$) 引力增强信号与 Banik et al. (2024) 的零结果之间的矛盾, 能否被归结为少数几个可指名的筛选/隐藏三体处理参数? 哪个参数在预注册网格上对统计量的翻转效应最大?

2 · 研究背景与空白

宽双星内部加速度可低至 10^{-10} m/s^2 , 是太阳系尺度外检验低加速度引力 (牛顿 vs MOND 类修改引力) 的独特探针。这是当前天体物理最尖锐的公开争论之一: Chae (2023, 2024, 2025, ApJ 含贝叶斯 3D 建模) 与 Hernandez 等报告低加速度区引力增强约 40%; Banik et al. (2024, MNRAS) 与 Pittordis & Sutherland 报告无偏离; 2026 年预印本 (arXiv:2602.24035) 提出质量框架并支持零结果。双方公认矛盾根源在方法论——尤其隐藏级联三体 (hidden hierarchical companion) 比例的自校准处理与样本纯度筛选, 而非数据本身。样本基础通常是 El-Badry & Rix (2021) 目录或等价的自建筛选。

空白在评价维度: 本课题不裁决谁对, 而是把"结论对筛选口径的敏感性"本身做成定量曲面: 在同一数据、同一统计量下, 系统扫描筛选参数, 测量统计量何时跨越牛顿/MOND 判别线。这是争论双方都缺少的中立地图。风险极高 (统计与建模都重), 回报也高 (题材前沿、公示期抗查重能力强)。

3 · 可检验假设

- H1: 在预注册的筛选参数网格 (RUWE 上限 $\times$ 投影分离范围 $\times$ 视差信噪比 $\times$ 视向速度可用性要求) 上, 仅改变 2 – 3 个参数即可使低加速度区归一化速度比统计量在"牛顿相容"与"增强 $\geq 20\%$ "两侧翻转。
- H2: 若在全合理口径下信号方向保持不变, 则"筛选敏感性解释矛盾"被否定——这同样是有效且重要的结论 (矛盾须另寻根源, 如建模差异)。

4 · 量化验收标准

1. 方法学校验 (硬门槛, 两级): (a) 自建 ADQL 筛选重建 El-Badry & Rix (2021) 类样本, 在相同名义判据下样本规模与其公布数偏差 $\leq 20\%$ (该目录仅作对照, 样本必须从 Gaia Archive 自建, 目录不计入本项目的数据贡献); (b) 自建前向蒙特卡洛 (轨道相位、倾角、偏心率、质量比抽样) 在注入纯牛顿模拟数据时恢复"无反常", 偏差 $\leq 1\sigma$ 。任一不过关, 后续全部结论无效。
2. 统计量预先写死: 投影相对速度 / 牛顿圆轨道速度的归一化比值分布 (文献标准量), 分加速度箱报告; 不引入事后新统计量。
3. 敏感性网格 $\geq 3^4 = 81$ 参数组合预注册, 全部结果无筛选地全量报告 (明确反 p-hacking 条款)。
4. 每个网格点误差用 bootstrap 置信区间; 隐藏三体污染用两种文献方案 (固定比例 vs 自校准) 分别处理并对照。
5. 全部 ADQL、管线、网格结果表开源。

用途	来源 / 工具
天测数据	Gaia Archive ADQL (gea.esac.esa.int)：自行实现宽双星配对查询（角距 + 视差一致 + 自行一致），分 HEALPix 天区分块执行；原始拉取 $\sim 10^6\text{--}10^7$ 行、数 GB 内，本地筛后 $\sim 10^4\text{--}10^5$ 对
对照目录	El-Badry & Rix (2021) via CDS—— 仅作硬门槛校验对照，不计入贡献
前向模型	自写 numpy 蒙特卡洛（开普勒轨道抽样投影，约 200 行）；MOND 类修改采用文献给出的有效引力增强参数化—— 仅实现统计量层面,不做贝叶斯 3D 轨道建模
统计	scipy/bootstrap 自写

算力（基于文献标尺）：Chae (2025) 的贝叶斯 3D 建模为其结果核心、计算重——**明确超出本项目范围**；本项目限于统计量层面。前向蒙特卡洛 10^5 系统 $\times 10^3$ 抽样 $\approx 10^8$ 轨道求值，numpy 向量化小时级；81 网格点的统计重算（筛选在本地表面上进行,无需重复查询）约 1–2 天，可行。**降规模版本**：网格 81 $\rightarrow$ 27、天区改为高银纬子集 $\rightarrow$ 敏感性曲面粗化，翻转判定框架保留。

6 · 方法路径

1. 精读双方方法论文各 1 篇（Chae 2024、Banik 2024）+ 2026 质量框架预印本，列出双方筛选与建模差异对照清单（本身即成果附录）。
2. 写通分块 ADQL 配对查询，完成硬门槛 (a)。
3. 实现前向蒙特卡洛与统计量计算，完成硬门槛 (b)（牛顿注入自治）。
4. 冻结预注册网格，跑全部 81 组合的统计量 + bootstrap。
5. 对隐藏三体的两种处理方案分别重跑关键网格点。
6. 绘制敏感性曲面，判定 H1/H2，指名最大翻转参数。
7. 独立交叉校验：在高纯度极限口径（视向速度齐全的小样本）复算统计量，与主网格对应点一致 $\leq 1\sigma$ 。

7 · 新颖性边界

- 本课题不声称裁决牛顿/MOND 之争，不声称重复 Chae 的贝叶斯 3D 建模或 Banik 的完整管线，不声称发现引力反常（正反两方结论均已发表，点名 Chae 2023–2025、Hernandez 2022–2024、Banik 2024、Pittordis & Sutherland、arXiv:2602.24035）。
- 已有工作：双方各自发表了完整分析与相互批评；2026 质量框架论文从纯度角度支持零结果。
- 本项目贡献（主结论）：**同一数据、同一统计量下筛选口径敏感性的中立定量曲面与最大翻转参数的指名**。这是双方论战中缺位的公共品：若敏感性大（H1），矛盾的方法论根源被定位；若敏感性小（H2），矛盾须归因于建模而非筛选——两种结果都推进争论。
- 与历届获奖作的关系：2020 铜奖 Hubble 张力暗相互作用模型为纯理论宇宙学，与本题（天测统计）无重叠，但说明评委接受前沿引力主题。

8 · 决策门槛 (go / no-go)

- **选择前提**：仅当队伍已具备较强统计与编程基础、指导老师明确接受"可能无法给出裁决性结论"、且第 1 周精读两篇方法论文无严重障碍时才启动本题；否则直接改选 P03/P09。

- 第 8 周末：硬门槛 (a) (样本规模 $\pm 20\%$)。不过关 → 具名降级 A：课题收窄为"Gaia 宽双星样本纯度对 RUWE/视差质量判据的依赖"的数据论文式研究（保留筛选敏感性框架，去掉引力判别层）。
- 第 16 周末：硬门槛 (b) (牛顿注入自治)。不过关 → 具名降级 B：放弃前向模型，改为对双方公开发表的筛选判据差异做逐项复现与样本层面的定量影响清单（敏感性框架仍在,统计量退为样本统计而非速度比分布）。
- 第 28 周末：主网格完成度 $\geq 60\%$ ，否则启动降规模版本 (81 $\rightarrow$ 27)。
- 预算裁剪顺序：隐藏三体双方案 $\rightarrow$ 网格密度 $\rightarrow$ 交叉校验小样本不可裁；全量报告条款任何情况下不可裁。
- 已知困难定位：视向速度缺失使三维速度不可得，只能用投影统计——这不是障碍而是双方共同的约束条件，本项目的敏感性框架正建立在其上。

P11 · 电涡流阻尼下落的终端速度测量：以壁厚—温度二维扫描检验薄壁标度并反演铜的电阻率温度系数

标签： 优先级:最高 · 桌面实验/电磁族 · 预算档位:≤700元 · Arduino+视频测量 · 技能取向:动手+数据拟合

1 · 研究问题

钕磁铁在非铁金属管中下落的终端速度（terminal velocity） v_t ，是否在薄壁极限下严格满足 $v_t \propto \rho_e/(w)$ （ ρ_e 为管材电阻率、 w 为壁厚），该标度在壁厚增大到何处失效？进一步：能否把该装置当作“无电极电阻率计”，仅凭 $v_t(T)$ 反演出铜的电阻率温度系数（temperature coefficient of resistivity） α ，并与文献值 0.0039 K^{-1} 在误差棒内一致？

2 · 研究背景与空白

磁铁穿过导尿管时，管壁感生电涡流（eddy current），其磁场对磁铁产生阻尼力，磁铁在几厘米内即达终端速度，典型量级 $0.03 - 0.2\text{ m/s}$ ，肉眼与手机 60 fps 视频即可分辨。薄壁偶极子模型给出 $v_t = C m g \rho_e a^4/(\mu_0^2 m_d^2 w)$ （ a 为管半径、 m_d 为磁偶极矩），是少数中学装置能逼近解析理论的电磁问题。

已有工作：MacLachy 等（Am. J. Phys., 1993）建立经典测量方案；Donoso 等（Eur. J. Phys., 2009）给出薄壁解析模型与实验对照；2021 年有论文报告了铜管温度对阻尼力的影响（终端速度随管温升高而增大），并指出此前文献只扫过厚度与材质、未扫温度。COMSOL 官方也有此算例。可见单变量版本已被反复做过。

空白在：没有公开工作在同一装置上做壁厚 $\times$ 温度二维扫描并用二维数据联合反演 α 、同时标定薄壁标度的失效边界（厚壁下 v_t 对 w 的依赖偏离 $1/w$ ）。该空白属“参数维度 + 评价维度”型，方法完全公开、结论正负都成立，适合一年期学生课题。

3 · 可检验假设

- H1: 对壁厚 $w/a \leq 0.15$ 的管， v_t 与 $1/w$ 的双对数斜率为 -1.00 ± 0.10 （95% CI）；当 $w/a \geq 0.3$ 时斜率显著偏离 -1 （CI 不含 -1 ）。
- H2: 在 $0 - 80\text{ }^\circ\text{C}$ 范围内由 $v_t(T)$ 反演的铜电阻率温度系数 α 与文献值 0.0039 K^{-1} 之差不超过 15%，且误差棒足以分辨 0 与 0.0039（即能否证“无温度效应”）。

4 · 量化验收标准

1. 方法学校验（硬门槛）：用同一视频测量链对无磁性等尺寸铜柱做自由落体，拟合 g 与 9.79 m/s^2 偏差 $\leq 2\%$ ；再用单摆重测 g 交叉验证，偏差 $\leq 1\%$ 。不过关则后续全部数据无效。
2. 终端速度判据预先写死：取管中段 60% 区间，位置 - 时间线性拟合残差的 $\text{RSD} \leq 2\%$ 方判定“已达终端”；每个 (w, T) 参数点重复下落 ≥ 10 次，报均值 $\pm$ 标准误。
3. 标度指数拟合用正交距离回归（orthogonal distance regression, ODR）与 Theil - Sen 稳健回归交叉验证，自助抽样（bootstrap， ≥ 2000 次）给 95% CI；不得只报双对数最小二乘。
4. 主要误差源清单必须量化：释放初态（用非磁性导管定心）、管温不均（管外保温棉 + 上下两只 DS18B20，温差 $> 2\text{ }^\circ\text{C}$ 的轮次作废）、磁矩随温度漂移（磁铁不入水浴，只加热管子，并用手机磁力计核查磁铁温度前后 B 变化 $< 1\%$ ）。

5. "差异不显著也是有效结论"条款：若某两档壁厚的 v_t 差异不显著，必须给出功效说明——误差棒需小到能分辨按模型预言的差值的 $1/2$ 。
6. 全部脚本（Tracker 导出 + Python 拟合）与原始视频开源，第三方可一键重跑。

5 · 数据与工具（含器材清单与估价）

用途	来源 / 工具	估价
铜管 4 根（同内径 ≈ 16 mm，壁厚 1/1.5/2/3 mm，长 0.5 m）+ 铝管 2 根	五金市场/淘宝定制切割	≈ 350 元
NdFeB 磁铁（D12–15 mm 圆片，叠 2–3 片）	淘宝	≈ 60 元
Arduino Nano + 2×DS18B20 温度探头 + 光电门（对射式，备选计时方案）	淘宝	≈ 80 元
保温水浴（家用保温桶 + 冰/热水）	家用	0 元
视频测量	手机 60/240 fps + Tracker（免费开源）/OpenCV	0 元
拟合与统计	Python（scipy.odr、numpy bootstrap），纯 CPU 秒级	0 元
对照基准	Donoso 2009 薄壁公式、铜 α 文献值（ 仅用于校验对比，不计入本项目贡献 ）	—

合计约 500 – 700 元。安全：磁铁夹伤与低温/热水烫伤为仅有风险，操作时戴棉手套、热水 $\leq 80^\circ\text{C}$ 。

6 · 方法路径

- 搭装置 + 完成第 4.1 条方法学校验（自由落体 g 、单摆 g ）。
- 标定终端速度测量链：同一条件连拍 15 次，确认 $\text{RSD} \leq 3\%$ ，否则先改进释放机构。
- 室温下扫壁厚（4 档铜 + 2 档铝），ODR 拟合双对数斜率与失效边界。
- 定一根薄壁铜管扫温度（0/20/40/60/80 $^\circ\text{C}$ 五点），反演 α 。
- 用铝管重复温度扫描做独立交叉校验（铝 $\alpha \approx 0.0043\text{ K}^{-1}$ ）。
- 汇总二维数据，与薄壁/厚壁模型全局对照，写清残差结构。
- 录制答辩视频：装置全景、带刻度尺的下落慢动作、温控读数同框，保留全部原始片段。

7 · 新颖性边界

本课题不声称发现新电磁物理，也不提出新模型； v_t 的薄壁公式、温度效应的存在均已发表（MacLatchy 1993；Donoso 2009；2021 年温度论文）。已有工作各自只扫单一变量。本项目的主结论是：(i) 同一装置上壁厚标度的定量失效边界；(ii) 由 $v_t(T)$ 反演 α 的精度评估——把演示实验升级为“无电极电阻率测量法”并给出其不确定度预算。若反演值与文献一致，是该方法可用性的独立支持；若系统偏离，偏离量本身即厚壁/端部效应的量度。两种结果都成立。

8 · 决策门槛（go / no-go）

- 第 3 周末：方法学校验未达标（ g 偏差 $> 2\%$ ）→ 停止扫描，先排查帧率标定与透视畸变；连续 2 周不过则改用光电门计时替代视频（备选方案已购）。

- 第 8 周末：单条件重复性 $RSD > 5\%$ → 降级路径 A：放弃温度轴，专做室温壁厚 + 材质（铜/铝/黄铜）标度与失效边界，主结论框架（标度检验 + 不确定度预算）不变。
 - 第 16 周末：若温度扫描中磁矩漂移无法控制在 1% 内 → 降级路径 B：改为"冰水 vs 室温"两点式差分测量，只检验 H2 的方向与量级，不报 α 的精确值。
 - 预算裁剪顺序：先砍铝管（-100 元），再砍 3 mm 厚壁管；两者砍掉后 H1 的失效边界结论弱化为"未观察到失效"，其余结论保留。
-

P12 · 纸条毛细渗吸的 Washburn $\sqrt{t}$ 定律与蒸发限长：以湿度受控扫描检验稳态爬升高度的标度

标签： 优先级:高 · 桌面实验/流体-多孔介质族 · 预算档位:≤400元 · 手机延时摄影+Tracker · 技能取向:细致操作+曲线拟合

1 · 研究问题

竖直纸条中水的爬升距离 $L(t)$ 在早期是否严格满足 Lucas – Washburn 定律 (Lucas – Washburn equation) $L=\sqrt{(Dt)}$ (指数 0.50) , 其最终停在远低于 Jurin 高度的稳态限长 $L_{\max}$ 是否由蒸发决定、并满足 $L_{\max}\propto\sqrt{(D/e)}$ (e 为单位面积蒸发率) ? 进一步: 改变环境相对湿度 (RH) 能否定量预测 $L_{\max}$ 的移动?

2 · 研究背景与空白

多孔介质毛细渗吸是纸基微流控 (paper-based microfluidics) 的物理基础。Lucas (1918) 与 Washburn (1921) 给出 $L\propto t^{1/2}$; ACS Omega (2020) "Liquid Wicking in a Paper Strip" 做了实验加数值, 指出长时间行为偏离 $\sqrt{t}$; arXiv:2407.20776 (2024) 研究了乙醇在多孔条上的蒸发限制铺展, 指出稳态高度远低于 Jurin 极限、蒸发是主控因素, 并给出理论。即: $\sqrt{t}$ 段与"存在蒸发限长"均已发表。

丘奖 2024 优胜奖 "Wetting Tracing Paper Curling Behavior and Penetration of Fiber Porous Media" (高雄中学) 研究的是湿纸卷曲力学与渗透, 未做蒸发限长的湿度标度——须在报告中点名并区分。

空白在: 公开文献中没有对水在常见纸条上的 $L_{\max}$ 随受控湿度的系统扫描及其与 $\sqrt{(D/e)}$ 标度的定量对照 (已发表工作用乙醇或未控湿度)。湿度控制可用饱和盐溶液密闭箱零成本实现, 结论正负都成立, 是典型"参数维度"空白。

3 · 可检验假设

- H1: 密闭近饱和湿度 ($RH > 95\%$) 下, $\log L - \log t$ 斜率为 0.50 ± 0.03 (95% CI) ; 开放环境下斜率在后期显著低于 0.5。
- H2: 在 $RH \approx 33\%$ ($MgCl_2$ 饱和液)、75% ($NaCl$)、>95% 三档下, $L_{\max}$ 随独立称重测得的蒸发率 e 满足 $L_{\max}\propto e^{-1/2}$, 幂指数 95% CI 含 -0.5。

4 · 量化验收标准

1. 方法学校验 (硬门槛): 用内径已知的玻璃毛细管做 Jurin 毛细上升, 反推水的表面张力, 与 72 mN/m 偏差 $\leq 8\%$; 同时在密闭条件下对基准滤纸复现 $\sqrt{t}$ 指数 0.50 ± 0.03 。不过关则后续数据全部无效。
2. 每个 (纸种 $\times$ RH) 条件重复 ≥ 8 条纸条, 取自同一批次并记录批号; 前沿位置读取用染色水 + Tracker, 读取误差 ≤ 0.5 mm, 判据 (取染色前沿最高连续点) 预先写死。
3. 蒸发率 e 独立测量: 0.01 g 电子秤称湿纸条失重 ≥ 60 min, 线性段斜率给 e , 每条件 3 次。
4. 幂指数拟合: ODR + Theil – Sen 交叉验证 + bootstrap 95% CI; 显著性检验用双侧 t 检验 ($\alpha = 0.05$) , 口径预先写死。
5. 误差源量化: 温度漂移 ($\pm 2^\circ C$ 内有效, 记录 DHT22)、纸条批间差 (以密闭条件 D 值的批间 RSD 报告)、悬挂角偏差 (铅垂 $\pm 2^\circ$) 。“三档湿度下 $L_{\max}$ 差异不显著”亦为有效结论, 但须证明误差棒能分辨标度预言差值的 $1/2$ 。

6. 全部图像、脚本开源可重跑。

5 · 数据与工具（含器材清单与估价）

用途	来源 / 工具	估价
滤纸/色谱纸/普通打印纸各一批（裁 10 mm 宽条）	实验耗材店	≈80 元
玻璃毛细管（内径 0.5/1.0 mm 标称）	淘宝	≈30 元
密闭整理箱 ×3 + MgCl ₂ 、NaCl 试剂级	超市/淘宝	≈100 元
DHT22 温湿度传感器 + Arduino Nano（记录环境）	淘宝	≈50 元
0.01 g 厨房电子秤	淘宝	≈60 元
手机延时摄影 + Tracker/OpenCV；食用色素	自有/超市	≈10 元
对照基准	Washburn 1921 理论、arXiv:2407.20776 模型（仅用于校验对比，不计入贡献）	—

合计约 350 元。安全：无风险试剂（氯化镁、食盐）。

6 · 方法路径

- 完成毛细管 Jurin 校验与密闭 $\sqrt{t}$ 基准（第 4.1 条）。
- 预实验确定单次爬升总时长与取样帧率（预计 30 – 120 min，延时 10 s/帧）。
- 三档湿度箱平衡 ≥ 24 h（DHT22 连续记录），逐条件跑纸条 8 条 + 蒸发率称重 3 次。
- 换第二、第三种纸重复，检验标度是否跨纸种成立。
- ODR/稳健拟合 + bootstrap，绘 $L_{\max} - e$ 双对数总图。
- 独立交叉校验：用称重法测同一纸条的吸水质量 $m(t)$ ，检验 $m \propto L$ 的一致性。
- 视频记录：湿度箱内跑一条全程延时（画面含刻度尺与温湿度屏），另录装置讲解，共 ≤ 3 min 成片，原始素材全保留。

7 · 新颖性边界

本课题不声称发现 $\sqrt{t}$ 定律或"蒸发限制爬升"现象——两者均已发表（Washburn 1921；ACS Omega 2020；arXiv:2407.20776 用乙醇建立理论）。丘奖 2024 已有湿纸卷曲课题（见上），主题为力学形变而非蒸发限长。本项目主结论是：水/常见纸体系中 $L_{\max} -$ 蒸发率标度的首次受控湿度检验，并给出跨纸种普适性判断。指数吻合 $-1/2$ 则理论获得新体系的独立支持；系统偏离则给出偏离量与候选解释（纸内扩散/滞后角），两种结果都有效。

8 · 决策门槛（go / no-go）

- 第 4 周末：Jurin 校验偏差 $> 8\%$ 或密闭 $\sqrt{t}$ 指数 CI 不含 0.5 $\rightarrow$ 排查纸条预处理（恒温恒湿平衡不足）与前沿判读；2 周内不过则换标准定量滤纸（Whatman 1 级）重做基准。

- 第 10 周末：若三档湿度下 $L_{\max}$ 的组间差 $<$ 组内 RSD（无分辨能力） $\rightarrow$ 降级路径 A：放弃湿度轴，改做"密闭 vs 开放"两态对照 + 纸种扫描，主结论换为 $\sqrt{t}$ 失效时刻的标度，框架不变。
 - 第 20 周末：批间 RSD $> 15\%$ $\rightarrow$ 降级路径 B：所有结论改为同批内配对设计（paired design），牺牲跨批普适性声明。
 - 预算裁剪顺序：先砍第三种纸，再砍 0.01 g 秤（改用毛细管失重估 e，精度降级须注明）。
-

P13 · 颗粒孔口流的 Beverloo 标度与堵塞统计：以连续称重判定 $5/2$ 指数与雪崩分布形态

标签：优先级：高 · 桌面实验/颗粒物族 · 预算档位： ≤ 600 元 · Arduino 称重传感器 · 技能取向：自动化采集+统计

1 · 研究问题

自制竖直料仓中颗粒经圆孔的稳态流量 Q 是否满足 Beverloo 定律 (Beverloo law) $Q = C_p \sqrt{g} (D-kd)^{5/2}$ ，拟合指数的 95% CI 是否含 2.5？进一步：当孔径比 D/d 降至 6 以下进入堵塞区时，两次堵塞之间的"雪崩规模" (avalanche size) 分布是否为指数型，其特征规模随 D/d 如何增长？

2 · 研究背景与空白

颗粒流量与孔径的 $5/2$ 幂关系 (Beverloo 1961) 是颗粒物质 (granular matter) 的基石经验律；沙漏计时与料仓设计都依赖"流量与料高无关"这一反直觉事实。Mankoc 等 (2007) 给出跨大范围孔径比的修正表达式；Zuriguél 等系统测量了堵塞概率与雪崩分布，发现雪崩规模呈指数分布、且"是否存在临界孔径比"仍有争议 (To 等 vs Zuriguél 等)；2024–2025 年仍有 DEM 论文在验证 Beverloo 适用界 (arXiv:2512.03698 指出 $D/d \approx 10$ 时流量还随料高变化)。即：定律本身与堵塞统计均为已发表结果。

空白在：以上实验多用精密球形颗粒与实验室天平；用一套 ≤ 600 元装置在同一实验中同时覆盖自由流区 (标度拟合) 与堵塞区 (雪崩统计)，并检验非球形厨房颗粒 (小米/绿豆/玻璃珠对照) 对指数与分布形态的影响，未见系统报道。这是"换样本 + 参数维度"型空白，正负结论均可发表级陈述。

3 · 可检验假设

- H1：对近球形玻璃珠， $Q - (D-kd)$ 双对数拟合指数为 2.50 ± 0.15 (95% CI)，且固定孔径下流量随料高 (3 倍变化) 漂移 $< 3\%$ 。
- H2： $D/d < 6$ 区雪崩规模 s 的生存函数在半对数图上为直线 (指数分布，K-S 检验 $p > 0.05$)，非球形颗粒 (绿豆) 的特征规模在同一 D/d 下显著小于球形颗粒。

4 · 量化验收标准

1. 方法学校验 (硬门槛)：HX711 + 称重传感器用标准砝码 (借学校物理组) 五点标定，线性偏差 $\leq 0.5\%$ ；再复现"流量与料高无关" ($D/d \approx 8$ ，料高 3 档，流量差 $\leq 3\%$)。不过关则全部数据无效。
2. 采样率 ≥ 20 Hz 连续记录 $m(t)$ ，稳态流量取中段线性拟合；每个孔径重复 ≥ 10 次装填，报均值 $\pm$ 标准误、RSD。
3. 指数拟合：ODR + Theil-Sen 交叉验证 + bootstrap 95% CI； k 与指数联合拟合并报告参数相关性；孔径 ≥ 7 档、跨度 ≥ 1 个数量级的 $(D-kd)$ 。
4. 雪崩统计：每个 D/d 采集 ≥ 200 次堵塞事件；分布形态判定预先写死为"指数 vs 幂律"的对数似然比 + K-S 检验 ($\alpha = 0.05$)。
5. 误差源量化：环境湿度 (颗粒吸潮改变摩擦，RH 记录且 $> 70\%$ 天次作废)、装填方式 (漏斗定高倾倒，写成 SOP)、传感器温漂 (每轮前后零点核查)。
6. 脚本 + 原始 $m(t)$ 数据全部开源。

5 · 数据与工具（含器材清单与估价）

用途	来源 / 工具	估价
透明亚克力管（内径 80 mm，长 400 mm）+ 底板支架	淘宝定制	≈150 元
孔板 ≥ 8 片（3D 打印或亚克力激光切孔 3–30 mm）	学校创客室/淘宝代打	≈100 元
玻璃珠（d≈2 mm 与 4 mm 各 2 kg）、小米、绿豆	淘宝/超市	≈120 元
HX711 + 5 kg 称重传感器 + Arduino Nano	淘宝	≈50 元
DHT22 湿度记录、振动排堵用小电机（堵塞后重启流动）	淘宝	≈40 元
手机 240 fps 慢动作（孔口流态定性记录；局限：无法解析单颗粒轨迹，仅作辅助）	自有	0 元
对照基准	Beverloo 1961、Mankoc 2007 修正式、Zuriguel 雪崩统计（仅校验对比，不计入贡献）	—

合计约 450 – 600 元。安全：无。

6 · 方法路径

- 搭料仓 + 完成称重标定与"料高无关性"复现（第 4.1 条）。
- 写 Arduino/Python 自动采集与堵塞自动判定（流量为零持续 > 5 s 判堵塞，判据写死）。
- 玻璃珠扫孔径 7 档以上，ODR 拟合 Beverloo 指数与 k。
- 进入堵塞区，逐 D/d 采 ≥ 200 次雪崩，做分布形态检验。
- 换绿豆/小米重复第 3 – 4 步，对照颗粒形状的影响。
- 独立交叉校验：用秒表 + 定量杯粗测两档流量，与传感器结果比对（系统误差核查）。
- 视频：装置全景、孔口 240 fps 流态、堵塞拱结构特写、数据屏同框；成片 ≤ 3 min。

7 · 新颖性边界

本课题不声称发现 Beverloo 定律或雪崩指数分布——均已发表（Beverloo 1961；Mankoc 2007；Zuriguel 等）。丘奖历届有颗粒相关获奖（2023 优胜 "Falsification of Sand Painting on Induction Cooker" 为振动颗粒成图；2024 金奖为听声计球），均不涉及孔口流标度。本项目主结论是：(i) 低成本连续称重装置对 5/2 指数的独立复测与不确定度预算；(ii) 厨房级非球形颗粒对指数、截距 k 与雪崩特征规模的形状效应定量对照。指数不漂移则普适性获得独立支持，漂移则量化形状依赖，两种结果都有效。"临界孔径比是否存在"的学术争议只作背景，不声称裁决。

8 · 决策门槛（go / no-go）

- 第 4 周末：称重链标定或料高无关性不过 → 检查支架共振与传感器量程；2 周不过改用"接液瓶置于电子秤 + 手机拍秤面读数"的备胎记录法（时间分辨率降至 1 Hz，只影响短时涨落分析）。

- 第 9 周末：单孔径流量 $RSD > 10\%$ → 降级路径 A：孔径档数减至 5、每档重复加倍，只做 H1，放弃雪崩统计；主结论框架（标度独立复测）不变。
 - 第 18 周末：堵塞事件采集速率 < 50 次/日（人工重启太慢）→ 启用振动电机自动排堵；仍不足则雪崩统计只做 2 个 D/d 档，分布形态结论降为"单点检验"。
 - 预算裁剪顺序：先砍第二种玻璃珠粒径，再砍振动电机；砍后 H2 的形状对照缩为"玻璃珠 vs 绿豆"一组，框架不变。
-

P14 · 多米诺倒塌波速的摩擦依赖：以受控表面摩擦扫描检验波速标度与"波死亡"概率

标签：优先级：高 · 桌面实验/刚体动力学族 · 预算档位：≤450元 · 手机240fps+Tracker · 技能取向：视频测量+实验设计

1 · 研究问题

多米诺倒塌波的稳态波速 v 在无量纲化 $v/\sqrt{gh}$ (h 为骨牌高) 后，是否只依赖间距比 s/h 与两类摩擦系数 (牌-牌 μ_d 、牌-地 μ_s)？进一步：改变牌-地摩擦能否按 PRApplied (2022) 模拟的预言，系统改变"波中途死亡" (toppling anomaly) 的发生概率？

2 · 研究背景与空白

多米诺波是碰撞传能的一维孤立波式过程：每张牌作为倒立摆倾倒、撞击下一张，稳态波速由能量注入 (重力) 与碰撞/摩擦耗散平衡决定。量纲分析给 $v=\sqrt{gh} f(s/h, w/h, \mu)$ 。已有工作：McLaren 等实测波速随间距的非单调曲线；"How Fast are Domino Waves?" (2022) 汇总实测；Song, Guo 等 Phys. Rev. Applied 17, 064021 (2022) 用离散元模拟系统扫了 μ_d 、 μ_s ，结论：波速主要受牌-牌摩擦影响，而牌-地摩擦决定是否出现波死亡；间距半厚时增摩擦使波变慢。该文以模拟为主，实验对照有限。

空白在：用受控表面 (同一批骨牌 + 更换不同粒度垫面) 对" μ_s 控制波死亡概率"这一模拟预言做系统实验检验，公开文献未见。这是"评价维度 + 换方法 (模拟→实验)"型空白；实验只需骨牌与手机，正负结论均有效。

3 · 可检验假设

- H1: 固定 $s/h=0.5$ ， $v/\sqrt{gh}$ 在 4 档垫面摩擦 (μ_s 约 0.2–0.8) 间的变化 $\leq 8\%$ ，即波速对 μ_s 不敏感 (与 PRApplied 模拟一致)。
- H2: 波死亡概率 P_{death} (每 50 张牌内至少一次中断) 随 μ_s 下降而单调上升，最低摩擦档与最高摩擦档差异显著 (Fisher 精确检验， $\alpha=0.05$)。

4 · 量化验收标准

1. 方法学校验 (硬门槛)：单张骨牌从静止倾倒的全程用 240 fps 拍摄，倒地时间与倒立摆刚体数值积分 (自写 Python, RK4) 偏差 $\leq 5\%$ ；同时以 240 fps 拍自由落体标定帧率与像素尺， g 偏差 $\leq 2\%$ 。不过关则全部波速数据无效。
2. 波速测量： ≥ 60 张牌链，取中段 40 张的前沿到达时刻线性拟合；每条件重复 ≥ 15 条链 (含重摆)，报均值 $\pm$ 标准误；单条件 RSD $\leq 5\%$ 为合格。
3. 摩擦系数独立测量：斜面起滑角法，每对界面重复 ≥ 20 次，报 μ 均值与 95% CI (bootstrap)。
4. $v(s/h)$ 曲线扫 ≥ 7 个间距点；标度/趋势拟合用 ODR + Theil–Sen + bootstrap CI；概率类结论 (H2) 用 Fisher 精确检验并报效应量与 CI，每档 ≥ 40 条链。
5. 误差源量化：摆牌间距公差 (用 3D 打印梳形定位规， $\pm 0.3\text{ mm}$)、首牌触发一致性 (定高摆锤触发)、垫面平整度 (水平仪 $\pm 0.2^\circ$)。"波速差异不显著"是 H1 的支持性结论，须证明误差棒可分辨 8% 的差。
6. 全部视频、逐帧数据与脚本开源。

5 · 数据与工具（含器材清单与估价）

用途	来源 / 工具	估价
标准号牌 200 张（同批次）	淘宝	≈120 元
垫面 4 种（亚克力板、木板、砂纸 P240/P1000 贴板）	五金店	≈100 元
3D 打印定位梳 + 摆锤触发器	学校创客室	≈50 元
数字水平仪、钢尺	淘宝	≈60 元
手机 240 fps 慢动作（局限：240 fps 对波速 ~1 m/s、牌距 ~2 cm 足够，无法解析碰撞瞬间接触力，接触细节不做定量声明）	自有	0 元
Tracker/OpenCV + Python（RK4 倒立摆、统计）	免费	0 元
对照基准	PRApplied 17, 064021 (2022) 模拟结果、McLaren 实测曲线（仅校验对比，不计入贡献）	—

合计约 350 – 450 元。安全：无。

6 · 方法路径

1. 完成帧率标定 + 单牌倒立摆校验（第 4.1 条）。
2. 建立摆牌 SOP（定位梳 + 触发器），在基准垫面复现 $v(s/h)$ 曲线形状并与文献汇总数据对照。
3. 斜面法测全部 8 个界面组合的摩擦系数。
4. 固定 $s/h=0.5$ ，扫 4 档垫面，各 ≥ 15 条链测波速（H1）。
5. 同一扫描中记录波死亡事件，各档补测至 ≥ 40 条链（H2）。
6. 独立交叉校验：用 phyphox 声学秒表听号牌撞击声测前沿到达时刻，与视频法比对系统差。
7. 视频：全链俯拍 + 侧拍 240 fps、波死亡实例集锦、装置讲解；成片 ≤ 3 min。

7 · 新颖性边界

本课题不声称发现多米诺波速标度—— $v(s/h)$ 实测与理论已发表数十年（McLaren；van Leeuwen 2004 理论；PRApplied 2022 模拟摩擦效应）。丘奖 2024 银奖 "The Inseparable Paper Vice"（北京十一学校）研究纸页间摩擦放大，与本题仅共享"摩擦"关键词，对象与机制不同。本项目主结论是：对 PRApplied 模拟的两条可证伪预言（波速对 μ_s 不敏感；波死亡受 μ_s 控制）的首次系统实验检验。证实则模拟获得实验支持；证伪则给出实验 – 模拟分歧的量化描述（可能源于牌形/恢复系数差异），两种结果都有效。

8 · 决策门槛（go / no-go）

- 第 4 周末：单牌校验偏差 $> 5\%$ → 排查号牌质心/尺寸测量与空气阻力假设；2 周不过则校验基准改为"双牌碰撞后角速度比"这一相对量。

- 第 8 周末：波速 $RSD > 5\%$ → 摆牌 SOP 迭代（加长定位梳）；仍不过则降级路径 A：所有比较改为同日配对设计 + 非参数检验（Wilcoxon），主结论框架不变。
 - 第 16 周末：若 4 档垫面的波死亡都过于罕见（每档 < 3 例/40 链，无统计功效）→ 降级路径 B：把间距推到大间距失稳区（ $s/h > 0.8$ ）重扫，把 H2 改为"死亡概率随 s/h 与 μ_s 的联合相图"，框架不变。
 - 预算裁剪顺序：垫面 4 档减为 2 档（极端对照），H2 从趋势检验降为两点比较；H1 不受影响。
-

P15 · 瓶子排空的"咕咚"振荡：以孔径—黏度—气腔柔度三轴扫描检验气弹簧模型的周期与排空时间标度

标签： 优先级：中高 · 桌面实验/两相流—声学族 · 预算档位：≤500元 · phyphox+压强传感器 · 技能取向：信号处理+机理检验

1 · 研究问题

倒置瓶经底孔排水时的"咕咚" (glug-glug) 振荡周期 T 与总排空时间 T_e ，是否满足 Clanet & Searby (J. Fluid Mech. 510, 2004) 的标度 $T_e/T_{e0}=(D_0/d)^{5/2}$ (D_0 为瓶径、 d 为孔径)？进一步：人为增大顶部气腔的柔度（接入刚性副气室）时，振荡频率是否按"气体压缩性 = 弹簧"模型预言的 $f \propto 1/\sqrt{V_{\text{gas}}}$ 移动——这是对机理的因果检验而非拟合检验。

2 · 研究背景与空白

倒置瓶排液是气液交替通过孔口的自激振荡：液柱下落使顶部气腔膨胀降压，直至气泡从孔口上吸补气，周而复始。Clanet & Searby (2004) 把它建为"气腔弹簧 + 液柱质量"振子，给出周期与排空时间标度并在一个数量级孔径范围内验证；Mer 等 (2019) 用压强测量细化了两相流细节；Kohira 等研究了"塑料瓶振荡器"的开关型振荡与双瓶耦合同步。即：模型与主标度均已发表。

空白在：已有实验均在固定气腔构型下做参数拟合；通过外接副气室直接操纵"弹簧常数"做因果检验，以及黏度轴（水 - 甘油）对周期与标度指数的影响，未见系统报道（Clanet 明确限于低黏度极限）。孔径与副气室改造仅需瓶盖钻孔与软管，正负结论均有效。

3 · 可检验假设

- H1: $T_e - d$ 双对数拟合指数为 -2.5 ± 0.2 (95% CI)，覆盖 $d = 4 - 25$ mm。
- H2: 接入副气室使有效气体体积 V_{gas} 增至 $2\times$ 、 $4\times$ 时，振荡频率按 $f \propto V_{\text{gas}}^{-1/2}$ 下移，实测指数 95% CI 含 -0.5 ；若黏度增至 20 倍（甘油水溶液）后指数显著偏离 -2.5 ，则记录偏离进入高黏修正讨论。

4 · 量化的验收标准

1. 方法学校验（硬门槛）：无孔限制的敞口圆筒排空时间与 Clanet 参考值 $T_{e0} \approx 3.0 \text{ L} / \sqrt{gD_0}$ 偏差 $\leq 10\%$ ；MPX5010 压强传感器用水柱压强计五点标定，偏差 $\leq 2\%$ 。不过关则全部数据无效。
2. 振荡频率双通道测量：phyphox 音频谱（咕咚声计数）与气腔压强时序（Arduino ≥ 100 Hz 采样）各自给 T ，两法差 $\leq 3\%$ 方采纳；每条件重复 ≥ 8 次。
3. 指数拟合：ODR + Theil - Sen 交叉验证 + bootstrap 95% CI；孔径 ≥ 7 档且跨度接近一个数量级；温度记录（水温影响黏度与气体状态， $\pm 2^\circ\text{C}$ 内有效）。
4. 误差源量化：孔口毛刺（钻孔后统一倒角，孔径用游标卡尺三向测量取均值 ± 0.05 mm）、初始液面高度复现（ ± 2 mm）、瓶壁形变（选硬质瓶并压强下检查形变 $< 1\%$ 体积）。
5. "黏度轴差异不显著"是有效结论，但须证明误差棒能分辨 10% 的周期移动。
6. 原始音频、压强时序与脚本全部开源。

5 · 数据与工具（含器材清单与估价）

用途	来源 / 工具	估价
硬质广口瓶/亚克力圆筒（ $D_0 \approx 80\text{ mm}$ ）+ 可换孔板瓶盖 ≥ 7 档	淘宝 + 创客室钻孔	≈ 150 元
MPX5010DP 压强传感器 + Arduino Nano + 软管三通	淘宝	≈ 70 元
副气室（500/1500 mL 硬质瓶各一，接头阀门）	淘宝	≈ 60 元
甘油 2.5 L（配 20—60% 水溶液）+ 温度计	淘宝	≈ 90 元
phyphox（音频谱/秒表）、Tracker（液面下降辅助）	免费	0 元
对照基准	Clanet & Searby 2004 标度、Mer 2019（仅校验对比，不计入贡献）	—

合计约 400 – 500 元。安全：甘油无毒；钻孔操作戴护目镜。

6 · 方法路径

- 搭装置，完成敞口排空基准与压强标定（第 4.1 条）。
- 双通道频率测量一致性核查（第 4.2 条），确定 SOP。
- 水介质扫孔径 7 档，拟合 $T_c(d)$ 与 $T(d)$ 。
- 固定中档孔径，接副气室扫 V_{gas} 三档，检验 $f \propto V_{\text{gas}}^{-1/2}$ （H2 前半）。
- 换 3 档甘油浓度重复孔径扫描（可减为 5 档），检验指数漂移（H2 后半）。
- 独立交叉校验：用称重法（瓶下接秤）测排空质量流量曲线，与压强/音频推算的周期结构比对。
- 视频：排空全程正拍 + 孔口气泡 240 fps 慢动作（局限：泡界面细节仅定性）、压强曲线屏幕同框；成片 $\leq 3\text{ min}$ 。

7 · 新颖性边界

本课题不声称发现咕咚振荡或其标度—— $(D_0/d)^{5/2}$ 律与气弹簧模型均为 Clanet & Searby (2004) 已发表结论；双瓶耦合同步也已有工作（Kohira 等）。丘奖历届有 2021 优胜 "Dynamic Stabilization of Water Bottles"（晃瓶行走稳定性）与 2023 铜奖 pop-pop 船（气腔两相振荡引擎），后者与本题共享"气腔弹簧"要素但对象为推进器件——均须点名。本项目主结论是：(i) 标度指数的独立复测；(ii) 副气室操纵下的机理因果检验——这是从"拟合一致"升级为"干预一致"的评价维度转换；(iii) 高黏区偏离的定量记录。频率若不按 $V_{\text{gas}}^{-1/2}$ 移动，即构成对气弹簧图像的有效反例，两种结果都成立。

8 · 决策门槛（go / no-go）

- 第 4 周末：敞口基准偏差 $> 10\%$ 或双通道频率不一致 $\rightarrow$ 排查瓶口涡流与传感器气路密封；2 周不过则频率只用压强通道并加大重复数。
- 第 9 周末：同条件周期 $\text{RSD} > 8\%$ $\rightarrow$ 降级路径 A：放弃黏度轴，专做水介质"孔径 + 副气室"两轴，主结论（标度复测 + 因果检验）不变。

- 第 18 周末：副气室接入后若振荡直接熄灭（进入连续流态）→ 把该现象转正为研究内容：测"振荡 - 连续流转变"的临界 $V_{\text{gas}}(d)$ 边界，仍在气弹簧框架内讨论，属降级路径 B。
 - 预算裁剪顺序：先砍甘油（-90 元，删 H2 后半），再减孔径至 5 档（CI 变宽，框架不变）。
-

P16 · 橡胶气球的非单调压强-体积曲线与双气球失稳：以循环充放气量化 Mullins 预处理对失稳阈值的影响

标签：优先级：中高 · 桌面实验/软物质-热力学族 · 预算档位：≤400元 · Arduino压强传感器 · 技能取向：细致标定+非线性物理

1 · 研究问题

乳胶气球的压强-体积曲线 $P(V)$ 的非单调形状（先峰后谷再上翘）经过多少次充放气循环（Mullins 预处理，Mullins effect）后收敛到稳定曲线——逐循环峰值压强的衰减是否呈指数收敛，特征循环数是多少？进一步：预处理程度不同的两只气球连通（双气球实验，two-balloon experiment）时，终态（等大/不等大）边界是否随预处理状态按 $P(V)$ 曲线预言移动？

2 · 研究背景与空白

气球膜是超弹性（hyperelastic）橡胶球壳，Laplace 关系叠加应变硬化给出非单调 $P(V)$ ：这是双气球"大者更大"反直觉失稳的根源。Merritt & Weinhaus (Am. J. Phys. 46, 1978) 给出双气球实验经典分析；Müller & Strehlow 系统测过气球压强曲线；已发表实验明确指出乳胶有显著迟滞——"气球须充放至少 10 次才能得到稳定曲线"，且隔天部分回复（Mullins 恢复）。即：曲线形状、双气球失稳、迟滞的存在均已发表。

空白在：文献把"预处理 ≥ 10 次"当经验操作规程，未见对收敛动力学（逐循环峰压衰减律、特征循环数、隔天回复量）的定量测量，也未见把预处理状态当自变量去移动双气球失稳边界的实验。这是"评价维度"型空白：把别人扔进 SOP 的步骤变成被测对象。器材只需压强传感器与注射器，正负结论均有效。

3 · 可检验假设

- H1：逐循环峰值压强 $P_{\text{peak}}(n)$ 相对首循环的衰减满足指数收敛 $P_{\text{peak}}(n)-P_{\infty} \propto e^{-n/n}$ ，特征循环数 n 的 95% CI 上界 ≤ 10 （与经验规程一致）；静置 24 h 后回复量可测且 > 3 倍测量误差。
- H2：一只"新"气球与一只充分预处理气球连通时，气体净流向由两者当前 $P(V)$ 曲线预测的压差方向决定，预测正确率 $\geq 90\%$ （ ≥ 20 组）。

4 · 量化验收标准

1. 方法学校验（硬门槛）：MPX5010DP 用自制水柱压强计在 0.3 – 6 kPa 五点标定，偏差 $\leq 2\%$ ；体积测量（大注射器分度 + 排水法交叉）互差 $\leq 3\%$ 。同批气球在充分预处理后复现文献 $P(V)$ 非单调形状（峰-谷结构清晰，峰压量级 3 – 6 kPa）。不过关则全部数据无效。
2. 充放气用 150 mL 注射器定容步进（每步 50 mL，静置 10 s 读稳态压），加载/卸载全程记录，迟滞环面积逐循环报告；每种气球型号 ≥ 5 只样本，报样本间 RSD。
3. 拟合：指数收敛用非线性最小二乘 + bootstrap 95% CI，并用 Theil – Sen 对 $\ln(P_{\text{peak}}-P_{\infty}) - n$ 做稳健交叉验证； P_{∞} 作为拟合参数并报告不确定度。
4. 误差源量化：温度（橡胶模量温敏，实验室 $\pm 2^\circ\text{C}$ 内有效并记录）、气体温升（缓慢步进准等温，步进后 10 s 弛豫）、漏气（每循环终点闭阀 60 s 压降 $< 1\%$ 方有效）、气球批间差（同袋同色）。
5. "隔天回复不显著"亦为有效结论，须给出可分辨 3% 峰压变化的误差棒证明。
6. 全部 $P - V$ 原始数据与脚本开源。

5 · 数据与工具（含器材清单与估价）

用途	来源 / 工具	估价
MPX5010DP ×2 + Arduino Nano + 三通/阀门/软管	淘宝	≈90元
150 mL 注射器 ×2、透明水柱管（标定用）	淘宝	≈50元
同批次乳胶气球 100 只（两种尺寸）	淘宝	≈60元
排水法量筒/量杯、温度计	淘宝	≈60元
手机延时 + 白板记录；Python 拟合（纯 CPU）	免费	0 元
对照基准	Merritt & Weinhaus 1978 双气球理论、已发表 P(V) 曲线与迟滞描述（仅校验对比，不计入贡献）	—

合计约 300 – 400 元。安全：气球爆破声与碎片，充气不超过标称直径 90%，戴护目镜。

6 · 方法路径

1. 搭气路，完成压强/体积标定与 P(V) 形状复现（第 4.1 条）。
2. 制定循环 SOP（步进速率、静置时间、终点体积固定），预实验确认漏气判据可执行。
3. 对 ≥ 5 只气球逐循环测 P(V) 至 20 循环，拟合收敛律与 n^* （H1 前半）。
4. 静置 24 h/72 h 复测，量化 Mullins 回复（H1 后半）。
5. 配对不同预处理程度的双气球连通实验 ≥ 20 组，逐组先用各自 P(V) 曲线做书面预测再开阀（预测判据预先写死），统计正确率（H2）。
6. 独立交叉校验：用视频测气球直径换算体积（球形近似，量化非球形误差），与注射器计量比对。
7. 视频：P(V) 采集全程延时、双气球开阀对决实拍（预测卡片同框）、装置讲解；成片 ≤ 3 min。

7 · 新颖性边界

本课题不声称发现非单调 P(V)、双气球失稳或 Mullins 效应——均已发表（Merritt & Weinhaus 1978；Müller & Strehlow；已发表实验明确记载" ≥ 10 次循环 + 隔天部分回复"）。丘奖 2022 铜奖 "Exploring the Thermodynamic Properties of Rubber Band and Related Refrigerator Design"（上海平和）研究橡皮筋熵弹性制冷，与本题同属橡胶弹性但对象、观测量不同，须点名。本项目主结论是：把预处理从操作规程变为被测对象——收敛动力学的首个定量描述（ n 、回复量），以及预处理状态对双气球终态的可预测操纵。若 n 远大于 10，则直接修正流行实验规程；若双气球终态不可由 P(V) 预测，则说明速率/迟滞效应不可忽略——两种结果都有效。

8 · 决策门槛（go / no-go）

- 第 4 周末：压强标定或漏气判据不过 → 更换阀门与硅胶管；2 周不过则改用凡士林密封 + 缩短单循环时长的快测方案（准等温性下降须在误差预算中加项）。

- 第 9 周末：样本间 $RSD > 15\%$ （批间差淹没循环效应）→ 降级路径 A：全部结论改为单球内配对（每球自身第 n 循环 vs 第 1 循环），放弃跨球平均，主结论框架不变。
 - 第 18 周末：双气球开阀后流动过快无法记录中间态 → 在连通管加毛细节流，把 H2 降级为仅判终态方向；判据不变。
 - 预算裁剪顺序：先砍第二种气球尺寸，再砍第二只传感器（双气球压差改为交替单点测量，时间分辨率降低）。
-

P17 · 热巧克力效应的定量检验：以同步测量气泡空隙率与共振频率判定 Wood 方程的适用边界

标签： 优先级：中 · 桌面实验/声学-两相流族 · 预算档位：≤350元 · phyphox频谱 · 技能取向：信号处理+物理建模

1 · 研究问题

搅动含气液体后敲击杯底所闻音调随时间爬升（热巧克力效应，hot chocolate effect），其瞬时共振频率 $f(t)$ 能否用独立测得的气泡体积分数（void fraction） $\varphi(t)$ 代入 Wood 方程（Wood's equation，气液混合物声速 $c(\varphi)$ ）定量预测——预测误差随 φ 如何变化，方程在多大 φ 、多大气泡尺寸下失效？

2 · 研究背景与空白

液柱敲击音的基频约为四分之一波长驻波 $f=c_{\text{eff}}/4h$ （ h 为液深）。少量气泡使混合物压缩性剧增而密度几乎不变，Wood (1930) 方程给出 c 随 φ 非线性骤降—— $\varphi \sim 1\%$ 即可把声速从 1480 m/s 拉到 ~ 100 m/s，音调因此低几个八度，气泡消退后回升。Crawford (Am. J. Phys. 50, 1982) 命名并解释了该效应；J. Sound Vib. (2012) 用谱图可视化了频率爬升；多篇教学文献（UT Austin 演示论文等）给出定性对照。即：现象、机理与谱图演示均已发表。

空白在：已有工作几乎都只测 $f(t)$ ，把 φ 留作解释性参数；同步、独立地测量 $\varphi(t)$ （由液面抬升高度换算）并与 $f(t)$ 逐时刻对照 Wood 预测的定量闭环，在低成本装置上未见系统报道。这是“评价维度转换”型空白（从演示到定量检验），正负结论均有效。

3 · 可检验假设

- H1：用碳酸水稀释液配制 $\varphi_0 = 0.2\% - 2\%$ 的初始空隙率，敲击基频的实测值与“液面抬升法 φ + Wood 方程 + 四分之一波长”链条的预测值在 $\varphi \leq 1\%$ 区偏差 $\leq 15\%$ 。
- H2：预测偏差随气泡半径增大（静置粗化后期 vs 初期）系统增大——当气泡共振频率接近驻波频率时 Wood 低频极限失效，偏差与 φ 无关而与泡径相关（方向性假设，偏差 - 时间斜率显著为正）。

4 · 量化验收标准

1. 方法学校验（硬门槛）：无泡清水中同一杯的敲击基频换算声速，与 20 °C 纯水 1482 m/s 偏差 $\leq 5\%$ （含杯壁弹性修正的讨论；若系统偏低，改用“频率比”口径：一切结论用 $f_{\text{bubbly}}/f_{\text{clear}}$ 表述，消去几何与杯壁因子）；空气中管共振测声速与 343 m/s 偏差 $\leq 1\%$ （phyphox 链路自检）。不过关则全部数据无效。
2. $f(t)$ 提取：等间隔电磁敲击器（Arduino 驱动，1 次/秒，力度固定）+ phyphox/audacity 谱图峰追踪，峰判据（最大谱峰且带宽 < 50 Hz）预先写死；每条件重复 ≥ 8 杯次。
3. $\varphi(t)$ 独立测量：微距镜头拍液面弯月面刻度，抬升分辨率 ≤ 0.2 mm（对应 $\varphi \sim 0.2\%$ 量级需液深 ≥ 10 cm 的窄口容器；此为装置设计约束，预先写死）；同步时间戳。
4. 拟合与统计： $f - \varphi$ 关系与 Wood 曲线的偏差用 bootstrap 95% CI；H2 的斜率检验用 Theil - Sen + bootstrap（对非高斯残差稳健）；温度记录（声速温漂 ~ 3 m/s/°C）。
5. 误差源量化：敲击力度（电磁敲击器一致性 RSD $< 5\%$ 预校）、泡沿壁聚集（换杯重复）、溶解气持续析出（用放置 24 h 的脱气水做零对照）。“偏差与 Wood 预测不可分辨”即 H1 成立，须证明误差棒 $\leq 15\%$ 判据的一半。

6. 音频原始文件与脚本全部开源。

5 · 数据与工具（含器材清单与估价）

用途	来源 / 工具	估价
高窄玻璃量筒/直筒杯（内径 ~40 mm，高 ~250 mm）×3	淘宝	≈80 元
电磁敲击器（电磁铁 + 继电器 + Arduino Nano）	淘宝	≈70 元
苏打水/碳酸水批量 + 量筒移液配比	超市	≈50 元
手机微距贴镜（液面读数）	淘宝	≈40 元
phyphox / Audacity 谱图、Python 峰追踪	免费	0 元
温度计、遮光背景板	自有	≈20 元
对照基准	Crawford 1982、Wood 方程、JSV 2012 谱图（仅校验对比，不计入贡献）	—

合计约 260 – 350 元。安全：无。

6 · 方法路径

- 完成声学链路自检与清水基准（第 4.1 条），确定是否采用"频率比"口径。
- 校准电磁敲击器一致性与谱峰提取脚本。
- 标定液面法 φ 测量：注入已知体积空气（注射器微量）验证换算，偏差 $\leq 10\%$ 。
- 主实验：4 档初始碳酸水配比 $\times$ 每档 ≥ 8 杯次，同步采 $f(t)$ 与 $\varphi(t)$ 。
- 与 Wood 预测逐时刻对照，量化偏差随 φ 、随时间（泡径代理）的结构（H1、H2）。
- 独立交叉校验：糖粉/可可粉成核方案重复 2 档（更接近原始"热巧克力"场景），检验结论不依赖气源方式。
- 视频：谱图实时爬升与液面同框录制、装置讲解、敲击慢动作；成片 ≤ 3 min。

7 · 新颖性边界

本课题不声称发现热巧克力效应或 Wood 方程——Crawford (1982) 已命名并解释，谱图演示已发表 (JSV 2012)。丘奖 2024 优胜 "Ups and Downs of Objects in Supersaturated Fluid"（重庆育才）研究过饱和液中气泡 – 固体耦合浮沉，与本题共享"析出气泡"要素但观测量（力学 vs 声学）不同；2024 优胜音乐盒扩音课题属结构声学，均须点名。本项目主结论是：低成本装置上 $f-\varphi$ 闭环对 Wood 方程的定量检验及其失效边界。若预测精确，给出该演示实验可作"声学空隙率计"的精度评估；若系统偏离，偏差 – 泡径关系即指向共振修正，两种结果都有效。

8 · 决策门槛 (go / no-go)

- 第 4 周末: 清水基准偏差 $> 5\%$ 且频率比口径仍不稳定 (同杯 $RSD > 3\%$) $\rightarrow$ 换厚壁玻璃容器并加装杯底硬垫; 2 周不过则整题止损转 P15 (器材共通)。
 - 第 10 周末: 液面法 φ 分辨率达不到 0.2% (弯月面抖动) $\rightarrow$ 降级路径 A: 改用精密称重法 (总质量不变、体积变化由溢流管读出) 或把最低 φ 档删去, 只在 $0.5\% - 2\%$ 区检验, 主结论框架不变。
 - 第 20 周末: H2 的泡径代理 (静置时间) 被证明与 φ 强耦合无法分离 $\rightarrow$ 降级路径 B: 放弃 H2, 加密 H1 的 φ 档数 (4 $\rightarrow$ 7 档), 把交付物聚焦为 "Wood 方程偏差 - φ 曲线", 框架不变。
 - 预算裁剪顺序: 先砍微距镜 (用普通微距模式, 分辨率降级入误差预算), 再砍第三只量筒。
-

P18 · 盐水密度振荡器的周期标度与双振子同步：以低成本电导探针长时序判定同步模式的耦合强度边界

标签： 优先级：中 · 桌面实验/非线性动力学族 · 预算档位： ≤ 300 元 · Arduino长时序采集 · 技能取向：非线性物理+自动化

1 · 研究问题

盐水密度振荡器（density/saline oscillator：装盐水的杯底小孔浸入淡水，盐水下流与淡水上流交替）的振荡周期 T 随孔径 d 与密度差 $\Delta\rho$ 的标度关系是什么？进一步：共享同一外槽的两只振子，其同步模式（同相/反相）随耦合强度（外槽水面积）变化的边界在哪里，相位差分布能否定量刻画该转变？

2 · 研究背景与空白

盐水振荡器是最廉价的化学－流体自激振子：Yoshikawa 等（J. Chem. Educ. 66, 1989）给出简明解释；Okamura & Yoshikawa 用三维 Navier－Stokes 数值证明其可约化为 Rayleigh 型方程；耦合实验（Physica D 1997/1998）报道了双振子的自同步与牵引，同步模式依赖外槽水面积（即耦合强度）。即：机理、模型化与"存在同步"均已发表。

空白在：已发表耦合实验以定性模式记录为主，未见用长时序（数千周期）统计给出同步模式转变的耦合强度边界与相位差分布的定量相图；周期－孔径标度也缺跨孔径十倍程的稳健拟合。Arduino 电导探针可无人值守采集数十小时，这正是学生时间充裕、器材无门槛的"评价维度＋参数维度"空白，正负结论均有效。

3 · 可检验假设

- H1：单振子周期 T 与孔径 d 的双对数斜率为负且量值稳定（预期 |斜率| 在 2－4 区间；具体值以预实验定，预注册后不得改），bootstrap 95% CI 宽度 ≤ 0.4 ； T 随 $\Delta\rho$ 单调下降。
- H2：双振子在大耦合（小外槽水面）呈反相锁定（相位差分布峰值在 $\pi \pm 0.3$ rad），随外槽面积增大出现去锁定或换相；转变边界在 ≥ 5 档面积扫描中可定位到相邻两档之间。

4 · 量化验收标准

1. 方法学校验（硬门槛）：电导探针链路先在标准盐水梯度杯中验证读数单调与时间常数 < 1 s；单振子在基准构型下连续记录 ≥ 200 周期，周期 RSD $\leq 10\%$ ，且与文献报道的周期量级（数十秒）一致。不过关则全部数据无效。
2. 周期与相位提取：电压时序带通滤波＋过零/峰检测，判据写死；相位差用 Hilbert 变换（scipy）计算并报告分布直方图与圆统计量（平均合矢量长度 R ）；锁定判据 $R > 0.8$ 预先写死。
3. 每个参数点 ≥ 3 次独立装填、每次 ≥ 100 周期；标度拟合 ODR＋Theil－Sen＋bootstrap 95% CI。
4. 误差源量化：孔径制作公差（针头规格化，用显微拍照实测）、水位下降的慢漂（记录并只取准稳段）、温度（ $\pm 2^\circ\text{C}$ ）、探针极化（交流激励＋定期换向）。
5. "在所有面积档均未见去锁定"亦为有效结论（边界在扫描范围外），须报告可分辨 R 变化 0.1 的统计功效。
6. 全部时序数据与脚本开源，装置图纸公开。

5 · 数据与工具（含器材清单与估价）

用途	来源 / 工具	估价
塑料杯/亚克力小筒 ×6、外槽整理箱 3 种规格 + 面积挡板	淘宝	≈80 元
规格化针头/毛细喷嘴 (0.5—2.5 mm ≥ 5 档)	淘宝	≈40 元
Arduino Nano + 不锈钢电极对 ×3 + 交流激励小电路（自搭）	淘宝	≈60 元
食盐、食用色素、比重计 (0.5 g/L 分度)	超市/淘宝	≈60 元
手机延时摄影（色柱交替可视化，佐证电信号）	自有	0 元
对照基准	Yoshikawa 1989 机理、Physica D 1997/1998 耦合模式（ 仅校验对比，不计入贡献 ）	—

合计约 240 – 300 元。安全：无。

6 · 方法路径

1. 搭单振子 + 电导链路，完成第 4.1 条校验并预实验定孔径可用范围。
2. 写自动采集与周期提取脚本，无人值守跑通 12 h。
3. 扫孔径 5 档 × 密度差 3 档，ODR 拟合标度（H1）。
4. 搭双振子共享外槽构型，扫外槽有效水面积 ≥ 5 档，每档 ≥ 3 次 × ≥ 100 周期，做相位差统计（H2）。
5. 用染色延时视频对 2 个代表工况做独立交叉校验（视觉周期 vs 电导周期，差 ≤ 3%）。
6. 汇总相图，与 Physica D 定性报道对照，写清一致与分歧。
7. 视频：色柱交替特写、双振子同步演示（同相/反相各一段）、数据屏同框；成片 ≤ 3 min。

7 · 新颖性边界

本课题不声称发现盐水振荡器、其 Rayleigh 方程约化或"耦合可同步"——均已发表（Yoshikawa 等 1989；Okamura & Yoshikawa；Physica D 1997/1998）。丘奖历届无盐水振荡器课题；2020 优胜耦合环摆（coupled looping pendulum）属耦合振子题材，对象不同，须点名。本项目主结论是：**同步模式 – 耦合强度定量相图与相位差圆统计刻画**（前人为定性模式记录），以及周期标度的稳健复测。若转变边界清晰，给出可复现的相图；若扫描范围内无转变，同样以功效分析陈述边界下限——两种结果都有效。

8 · 决策门槛（go / no-go）

- 第 5 周末：单振子 200 周期 RSD > 10% → 排查孔口毛刺与液位漂移（加恒液位溢流）；2 周不过则孔径改用注射针头全系列（公差更小），重跑校验。
- 第 10 周末：电导探针长时漂移淹没信号 → 降级路径 A：改为染色 + 手机延时视频自动图像判相（OpenCV 色柱检测），采样率足够（周期数十秒），主结论框架不变。
- 第 20 周末：双振子档间差异无统计功效 → 降级路径 B：缩为"强耦合 vs 弱耦合"两点对照 + 加倍重复数，相图降级为两态判定，框架不变。

- 预算裁剪顺序：密度差 3 档减为 1 档（只保孔径轴），H1 缩窄但 H2 完整保留。
-

P19 · 链条喷泉的高度标度与"基底推力"机制判别：以基底刚度扫描检验 Biggins—Warner 反冲假说

标签： 优先级：中低（机制判别有争议风险） · 桌面实验 / 链绳动力学族 · 预算档位：≤500元 · 手机 240fps+Tracker · 技能取向：力学建模+精细视频测量

1 · 研究问题

珠链从高处容器自发跃出形成的链条喷泉（chain fountain / Mould effect）高度 h_f 是否随落差 h 线性增长（ $h_f = \alpha h$ ），比例系数 α 对链型如何依赖？进一步：Biggins & Warner (Proc. R. Soc. A, 2014) 主张喷泉源于堆内基底对被拾起链段的反冲推力——若把链堆置于逐渐变软的基底（硬板→泡棉→悬空网）， α 是否按该假说预言系统下降？

2 · 研究背景与空白

链条喷泉由 Steve Mould 视频引爆，Biggins & Warner (2014) 建立动量流模型：仅靠张力拾取无法产生喷泉，需堆与基底提供额外向上推力，模型预言 $h_f \approx 0.14h$ 量级的线性标度并被其实验支持。其后有多篇再检验：Pantaleone (Am. J. Phys.)、"Reexamining the Chain Fountain" (arXiv:1810.13008) 质疑或细化推力来源；Martins (arXiv:1612.09319) 与 2020 年稳态喷泉工作 (arXiv:2009.00047) 给出替代/补充机制。即：现象、线性标度与机制争论均已发表，机制尚无定论。

空白在：各方都承认"推力来自堆-基底相互作用"可由改变基底力学性质直接检验，但公开文献中系统的基底刚度扫描实验很少（多为链型与落差扫描）。这是"参数维度+机制判别"空白；风险在于效应量可能小于测量散布，故排在 P19。

3 · 可检验假设

- H1: 标准球珠链在硬基底上 $h_f - h$ ($h = 1 - 4\text{ m}$, ≥ 6 档) 线性拟合优良 (ODR 斜率 CI 不含 0, 截距与 0 相容), α 落在文献报道量级 0.1 - 0.2。
- H2: 基底从硬亚克力换为软泡棉 (模量差 ≥ 100 倍) 再换为张紧弹性网时, α 单调下降且最硬/最软两档差异显著 (双侧 t 检验, $\alpha=0.05$)；若 α 不变, 则构成对"基底推力主导"的反例证据。

4 · 量化验收标准

1. 方法学校验 (硬门槛)：同一视频链对自由落体小球拟合 g 偏差 $\leq 2\%$ ；再复现链下落端稳态速度 $v \approx \sqrt{2gh}$ (偏差 $\leq 5\%$, Biggins - Warner 亦以此为基准)。不过关则全部数据无效。
2. h_f 判读判据写死：喷泉弧顶最高点相对容器口沿的高度，取稳态段 (起动后 1 s 至链余 1/3 前) 逐帧最高点的中位数；每条件重复 ≥ 12 次，报中位数 $\pm$ bootstrap CI。
3. 标度拟合：ODR + Theil - Sen 交叉验证 + bootstrap 95% CI；跨链型 (≥ 2 种珠距/珠径) 重复。
4. 误差源量化：链的初始堆叠形态 (SOP: 绕环式装堆, 拍照存档)、容器口沿摩擦 (统一光滑口沿)、气流 (室内关窗, 烟线检查)；240 fps 手机慢动作用于拾取区定性观察，明确写明局限：无法解析珠间碰撞接触时间 (亚毫秒级)，机制判别只依赖 α 的宏观统计，不依赖接触细节。
5. " α 各基底无显著差异"亦为有效结论 (反例证据)，须证明误差棒可分辨 20% 的 α 变化。
6. 全部视频与逐帧数据开源。

5 · 数据与工具（含器材清单与估价）

用途	来源 / 工具	估价
金属球珠链（2.4 mm 与 4.5 mm 珠径，各 15 m）	淘宝	≈180 元
烧杯/塑料筒容器 + 光滑口沿套环	淘宝	≈40 元
基底三件：亚克力硬板 / EVA 泡棉（两种硬度） / 弹性网框	淘宝	≈100 元
落差架设：楼梯间 + 可调挂点横杆、卷尺、铅垂线	五金店	≈80 元
手机 240 fps + 三脚架 + 大字标尺背景板	自有 + 淘宝	≈60 元
Tracker/OpenCV + Python 统计	免费	0 元
对照基准	Biggins & Warner 2014 模型与 α 量级、arXiv:1810.13008 再检验（仅校验对比，不计入贡献）	—

合计约 400 – 500 元。安全：高处作业在楼梯间护栏内侧完成，链坠落区设警戒；不登梯攀高。

6 · 方法路径

- 完成视频链校验与 $\sqrt{2gh}$ 基准（第 4.1 条），确立装堆 SOP。
- 硬基底扫落差 6 档 $\times$ 12 次，确立 $h_f - h$ 线性与 α （H1）。
- 固定最佳落差（喷泉最显著档），扫基底 4 档 $\times \geq 12$ 次（H2）。
- 换第二种珠链重复第 2 – 3 步，检验结论跨链型稳健性。
- 拾取区 240 fps 观察：记录链离堆瞬间的形态学差异（硬 vs 软基底），只作定性佐证。
- 与 Biggins – Warner 模型及各再检验文献逐条对照，写清支持/反对证据链。
- 视频：全景喷泉 + 标尺、基底对照并排剪辑、拾取区慢动作；成片 ≤ 3 min。

7 · 新颖性边界

本课题不声称发现链条喷泉或线性标度——Biggins & Warner (2014) 已给出模型与实验，机制争论（Pantaleone；arXiv:1810.13008 等）也已公开。丘奖 2020 金奖 "Spontaneous Knotting of Strings"（南外）属链绳统计力学，对象不同，须点名。本项目主结论是：以基底刚度为自变量的机制判别实验—— α 随基底软化显著下降则为"基底推力"假说提供直接干预证据；不变则为反例证据。本项目不声称终结机制争论，只声称提供一条此前缺失的、可复现的干预性证据链；两种结果都有效。

8 · 决策门槛（go / no-go）

- 第 5 周末： h_f 判读的单条件 RSD $> 15\%$ （装堆形态不可控）→ 改用定制绕线轴喂链（稳态喷泉方案，参照 arXiv:2009.00047）；该改造预算 +100 元，已在总预算内。
- 第 12 周末：H1 线性都无法稳定复现 → 降级路径 A：课题重心移至" α 对链型（珠距/珠径/线密度）的依赖"这一纯标度研究，放弃机制判别，主结论框架（标度 + 不确定度预算）保留。

- 第 24 周末：基底效应量 $<$ 可分辨阈（功效不足） $\rightarrow$ 降级路径 B：把基底轴极端化（刚性板 vs 完全悬空网两点）并把重复数加到 ≥ 30 ，结论降为两点判定；仍不显著则如实报告反例证据并给功效分析——该结果本身可写。
 - 预算裁剪顺序：先砍第二种珠链（-90 元，牺牲跨链型稳健性），基底件与落差架不可砍。
-

P20 · 倾斜管中 Kelvin—Helmholtz 失稳的起始时间与波长：以倾角—密度差二维扫描对照 Thorpe 理论

标签：优先级：低（本组风险最高） · 桌面实验/分层流体族 · 预算档位：≤1200元 · 手机60–240fps+图像测量 · 技能取向：流体实验+稳定性理论

1 · 研究问题

密封长直管内两层不同密度液体在管子突然倾斜后产生加速剪切流，界面上 Kelvin – Helmholtz 失稳（Kelvin – Helmholtz instability, KHI）的起始时间 t_c 与初现波长 λ_c ，随倾角 θ 与密度差（Atwood 数 $A=\Delta\rho/2\bar{\rho}$ ）如何变化？实测 $t_c(\theta,A)$ 是否符合 Thorpe（J. Fluid Mech. 39, 1969）的加速流临界条件， λ_c 是否如 Thorpe 报道般系统大于线性理论预言？

2 · 研究背景与空白

倾斜管实验（Thorpe experiment）是产生可控 KHI 的经典桌面构型：倾斜后上下层反向加速，剪切超过临界值时界面长出近驻定的波列。Thorpe (1969) 实测淡水/盐水界面初现波长 $\lambda_c \approx 3$ cm、失稳发生于倾斜后约 1.9 s，理论对 t_c 预测良好、对 λ_c 系统偏小；Fluids 8, 318 (2023) 发表了“低成本 KHI 演示装置 + CFD”，以演示与单工况对照为主。即：构型、理论与低成本化均已发表。

空白在：低成本装置上系统的 (θ, A) 二维扫描并对 t_c 、 λ_c 两个观测量分别做统计对照（Thorpe 原文扫描有限，2023 年低成本文以演示为目的）。风险在于每次实验后两层液体部分混合、需重新配液，实验通量低；且 t_c 、 λ_c 判读主观性强，须靠预写死判据压制——故列为 P20。

3 · 可检验假设

- H1: t_c 随 $\sin\theta$ 的标度与 Thorpe 加速流理论一致：理论预言 $t_c \propto (\sin\theta)^{-1/2}$ 量级依赖（以自写理论计算为准，预实验前完成推导并预注册），实测指数 95% CI 覆盖理论值。
- H2: λ_c 与理论预言之比 >1 且在扫描范围内近似常数（复现并推广 Thorpe 的“波长系统偏大”结论）；若该比值随 A 漂移，则记录漂移作为界面厚度效应的证据。

4 · 量化验收标准

1. 方法学校验（硬门槛）：先在基准工况（贴近 Thorpe 原文参数： $A \sim 10^{-3} - 10^{-2}$ 、 $\theta \approx 4^\circ$ ）复现 λ_c 与 t_c 的文献量级—— λ_c 在 2 – 4 cm、 t_c 在 1 – 3 s 内，且 5 次重复的 $RSD \leq 15\%$ ；同时用静水压/比重计核密度配制误差 $\leq 5\%$ 。不过关则不得进入扫描，全部后续数据无效。
2. 判据写死： t_c = 界面扰动幅值超过界面厚度 2 倍的首帧（图像灰度剖面自动判定，OpenCV 脚本，不许人工挑帧）； λ_c = 首个可辨波列相邻 ≥ 4 个波峰间距的均值（FFT 与直接量测双口径互核，差 $\leq 10\%$ ）。
3. 扫描规模： $\theta \geq 4$ 档 $\times A \geq 3$ 档，每格 ≥ 5 次重复；指数拟合 ODR + Theil – Sen + bootstrap 95% CI。
4. 误差源量化：初始界面厚度（缓注法 + 静置时间标准化，界面厚度由染色灰度剖面实测并入模型讨论）、倾斜动作一致性（限位挡块 + 配重杠杆，一次到位，动作时间 < 0.3 s 且录像核查）、温度（密度温漂， $\pm 1^\circ\text{C}$ ）。

- "指数与理论不可分辨"即 $H1$ 成立；显著偏离亦为有效结论（加速流近似或界面厚度失效的量化），均须给误差棒功效说明。
- 全部视频、判定脚本与理论推导笔记开源。

5 · 数据与工具（含器材清单与估价）

用途	来源 / 工具	估价
透明矩形亚克力密封长槽（内腔约 1000×100×30 mm，带注液口，定制）	淘宝亚克力定制	≈450 元
倾斜机构：铰链底座 + 限位挡块 + 配重杠杆 + 数显角度计	五金 + 淘宝	≈150 元
盐、比重计（1.000—1.100，分度 0.001）、食用色素、缓注漏斗	淘宝	≈100 元
照明：LED 平板背光 + 描图纸匀光	淘宝	≈100 元
手机 60/240 fps + 三脚架（240 fps 仅用于失稳瞬间；波列生长时间尺度 ~1 s，60 fps 已足； 不假设高速相机 ）	自有	0 元
OpenCV/Tracker + Python（含自写线性稳定性数值求解，纯 CPU）	免费	0 元
对照基准	Thorpe 1969 实测与理论、Fluids 8, 318 (2023) 低成本装置（仅校验对比，不计入贡献）	—

合计约 800 – 1200 元（本组最高，仍远低于 3000 元上限）。安全：仅盐水与亚克力，槽体密封后倾斜时防滑落砸脚，双人操作。

6 · 方法路径

- 完成理论功课：推导/数值求解加速两层流的临界条件与最不稳定波长（预注册理论预言）。
- 定制槽体到货后做密封与注液排气工艺定型；建立缓注双层液 SOP（预计每轮配液 + 静置 40 min，评估日通量）。
- 完成第 4.1 条基准复现，锁定判定脚本。
- 执行 (θ, A) 扫描（ ≥ 60 轮，跨 8 – 10 周），中途每 10 轮插入一次基准工况回测（漂移核查）。
- ODR/稳健拟合 t_c 、 λ_c 标度，与预注册理论对照（ $H1$ 、 $H2$ ）。
- 独立交叉校验：选 2 个工况用染色浓度剖面反演界面厚度，检验厚度修正能否解释 λ_c 偏大。
- 视频：倾斜瞬间至波列破碎全程 60 fps + 失稳瞬间 240 fps、背光波列特写、装置讲解；成片 ≤ 3 min。

7 · 新颖性边界

本课题不声称发现 KHI 或倾斜管构型——Thorpe (1969) 是奠基实验，"波长偏大"也是其原文结论；低成本化已由 Fluids (2023) 发表（该文以演示 + CFD 对照为主）。丘奖 2025 铜奖 "Study on Liquid Sloshing"（复旦二附中）属容器晃动非线性动力学，2020 优胜两相液滴下落属界面流，对象均不同，须点名。本项目主结论是：低成本装置上 t_c 、 λ_c 对 (θ, A) 的二维系统扫描与预注册理论的统计对照——把一次性的经典对照升级为参数化检验。指数吻合则理论在弱分层桌面区间获得独立支持；"波长偏大比"若随 A 漂移则给出界面厚度效应的定量证据——两种结果都有效。

8 · 决策门槛 (go / no-go)

- 第 6 周末：槽体密封或注液工艺不成（界面注成即混）→ 换更小截面槽（混合减弱）或改用糖水/盐水更小 A ；2 周不过则止损转 P15/P17（预算未大额支出前定制槽先不下单，第 2 周才订）。
- 第 10 周末：基准工况 $RSD > 15\%$ 或判定脚本与人工判读系统不一致 → 判据迭代一轮；仍不过则降级路径 A：放弃 t_c ，只做 $\lambda_c - (\theta, A)$ （波长对判读噪声更稳健），主结论框架（参数化理论对照）不变。
- 第 24 周末：扫描通量不足（ < 30 轮完成）→ 降级路径 B：A 轴砍到 1 档，做纯倾角扫描 + 加密重复，结论范围收窄，框架不变。
- 选择前提：仅当学生接受"配液 - 静置 - 单次拍摄"的低通量节奏且有稳定实验场地时才启动本题；否则顺位改选 P15 - P18。
- 预算裁剪顺序：先砍 LED 背光（自然光 + 台灯，图像信噪比降级入误差预算），角度计改手机倾角仪；定制槽不可砍。

关于第二批课题（P21—P40）

本赛道的后 20 则出自独立撰写的第二批方案，与前 20 则同样遵循八块结构、方法学校验硬门槛与 go/no-go 要求，选题角度则明显不同——可以当作互补的第二组视角来读。

但有一处差别必须先说清楚：第二批在撰写时未做英文文献检索。它的新颖性边界依据的是历届获奖图谱与既有知识，而不是当轮查新。这意味着"这个想法是否已被发表"这一关尚未过。每则的正文里都重复标注了这一点。若选中第二批的某一则，第一件事是补做英文检索协议（把想法写成一句英文摘要，抽 4—6 个关键词组合检索，并专门搜"最显而易见的那个做法"），再决定是否启动。

P21 · 「热巧克力效应」中气泡液体声速的时变规律：手机频谱测量与 Wood 方程检验

★★★ · A 族·桌面实验 · 物理赛道·声学/流体 · 手机 + 厨房器材 <100 元 · 声学测量 + Python 数据分析

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

搅拌后含气泡的热液体中，敲击杯底音调随时间恢复的频率曲线 $f(t)$ ，能否由 Wood 方程(气泡液体声速公式)加上单一气泡消除时间常数的模型定量预测？液体温度与气泡引入方式如何改变该时间常数？

2 · 研究背景与空白

含微小气泡的液体声速远低于纯水：气泡提供压缩性、液体提供惯性，Wood 方程给出体积分数 $\varphi \sim 10^{-3}$ 量级的气泡即可把声速压低数倍。杯中驻波频率正比于声速，因此敲击音调是气泡含量的实时探针——这就是冲热巧克力/速溶咖啡时音调在几秒内爬升的原因(hot chocolate effect)。

Crawford(1982, Am. J. Phys.)对该效应做了经典的定性与半定量解释，报告音调爬升可达约 3 个八度，并指出 Wood 方程可解释量级。此后该现象常作为教学演示出现。

空白在于：该效应缺乏系统的参数扫描(温度 $\times$ 气泡引入方式 $\times$ 容器几何)与配套的公开数据集和不确定度预算。手机频谱 app 使逐秒追踪 $f(t)$ 变得零成本，一年期学生课题恰好能补上这块——且"模型拟合好/不好"两种结果都是有效结论。

3 · 可检验假设

- H1: $f(t)$ 恢复曲线可由"Wood 方程 + 气泡体积分数单指数衰减 $\varphi(t) = \varphi_0 e^{-(t/\tau)}$ "模型拟合，决定系数 $R^2 \geq 0.95$ ；时间常数 τ 随水温升高($50 \rightarrow 90^\circ \text{C}$, ≥ 3 档)单调下降。
- H2: 由频率反演的初始气泡体积分数 φ_0 与独立几何测量(搅拌前后液面体积差)在同一数量级(比值落在 $0.3 - 3$ 之间)；若系统偏离，偏离方向可由气泡尺寸分布假设解释(该解释须标注为推测)。

4 · 量化验收标准

1. 方法学校验(硬门槛)：静置纯水杯的敲击基频与一维驻波模型($f=c/4h$ ，水深 h 实测)预测偏差 $\leq 5\%$ ，且用信号发生器 app 播放 440/1000 Hz 标准音校验手机频谱管线的频率读数误差 $\leq 1\%$ 。此条不过关，后续全部结论无效。
2. 每组实验条件重复 ≥ 10 次； $f(t)$ 曲线提取判据(功率谱主峰，窗长 0.1 s，峰高阈值)预先写死并写入脚本，不许逐条人工挑峰。
3. 参数扫描 ≥ 3 个温度 $\times \geq 3$ 种气泡引入方式(搅拌/倾倒/吸管注气)，合计 ≥ 90 条有效曲线； τ 的拟合给自助法(bootstrap, ≥ 1000 次重抽)95% 置信区间。
4. H1 的"单调下降"用 Spearman 秩相关 + 置换检验报 p 值，不用肉眼判断。
5. 物理自洽检查：反演 $\varphi_0 \leq 5 \times 10^{-2}$ (超出则说明模型失效区，需报告)。

6. 可复现性：全部原始音频、脚本与处理流程开源(GitHub/Gitee)，第三方可一键重跑出全部图表。

5 · 数据与工具

用途	来源 / 工具
声学采集	手机 + Phyphox(开源,phyphox.org)或任意录音 app，采样 44.1 kHz；纯手机即可
频谱分析	Python:numpy / scipy.signal(STFT)；全程纯 CPU，单条曲线处理 < 1 s
温度测量	学校实验室温度计或 <30 元电子测温计；冰水/沸水两点校准
理论对照	Wood 方程(标准两相介质声学公式，教材级);Crawford 1982 结果仅用于校验与对比，不计入本项目的数据贡献
器材	玻璃杯/量筒、吸管、可可粉或洗洁精微泡液，合计 <100 元

6 · 方法路径

1. 搭建测量链：手机固定支架 + 标准敲击器(定高落锤小球)，用标准音完成频率校准(验收 1)。
2. 跑通纯水基线：验证驻波模型、确定敲击可重复性(基频标准差/均值 $\leq 2\%$)。
3. 编写自动峰提取脚本，判据写死；用 3 条人工标注曲线验证脚本一致性。
4. 执行 3×3 参数扫描，每格 ≥ 10 次；当日备份原始音频。
5. 拟合 Wood + 指数衰减模型，输出 $\tau(T)$ 与自助法置信区间。
6. 独立交叉校验：用液面体积差反演 φ_0 ，与声学反演对照(H2)。
7. 整理开放数据包 + 双语术语表(bubbly liquid, Wood equation, standing wave)。

7 · 新颖性边界

- 本课题不声称发现新物理：效应机制与 Wood 方程解释已由 Crawford(1982)发表，本项目不得声称为自己的发现。
- 已有工作:Crawford 给出机制与量级估计；教学文献中多为演示性复现。
- 本项目贡献(主结论)：首个系统的温度 $\times$ 气泡引入方式参数扫描 + 时间常数 τ 的定量模型检验 + 公开可复现数据集。属"评价维度的转换 + 参数维度补全"。
- 价值：把一个演示级现象升级为带误差棒的定量检验；若模型在某参数区失效，失效边界本身就是结论。
- 核对获奖清单：声学类相邻获奖有 "Singing Tube"(2020 优胜，波纹管气流激励)、"An Investigation into the Sound Amplification in a Music Box"(2024 优胜)、"The Whistling Metal Plate"(2025 铜奖，结构振声)——均为发声机制研究；本题对象是声速作为气泡含量探针，测量量与物理问题不同。
- " τ 与温度无显著关系"同样是有效结论，但须给出 τ 的置信区间证明有能力分辨 2 倍以内的差异。

8 · 决策门槛(go / no-go)

⚠ 第 4 周末：完成验收 1 校准 + 10 条热液体预实验曲线。可测阈值：敲击频移动态范围 ≥ 1 个八度、单条曲线可追踪时长 ≥ 5 s。若热液体气泡消除过快(<2 s)导致追踪失败,立即降级：改用室温水 + 吸管定量注气体系(气泡量可控、消除更慢),Wood 方程检验的主结论框架完全保留，仅损失"温度维度"。⚠ 第 8 周末：核实项——手机麦克风在 200 – 4000 Hz 的响应是否足够平直(双 app 交叉验证)；若否，改报相对频移而非绝对频率，

分析框架不变。⚠️ 预算裁剪顺序：容器几何维度可砍(不影响主结论)；温度与气泡方式两维度砍一保一仍可成文。

P22 · 揉皱纸张的声发射「雪崩」：脉冲能量幂律指数的稳健估计与材料依赖性

★★★ · A 族·桌面实验 · 物理赛道·统计物理 · 手机 + 纸材 <50 元 · 信号处理 + 稳健统计

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

揉皱纸张时的声发射脉冲能量是否服从幂律分布？其指数在不同材料(打印纸/牛皮纸/铝箔)与不同压缩速率下是否在置信区间内不变？

2 · 研究背景与空白

薄片被揉皱时，能量集中在折痕与顶点，以间歇"咔嗒"脉冲释放——这是实验室里最便宜的"雪崩动力学"系统，与地震、磁畴翻转共享无标度统计的表象。脉冲能量分布若为幂律，说明系统缺乏特征能量尺度。

Houle & Sethna(1996, Phys. Rev. E)测量了揉纸声发射，报告脉冲能量在约两个数量级范围内近似幂律(指数具体值：需核实，搜索关键词 "Houle Sethna crumpling acoustic emission power law exponent")。Kramer & Lobkovsky(1996)研究了单次折痕的声发射。这些工作用实验室麦克风与示波器完成。

空白在于：材料与加载速率维度的系统比较，以及用现代稳健方法(Clauset – Shalizi – Newman 最大似然 + KS 检验定 $x_{\min}$, 2009)替代当年的双对数直线拟合。手机 + Python 即可完成全链条，数据公开、正负结论均成立，适合一年期课题。

3 · 可检验假设

- H1：三种纸材的脉冲能量分布均可用幂律描述(有效标度区 ≥ 1.5 个数量级)，打印纸指数的 95% 自助法置信区间宽度 ≤ 0.3 。
- H2：铝箔(塑性主导)与打印纸(弹塑性混合)的指数差 ≥ 0.2 且置信区间不重叠；若无显著差异,"材料不敏感"即为主结论(须证明分辨力达 0.2)。

4 · 量化验收标准

1. 方法学校验(硬门槛)：用节拍器/信号发生器播放已知间隔、已知幅度比的标准脉冲序列，自建脉冲检测管线的事件检出率 $\geq 95\%$ 、假检率 $\leq 5\%$ ，幅度比读数误差 $\leq 10\%$ ；并用双源法核查手机麦克风在工作声压范围内的线性(两源同放时功率相加偏差 $\leq 15\%$)。此条不过关，后续全部结论无效。
2. 统计口径预先写死：幂律拟合以 Clauset – Shalizi – Newman 最大似然为主方法, $x_{\min}$ 由 KS 距离最小化确定，置信区间用 ≥ 1000 次自助重抽；双对数最小二乘只作对照展示，不作为主结果；幂律 vs 对数正态 vs 截断幂律用似然比/AIC 比较并全部报告。
3. 每种材料 ≥ 15 张纸、单张检出脉冲 ≥ 300 个，单材料合并样本 ≥ 5000 个脉冲。
4. 压缩速率两档(慢速手压 ~ 10 s/张、快速 ~ 2 s/张)，速率由视频帧计时，组间用同一管线。
5. 环境噪声门槛：安静房间本底比最小计入脉冲低 ≥ 20 dB，每次实验前实测并记录。
6. 可复现性：全部音频原始文件 + 检测脚本 + 拟合脚本开源，第三方一键重跑。

5 · 数据与工具

用途	来源 / 工具
声学采集	手机录音 44.1 kHz; 泡沫垫隔振,<20 元
脉冲检测与拟合	Python:numpy / scipy; 幂律拟合用开源 <code>powerlaw</code> 包(Alstott et al., 实现 CSN 方法)+ 自写自助法; 纯 CPU, 全数据集处理 < 10 min
纸材	A4 打印纸、牛皮纸、厨房铝箔,<50 元
理论对照	Houle & Sethna 1996 的指数与标度区间,仅用于校验与对比, 不计入本项目的数据贡献

6 · 方法路径

1. 搭建录音环境并完成标准脉冲序列校验(验收 1)。
2. 预实验 3 张纸, 确定脉冲检测阈值与死时间参数, 判据写死进脚本。
3. 正式采集:3 材料 $\times$ 2 速率 $\times \geq 15$ 张, 原始音频当日归档。
4. 运行 CSN 拟合 + 模型比较 + 自助法置信区间。
5. 核实工具边界: `powerlaw` 包对截断幂律的实现是否与原文一致(读文档 + 合成数据注入测试)。
6. 合成数据闭环测试: 向管线注入已知指数的人工幂律脉冲, 验证指数回收偏差 ≤ 0.05 。
7. 独立交叉校验: 用脉冲计数-幅度阈值法(Zipf 型)复算指数, 与 CSN 结果对照。

7 · 新颖性边界

- 本课题不声称发现揉纸声发射幂律——Houle & Sethna(1996)已发表; 不声称提出新统计方法。
- 已有工作:1996 年实验覆盖单一材料、实验室麦克风、双对数直线拟合时代的统计。
- 本项目贡献(主结论): 材料 $\times$ 速率维度的系统比较 + 现代稳健统计(CSN + 模型比较)下指数是否真的"无标度"的再评估 + 开放数据集。属"评价维度的转换"。
- 价值: 若对数正态在 AIC 下击败幂律, 则纠正一个流传的教学案例; 若幂律稳健, 则给出首个材料依赖性数据。两种结果都成立。
- 核对获奖清单:"Physical Mechanism and Governing Factors of Spontaneous Knotting of Strings"(2020 金奖)证明日常物体的统计物理是本赛道金奖级题材;"The Inseparable Paper Vice"(2024 银奖)同为纸张力学, 但研究的是层间摩擦放大, 与本题的折痕声发射统计无重叠——须在报告中点名并写清差异。

8 · 决策门槛(go / no-go)

⚠ 第 5 周末: 预实验判定。阈值: 单张打印纸可检出脉冲 ≥ 300 个、标度区 ≥ 1.5 decade。若脉冲数不足,降级
路径一: 改用铝箔为主材料(脉冲更清晰响亮);降级路径二: 单张不足则按张合并统计并检验张间同分布性(KS 两样本检验)。两条路径均保留"分布形态 + 材料依赖"主结论框架。⚠ 第 10 周末: 若三模型比较中幂律标度区始终 < 1.5 decade, 把课题主结论转为"揉纸声发射的分布形态判别"(幂律/对数正态/截断幂律), 这本身是统计物理界承认的问题, 框架不变。⚠ 选择前提: 无特殊前提, 适合任何能坚持规范采数的学生; 预算裁剪: 速率维度可砍。

P23 · 颗粒物料孔口流量的 Beverloo 标度:5/2 指数与截止参数对颗粒形状的稳健性检验

★★★ · A 族·桌面实验 · 物理赛道·颗粒物质 · 五金件 + 粮食颗粒 <100 元 · 实验设计 + 非线性拟合

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

平底容器孔口直径 D 变化时，颗粒质量流量 W 是否服从 Beverloo 律 $W = C \rho \sqrt{g} (D - kd)^{5/2}$ ？指数 5/2 与截止参数 k 在从近球形(小米)到强非球形(大米、红扁豆)的颗粒形状序列上是否稳健？

2 · 研究背景与空白

与液体不同，颗粒从孔口流出的流量几乎不依赖料位高度——孔口上方形成自由落体拱，特征速度 $\sim \sqrt{gD}$ ，由此量纲分析给出 $W \propto D^{5/2}$ 。Beverloo(1961)引入有效孔径修正 $(D - kd)$ (d 为颗粒直径, $k \approx 1.4$ 对球形颗粒，需核实精确值：搜索关键词 "Beverloo correlation k value spheres")，该律是粮仓、制药下料设计的基石。

已有文献对球形颗粒(玻璃珠)验证充分；非球形颗粒的系统数据较少且分散, k 与指数是否漂移在教学与工程文献间说法不一。

空白在于：用同一装置、同一拟合口径横跨形状序列的对照数据很少，而这恰是低成本、高可重复、正负结论均成立的一年期课题：装置是一块钻孔板加电子秤，难点全在实验纪律与统计。

3 · 可检验假设

- H1: 近球形小米的拟合指数与 2.5 的偏差 ≤ 0.1 (95% 自助法置信区间覆盖 2.5)。
- H2: 非球形颗粒(长径比 ≥ 1.5 的大米)的有效 k 显著大于球形颗粒(置信区间不重叠)；若指数本身漂移超过 0.15，则"Beverloo 形式失效"即为主结论。

4 · 量化验收标准

1. 方法学校验(硬门槛)：计量链校准——电子秤用标准砝码(或整包定量食盐)校准误差 $\leq 1\%$ ，计时用手机视频帧号(30 fps, ± 1 帧)校准；随后用小米在单一孔径下重复 10 次，流量变异系数 $\leq 3\%$ ，证明装置可重复。此条不过关，后续全部结论无效。
2. 孔径序列 ≥ 7 个点，覆盖 $D/d = 5 - 30$ (下限避堵塞、上限避容器壁效应，容器直径 $\geq 5D$)。
3. 统计口径预先写死：主拟合为非线性最小二乘同时估计 (C, k , 指数)；另做固定指数 2.5 只估 (C, k) 的对照；两套都报，参数置信区间用 ≥ 1000 次自助法(以"次"为重抽单元)。
4. 每个孔径重复 ≥ 5 次；流量取秤读数-时间曲线线性段斜率，线性段判据($R^2 \geq 0.999$ 的最长窗)写死。
5. 颗粒表征：每种颗粒随机取 ≥ 50 粒，游标卡尺测三轴尺寸，报等效径与长径比分布。
6. 可复现性：数据表(每次放流的原始秤曲线)、拟合脚本开源。

5 · 数据与工具

用途	来源 / 工具
装置	大口塑料瓶/亚克力筒 + 可换钻孔底板(孔径 4–25 mm, 五金店加工)+ 支架,<80 元
计量	厨房电子秤(0.1 g 分度,<40 元)+ 手机视频计时
颗粒	小米、大米、红扁豆、绿豆(超市),<20 元
分析	Python:scipy.optimize.curve_fit + 自写自助法; 纯 CPU, 秒级
理论对照	Beverloo 1961 关联式与球形颗粒 k 值,仅用于校验与对比, 不计入本项目的数据贡献

6 · 方法路径

1. 加工孔板序列, 完成秤/计时校准与小米重复性验证(验收 1)。
2. 表征全部颗粒的尺寸分布与堆积密度。
3. 正式测量:4 颗粒 $\times \geq 7$ 孔径 $\times \geq 5$ 次, 随机化测量顺序以摊薄湿度漂移。
4. 提取流量(判据写死), 拟合两套模型 + 自助法置信区间。
5. 核实边界效应: 最大孔径下换更大容器复测 2 点, 验证壁效应 $< 2\%$ 。
6. 独立交叉校验: 用视频颗粒面计数法对小米复测 2 个孔径的流量, 与秤法偏差 $\leq 5\%$ 。
7. 汇总形状序列上的 (指数, k) 图, 写双语术语表(orifice flow, free-fall arch)。

7 · 新颖性边界

- 本课题不声称发现 Beverloo 律(1961 年已发表), 不声称建立新的颗粒流理论。
- 已有工作: 球形玻璃珠体系的指数与 k 已被多次验证; 非球形颗粒有零散数据但口径不一。
- 本项目贡献(主结论): 同装置同口径下跨形状序列的 (指数, k) 稳健性数据 + 公开原始数据。属"参数维度补全 + 可复现性贡献", 定位须讲清, 否则会被误读为验证性实验。
- 价值:"k 随长径比上升而指数不动"或"Beverloo 形式失效"都是可发表级信息; 负结果同样成立, 但须报告参数置信区间证明分辨力(能分辨 k 的 20% 差异)。
- 核对获奖清单: 颗粒/沉降类相邻获奖有 "Boycott Effect"(2022 优胜, 悬浮液沉降)与 "Falsification of Sand Painting on Induction Cooker"(2023 优胜, 振动颗粒图案)——均不涉及孔口流; 本赛道尚无 Beverloo 类获奖题。

8 · 决策门槛(go / no-go)

⚠ 第 4 周末: 装置重复性达标(变异系数 $\leq 3\%$)且最小孔径堵塞率 $\leq 30\%$ 。若小孔频繁堵塞, 收窄孔径范围至 $D/d \geq 7$ 并增加大孔端点; 若非球形颗粒出现强间歇流(秤曲线台阶化),降级: 把间歇性本身升级为研究对象(流量涨落的功率谱与堵塞概率 vs D/d), 保留"形状如何改变孔口流"的主结论框架。⚠ 第 8 周末: 核实湿度影响——同批颗粒烘干前后流量差若 $> 3\%$, 全部实验加干燥剂储存流程并记录湿度。⚠ 预算裁剪: 颗粒种类可从 4 减到 2(球形 + 一种非球形), 主结论仍成立。

P24 · 磁体在非铁磁金属管中的涡流制动：终端速度对管材电导率-壁厚标度的定量检验

★★★ · A 族·桌面实验 · 物理赛道·电磁学 · 金属管 + 磁体 <250 元 · 电磁建模 + 手机传感器

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

钕磁体在铝、铜、黄铜管中下落的终端速度 v_∞ ，是否与磁偶极-薄壁管模型预言的 $v_\infty \propto 1/(\sigma w)$ (σ 电导率、 w 壁厚) 成反比标度一致？几何因子对磁体长径比的依赖能否被同一模型捕捉？

2 · 研究背景与空白

磁体在导体管中下落时，变化磁通在管壁感应涡流，涡流磁场反作用产生速度成正比的阻尼力，磁体在亚秒内达到终端速度——这是本科电磁学最干净的"可测量的楞次定律"。偶极近似 + 薄壁近似下阻尼系数可解析写出，预言 v_∞ 与 σw 成反比。

该体系在物理教学文献中有解析与实验工作(如 Donoso 等人在欧洲物理教学期刊的磁体-导体管系列，需核实：搜索关键词 "magnet falling conducting pipe terminal velocity Donoso European Journal of Physics")，多为单管演示或单变量验证。

空白在于：跨材料(σ 变化 ~ 2 倍) $\times$ 跨壁厚 $\times$ 跨磁体长径比的同口径标度检验很少见，且手机磁力计使无接触测速零成本化。装置极其可靠、正负结论均成立(偏离标度即说明薄壁/偶极近似失效边界)，适合一年期课题。

3 · 可检验假设

- H1: v_∞ 对 $1/(\sigma w)$ 的线性回归决定系数 $R^2 \geq 0.98$ ，截距与 0 的偏差在 95% 置信区间内不显著。
- H2: 同一磁体的几何因子(回归斜率换算)在不同管材间一致($\pm 10\%$)；长径比 0.5 $\rightarrow$ 2 的磁体序列上，几何因子变化方向与偶极模型数值积分预测一致。

4 · 量化验收标准

1. 方法学校验(硬门槛)：测速链校准——磁体在无管自由下落同一高度，手机磁力计/视频测得的下落时间与运动学预测偏差 $\leq 3\%$ ；管内测速用"磁力计峰间隔法"(管外等距贴双手机或单手机移动重复)与视频法交叉一致 $\leq 5\%$ 。此条不过关，后续全部结论无效。
2. 材料-壁厚矩阵 ≥ 5 个 (σ, w) 组合(纯铝、纯铜、黄铜；壁厚两档)， σ 用文献牌号值并标注牌号不确定度 ($\pm 5\%$)。
3. 每组合重复 ≥ 10 次落体，报均值 $\pm$ 标准差；终端性验证：管中段与下段分段速度差 $\leq 2\%$ (否则加长管)。
4. 统计口径预先写死：回归用加权最小二乘(权重 = $1/\text{方差}$)，斜率与截距置信区间用自助法 ≥ 1000 次；并报稳健回归(Theil-Sen)对照。
5. 温度记录：涡流致管壁温升与环境温度变化 $\leq 3^\circ\text{C}$ (σ 温漂 $< 1.5\%$)，否则组间归一。
6. 可复现性：全部磁力计原始时序 + 脚本开源。

5 · 数据与工具

用途	来源 / 工具
管材	五金/淘宝：铝管、铜管、黄铜管(内径 ~16 mm, 长 ≥ 0.5 m, 两种壁厚), 合计 <180 元
磁体	钕磁体圆柱序列(直径 12 mm, 长 6/12/24 mm), <50 元(注意夹伤与远离银行卡)
测速	手机磁力计(Phyphox 记录, 采样约 100 Hz, 具体上限依机型, 需核实自己手机实测)+ 手机 240 fps 慢动作交叉法
建模	Python: 偶极场 + 薄壁环感应的数值积分(scipy.integrate), 纯 CPU 秒级
理论对照	教学文献解析式, 仅用于校验与对比, 不计入本项目的数据贡献

6 · 方法路径

1. 完成测速链双法交叉校准(验收 1), 确定磁力计采样率是否足以分辨峰(见 go/no-go)。
2. 单管基线: 验证终端性(分段速度恒定)与重复性。
3. 执行 (σ, w) 矩阵 $\times$ 磁体长径比序列的正式测量。
4. 自写偶极-薄壁数值积分模型, 预测几何因子; 与教学文献解析式对照(核实其近似条件)。
5. 回归分析 + 自助法置信区间, 检验 H_1/H_2 。
6. 独立交叉校验: 能量法——用磁力计峰宽反演通过时间并与秒表总下落时间自洽(闭合 $\leq 5\%$)。
7. 撰写失效边界讨论: 厚壁/近壁情形偏离量与近似破坏的对应(标注为模型内推断)。

7 · 新颖性边界

- 本课题不声称发现涡流制动或其解析模型(教学文献已充分发表), 不声称新电磁理论。
- 已有工作: 单管终端速度演示与偶极解析模型; 多为单变量、无系统误差预算。
- 本项目贡献(主结论): 跨 $\sigma \times w \times$ 长径比的同口径标度检验 + 手机磁力计无接触测速方法的误差预算 + 开放数据。属"评价维度转换(演示 $\rightarrow$ 带误差棒的标度检验)"。
- 价值: 标度成立给出教学级基准数据; 偏离则标出偶极/薄壁近似的定量失效边界。两种结果均成立, 须给出斜率置信区间证明能分辨 10% 的几何因子差异。
- 核对获奖清单: 磁类相邻获奖有 "Dynamics and Abnormal Sway Precession of Euler's Magnetic Pendulum"(2023 银奖)与 "Nonlinear Dynamic Modelling and Analysis of Magnetic Suspension System"(2025 优胜)——均为磁体动力学/悬浮的非线性问题, 不涉及涡流制动标度; 本赛道尚无涡流课题。

8 · 决策门槛(go / no-go)

⚠ 第 4 周末: 核实手机磁力计实测采样率与磁体经过特征时间。阈值: 峰宽 ≥ 5 个采样点。若磁体太快(经过 < 50 ms), 降级路径一: 换更强磁体/更薄铝管把 v_{∞} 压到 < 0.3 m/s; 降级路径二: 全部改用 240 fps 视频测速(标尺标定)。两条路径的标度分析框架不变。⚠ 第 6 周末: 管材电导率标称值核实。若牌号不明(σ 不确定 $> 10\%$), 增加"四线法电阻测量"小实验(学校电学器材)自测 σ , 或把回归改为"同材料内壁厚标度 + 跨材料定性排序", 主框架保留。⚠ 预算裁剪: 黄铜管可砍(保铝/铜两点 σ 对比); 长径比序列可从 3 减到 2。

P25 · Leidenfrost 液滴寿命曲线与临界温度：从沸腾曲线测量到蒸汽膜标度模型的检验

★★★ · A 族·桌面实验 · 物理赛道·热学/流体 · 加热板 + 测温 <150 元 · 热学实验 + 图像分析

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

水滴在加热金属板上的总蒸发寿命曲线 $\tau(T)$ 的峰值(Leidenfrost 点)位于何温度、随液滴体积与表面粗糙度如何移动？膜沸腾平台区的寿命对液滴半径的依赖是否符合蒸汽膜润滑模型的幂律预期？

2 · 研究背景与空白

板温超过某临界值后，液滴底部形成连续蒸汽膜，液滴悬浮其上，传热骤降、寿命反而暴增至分钟级——Leidenfrost 效应(1756)。寿命-板温曲线 $\tau(T)$ 是沸腾传热学的标准表征：峰值位置即 Leidenfrost 点 T_L ，其随表面状态与液滴尺寸移动。

Biance, Clanet & Qu  r  (2003, Phys. Fluids)给出了蒸汽膜厚度与蒸发速率的标度理论并实验验证；工程文献中 T_L 对粗糙度的敏感性有大量报告但分散(具体标度：需核实，搜索关键词 "Leidenfrost point roughness dependence lifetime curve")。

空白在于：中学可及器材下的系统 $\tau(T; V, \text{粗糙度})$ 曲面测量几乎没有公开数据集；该测量只需温控、计时与耐心，且 " T_L 对粗糙度敏感/不敏感"两个方向都是有效结论，适合一年期课题。

3 · 可检验假设

- H1: $\tau(T)$ 在 $150 - 350^\circ\text{C}$ 区间呈单峰, T_L 的 95% 置信区间宽度 $\leq 15^\circ\text{C}$ ；粗糙面(砂纸打磨)相对抛光面的 T_L 移动 $\geq 20^\circ\text{C}$ (方向：升高)。
- H2：平台区($T = T_L + 50^\circ\text{C}$)寿命随初始体积的幂律 $\tau \propto V^\alpha$ 的指数 α 落在蒸汽膜模型预期区间(约 0.3 – 0.6，准静态蒸发标度；精确预期须由模型推导给出，推导为本项目内容)。

4 · 量化验收标准

1. 方法学校验(硬门槛)：温度计量校准——测温装置(热电偶或红外枪)用冰水混合物(0°C)与当地气压修正后的沸水两点校准，偏差 $\leq 3^\circ\text{C}$ ；红外枪须用黑胶带发射率标定(与热电偶同点对照 $\leq 5^\circ\text{C}$)。板面温度均匀性:9 点扫描极差 $\leq 15^\circ\text{C}$ 。此条不过关，后续全部结论无效。
2. 温度扫描 ≥ 10 个板温点($120 - 350^\circ\text{C}$)，每点 ≥ 8 滴；液滴体积用定量移液(1 mL 注射器, $\pm 0.02\text{ mL}$)。
3. 统计口径预先写死：寿命分布偏态，报中位数 $\pm$ 四分位距(不报均值 $\pm$ 标准差为主结果); T_L 由 $\tau(T)$ 局部极大 + 自助法(对滴重抽)95% 置信区间；幂律指数 α 用对数空间最小二乘 + 自助法，并报 Theil – Sen 稳健对照。
4. 表面状态量化：抛光/砂纸目数(400/80)三档，并附手机微距照片存档。

5. 安全与操作规程成文(高温、防溅)，指导老师签字确认后方可开始。
6. 可复现性：每滴的计时视频索引表 + 全部数据表 + 脚本开源。

5 · 数据与工具

用途	来源 / 工具
加热	可控温电炉/电磁炉 + 厚铝板或铸铁板(均热), <80 元(或用学校酒精喷灯 + 铁板, 需教师监护)
测温	学校热电偶数字温度计; 或 <60 元 K 型热电偶表; 红外枪仅作辅助(发射率校准后)
计时/成像	手机视频(30/240 fps); 寿命 = 液滴消失帧 - 落板帧, 判据写死
液滴	蒸馏水 + 1 mL 注射器; 表面: 铝板抛光/砂纸序列, <20 元
分析	Python: 数据表 + 拟合, 纯 CPU 秒级; Biance et al. 2003 标度仅用于校验与对比, 不计入本项目的数据贡献

6 · 方法路径

- 搭建加热-均热-测温链并完成两点校准与均匀性扫描(验收 1)。
- 预实验：粗扫 $\tau(T)$ 找到峰的大致位置，确定精扫网格。
- 正式测量抛光面 $\tau(T; V)$ 曲面: 10 板温 $\times$ 3 体积 $\times$ 8 滴。
- 更换砂纸粗糙面，重复温度扫描(单体积)。
- 由蒸汽膜润滑模型推导 $\tau(V)$ 幂律预期(书面推导，附于报告)，与平台区数据对照。
- 独立交叉校验：用 240 fps 视频测液滴悬浮时的振荡频率，与文献气膜液滴本征频率量级对照(标注为辅助证据)。
- 汇总 T_L (粗糙度) 与 α 的置信区间，写双语术语表(film boiling, Leidenfrost point)。

7 · 新颖性边界

- 本课题不声称发现 Leidenfrost 效应或蒸汽膜标度理论(Biance – Clanet – Quéré 2003 已发表)，不声称新传热机理。
- 已有工作：寿命曲线与膜厚标度在实验室级仪器上已充分测量；粗糙度效应有分散报告。
- 本项目贡献(主结论)：中学器材可复现的 $\tau(T; V, \text{粗糙度})$ 公开数据集 + T_L 移动量的置信区间 + α 指数检验。属"参数维度补全 + 可复现性贡献"。
- 价值：为教学与低成本传热实验提供带误差棒的基准;"粗糙度不移动 T_L "也是有效结论，须证明分辨力达 15 ° C。
- 核对获奖清单：液滴类是本赛道最拥挤的获奖族——"Kinetic study of viscous droplet impinging on horizontally moving surface"(2021 铜)、"Hydrodynamic Behavior of Non-viscous Droplet Retraction"(2024 优胜)、"Liquid Droplet Trajectories: Harnessing Sound"(2025 铜)等——但均不涉及沸腾与蒸汽膜；热学相邻题 "Pop pop boat 相变传热"(2023 铜)研究的是引擎脉动。本题以沸腾曲线与 T_L 为对象，与上述均无重叠，报告中逐一点名区分。

8 · 决策门槛(go / no-go)

⚠ 第 5 周末：板温均匀性与稳定性达标(9 点极差 $\leq 15^\circ\text{C}$ 、10 min 漂移 $\leq 5^\circ\text{C}$)。若加热器控温不达标,降级：放弃全曲线，固定两个板温(膜沸腾区一点 + 过渡区一点)，把课题收缩为"平台区 $\tau(V)$ 幂律 + 两温度对照"，蒸汽膜模型检验主框架保留。⚠ 第 8 周末：核实液滴投放高度一致性影响($\pm 2\text{ mm}$ 内寿命差 $< 5\%$)，否则加装定高滴管支架。⚠ 选择前提：涉及 $>150^\circ\text{C}$ 高温操作，须实验室环境 + 教师在场；不接受此条件的学生应改选 T01/T02。⚠ 预算裁剪：粗糙度维度可砍(保体积标度)，主结论(模型检验)仍成立。

P26 · 竖直肥皂膜的排液动力学：干涉条纹追踪测厚与重力排液标度的检验

★★★ · A 族·桌面实验 · 物理赛道·软物质/光学 · LED + 手机 <100 元 · 光学图像处理 + 幂律拟合

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

竖直肥皂膜固定高度处的厚度 $h(t)$ 是否服从重力排液模型预期的幂律衰减？表面活性剂浓度与甘油添加量如何改变衰减指数与黑膜(厚度 < 可见光四分之一波长)出现时间？

2 · 研究背景与空白

竖直皂膜在重力下变薄，膜厚决定反射光的薄膜干涉色，因此彩色横条纹就是"等厚线"，手机相机即是测厚仪。排液动力学取决于气-液界面边界条件：界面近似刚性时为 Poiseuille 型缓慢排液，界面可流动或存在边缘再生(marginal regeneration, Mysels 等 1959 提出)时排液更快，预期的 $h(t)$ 幂律指数不同(各机制的具体指数：需核实，搜索关键词 "vertical soap film drainage law marginal regeneration exponent")。

经典工作(Mysels、后续流体文献)建立了机制分类，但多用干涉仪/单色仪级设备；定量指数在不同配方间的报告分散。

空白在于：用消费级器材对"配方 → 排液指数"做同口径扫描并公开数据。测量链便宜可靠，机制判别(指数落入哪一机制区间)正负结论均成立，适合一年期课题。

3 · 可检验假设

- H1: 固定高度处 $h(t)$ 在膜生成后 5 – 60 s 内服从幂律 $h \propto t^{(-\beta)}$ ，标准配方(洗洁精 2%)的 β 的 95% 自助法置信区间宽度 ≤ 0.15 。
- H2: 加入甘油(20% 体积比)使 β 显著减小(置信区间不重叠)，方向与粘度增大抑制排液一致；黑膜出现时间相应推迟 ≥ 2 倍。

4 · 量化验收标准

1. 方法学校验(硬门槛)：测厚管线校验——用两玻片夹薄纸形成的空气楔重现牛顿等厚条纹，条纹间距与几何预测偏差 $\leq 5\%$ ；并用单色 LED(标称波长 ± 10 nm)确认条纹计数 → 厚度换算的自洽(相邻暗纹厚度差 $= \lambda/2n$)。此条不过关，后续全部结论无效。
2. 每配方 ≥ 15 张膜; $h(t)$ 由固定像素行的条纹计数自动提取，识别判据(灰度极值检测参数)写死；人工复核抽样 10%，与自动计数差 ≤ 1 条纹。
3. 统计口径预先写死：幂律拟合在对数空间最小二乘为主, Theil – Sen 稳健回归交叉验证; β 置信区间用对膜重抽的自助法 ≥ 1000 次；拟合时间窗预先定义(成膜后 5 s 起至条纹速率低于分辨极限)。
4. 环境控制：温度记录 $\pm 2^\circ \text{C}$ ，防风罩(透明整理箱)内成膜；湿度记录(手机可读的 <30 元温湿度计)。
5. 黑膜出现时间的判据写死(反射强度降至本底 + 3σ 以下的首帧)。
6. 可复现性：全部原始视频 + 提取脚本 + 数据表开源。

5 · 数据与工具

用途	来源 / 工具
成膜	3D 打印/铁丝矩形框(宽 3—5 cm)+ 提拉槽；洗洁精、甘油(药店),<40 元
照明	单色 LED(如 590 nm 黄光,<20 元)+ 白光 LED 对照；黑色背景布
成像	手机固定支架,1080p/30 fps 视频
图像处理	Python:OpenCV(灰度剖面提取)+ numpy/scipy；纯 CPU，单视频处理 < 1 min
理论对照	排液机制分类与指数区间(经典皂膜文献，具体值需核实),仅用于校验与对比，不计入本项目的数据贡献

6 · 方法路径

1. 搭建成膜-照明-成像装置，完成空气楔校验(验收 1)。
2. 预实验确定标准配方与可追踪时长(目标 $\geq 60\text{ s}$)，必要时调甘油保膜。
3. 编写条纹自动计数脚本并做人工一致性抽检。
4. 正式采集: ≥ 3 配方 $\times \geq 15$ 膜，同批溶液当日用完并记录配制时间。
5. 拟合 β 与黑膜时间，自助法置信区间；对照机制预期区间判别排液机制。
6. 核实工具边界：手机自动曝光是否破坏灰度剖面(锁定曝光/ISO 后复测 2 膜验证)。
7. 独立交叉校验：白光干涉色序(与标准牛顿色序表比色)对单膜复算一次 $h(t)$ ，与单色计数法偏差 $\leq 15\%$ 。

7 · 新颖性边界

- 本课题不声称发现皂膜排液机制或边缘再生(Mysels 等 1959 已发表)，不声称新流体理论。
- 已有工作：机制分类与实验室级干涉测量已发表；配方-指数的系统消费级数据缺乏。
- 本项目贡献(主结论)：配方序列上排液指数 β 与黑膜时间的同口径测量 + 机制区间判别 + 公开数据。属"参数维度补全 + 评价维度转换"。
- 价值：指数落在哪个机制区间是明确可判的问题;"各配方 β 无显著差"也是有效结论，须证明分辨力达 0.15。
- 核对获奖清单：薄液层相邻获奖为 "Study on the Stability of Vertically Vibrating Liquid Layers"(2025 入围，振动悬浮不稳定性)与大量液滴课题——均不涉及皂膜排液；本赛道尚无皂膜干涉课题。

8 · 决策门槛(go / no-go)

⚠ 第 5 周末：标准配方膜寿命 $\geq 60\text{ s}$ 且条纹自动计数通过抽检。若膜寿命不足，加甘油/换配方；若自动计数始终失败,降级：改为人工逐帧计数(抽帧 1 fps)并给出双人独立计数一致性 ≤ 1 条纹，分析框架不变。⚠ 第 9 周末：若幂律窗口 < 1 个 decade 导致 β 不稳定,降级：主结论改为"黑膜出现时间与初始排液速率 vs 配方"(不依赖幂律形式)，保留机制判别框架。⚠ 预算裁剪：配方维度从 3 减到 2 仍可成文；单色/白光双照明可砍白光。

P27 · 第 25 太阳活动周耀斑等待时间分布的独立检验：非平稳泊松模型与幂律尾指数

★★★ · B 族·公开数据 · 物理赛道·太阳物理 · 纯 CPU + 公开耀斑目录 · 统计推断 + Python

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

第 25 太阳周(2020 – 2026)GOES 软 X 射线耀斑目录的等待时间分布，其幂律尾指数与已发表的早期周期结果(Wheatland 2000 等报告的约 -2.1 至 -2.4 区间，精确值需核实)是否在 95% 置信区间内一致？偏离能否由非平稳泊松模型(活动率随周期演化)完全解释？

2 · 研究背景与空白

太阳耀斑是磁能突然释放事件,GOES 卫星自 1970 年代起连续记录全日面软 X 射线流量，耀斑目录(时间、等级)是太阳物理最长的均匀事件序列之一。相邻耀斑的"等待时间"分布形态是区分触发机制的探针：若耀斑为速率缓变的泊松过程，分布应为活动率分布的混合指数，其长时尾呈幂律——这是 Wheatland(2000, ApJ)非平稳泊松图像的核心预言；自组织临界(SOC)框架(Aschwanden 等)也对该分布有预期。

已有工作覆盖第 21 – 23 周，报告长时尾幂律并支持非平稳泊松解释；第 24 周有零星分析。

空白在于 时间覆盖：第 25 周(2019 年底起,2024 – 2025 达极大)刚刚完整，尚无(或极少，需核实：搜索关键词"solar flare waiting time distribution cycle 25")用同一口径对其完整重做的公开分析。方法成熟、数据公开、指数漂移与否两个结果都有物理含义，是教科书级的"新样本独立检验"型学生课题。

3 · 可检验假设

- H1：第 25 周 M 级及以上耀斑等待时间分布长时尾(> 1 天)幂律指数与本项目在第 23 周复现值的差 ≤ 0.3 (置信区间重叠)。
- H2：用当周耀斑发生率曲线驱动的非平稳泊松蒙特卡洛，能在 KS 检验 $p > 0.05$ 水平上再现观测分布；若被拒绝，超出部分(如短时聚集)即为主要结论(同震余震式聚集的证据，标注为待解释现象)。

4 · 量化验收标准

1. 方法学校验(硬门槛)：用与已发表工作相同的历史时段(如第 22 – 23 周)重跑等待时间分布，幂律尾指数与文献值偏差 $\leq 10\%$ ，分布形态图与文献图目视逐段一致。此条不过关，后续全部结论无效。
2. 统计口径预先写死：幂律尾用最大似然(CSN 方法)估计 + 自助法 95% 置信区间; $x_{\min}$ 由 KS 距离定；不用双对数直线拟合作主结果。事件阈值敏感性:C/M/X 三档阈值分别报告全套结果。
3. 数据完整性预算：仪器换代(GOES-8→16 系列)与目录缺口时段列表化，剔除缺口 > 1 天的区间并报告剔除比例($< 5\%$)。
4. H2 蒙特卡洛 $\geq 10^4$ 次合成序列,KS 统计量的零分布由合成序列自举。
5. 样本量下限：第 25 周 M 级以上事件数 ≥ 500 (预计满足，准确数第 2 周核实)。

6. 可复现性：下载脚本 + 清洗规则 + 分析脚本全部开源，从原始目录到全部图表一键重跑。

5 · 数据与工具

用途	来源 / 工具
耀斑目录	NOAA SWPC / NCEI 的 GOES X 射线耀斑事件表(公开,swpc.noaa.gov 与 ngdc.noaa.gov 事件列表；具体文件路径以当前官网为准，需核实)；备选：通过 SunPy 的 Fido/HEK 接口检索耀斑事件(sunpy.org，开源)
覆盖范围	1975 至今全日面记录；本项目用 1996—2026(第 23—25 周)；表格级数据 < 100 MB，纯 CPU
分析	Python:pandas + numpy + powerlaw 包 + 自写蒙特卡洛；全套分析笔记本级 CPU < 1 小时
理论对照	Wheatland 2000 非平稳泊松框架与早期周期指数,仅用于校验与对比，不计入本项目的数据贡献

6 · 方法路径

1. 装 SunPy/pandas 环境，下载并解析全部目录；列出仪器换代与缺口清单。
2. 在第 22 – 23 周历史段复现已发表分布与指数(验收 1)。
3. 固化清洗与阈值规则，生成第 25 周等待时间序列。
4. CSN 拟合 + 自助法，三阈值敏感性矩阵。
5. 构建非平稳泊松蒙特卡洛(速率 = 平滑月发生率)，做 KS 检验(H2)。
6. 核实边界:GOES 换代前后目录定标是否一致(读 NCEI 说明文档；不一致则分段分析)。
7. 独立交叉校验：用秩序统计(去趋势后的归一化等待时间)复核幂律尾稳健性。

7 · 新颖性边界

- 本课题不声称提出等待时间统计新方法，不声称判定耀斑触发机制；非平稳泊松框架与早期指数为 Wheatland(2000)等已发表结论。
- 已有工作：第 21 – 23 周分布与指数、SOC 解释(Aschwanden 等综述)。
- 本项目贡献(主结论)：第 25 周完整数据的同口径独立检验——指数漂移则提示分布依赖周期强度(第 25 周弱于 23 周)，不漂移则普适性获得一次独立支持。两种结果都构成有效结论。属"时间覆盖空白 + 新样本独立检验"。
- 核对获奖清单：本赛道尚无太阳统计课题；最接近的公开数据获奖题为 "Developing an Automated Pipeline for Identifying False Positives Among Released TESS Objects of Interest"(2021 铜)与 "Investigating Physical Conditions across the Center of M82"(2024 铜)，均非太阳物理，说明评委接受档案数据管线类工作。

8 · 决策门槛(go / no-go)

⚠ 第 2 周末：核实数据入口与事件数。阈值：能取到 1996 – 2026 连续目录、第 25 周 M+ 事件 ≥ 500 。若 NCEI 文件接口变动，改走 SunPy/HEK；若 M+ 不足 500,降级：阈值降至 C 级(事件数增一个数量级)，框架不变。⚠ 第 6 周末：完成验收 1 历史复现，指数偏差 $\leq 10\%$ ；不达标则冻结新分析，排查清洗规则差异(常见原因：重复事件、阈值口径)，直到达标。⚠ 第 20 周检查点：若第 25 周分布与非平稳泊松完全一致且与历史无

差(最平淡结果)，按预案把"三阈值 + 跨周期指数表 + 完全可复现管线"作为主交付，结论定位为普适性支持——仍是完整论文。

P28 · 基于 SDO/HMI SHARP 磁场参数的耀斑预报模型在第 25 周新样本上的独立检验

★★★ · B 族·公开数据 · 物理赛道·太阳物理/机器学习 · 纯 CPU + JSOC 参数表 · 不平衡分类 + 特征分析

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

用 Bobra & Couvidat(2015)的 SHARP 磁场特征集与同型线性分类器，在其训练年代(第 24 周)之外的第 25 周(2020 - 2026)数据上重新训练与测试，预报 M 级以上耀斑的技能(TSS、PR-AUC)相比原文是否显著下降？下降幅度能否由类别不平衡率与活动区群体分布的变化解释？

2 · 研究背景与空白

SDO/HMI 自 2010 年起对每个活动区输出 SHARP(Space-weather HMI Active Region Patch)参数——总无符号磁通、电流螺度代理等约 20 个标量，逐 12 分钟更新，是耀斑物理条件的标准指纹。耀斑预报因此可表述为极不平衡二分类：未来 24 h 内产生 M+ 耀斑的活动区时刻仅占百分之几。

Bobra & Couvidat(2015, ApJ)用 SHARP 特征 + SVM 在第 24 周数据上得到 $TSS \approx 0.76$ (准确数值需核实)，成为该领域引用最多的基线；后续深度学习工作层出，但多数仍在相近年代的数据上交叉验证，存在"周期内乐观偏差"的公开讨论。

空白在于 时间覆盖与评价维度：第 25 周构成一个真正独立的群体(极大期强度、活动区分布均不同)，对基线模型做严格的跨周期外推检验、并把性能变化分解为"不平衡率变化 vs 特征分布漂移"的公开分析很少(需核实：搜索关键词 "SHARP flare prediction solar cycle 25 cross-cycle generalization")。表格级数据 + 线性模型，纯 CPU 完全可行。

3 · 可检验假设

- H1：在第 25 周按时间划分的测试集上，同型模型 TSS 相比原文报告值下降 ≥ 0.05 (即跨周期外推存在可测退化)；PR-AUC 相对第 24 周复现值下降 $\geq 10\%$ 。
- H2：对特征逐一做第 24/25 周分布比较(KS 检验)，漂移最大的前 3 个特征被剔除后，跨周期性能差至少缩小一半；若 H1 不成立(无退化)，"基线跨周期稳健"即为主结论。

4 · 量化验收标准

1. 方法学校验(硬门槛)：在原文时间范围与口径(样本定义、正负例窗口、特征集)内重建数据集并复现其分类性能，TSS 与原文差 ≤ 0.05 。此条不过关，后续全部结论无效。
2. 统计口径预先写死：极不平衡分类，主指标为 PR-AUC、召回、精确率与 TSS；明确不报 accuracy(不平衡率 $\sim 2\%$ 时全负预测即可得 98% accuracy，毫无意义)。训练/测试按时间划分(逐活动区整段划分防泄漏)；另做一次随机划分对照，量化随机划分带来的乐观高估幅度并写入正文。
3. 性能置信区间：对测试集按活动区整块自助重抽 ≥ 1000 次。

4. 消融矩阵：全特征 / 剔除漂移特征 / 单特征基线(总无符号磁通)三档全报。
5. 样本规模：第 25 周正例(M+ 前 24 h 时刻的活动区样本) ≥ 300 (第 4 周核实实际数)。
6. 可复现性：特征提取查询语句、样本定义文档、训练脚本全部开源。

5 · 数据与工具

用途	来源 / 工具
SHARP 参数	斯坦福 JSOC(jsoc.stanford.edu), 通过 Python <code>drms</code> 包只拉取关键字表头(不下载磁图), 第 25 周全表约 GB 级以下, 纯 CPU; <code>drms</code> /SunPy 均开源
耀斑标签	同 T07 的 GOES 耀斑目录(NOAA, 公开), 按 NOAA 活动区号关联
建模	Python:scikit-learn(SVM/逻辑回归), 表格级训练秒-分钟级 CPU
基线对照	Bobra & Couvidat 2015 的特征定义与性能,仅用于校验与对比, 不计入本项目的数据贡献

6 · 方法路径

1. 装 `drms/scikit-learn` 环境, 小批量拉取 SHARP 关键字验证查询通路。
2. 按原文口径重建第 24 周数据集, 复现基线性能(验收 1)。
3. 用完全相同管线生成第 25 周数据集, 冻结样本定义文档。
4. 跨周期实验矩阵:24 训 25 测 / 25 训 25 测(时间划分)/ 随机划分对照。
5. 特征分布漂移分析(逐特征 KS + 效应量), 做消融(H2)。
6. 自助法置信区间与结果表; 负例欠采样敏感性检查。
7. 独立交叉校验: 用逻辑回归替换 SVM 复跑主矩阵, 确认结论不依赖分类器选择。

7 · 新颖性边界

• 本课题不提出新预报方法、不声称超越深度学习现状;SHARP 特征集与 SVM 基线是 Bobra & Couvidat(2015)的已发表工作。 • 已有工作: 原文在第 24 周数据上定义特征与基线; 后续文献多在相近年代交叉验证。 • 本项目贡献(主结论): 第 25 周真正跨周期的独立检验 + 性能变化的"不平衡率 vs 特征漂移"分解 + 随机划分乐观偏差的量化。属"新样本独立检验 + 评价维度转换"。 • 价值: 预报模型的跨周期泛化是空间天气业务的真实痛点;"不退化"与"退化并可归因"都是有效结论, 由自助法置信区间保证可分辨 0.05 的 TSS 差。 • 核对获奖清单:"Development of a High-Efficiency Objective-Prism Stellar Spectrograph And Construction of its Dedicated AI Classification Model"(2025 银奖)表明"物理 + ML"路线被评委接受; 该题是自制仪器 + 分类, 本题是公开档案 + 跨周期泛化检验, 对象与贡献类型均不同。2021 铜奖 TESS 假阳性管线同为"档案 + 分类", 但对象是凌星候选体, 与太阳磁参数无重叠; 报告中点名两者并写清差异。
--

8 · 决策门槛(go / no-go)

⚠ 第 4 周末: 核实 JSOC/drms 通路与第 25 周正例数(阈值 ≥ 300)。若 JSOC 批量查询受限,降级: 改用社区整理的公开 SHARP 机器学习基准数据集(存在多个, 具体可用性需核实: 搜索关键词 "SHARP flare prediction"

benchmark dataset Liu 2019")作为特征源，跨周期检验框架不变。⚠️ 第 8 周末：验收 1 复现达标(TSS 差 ≤ 0.05)；不达标则冻结新实验，按"样本窗口定义 $\rightarrow$ 特征版本 $\rightarrow$ 类权重"顺序排查。若 12 周仍不达标,降级：放弃复现原文绝对值，改为"同一自建口径下第 24 vs 25 周对照"(内部一致比较)，独立检验框架保留。⚠️ 选择前提：学生需已具备 scikit-learn 基础；英文答辩须能解释 TSS/PR-AUC 为何优于 accuracy——这本身是加分点。

P29 · 用 Gaia DR3 自行与视差测量疏散星团的质量分层程度及其随动力学年龄的变化

★★★ · B 族·公开数据 · 物理赛道·天体物理 · 纯 CPU + Gaia 档案 · 天文数据 + 聚类/秩统计

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

对 ≥ 10 个邻近疏散星团，用最小生成树质量分层比 Λ_{MSR} 在同一管线下量化的质量分层程度，是否随动力学年龄(星团年龄/弛豫时间)单调增强？

2 · 研究背景与空白

疏散星团中重星通过两体弛豫把动能传给轻星、自身下沉——质量分层(mass segregation)是恒星动力学最干净的可观测预言，其程度应由星团经历的弛豫次数(动力学年龄)决定。Gaia DR3 提供毫角秒级视差与自行，使邻近星团的成员判定与质量排序变成表格运算。

Allison et al.(2009, MNRAS)提出最小生成树方法：比较最重 N 颗星的 MST 边长与随机同数星的 MST 边长之比 Λ_{MSR} ，无需假设星团中心与对称性，已成标准工具。单星团的分层测量已大量发表；Cantat-Gaudin et al. (2018/2020)基于 Gaia 的星团成员表覆盖上千星团。

空白在于 评价维度的统一：文献中各星团的 Λ_{MSR} 出自不同数据版本、不同成员判定与不同 N 选择，横向不可比；用 DR3 + 单一冻结管线跨样本重测并对动力学年龄作秩相关检验的公开工作少(需核实：搜索关键词 "mass segregation open clusters Gaia DR3 Lambda MSR sample")。数据均为表格级，纯 CPU。

3 · 可检验假设

- H1: Λ_{MSR} (最重 20 颗)与动力学年龄的 Spearman 秩相关 $\rho \geq 0.5$ ，置换检验 $p < 0.05$ 。
- H2: 动力学年龄 < 1 (未经历一个弛豫时间)的星团中，与 $\Lambda_{\text{MSR}} = 1$ (无分层)不相容的比例 $\geq 30\%$ ——即存在"原初分层"候选(诞生即分层，非动力学产物)；若为 0%，"分层皆动力学起源"即为结论。

4 · 量化验收标准

1. 方法学校验(硬门槛)：对昴星团(Pleiades)重跑成员判定，成员数与已发表表(Cantat-Gaudin et al. 2018)交叉重合率 $\geq 80\%$ ，色-星等图主序形态一致；并用合成星团(已知注入分层程度的模拟数据)验证 Λ_{MSR} 管线能以 95% 置信区间回收注入值。此条不过关，后续全部结论无效。
2. 统计口径预先写死： Λ_{MSR} 不确定度由随机参照集 ≥ 500 次重抽给出；相关性用 Spearman + 置换检验 ($\geq 10^4$ 次)； N 敏感性($N = 10/20/50$)三档全报。
3. 成员判定判据冻结：视差-自行空间的截断参数(或 HDBSCAN 超参数)在校验星团上定死后不得逐团调整；确需调整的星团单独列表并做敏感性分析。
4. 样本 ≥ 10 团，每团成员 ≥ 100 颗；年龄与弛豫时间的来源目录与公式写死(年龄取已发表星团目录值，标注不计入本项目贡献)。

5. 距离系统差检查：视差反演用同一先验，报告 zero-point 修正($-17\ \mu\text{as}$ 量级，需核实 DR3 推荐值)开/关两套结果。
6. 可复现性:ADQL 查询语句 + 全管线脚本开源。

5 · 数据与工具

用途	来源 / 工具
天测/测光	ESA Gaia 档案(gea.esac.esa.int/archive),ADQL 锥形检索,Python <code>astroquery.gaia</code> ; 每团 $< 10^5$ 行, 总量 $< 1\ \text{GB}$, 纯 CPU
成员对照表	Cantat-Gaudin et al. 星团成员目录(CDS/VizieR, vizier.cds.unistra.fr , 公开),仅用于校验与对比, 不计入本项目的数据贡献
星团年龄	已发表星团参数目录(VizieR 同上), 同样仅作输入并声明
分析	Python:astropy + scipy(MST 用 <code>scipy.sparse.csgraph</code>)+ networkx 可选; 单团 Λ_{MSR} 计算秒级

6 · 方法路径

1. 装 astroquery/astropy, 跑通单团 ADQL 检索与色-星等图。
2. 完成昴星团成员判定复现 + 合成星团 Λ_{MSR} 回收测试(验收 1)。
3. 冻结成员判定与 Λ_{MSR} 参数, 写入管线文档。
4. 批量处理 ≥ 10 团(候选池 20 团, 按成员数与拥挤度筛选)。
5. 质量代理排序(G 星等 + 主序假设, 判据写死), 计算 Λ_{MSR} (N 三档)。
6. 秩相关 + 置换检验,zero-point 开关敏感性。
7. 独立交叉校验: 对 2 个团改用已发表成员表重算 Λ_{MSR} , 与自建成员结果偏差 $\leq 1\sigma$ 。

7 · 新颖性边界

- 本课题不声称提出 Λ_{MSR} 方法(Allison et al. 2009 已发表), 不声称首次发现质量分层; 单团分层现象为已发表事实。
- 已有工作: 方法论文 + 若干单团测量 + Gaia 成员目录(他人贡献, 已声明)。
- 本项目贡献(主结论):DR3 下单一冻结管线的跨样本 Λ_{MSR} 一致重测与"分层-动力学年龄"关系的统一检验。属"评价维度统一 + 新数据版本独立检验"。
- 价值: 相关显著支持动力学起源图像; 不显著或出现年轻分层团则指向原初分层, 两个方向都有文献辩论价值。误差棒由参照集重抽保证。
- 核对获奖清单: 天体物理相邻获奖为 "Investigating Physical Conditions across the Center of the Galaxy M82"(2024 铜, 星系中心多波段条件)与 "Discovery of a galactic fountain"(2020 优胜)——均为星系尺度气体, 本题是星团恒星动力学, 数据与物理量无重叠。

8 · 决策门槛(go / no-go)

⚠ 第 3 周末:Gaia 档案检索与单团管线跑通。 ⚠ 第 6 周末: 验收 1 双校验达标。若成员重合率 $< 80\%$, 排查视差质量切割; 若合成回收失败, 修 MST 实现后再前进——此门槛前禁止批量分析。 ⚠ 第 12 周末: 批量样本核查。若可用团(成员 ≥ 100)不足 10 个,降级: 固定 5 - 6 个经典星团(昴、鬼宿、毕宿等)做深入分析 + N 敏感

性 + 双成员表交叉，主结论改为"分层程度的稳健测量与年龄趋势的初步检验"——框架保留，牺牲样本量换稳健性。⚠ 已知困难定位：场星污染是本题核心系统误差，把"污染率对 Λ_{MSR} 的偏置"做成注入-回收实验，写成研究内容而非障碍。

P30 · 系外行星半径谷位置随轨道周期的标度：在 2025—2026 扩充样本上的独立重测

★★★ · B 族·公开数据 · 物理赛道·行星科学 · 纯 CPU + 系外行星档案 · 样本统计 + 误差传播

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

在 NASA 系外行星档案的最新确认行星样本(较原发现样本扩充数倍)中, $1.5 - 2 R_{\oplus}$ 半径谷中心位置随轨道周期的幂律指数 $d \log R_{\text{valley}} / d \log P$ ，与已发表值(Van Eylen et al. 2018 报告约 -0.10 ，精确值与误差需核实)是否在置信区间内一致？

2 · 研究背景与空白

小型系外行星的半径分布在 $\sim 1.8 R_{\oplus}$ 处有一条显著的低谷("半径谷")，把超级地球与亚海王星分成双峰——Fulton et al.(2017, AJ)在 California-Kepler Survey(CKS)约 2000 颗行星中首次分辨。谷的位置随轨道周期负向移动，其斜率是区分成因机制(光致蒸发 vs 核动力质量损失 vs 原初无气包层)的关键判别量，不同机制预言的指数符号与数值不同。

Van Eylen et al.(2018)用星震学精化的小样本测得负斜率；此后确认行星数与恒星参数精度(Gaia 视差)都大幅提升。

空白在于 样本覆盖：用 2025 – 2026 时点的全档案(含大量 TESS 确认行星)按同一口径重测谷位置与斜率的公开工作有限(需核实：搜索关键词 "radius valley slope period reanalysis 2024 2025")。数据是一张 CSV，纯 CPU；斜率一致与否两个结果都直接对话成因机制，是理想的"新样本独立检验"。

3 · 可检验假设

- H1: 扩充样本(半径精度优于 10% 的子样本)中谷位置斜率的 95% 置信区间与 Van Eylen et al. 2018 值重叠；区间宽度较其减半(样本增大红利)。
- H2: 斜率符号在"仅 Kepler 子样本"与"含 TESS 子样本"之间一致；若不一致，差异可由两巡天目标星群体(光谱型分布)差异定性解释(标注为待检验推测)。

4 · 量化验收标准

1. 方法学校验(硬门槛)：用原始 CKS 样本(公开表)重现 Fulton et al. 2017 的双峰半径分布与谷位置，谷中心偏差 $\leq 5\%$ 。此条不过关，后续全部结论无效。
2. 统计口径预先写死：谷位置判据冻结为"对 $\log R - \log P$ 平面做核密度估计后取谷脊线"与"分周期箱内双高斯拟合取谷心"两法，主结果用前者、后者作交叉；测量误差用对每颗行星按报告误差重抽的蒙特卡洛(≥ 1000 次全样本重抽)传播；斜率置信区间用自助法。
3. 样本筛选规则(半径相对误差 $< 10\%$ 、周期 < 100 天、剔除争议标记行星)在分析前冻结成文档。

4. 明确声明不做巡天完备性修正，结论限定为"观测样本内的谷位置"，并单列一节讨论选择效应可能的偏置方向(标注为定性推断)。
5. 两法谷位置差 > 各自误差棒时必须报告并分析，不得只报好看的一法。
6. 可复现性：档案查询 URL/TAP 语句 + 筛选文档 + 脚本开源。

5 · 数据与工具

用途	来源 / 工具
行星参数总表	NASA Exoplanet Archive(exoplanetarchive.ipac.caltech.edu), Planetary Systems 复合表，网页 CSV 导出或 TAP/ <code>astroquery</code> ；单表 < 100 MB，纯 CPU
校验样本	CKS 公开数据表(Fulton et al. 2017 随文发布,CDS/VizieR 可取), 仅用于校验与对比，不计入本项目的数据贡献
分析	Python:pandas + scipy(KDE、拟合)+ 自写自助法；全分析 CPU < 10 min
文献对照	Fulton 2017 / Van Eylen 2018 谷参数, 仅用于校验与对比，不计入本项目的数据贡献

6 · 方法路径

1. 下载复合表与 CKS 表，建立字段映射与清洗笔记。
2. 复现 CKS 双峰与谷位置(验收 1)。
3. 冻结筛选规则与两套谷位置判据，写入预注册式文档(报告附录)。
4. 生成扩充样本，重测谷位置-周期关系，蒙特卡洛误差传播。
5. 子样本分解(Kepler/TESS、恒星光谱型)检验 H2。
6. 核实边界：复合表同一行星多解的取值规则(默认参数集标志)是否引入偏置——换"最新发表参数"规则复跑一遍对照。
7. 独立交叉校验：把半径换成用 Gaia DR3 恒星半径重新标定的值(档案已含)，对照斜率稳健性。

7 · 新颖性边界

- 本课题不声称发现半径谷(Fulton et al. 2017)或其周期依赖(Van Eylen et al. 2018)，不声称判定成因机制。
- 已有工作：原发现 + 星震学小样本斜率 + 各机制理论预言(均已发表)。
- 本项目贡献(主结论):2025 – 2026 时点扩充样本上同口径的谷位置与斜率独立重测 + 双判据交叉 + 误差蒙特卡洛。斜率漂移则提示样本群体依赖，不漂移则机制判别量获得更强约束——两种结果都构成有效结论。属"时间/样本覆盖空白"。
- 核对获奖清单：凌星相邻获奖为 TESS 假阳性管线(2021 铜)——处理光变曲线证认；本题不碰光变曲线，只做确认样本的统计几何，两者在数据层级与问题上均不同，报告中点名区分。

8 · 决策门槛(go / no-go)

⚠ 第 3 周末：档案表下载与字段核实完成(半径/周期/误差列齐全)。⚠ 第 5 周末：验收 1 复现达标；不达标先查 CKS 样本切割(星等、撞击参数过滤)。⚠ 第 10 周末：扩充样本量核查。阈值：精度 < 10% 子样本 ≥ 1500 颗行星；不足则放宽至 15% 并把精度阈值作为敏感性维度报告,降级不改变主框架。⚠ 最平淡结果预案：

斜率与文献完全一致——交付定位为"更强约束 + 完全可复现管线", 在讨论中给出对三类成因模型的排除力度对比(引用其预言区间, 标注需核实各模型预言精确值: 搜索关键词 "photoevaporation core-powered mass loss radius valley slope prediction")。

P31 · 用 LIGO 开放数据重建双黑洞并合信噪比，并检验 O3 事件啁啾质量分布的幂律描述

★★ · B 族·公开数据 · 物理赛道·引力波天文 · 纯 CPU + GWOSC · 信号处理 + 群体统计

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

用 GWOSC 公开应变数据与开源匹配滤波管线重现 GW150914 的信噪比(文献值 $\text{SNR} \approx 24$)之后,GWTC-3 目录中双黑洞事件的啁啾质量分布用单幂律描述的最优指数是多少？不做选择效应修正的"朴素"指数与已发表群体分析结果的偏差有多大、方向是否与探测偏好(重系统更易被探测)一致？

2 · 研究背景与空白

引力波应变数据中，双黑洞并合信号淹没在噪声下，靠与理论波形模板做匹配滤波提取；信噪比(SNR)是探测显著性的核心量。LIGO-Virgo 合作组通过 GWOSC(Gravitational Wave Open Science Center)公开了全部事件段应变数据、事件目录(GWTC 系列)与参数后验样本，并配有官方教程——这是所有公开物理数据中文档最完善的档案之一。

GWTC-3 群体分析(Abbott et al.)用含选择效应修正的层级贝叶斯推断给出质量分布(幂律 + 峰结构);"探测到的样本"与"天体物理真实分布"之间隔着强烈的探测偏好，这一点在文献中反复强调但很少被定量地"演示"出来。

空白在于 评价维度：把"朴素分布(不修正)与已发表修正后分布差多少、差的方向和大小能否用简单 SNR 阈值模型半定量再现"做成一个可复现的教学级定量对照，公开工作很少。全部计算(单事件匹配滤波 + 目录级统计)在笔记本 CPU 上分钟-小时级。

3 · 可检验假设

- H1: 自建管线对 GW150914 的匹配滤波 SNR 与文献值偏差 $\leq 10\%$ ；对另 2 个 O1/O2 高显著事件亦达标。
- H2: 观测样本啁啾质量的朴素幂律指数与 GWTC-3 已发表群体幂律指数之差为正方向(朴素分布偏重)，且用" $\text{SNR} \propto M_{\text{chirp}}^{5/6}/\text{距离}$ 的阈值探测玩具模型"加权后，差值缩小 $\geq 50\%$ 。

4 · 量化验收标准

1. 方法学校验(硬门槛)：跟随 GWOSC 官方教程完成 GW150914 的数据读取、白化、带通与匹配滤波,SNR 复现偏差 $\leq 10\%$ ；白化后时域波形图与教程结果目视一致。此条不过关，后续全部结论无效。
2. 统计口径预先写死：啁啾质量取 GWTC-3 公开后验中位数，测量不确定度用对每事件后验重抽的蒙特卡洛；幂律指数用 CSN 最大似然 + 对事件自助法 ≥ 1000 次给 95% 置信区间；样本定义(FAR $< 1/\text{yr}$ 的 BBH)冻结成文档。
3. 玩具选择函数的假设(SNR 阈值 8、探测体积 $\propto$ 距离立方内均匀)全部显式列出，并做阈值 ± 2 的敏感性。
4. 明确声明：本项目不做完整的层级贝叶斯选择效应修正，所有"天体物理分布"结论仅为半定量对照(这是本题的能力边界，写进报告)。

5. 事件数下限：样本 ≥ 60 个 BBH(GWTC-3 满足)。
6. 可复现性：环境配置(conda 环境文件)+ 笔记本脚本开源，官方教程数据段一键重跑。

5 · 数据与工具

用途	来源 / 工具
应变数据/教程	GWOSC(gwosc.org), 事件段 hdf5 公开下载(单事件 ~百 MB), 官方 Jupyter 教程公开
事件目录与后验	GWTC-3 目录与参数后验样本(GWOSC/Zenodo 公开发布)
信号处理	Python:GWpy、PyCBC(均开源,pip 可装;PyCBC 匹配滤波单事件 CPU 分钟级)
统计	numpy/scipy + powerlaw 包 + 自写蒙特卡洛,CPU 分钟级
文献对照	GW150914 SNR、GWTC-3 群体幂律参数(Abbott et al. 系列, 公开),仅用于校验与对比, 不计入本项目的数据贡献

6 · 方法路径

1. 装 GWpy/PyCBC, 跑通 GWOSC 官方 GW150914 教程(验收 1 第一半)。
2. 独立重写匹配滤波脚本(不复制教程代码), 对 3 个事件复现 SNR(验收 1)。
3. 下载 GWTC-3 目录与后验, 冻结样本定义。
4. 朴素啁啾质量分布拟合 + 蒙特卡洛误差 + 自助法。
5. 构建玩具选择函数并加权重拟合, 与已发表群体指数对照(H2)。
6. 敏感性矩阵:SNR 阈值、样本 FAR 切割、幂律下限 $x_{\min}$ 。
7. 独立交叉校验: 用秩和检验比较 O1/O2 与 O3 子样本分布, 确认合并样本的合法性。

7 · 新颖性边界

- 本课题不声称探测新事件、不声称测得天体物理质量分布(GWTC-3 群体分析为已发表结论, 选择效应修正非本项目能力范围, 已显式声明)。
- 已有工作: 事件目录、后验、含修正的群体推断全部由 LVK 合作组发表。
- 本项目贡献(主结论): 可复现的匹配滤波重建 + "选择效应到底把分布扭曲了多少"的半定量对照(评价维度转换 + 方法可复现性贡献)。这类定位在计算/数据领域是被承认的, 但必须讲清楚, 否则会被误读为重复群体分析。
- 价值: 把文献里一句"存在探测偏好"变成有数字、有敏感性分析的对照; 玩具模型解释不了的残差本身是讨论点。
- 核对获奖清单: 本赛道 2020 – 2025 无引力波课题; 最接近的是宇宙学理论题 "Dark Sector Interaction Model / Hubble Tension"(2020 铜), 数据驱动的引力波分析属未被占据的题材空间。

8 · 决策门槛(go / no-go)

⚠ 选择前提: 概念门槛高(匹配滤波、后验、选择效应), 仅适合已自学过信号处理基础、且愿意先花 3 周啃 GWOSC 教程的学生; 英文答辩需能徒手解释匹配滤波, 讲不清则反噬。⚠ 第 8 周末: 验收 1 三事件 SNR 复

现达标。不达标,降级: 砍掉自建匹配滤波, 课题收缩为"GWTC-3 目录的啁啾质量分布 + 玩具选择函数对照"(纯目录统计, 主结论框架保留, 失去管线复现亮点但仍完整)。⚠ 第 16 周末: 玩具模型若与已发表指数差距无法缩小(H2 失败), 把"简单阈值模型不足以解释探测偏好"作为主结论, 加上残差的定量刻画——负结果成立。⚠ 预算裁剪:PyCBC 若安装受阻(平台兼容, 需核实 Apple Silicon 支持现状), 全程用 GWpy + 自写频域滤波, 教程等价路径。

P32 · 低轨卫星轨道衰减速率与太阳活动指数 F10.7 的相关性：基于公开 TLE 的第 25 周检验

★★ · B 族·公开数据 · 物理赛道·空间物理/轨道力学 · 纯 CPU + CelesTrak · 轨道力学 + 时间序列

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

由公开 TLE(两行轨道根数)序列导出的低轨卫星半长轴衰减速率，与太阳 10.7 cm 射电流量指数 F10.7 的相关性有多强、滞后多少天？第 25 周极大期(2024 前后)衰减率相对静期的增幅，能否用标准热层密度模型半定量解释？

2 · 研究背景与空白

400 – 600 km 高度的卫星受稀薄热层大气拖曳持续降轨；热层密度对太阳 EUV 辐照极端敏感，而 F10.7 是 EUV 的经典代理指数。太阳活动上升会使同一卫星的衰减速率成倍变化——2022 年一批刚发射的 Starlink 卫星因地磁暴密度激增而再入，是该效应的著名实例。TLE 是北美防空司令部对全部编目天体持续发布的轨道根数，配合 SGP4 模型可还原平均半长轴的时间序列：公开、免费、更新至今。

热层密度-太阳活动关系有成熟的经验模型(NRLMSIS 系列)与大量文献；用 TLE 反演大气密度也是航天动力学的成熟技术。

空白在于 时间覆盖：第 25 周极大期(2024 – 2025)是 20 年来最强的一段热层加热，用公开 TLE 对"多颗同类卫星 × 完整极大期"做同口径相关与滞后分析、并与密度模型对照的公开学生级工作很少(需核实：搜索关键词 "TLE orbital decay F10.7 solar cycle 25 thermosphere")。全部为文本数据，纯 CPU。

3 · 可检验假设

- H1：去趋势后的衰减速率与 F10.7(81 天平滑)的互相关峰值 ≥ 0.6 ，滞后在 0 – 5 天量级(热层响应时标)。
- H2：极大期(2024)平均衰减速率相对静期(2020 – 2021)的比值，与 NRLMSIS 模型在相应 F10.7 下的密度比一致到 2 倍以内；若系统偏离，偏离方向与地磁活动(Ap 指数)残余相关(标注为待检验推测)。

4 · 量化验收标准

1. 方法学校验(硬门槛):SGP4 传播链校验——用 ISS 公开 TLE 导出的高度序列与 NASA 公开的 ISS 轨道再推升(reboost)事件时间线对齐，全部推升事件(高度突增)识别率 100%，推升间隙的自然衰减段高度与公开报道值偏差 ≤ 2 km。此条不过关，后续全部结论无效。
2. 统计口径预先写死：时间序列一律按时间划分做任何拟合的训练/验证；相关性在双序列去趋势(21 天高通)后计算，显著性用块自助法(block bootstrap，块长 27 天 $\approx$ 太阳自转周期) ≥ 1000 次；滞后由互相关峰 + 自助法置信区间。
3. 卫星样本 ≥ 10 颗：选无主动轨道维持的低轨目标(如已停止维持的退役卫星/火箭体,450 – 600 km)，筛选规则冻结；含机动的目标用突变检测剔除机动段(判据写死)。
4. 半长轴提取规则冻结:TLE 平均运动 $\rightarrow$ 平均半长轴，历元不均匀采样用逐历元差分 + 30 天滑动中位数。

5. H2 模型对照:NRLMSIS(Python `pymgis` 包, 开源)在卫星高度、对应 F10.7/Ap 下的密度比; 模型参数全部列表。
6. 可复现性:TLE 快照存档(下载即存, 避免源变动)+ 全脚本开源。

5 · 数据与工具

用途	来源 / 工具
TLE 当前与近期	CelesTrak(celestrak.org, 公开无需注册,GP 数据接口)
TLE 历史序列	Space-Track(space-track.org, 免费注册; 批量历史下载受 API 条款约束, 需核实当前限速与允许用量; 不可用时见 go/no-go 降级)
轨道传播	Python <code>sgp4</code> 包(官方算法移植, 开源); 全部计算 CPU 秒-分钟级
太阳/地磁指数	F10.7: 加拿大自然资源部/NOAA SWPC 公开日值;Ap: 德国 GFZ 公开(kp.gfz-potsdam.de)
大气模型	<code>pymgis</code> (NRLMSIS 官方 Python 封装, 开源),仅用于校验与对比, 不计入本项目的数据贡献

6 · 方法路径

- 装 `sgp4/pymgis`, 取 ISS 历史 TLE 完成推升事件校验(验收 1)。
- 冻结目标筛选规则, 建立 ≥ 10 颗无维持目标清单并核实其状态(公开编目信息)。
- 批量拉取 2020 – 2026 TLE, 生成半长轴与衰减速率序列, 机动剔除。
- 与 F10.7 做去趋势互相关 + 块自助法显著性 + 滞后估计(H1)。
- NRLMSIS 密度比对照极大/静期衰减比(H2),Ap 残差分析。
- 核实工具边界:TLE 平均根数的类型(Brouwer 平均)与直接差分的系统差——读 `sgp4` 文档并用短弧数值传播对照一次。
- 独立交叉校验: 换一组不同高度(~500 vs ~800 km)目标, 验证相关强度随高度降低而增强的预期。

7 · 新颖性边界

- 本课题不声称发现热层-太阳活动耦合(教科书级已知), 不声称改进大气模型;NRLMSIS 与 TLE 反演技术均为他人已发表工作。
- 已有工作:TLE 反演密度的方法学文献、极大期卫星再入的工程报告。
- 本项目贡献(主结论): 第 25 周完整极大期 $\times$ 多目标同口径的相关-滞后定量分析 + 与 NRLMSIS 的偏差表。属"时间覆盖空白 + 新样本独立检验"。
- 价值: 相关强、滞后短且模型符合 $\rightarrow$ 公开数据即可监测热层响应的可复现范例; 模型系统偏离 $\rightarrow$ 有文献价值的极大期偏差记录。两个结果都成立, 块自助法保证显著性可信。
- 核对获奖清单: 本赛道无轨道力学/空间天气课题; 与 T07/T08 同属太阳影响链但观测端不同(X 射线事件 vs 大气拖曳), 三题互不重叠。

⚠ 第 4 周末：数据入口判定。阈值:Space-Track 历史 TLE 批量获取跑通(或确认受限)。若受限,降级路径一：用 CelesTrak 近实时数据前瞻式采集(第 4 – 40 周每周存档，累计 36 周仍覆盖一段活动变化)+ 公开的历史衰减案例(如 2022 Starlink 事件公开数据，需核实可得性);降级路径二：改用 ISS 等 NASA 直接公布轨道数据的少数目标做深入单星分析。两条路径保留"衰减-F10.7 相关"主框架。⚠ 第 8 周末：验收 1 达标；不达标先查 TLE 历元解析与时区。⚠ 第 20 周检查点：若互相关 < 0.3 (H1 大幅失败)，检查目标是否有未公开机动；确认后把"TLE 序列中机动/异常检测方法"升级为并列结论，框架不散。⚠ 选择前提：需接受"数据获取端存在条款不确定性"这一风险；不接受者选 T07/T10。

P33 · 二维 Ising 模型临界指数的有限尺寸标度：跨三种蒙特卡洛算法的收敛性与临界慢化对照

★★★ · C 族·数值模拟 · 物理赛道·统计力学 · 纯 CPU(累计小时级) · 蒙特卡洛 + 有限尺寸标度

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

Metropolis、Wolff 团簇与热浴三种蒙特卡洛算法，在 $L = 16 - 128$ 的二维 Ising 模型上给出的临界温度与指数比 β/v 、 γ/v 是否在各自误差棒内一致并符合 Onsager 精确解？三者的自关联时间随尺寸的动力学标度指数 z 相差多少、对给定精度目标各需多少墙钟时间？

2 · 研究背景与空白

二维 Ising 模型是唯一同时具备"非平庸相变 + 精确解"的多体系统:Onsager(1944)给出 $T_c = 2/\ln(1+\sqrt{2}) \approx 2.269$ (以 J/k_B 为单位)与 $\beta = 1/8$ ，是检验一切蒙特卡洛管线的黄金标准。局部更新算法(Metropolis、热浴)在临界点遭遇临界慢化($z \approx 2.2$),Wolff(1989)团簇算法把 z 压到接近 0——这是计算物理教学的经典叙事。

各算法的指数与 z 值在文献与教材(Newman & Barkema 《Monte Carlo Methods in Statistical Physics》)中均有发表；但教学实现常常只跑单一算法、不给误差棒。

空白在于 可复现性维度：同一代码框架、同一随机数流水线、同一误差估计口径下三算法的"精度-成本"完整对照表，配开源代码，公开资料中意外地少。物理结论全部已知，正因如此代码对错可被无情检验——是训练科学计算规范的理想课题，且"对照表"本身即是被计算领域承认的贡献形式。

3 · 可检验假设

- H1: 三算法由 Binder 累积量交叉法得到的 T_c 互相一致且与精确值偏差 $\leq 0.5\%$ ；有限尺寸标度得到的 β/v 与精确值 $1/8 \div 1 = 0.125$ 偏差 $\leq 5\%$ 。
- H2:Metropolis 的动力学指数 z 落在 $2.0 - 2.4$,Wolff 的有效 $z \leq 0.5$ ；等目标误差下 $L = 128$ 临界点处 Wolff 的墙钟时间优势 ≥ 30 倍。

4 · 量化验收标准

1. 方法学校验(硬门槛)：与解析解极限对照 + 分辨率(尺寸)收敛扫描—— $L = 4$ 小格子的能量/磁化与穷举精确求和(2^{16} 态，直接计算)逐点一致到浮点精度; T_c 与 Onsager 值偏差 $\leq 0.5\%$ ；比热峰位随 L 的移动符合 $1/L$ 修正趋势。此条不过关，后续全部结论无效。
2. 统计口径预先写死：一切观测量误差用分块 jackknife(块长 ≥ 10 倍积分自关联时间)；自关联时间用积分法，窗口截断判据(自洽窗口 $c = 6$)写死；指数拟合的置信区间用对 bin 的自助法 ≥ 1000 次。
3. 尺寸序列 $L = 16/24/32/48/64/96/128$ ，每尺寸临界区 ≥ 10 个温度点；热化丢弃 ≥ 20 倍自关联时间并报告。
4. 随机数:PCG64(numpy 默认)，种子全部记录；同种子重跑结果逐位一致。

5. 三算法共用观测量与分析代码，只有更新核不同(保证对照公平)，代码结构在报告中说明。
6. 可复现性：代码 + 运行配置 + 原始时间序列开源，单尺寸算例可在 1 小时内重跑验证。

5 · 数据与工具

用途	来源 / 工具
实现	Python + numba(JIT; 开源)自写三算法; 参考实现见 Newman & Barkema 教材伪代码
运行量估算	numba 加持的 Metropolis 单自旋更新 $\sim 10^7\text{--}10^8$ 次/秒(常见笔记本 CPU 公开基准量级); $L=128$ 、 10^6 全格扫描 $\approx 1.6\times 10^{10}$ 次更新 $\approx$ 数十分钟/温度点; 全项目累计 ≤ 30 CPU 小时, 可夜间后台完成
精确解对照	Onsager 1944 解析结果(教材级公式),仅用于校验与对比, 不计入本项目的数据贡献
分析	numpy/scipy + matplotlib;jackknife/自助法自写

6 · 方法路径

1. 写 Metropolis 单核 + 观测量框架, 先过 $L = 4$ 穷举对照(验收 1 前半)。
2. 实现 Wolff 与热浴, 三算法过同一穷举对照。
3. 小尺寸($L \leq 32$)扫温,Binder 交叉定 T_c , 对照 Onsager(验收 1 后半)。
4. 全尺寸生产运行(夜间后台), 原始序列落盘。
5. 自关联时间与 z 的测量(H2), 等误差成本核算(墙钟实测)。
6. 有限尺寸标度提取 β/ν 、 γ/ν + 自助法置信区间(H1)。
7. 独立交叉校验: 用重加权(single-histogram reweighting)在单温度点数据上复算峰位, 与直接扫温一致 $\leq 1\sigma$ 。

7 · 新颖性边界

- 本课题不声称任何新物理: T_c 、临界指数、临界慢化与团簇算法优势全部为已发表结论(Onsager 1944;Wolff 1989; 教材)。物理结论已发表, 不得声称为本项目发现。
- 已有工作: 算法原始论文 + 教材级实现; 分散的 z 测量。
- 本项目贡献(主结论): 同框架、同误差口径下三算法的收敛性与精度-成本对照表 + 一条可操作的"给定精度选算法与格子大小"判据 + 开源全套。属"方法层面的可复现性贡献"——在计算领域被承认, 但定位必须讲清楚, 否则会被误读为课后作业。答辩时的叙事重点是误差分析与成本核算的严谨性。
- 核对获奖清单:"Phase Transition' in a Mechanical System: Rotation-Induced Spontaneous Symmetry Breaking"(2025 金奖)是宏观力学系统的对称破缺实验, 与统计力学蒙特卡洛无方法重叠, 但说明"相变"叙事受评委欢迎——报告可借该题说明相变普适性话题的听众基础, 并点名区分。

8 · 决策门槛(go / no-go)

⚠ 第 4 周末: $L = 4$ 穷举对照通过(逐位级一致)。不通过禁止前进——此关卡专为暴露更新核 bug 设计。⚠ 第 8 周末: 小尺寸 T_c 达标($\leq 0.5\%$)。若 numba 性能不足以支撑 $L = 128$ (单温度点 > 2 小时),降级: 尺寸上限降至 $L = 64$ (文献与教材表明该范围足以提取指数, 误差棒略宽), 全部分析框架不变; 同时把"Python/numba 的性能边界实测"写进成本对照一节, 变损失为内容。⚠ 第 20 周检查点: 三算法一致性若出现 $> 2\sigma$ 分歧, 最可能是自关

联低估——加倍运行长度复核；分歧被修复的过程写入"误差分析"一节。⚠️ 预算裁剪顺序：热浴算法可砍(保 Metropolis vs Wolff 二元对照)；温度点可从 10 减到 7。

P34 · 格子 Boltzmann 方法模拟圆柱绕流:Strouhal—Reynolds 关系的复现与每直径格点数收敛判据

★★★ · C 族·数值模拟 · 物理赛道·计算流体 · 纯 CPU(小时级/工况) · LBM 编程 + 收敛性分析

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

自研二维 D2Q9 格子 Boltzmann(LBM)求解器模拟的圆柱绕流，在 $Re = 60 - 180$ 区间给出的涡脱落 Strouhal 数 $St(Re)$ 与已发表实验关联式(Roshko 1954 及后续，需核实拟合式精确系数)的偏差是多少？要把偏差压到 $\leq 3\%$ ，每圆柱直径所需的最少格点数(分辨率判据)是多少？

2 · 研究背景与空白

圆柱绕流是流体力学的果蝇： Re 超过约 47 后尾流失稳，形成卡门涡街，无量纲脱落频率 $St = fD/U$ 随 Re 的关系被百年实验钉死，是一切新 CFD 代码的标准考题。LBM 用分布函数的迁移-碰撞代替直接求解 Navier - Stokes，实现只需百行代码、天然规则网格，是学生自写 CFD 的最低门槛路线；开源教学实现(如公开的 Python 教学脚本 `lbmFlowAroundCylinder`)广为流传。

$St(Re)$ 实验关联式(Roshko 1954;Williamson 对层流脱落区的精化，系数需核实)与 LBM 方法论文献均已充分发表。

空白在于 可复现性与判据：教学实现大多"能看到涡街就停"，不做系统的分辨率收敛、边界实现对照与定量误差表；"要多少格点才可信"这条对后来者最有用的判据，很少被明确给出。物理全部已知，贡献定位为收敛性对照与判据——与 T13 同类，合法但必须讲清。

3 · 可检验假设

- H1：分辨率 $\geq N$ 格点/直径时， St 与实验关联式偏差 $\leq 3\%$ (N 为本项目待测的判据，预期落在 20 - 60，依据教学文献常用值，需核实)；偏差随分辨率按接近二阶的速率收敛(LBM 理论精度)。
- H2：反弹(bounce-back)与插值曲面两种圆柱边界实现，在低分辨率端 St 偏差差 ≥ 1 个百分点、高分辨率端趋同——即边界实现主导低分辨率误差。

4 · 量化验收标准

1. 方法学校验(硬门槛)：与解析解极限对照 + 分辨率收敛扫描——先在平面 Poiseuille 流上与解析抛物线速度剖面对照， L_2 误差 $\leq 1\%$ ，且网格加密时按二阶收敛；再做圆柱工况 ≥ 4 档分辨率(如 10/20/40/80 格点/直径)的收敛扫描。此条不过关，后续全部结论无效。
2. 守恒量漂移强制报告：全域质量漂移 $\leq 10^{-6}$ 相对值/ 10^4 步；可压缩误差控制：马赫数 ≤ 0.1 并报告。
3. 统计口径预先写死： St 由升力系数时间序列的功率谱主峰读出，谱分辨率(采样时长 ≥ 50 个脱落周期)决定的频率不确定度随结果一并报告；瞬态丢弃判据(升力振幅达到饱和值 95% 后再等 10 周期)写死。
4. Re 序列 ≥ 7 点(60 - 180)，阻塞比(圆柱直径/通道宽) $\leq 1/8$ 并做一次半宽对照验证阻塞效应 $< 1\%$ 。

- 有效物理参数(Re 、 Ma 、松弛时间 τ 与格子粘度对应)每工况列表, 确认达到设定值。
- 可复现性: 求解器代码 + 全部工况配置文件开源, 最低分辨率工况可在 10 分钟内重跑。

5 · 数据与工具

用途	来源 / 工具
求解器	Python + numpy(向量化)自写 D2Q9 BGK; 参考公开教学实现的结构但独立编码
运行量估算	向量化 numpy LBM 公开教学基准约 1–10 MLUPS(百万格点更新/秒)量级(需核实自己机器实测); 最大工况 $1600 \times 400 = 6.4 \times 10^5$ 格点 $\times 2 \times 10^5$ 步 $\approx 1.3 \times 10^{11}$ 更新 $\approx$ 数小时–半天/工况, 全项目累计 ≤ 5 天 CPU, 夜间后台可行; 若不足转 numba(预期再快 5–20 倍, 需实测)
实验对照	Roshko 1954 及后续 $St-Re$ 关联式(系数需核实原文), 仅用于校验与对比, 不计入本项目的数据贡献
分析	scipy.signal(功率谱)+ matplotlib

6 · 方法路径

- 写 D2Q9 BGK 核 + 周期/无滑移边界, 过 Poiseuille 解析对照(验收 1 前半)。
- 实现圆柱反弹边界与升阻力的动量交换法测量。
- 单工况($Re = 100$)出涡街, 确认 St 落在文献 $0.16 - 0.17$ 邻域(粗检)。
- 分辨率收敛扫描(≥ 4 档)与阻塞比对照, 确定误差主导项。
- 全 Re 序列生产运行(后台), $St(Re)$ 曲线 + 误差棒。
- 实现插值边界, 做 $H2$ 的边界实现对照。
- 独立交叉校验: 对 $Re = 100$ 用涡量场的时空图法复测脱落频率, 与升力谱法一致 $\leq$ 谱分辨率。

7 · 新颖性边界

- 本课题不声称发现涡街或 $St-Re$ 关系(实验文献已钉死), 不声称提出新数值方法; LBM 与边界处理均为已发表方法。物理结论已发表, 不得声称为本项目发现。
- 已有工作: $St-Re$ 实验关联式; LBM 方法论与教学实现。
- 本项目贡献(主结论): 跨分辨率、跨边界实现的系统收敛对照 + 一条可操作的"格点/直径"分辨率判据 + 开源可一键复跑的求解器。属"方法层面的可复现性贡献", 定位必须在摘要里就讲清楚。
- 价值: 给所有用教学级 LBM 做课题的后来者一张"多少格点才可信"的表; 若二阶收敛在某分辨率段破坏, 该破坏点本身是结论。
- 核对获奖清单: "PINN-LOCK: Efficient Density Current Simulation via Physics-Constrained Loss"(2025 优胜)是本赛道最近的 CFD 题——该文用物理约束损失做密度流模拟; 本题不用神经网络、对象是涡街基准与收敛判据, 方法与问题均不同, 报告中点名区分。流体实验类获奖众多(水跃 2022、射流 2022/2023 等)均为实验路线, 无方法重叠。

8 · 决策门槛(go / no-go)

⚠ 第 5 周末:Poiseuille 校验达标($\leq 1\%$ 且二阶收敛)。不达标禁止上圆柱——先修碰撞核/边界。⚠ 第 9 周末:Re = 100 粗检达标 + 实测本机 MLUPS。若性能低于 1 MLUPS 且 numba 迁移后仍不足以完成最大分辨率,降级:最高分辨率降为 60 格点/直径、Re 上限降至 150,用 Richardson 外推补偿最高档缺失——收敛判据框架完整保留。⚠ 第 24 周检查点:H2 边界对照若因插值边界实现复杂而进度落后,砍 H2 保 H1(判据主线),把边界对照列为"展望"。⚠ 选择前提:学生需接受"结果是一张收敛表而非炫目现象"的叙事,并在答辩中能讲清 LBM 与 NS 方程的关系(Chapman – Enskog 至少到能说清量级对应)。

P35 · 紧密三行星系统失稳时间随轨道间距的标度：用 REBOUND 在偏心率维度上的独立检验

★★ · C 族·数值模拟 · 物理赛道·天体力学 · 纯 CPU(累计数天后台) · N 体积分 + 生存分析

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

等质量三行星系统的失稳时间对数 $\log_{10}(t_{\text{inst}}/P_1)$ 随以互 Hill 半径为单位的轨道间距 Δ 的近线性增长关系 (Chambers et al. 1996 建立;Obertas et al. 2017 在五行星零偏心系统精测)，在三行星、初始偏心率 $e = 0$ 与 $e = 0.02$ 两种群体下的斜率与截距是多少？偏心率是否如共振重叠图像预期地系统性缩短失稳时间？

2 · 研究背景与空白

多行星系统的长期稳定性没有解析判据，只能数值积分：间距(以互 Hill 半径 R_H 计)每增加一点，失稳时间指数式增长—— $\log t$ 对 Δ 近线性，是紧密系统(如 Kepler 多行星系)形成史约束的基石。Chambers et al.(1996, Icarus) 首次系统测绘该关系;Obertas et al.(2017, Icarus)用五行星零偏心系统把它精化到 $\Delta \leq 13$ 、每点大样本。开源 N 体框架 REBOUND(Rein & Liu 2012,A&A)的 WHFast 辛积分器让这类实验在笔记本上可行。

空白在于 参数维度：三行星(最小非平庸数目)+ 小初始偏心率的系统性标度测绘，在公开文献中不如五行星零偏心情形完整(近年机器学习稳定性预测工作侧面覆盖，需核实：搜索关键词 "three planet stability time spacing Hill radius eccentricity scaling")。诚实定位：这是"参数维度补全"型空白，可靠性次于时间覆盖型，但计算完全可控、正负结论均成立。

3 · 可检验假设

- H1: $e = 0$ 三行星系统在 $\Delta = 4 - 9$ 区间的斜率 $d \log t / d\Delta$ 落在已发表多行星结果的 $0.7 - 1.5$ 量级区间内(精确文献值需核实)，且本项目置信区间宽度 ≤ 0.1 。
- H2: $e = 0.02$ 群体在同一 Δ 下失稳时间中位数系统性缩短，等效于 Δ 平移 $\geq 0.5 R_H$ (方向与共振重叠加剧一致)；若平移不显著,"小偏心不敏感"即为有效结论。

4 · 量化验收标准

1. 方法学校验(硬门槛)：与解析极限对照 + 收敛扫描——两行星圆轨道系统在 Δ 大于临界值 $2\sqrt{3} R_H$ (Hill 稳定解析判据)时 10^7 轨道不失稳、小于时快速失稳，复现该解析边界 $\pm 5\%$ ；单条轨道的相对能量误差 $|\Delta E/E| \leq 10^{-8}$ (WHFast 辛特性)，超限的运行作废重跑；时间步长减半的对照子集(30 组)失稳时间中位数漂移 $\leq 0.3 \text{ dex}$ 。此条不过关，后续全部结论无效。
2. 统计口径预先写死：截尾数据(至 $10^7 P_1$ 未失稳)用生存分析口径处理——报 Kaplan - Meier 中位数与截尾比例，回归用含删失的最大似然，不得把截尾值当实测值；斜率置信区间用对系统的自助法 ≥ 1000 次；同一 Δ 内初始相位随机化 ≥ 20 组。
3. Δ 网格 ≥ 11 点(4.0 - 9.0 步长 0.5)，两偏心群体，总运行 $\geq 2 \times 11 \times 20 = 440$ 组。

4. 失稳判据写死：任意行星对最近距离 $< 1 R_H$ 或半长轴变化 $> 20\%$ ，以先到者为准。
5. 守恒量漂移(能量、角动量)每组落盘并汇总报告分布。
6. 可复现性：全部初条件种子 + REBOUND 版本锁定 + 脚本开源，任意单组可精确重跑。

5 · 数据与工具

用途	来源 / 工具
N 体积分	REBOUND(rebound.readthedocs.io,pip 开源),WHFast 辛积分器
运行量估算	WHFast 四体系统步进 $\sim 10^5-10^6$ 步/秒(官方文档基准量级，需本机实测)；单组 10^7 轨道 $\times 20$ 步/轨道 = 2×10^8 步 $\approx 5-30$ 分钟;440 组 $\approx 40-200$ CPU 小时，多进程 4 核夜间后台 1-2 周内完成
文献对照	Chambers 1996 / Obertas 2017 的标度参数,仅用于校验与对比，不计入本项目的数据贡献
分析	Python:numpy + lifelines(生存分析，开源)或自写 KM 估计；秒级

6 · 方法路径

1. 装 REBOUND，跑官方示例，完成两行星 Hill 边界与能量误差校验(验收 1)。
2. 写批量实验框架(初条件生成、种子管理、多进程调度、断点续跑)。
3. 小规模试跑(2 个 $\Delta \times 10$ 组)实测单组耗时，锁定总预算。
4. $e = 0$ 全网格生产运行,KM 分析 + 斜率拟合(H1)。
5. $e = 0.02$ 群体运行与平移量估计(H2)。
6. 步长减半对照子集，确认数值收敛(验收 1 收尾)。
7. 独立交叉校验：抽 10 组用 IAS15(高精度非辛积分器)重跑，失稳时间与 WHFast 一致 ≤ 0.3 dex。

7 · 新颖性边界

- 本课题不声称发现"间距-失稳时间"标度(Chambers 1996 已建立)，不声称提出稳定性新判据;REBOUND 为他人开源工具。
- 已有工作：五行星零偏心的精测(Obertas 2017)、两行星解析判据、近年 ML 稳定性预测(侧面覆盖参数空间，需核实其三行星覆盖密度)。
- 本项目贡献(主结论)：三行星 + 偏心维度的同口径标度测绘与"偏心等效平移量"的量化。属"参数维度补全的独立检验"——若斜率与五行星情形一致，标度的行星数普适性获得支持；若偏心平移显著，给出一个可被 Kepler 系统统计使用的修正量。两种结果都构成有效结论。
- 核对获奖清单：本赛道无 N 体动力学课题;"The Starry Sky on the Elastic Membrane"(2021 铜)是引力类比演示，与数值天体力学无重叠。
- 风险声明：若近年文献已密集覆盖本参数区(第 6 周核实)，贡献降为"独立复现 + 开源"，定位仍合法但上限下降——见 go/no-go。

⚠ 第 6 周末：文献核实关卡——检索三行星偏心标度的最新覆盖(关键词见上)。若发现完全重合的已发表网格，降级：把偏心维度改为质量比维度(等质量 vs 质量递增序列，文献覆盖更稀)。⚠ 第 8 周末：验收 1 达标 + 单组耗时实测。若单组 > 30 分钟导致总预算超 200 CPU 小时，降级：截尾缩短到 10^6 轨道并把 Δ 上限收到 7.5(标度在该窗口已可测,Obertas 2017 图示支持，需核实)，点数不减——框架保留。⚠ 第 30 周检查点：生产运行完成率 $\geq 80\%$ ，否则砍偏心群体样本量(20→12 组/点)保网格完整。⚠ 选择前提：需要能长期无人值守跑后台任务的电脑习惯(散热、电源、断点续跑)；笔记本夜间运行需家长知情。

P36 · 物理信息神经网络(PINN)求解一维 Burgers 方程：与谱方法/有限差分的双成本口径对照(高风险高上限)

★(高风险) · C 族·数值模拟 · 物理赛道·计算方法评测 · 纯 CPU + PyTorch · 深度学习 + 数值分析

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

选择前提(启动即确认)：学生已具备 PyTorch 基础与偏微分方程数值解初步；明确接受两点——其一，主结论很可能是"CPU 上 PINN 全面劣于传统方法"(负结论有效但不炫)；其二，答辩需用英文讲清自动微分与损失构造，讲不清则整题反噬。满足才启动，否则选 T13/T14。

1 · 研究问题

在纯 CPU 笔记本上,PINN 求解一维粘性 Burgers 方程到给定 L2 误差(相对解析参考解)所需的墙钟时间与自由参数量,相比 Chebyshev 谱方法与二阶有限差分各差多少倍? 误差-成本曲线是否存在交叉点(某精度以下 PINN 反而占优)?

2 · 研究背景与空白

PINN(Raissi et al. 2019, J. Comput. Phys.)把 PDE 残差写进神经网络损失,以采样代替网格,近年应用爆发;但对"它到底比传统方法贵多少"的严格核算,文献中相对稀缺且结论分歧,已有批评性工作指出其训练病态(Krishnapriyan et al. 2021, NeurIPS,"failure modes"分析)。一维粘性 Burgers 方程兼具非线性与陡峭梯度,且经 Cole-Hopf 变换存在解析参考解——是做严格误差核算的标准靶。

空白在于评价维度:绝大多数对照要么只报精度不报成本,要么只在 GPU 上核算;纯 CPU、双成本口径(等自由参数量 / 等墙钟时间)、多随机种子散布全报的小规模严格评测,恰好是笔记本学生能做、且对"CPU 用户该不该用 PINN"这一真实问题有用的补缺(相近评测是否已存在,需核实:搜索关键词 "PINN versus spectral method cost accuracy comparison CPU Burgers")。

3 · 可检验假设

- H1: 达到 L2 相对误差 10^{-3} 时,PINN 的墙钟成本 $\geq$ 传统方法的 100 倍(方向:PINN 显著劣势);等墙钟口径下 PINN 最优可达误差比谱方法差 ≥ 2 个数量级。
- H2:PINN 的误差-成本曲线在低精度端(相对误差 $\geq 5\%$)与差分法的差距缩小到 10 倍以内——即存在"够用就好"区间;若连该区间都不存在,亦为有效结论。

4 · 量化验收标准

1. 方法学校验(硬门槛):与解析解极限对照+分辨率收敛扫描——差分与谱方法先对 Cole-Hopf 解析参考解验证:差分呈二阶收敛、谱方法呈指数收敛,最细网格 L2 误差 $\leq 10^{-4}$;PINN 复现 Raissi 原文 Burgers 算例的定性解形态且 $L2 \leq 10^{-2}$ 。此条不过关,后续全部结论无效。
2. 统计口径预先写死(双成本口径强制):等自由度(网格点数/谱模数 vs 网络参数量)与等墙钟(同一台断网、插电、固定电源模式的笔记本实测)两种口径都要报;PINN 每配置 ≥ 5 个随机种子,报中位数与全距(训练结果散布大,不许只报最好种子);墙钟计时重复 3 次取中位数。

3. 配置矩阵:PINN(3 档网络规模 $\times$ 3 档采样点数)、差分(5 档网格)、谱(5 档模数), 全部列表。
4. 软件栈锁定并报告(PyTorch 版本、线程数固定、BLAS 后端), 排除多线程漂移。
5. 明确声明: 结论限定于"一维 Burgers + 笔记本 CPU + 所试超参", 不得外推到高维/GPU(外推仅作展望并标注为推测)。
6. 可复现性: 全部配置文件 + 种子 + 环境文件开源, 单配置可 30 分钟内重跑。

5 · 数据与工具

用途	来源 / 工具
PINN	PyTorch(CPU 版,pip 开源); 网络 ≤ 4 层 $\times$ 64 宽, 单次训练 10^4 — 10^5 步,CPU 数十分钟/次(Raissi 原文问题规模本就小,CPU 可行); 全矩阵 ≈ 45 次训练 ≈ 1 —3 天后台
参考解	Cole–Hopf 变换解析式自写实现(含数值积分的高精度求值,scipy.integrate.quad)
传统方法	numpy 自写二阶差分(Crank–Nicolson 或 RK4)+ Chebyshev 谱(numpy.polynomial 或自写 DCT 求导), 秒–分钟级
文献对照	Raissi et al. 2019 算例设置、Krishnapriyan et al. 2021 失败模式分类,仅用于校验与对比, 不计入本项目的数据贡献

6 · 方法路径

1. 实现解析参考解与两种传统求解器, 完成收敛性校验(验收 1 前半)。
2. 实现 PINN 基线, 复现 Raissi 算例形态(验收 1 后半)。
3. 冻结配置矩阵与计时规程(断网、电源模式、线程数)文档。
4. 生产运行全矩阵(PINN 部分夜间后台), 原始日志落盘。
5. 绘制误差-成本双口径曲线, 检验 H1/H2; 失败种子按 Krishnapriyan 分类归因。
6. 核实边界:PyTorch CPU 线程数对墙钟的影响曲线(1/2/4/8 线程), 选定并声明。
7. 独立交叉校验: 换一个粘度参数(更陡激波)复跑三方法中档配置, 确认结论方向不变。

7 · 新颖性边界

• 本课题不提出新网络结构或新训练技巧, 不声称"PINN 好/不好"的普适判断;PINN 方法(Raissi 2019)与其失败模式分析(Krishnapriyan 2021)均为已发表工作。 • 已有工作: 原始方法论文含 Burgers 算例(报精度、基本不报成本口径); 批评性文献报训练病态;GPU 上的若干评测。 • 本项目贡献(主结论): 纯 CPU、双成本口径、全种子散布的严格小规模评测 + 可复现基准包。属"评价维度的转换"——已有工作评估的是可行性与精度, 本项目评估的是消费级 CPU 场景下的成本-精度权衡。 • 价值: 对没有 GPU 的学生/教学用户,"该不该用 PINN"缺少诚实的数字; 本项目就是给这个数字。H1 的负方向结论(PINN 贵百倍)与正方向(存在够用区间)都成立。 • 核对获奖清单:"PINN-LOCK"(2025 优胜)——用物理约束损失加速密度流模拟, 是应用改进路线; 本题是严格评测路线, 不与其竞争同一贡献, 报告中必须点名并把差异写在摘要级。

⚠ 第 6 周末：传统方法校验达标(收敛阶正确)。这部分独立成立，是本题的保底资产。⚠ 第 10 周末:PINN 基线达 $L2 \leq 10^{-2}$ 。若反复失败(训练不收敛),立即降级：课题改为"Chebyshev 谱 vs 有限差分 vs Cole – Hopf 解析解的收敛性与成本对照 + PINN 失败模式的记录性附录"——双成本口径评测的主框架完整保留，只是被评测对象从三个减为两个，仍是完整论文(PINN 失败记录本身有参考价值)。⚠ 第 20 周检查点：全矩阵完成率 $\geq 60\%$ ，否则砍 PINN 配置矩阵到 2×2 。⚠ 预算裁剪顺序：第二粘度参数 $\rightarrow$ PINN 网络规模档 $\rightarrow$ 谱方法模数档(最后砍，它最便宜)。

P37 · 链条喷泉(Mould 效应)的高度标度：提升力来源之争的定量判别实验

★★ · D 族·理论与标度 · 物理赛道·经典力学 · 珠链 + 手机慢动作 <150 元 · 动量流建模 + 视频分析

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

珠链从容器中被拉出下落时形成的喷泉拱高度 h 与下落高度 H 之比是否如 Biggins & Warner(2014)的"堆推力"模型预言近似恒定，且该比值随链的离散结构(珠径/珠间距比)按其模型预测的方向变化？

2 · 研究背景与空白

珠链一端垂出容器下落时，链会先升到容器口以上形成拱——"链条喷泉"(Mould 效应,2013 年因科普视频引发关注)。朴素动量流分析给不出上升拱：若链仅被"拉"出，能量核算显示喷泉不该存在。Biggins & Warner(2014, Proc. R. Soc. A)提出关键机制：链是离散刚性单元的串联，从堆中被拽起时，堆对链节施加一个额外向上的推力，理论预言 h/H 为常数，且该常数由链的离散结构(单元能否被视为刚杆)决定；他们的实验测得 h/H 约百分之几到百分之十几的量级(精确数值需核实原文)。此后有后续实验与数值工作探讨推力来源细节，结论并不完全一致(需核实：搜索关键词 "chain fountain mechanism experiments after Biggins Warner")。

空白在于：争论的判别点——"改变链的离散结构应按可预测方向改变 h/H "——恰好是低成本实验能系统扫描的维度(不同珠径、珠间距、软绳对照)，而系统的结构序列数据在公开文献中不充分。文献之争本身就是研究内容，正负结论均成立。

3 · 可检验假设

- H1：同一珠链的 h/H 在 $H = 1 - 4\text{ m}$ 范围内为常数(线性回归斜率与 0 在 95% 置信区间内相容)，常数值落在 0.01 - 0.2 量级。
- H2：珠间距/珠径比更大的链(更接近刚杆串) h/H 更大；连续软绳(无离散结构) h/H 与 0 相容——两条与堆推力模型方向一致。任一被否定即构成对模型的有效质疑。

4 · 量验收标准

1. 方法学校验(硬门槛)：测量链校准——用已知长度标尺完成像素-米标定，误差 $\leq 1\%$ ；对自由落体小球用同一视频链路测 g ，得 $9.8 \pm 0.3\text{ m/s}^2$ ；喷泉稳态判据(拱顶高度在 $\geq 1\text{ s}$ 窗口内波动 $\leq 10\%$)预先写死。此条不过关，后续全部结论无效。
2. 每种链 $\times$ 每个 H 重复 ≥ 8 次； h 取稳态窗口中位数，误差取四分位距。
3. 统计口径预先写死： h 对 H 的回归用加权最小二乘 + 自助法 ≥ 1000 次置信区间；链间比较用置换检验；链速 v 由逐帧追踪($\geq 120\text{ fps}$)，动量流核算 $T \approx \lambda v^2$ 的自洽性单独报告(λ 线密度实测)。
4. 链结构表征：每链实测珠径、间距、线密度(≥ 3 段取平均)、最小弯曲半径(绕棒法，判据写死)。
5. 容器与堆放方式固定(盘绕方式写入规程并拍照存档)——堆构型是已知的强扰动源，作为控制变量。
6. 可复现性：全部视频 + 追踪脚本 + 数据表开源。

5 · 数据与工具

用途	来源 / 工具
链材	金属珠链 ≥ 3 种规格(窗帘/工艺珠链, 珠径 2.4/3.2/4.5 mm 等)+ 连续软绳对照, 合计 <100 元, 总长各 ≥ 10 m
场地	学校楼梯间/体育馆(落差 1—4 m, 下方软垫防砸)
成像	手机 120/240 fps 慢动作 + 三脚架 + 标尺;<50 元附件
追踪分析	Python + OpenCV(拱顶检测, 判据写死)或 Tracker 开源软件(physlets.org/tracker)
理论对照	Biggins & Warner 2014 模型与数据,仅用于校验与对比, 不计入本项目的数据贡献

6 · 方法路径

1. 搭建视频链路并完成标定与 g 校验(验收 1)。
2. 预实验: 单链找出能形成稳定喷泉的 H 下限与堆放规程。
3. 正式扫描: ≥ 3 链 + 软绳对照 $\times \geq 4$ 个 H $\times \geq 8$ 次。
4. 逐帧提取拱顶高度与链速, 做 H1 回归与 H2 组间检验。
5. 理论侧: 独立推导动量流模型(含/不含堆推力两版), 给出 h/H 与推力份额 α 的关系式(书面推导入报告)。
6. 用实测 h/H 反演各链的 α , 对照结构序列趋势。
7. 独立交叉校验: 对一种链改用不同容器高度(桌面 vs 手持杯)复测, 确认堆构型控制有效(组间差 $\leq$ 误差棒)。

7 · 新颖性边界

- 本课题不声称发现链条喷泉或提出其首个理论(Mould 现象 + Biggins & Warner 2014 模型均已发表), 不声称终结机制之争。
- 已有工作: BW 模型与其原始实验; 后续争鸣文献(覆盖范围需核实, 见上关键词)。
- 本项目贡献(主结论): 离散结构序列 + 软绳对照的同口径 h/H 数据集与推力份额 α 的反演——为"堆推力"图像提供结构维度的系统检验。属"参数维度补全的判别实验"。
- 价值: H2 任一方向的结果都直接对话文献之争;"软绳 h/H 与 0 相容"这一对照本身在公开文献中少见系统报告(需核实)。
- 核对获奖清单: 日常力学获奖密集——"Coupled Looping Pendulum"(2020 优胜)、"Dynamic Stabilization of Water Bottles"(2021 优胜)、"Beijing Mane Man 动力学"(2021 优胜)——但链条喷泉未出现过;2020 金奖"绳结自发形成机制"同为"绳/链类日常对象 + 机制判别", 证明该题材有金奖先例, 且对象与本题(静态拓扑统计 vs 动力学喷泉)完全不同, 报告中点名区分。

8 · 决策门槛(go / no-go)

⚠ 第 5 周末: 预实验判定。阈值: 至少一种链在 $H \leq 4$ m 下形成可稳定测量的喷泉($h \geq 3$ cm、稳态 ≥ 1 s)。若喷泉过低, 换更重珠链 + 最大可用落差; 若仍 < 3 cm, 降级: 研究对象改为"链条拉出动力学"——测链速 $v(t)$ 逼近极限速度的规律与张力核算(同一装置、同一模型框架的另一可测端), BW 模型检验的主框架保留。⚠ 第 10 周末: 拱顶自动追踪与人工抽检一致 $\leq 5\%$; 不达标改人工逐帧 + 双人一致性判据。⚠ 已知困难定位: 堆构型敏感性是本题最大系统误差, 方案已把它升级为控制变量 + 交叉校验步骤——若它仍主导散布, 把"喷泉高度

对初始堆构型的敏感性"本身写成结论的一部分(这正是后续文献争论点之一)。⚠ 预算裁剪:H 档数可从 4 减到 3; 链种类不可少于 $2 + \text{软绳}$ (否则 H2 失去判别力)。

P38 · 落球撞击颗粒床的成坑直径-能量标度:1/4 幂律在弹体密度与床压实度维度上的检验

★★ · D 族·理论与标度 · 物理赛道·颗粒物质/量纲分析 · 颗粒床 + 落球 <100 元 · π 定理建模 + 图像测量

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

小球低速撞击松散颗粒床形成的坑径 D_c 是否服从已发表的能量 1/4 幂律 $D_c \propto E^{(1/4)}$ (Uehara et al. 2003)? 该指数在弹体密度(钢/玻璃/木)与床压实度(松装 vs 振实)改变时是否漂移，漂移方向能否由量纲分析的 π 组合预判?

2 · 研究背景与空白

行星表面撞击坑的直径-能量标度无法在实验室全尺度复现，但低速颗粒撞击提供了受控类比:Uehara, Ambroso, Ojha & Durian(2003, PRL)测得松散玻璃珠床上坑径 $D_c \propto E^{(1/4)}$ 、坑深另有独立标度，并指出该指数可由能量-重力标度的量纲分析合理化(与行星撞击的重力区标度联系)。该文成为颗粒撞击的标准参照，后续文献扩展了弹体与床的参数(覆盖程度需核实：搜索关键词 "granular impact crater scaling projectile density packing fraction")。

空白在于 参数维度的同口径覆盖：床压实度(体积分数)是理论上应进入 π 组合、实验上又最难控制的变量，系统数据分散。本题的理论侧(π 定理推导可能的标度族、并指出各机制区的判别量)与实验侧(可控扫描)恰好构成一年期的完整闭环；指数漂移与否均为有效结论。

3 · 可检验假设

- H1: 松装玻璃珠床 + 钢球的 $D_c - E$ 幂律指数为 0.25 ± 0.02 (95% 置信区间覆盖 0.25)，复现 Uehara et al. 基准。
- H2: 振实床(体积分数提高 ≥ 3 个百分点)的指数不变但前置系数显著减小(坑更小)；弹体密度进入前置系数而非指数。任一"指数漂移 > 0.03 "的观测都构成对纯能量-重力标度的有效偏离证据。

4 · 量化验收标准

1. 方法学校验(硬门槛)：重跑已发表算例——在与 Uehara et al. 2003 尽可能一致的工况(松散玻璃珠、密实球、低速)下复现指数 0.25，偏差 ≤ 0.02 ；坑径测量的双人独立测量一致性 $\leq 5\%$ ；像素标定误差 $\leq 1\%$ 。此条不过关，后续全部结论无效。
2. 统计口径预先写死：指数用对数空间最小二乘为主 + Theil - Sen 稳健回归交叉 + 对次重抽自助法 ≥ 1000 次置信区间;E 覆盖 ≥ 1.5 个数量级(落高 5 cm - 2 m, ≥ 8 档)；每档 ≥ 10 次。
3. 床制备规程写死：松装用"筛雨法"(固定筛高缓慢撒落)，振实用固定次数敲击；每次撞击前重新制床；体积分数每 10 次实测一次(称重/体积)，批内波动 $\leq 1\%$ 。
4. 坑径判据写死：侧光俯拍，坑缘取环形亮暗边界的拟合圆直径；附判据示例图。
5. 低速条件核查：弹体入射速度($\sqrt{2gH}$)与 240 fps 视频抽查一致 $\leq 3\%$ 且远低于颗粒声速，标注适用区。

6. 可复现性：全部俯拍原图 + 测量脚本 + 数据表开源。

5 · 数据与工具

用途	来源 / 工具
颗粒床	玻璃珠(0.3—0.5 mm, 工艺品店/淘宝,~3 kg,<60 元)或过筛干沙；容器直径 ≥ 10 倍最大坑径
弹体	钢球/玻璃球/木球各 ≥ 2 个直径(1—2.5 cm),<30 元
释放与成像	电磁铁或机械夹释放(定高, 无初速)；手机俯拍 + 侧光台灯;240 fps 抽查入射速度
理论对照	Uehara et al. 2003(PRL)的指数与工况,仅用于校验与对比, 不计入本项目的数据贡献
分析	Python + OpenCV(圆拟合)+ scipy；纯 CPU 秒级

6 · 方法路径

- 搭装置，完成标定、双人一致性与筛雨制床规程(验收 1 准备)。
- 基准工况复现 Uehara 指数(验收 1)。
- 理论推导: π 定理列出无量纲组合，写出能量-重力区预期与可能偏离机制(书面推导入报告)。
- 正式扫描:2 压实度 $\times$ 3 弹体密度 $\times$ 8 能量档 $\times$ 10 次。
- 拟合矩阵(指数 + 前置系数)与自助法置信区间，检验 H2。
- 核实边界：容器尺寸效应——最大能量档换大容器复测 2 档，边界影响 $\leq 2\%$ 。
- 独立交叉校验：对基准工况用坑深-能量独立标度(Uehara 文中另一关系)复核数据自洽。

7 · 新颖性边界

- 本课题不声称发现颗粒坑标度(Uehara et al. 2003 已发表)，不声称直接外推到行星撞击(类比仅作背景，标注为定性联系)。
- 已有工作：基准指数 + 若干参数扩展(覆盖需核实)。
- 本项目贡献(主结论)：压实度 $\times$ 弹体密度矩阵的同口径指数/系数分解 + 公开原始图像数据集。属"参数维度补全"，与 T03 同类定位，但本题以量纲分析推导为理论主轴(D 族属性)。
- 价值:"压实度只进系数不进指数"若成立，是对能量-重力标度稳健性的干净支持；若指数漂移，偏离机制(如塑性耗散占比变化)即为讨论主线。误差棒设计保证能分辨 0.03 的指数差。
- 核对获奖清单：颗粒类相邻获奖同 T03(Boycott 2022、沙画 2023)；行星类比相邻题 "Geosynthetic-Reinforced Soil in Martian Simulant"(2025 优胜)为土工结构强度，与撞击标度无重叠。本赛道无成坑课题。

8 · 决策门槛(go / no-go)

⚠ 第 4 周末：双人坑径一致性与制床重复性达标(体积分数批内 $\leq 1\%$)。不达标先改测量/制床规程，不许进数据生产。⚠ 第 7 周末：基准指数复现达标(± 0.02)。若系统偏离(常见原因：颗粒过湿、床太浅 < 10 倍弹径、容器太小)，逐项排除；若确认无法复现,降级：放弃对接文献绝对值，改为"自建体系内的压实度/密度对照"(内部一致比较)，标度检验框架保留，但须在报告中如实说明与 Uehara 工况的差异清单。⚠ 已知困难定位：静电(干燥玻璃珠)与湿度是隐藏变量——全程记录湿度，雨天不采数；把"湿度敏感性"做成附录小实验。⚠ 预算裁剪：弹体密度档从 3 减到 2；压实度两档不可砍(H2 主判别)。

P39 · 用 SPARC 公开旋转曲线独立重建径向加速度关系(RAR)并检验其内禀散度

★★ · D 族·理论与标度 · 物理赛道·星系动力学/引力 · 纯 CPU + SPARC 表格 · 贝叶斯/误差预算 + 引力理论

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

1 · 研究问题

用 SPARC 数据库(~175 个盘星系，准确数需核实)的旋转曲线与光度轮廓重建的径向加速度关系(RAR：观测加速度 g_{obs} 对重子预期加速度 g_{bar} 的关系)，其特征加速度尺度 g_t 与 McGaugh, Lelli & Schombert(2016, PRL)报告的 $1.2 \times 10^{-10} \text{ m/s}^2$ 是否一致？扣除测量误差后的内禀散度与零相容，还是显著非零？

2 · 研究背景与空白

盘星系外围旋转速度不随半径下降，是暗物质假说的经典证据；而 RAR 把上百个星系的数千个测点压到一条极窄的曲线上——观测加速度只由重子(恒星 + 气体)加速度决定，散度小到接近测量误差。这条关系是暗物质晕模型与修改引力(MOND 类)争论的核心战场：内禀散度是否为零直接区分"紧密的动力学定律"与"晕-星系共演化的统计巧合"，McGaugh et al.(2016)测得观测散度约 0.13 dex 并主张几乎无内禀散度，随后多篇文献用不同统计口径得出不同结论(如 Rodrigues et al. 2018 质疑 g_t 的普适性，后续争论持续，需核实文献链：搜索关键词 "radial acceleration relation intrinsic scatter debate")。

空白在于 评价维度：这场争论很大程度上是统计口径之争(误差模型、星系间 vs 星系内散度、质光比先验)。用单一冻结口径 + 误差注入-回收实验重建 RAR、并把"各口径选择对内禀散度结论的影响"做成敏感性矩阵，是学生可完成、且对读懂这场著名争论真正有用的工作。数据为发布的表格，纯 CPU。

3 · 可检验假设

- H1: 重建的 g_t 与 $1.2 \times 10^{-10} \text{ m/s}^2$ 偏差 $\leq 20\%$ (置信区间重叠)；观测总散度复现 0.10 – 0.16 dex 区间。
- H2: 误差反卷积后的内禀散度置信区间要么与 0 相容、要么下界 $\geq 0.03 \text{ dex}$ ——两种结果分别支持争论两方，均为有效结论；质光比先验从 0.5 变到 0.7($3.6 \mu\text{m}$ 波段常用区间，需核实文献取值)时，内禀散度结论方向不变(稳健)或反转(口径敏感，亦为主结论)。

4 · 量化验收标准

1. 方法学校验(硬门槛)：用 SPARC 官方发布表按 McGaugh et al. 2016 的口径重建 RAR，拟合 g_t 偏差 $\leq 10\%$ 、总散度偏差 $\leq 0.02 \text{ dex}$ 。此条不过关，后续全部结论无效。
2. 统计口径预先写死：主拟合为含测点误差的最大似然(正交距离回归口径)， g_t 置信区间用对星系(非对测点)自助法 ≥ 1000 次——测点在星系内强相关，对测点重抽会假性收窄区间，此点在报告中显式说明；内禀散度用注入-回收法标定(合成数据注入已知散度，验证管线回收无偏， ≥ 20 个注入水平)。
3. 样本切割规则冻结(倾角 $> 30^\circ$ 、质量测点 ≥ 5 等，依 SPARC 推荐旗标，需核实旗标定义)。
4. 敏感性矩阵强制全报：质光比 $\times 2$ 档、距离误差开/关、倾角误差开/关、星系内/星系间散度分解。
5. 明确声明：本项目不判定暗物质 vs 修改引力(超出能力范围)，只交付"口径 $\rightarrow$ 内禀散度结论"的映射表。

6. 可复现性：下载链接 + 清洗脚本 + 全拟合代码开源。

5 · 数据与工具

用途	来源 / 工具
旋转曲线与光度	SPARC 数据库(astroweb.cwru.edu/SPARC,Lelli, McGaugh & Schombert 2016 发布, 公开纯文本表,< 50 MB)
分析	Python:pandas + scipy + emcee(可选,MCMC 开源)或自写最大似然; 全分析 CPU < 1 小时
文献对照	McGaugh et al. 2016 的 $g\ddagger$ 与散度、后续争论文献的口径清单,仅用于校验与对比, 不计入本项目的数据贡献

6 · 方法路径

1. 下载 SPARC 表, 读数据说明, 建立字段字典与样本切割文档。
2. 复现 RAR 与 $g\ddagger$ (验收 1)。
3. 写注入-回收管线, 标定内禀散度估计的偏差与分辨力。
4. 主拟合 + 对星系自助法, 得 $g\ddagger$ 与内禀散度基准结果。
5. 跑敏感性矩阵(质光比/距离/倾角/分解口径)。
6. 理论侧: 书面推导 RAR 在牛顿 + 晕模型与 MOND 插值函数下各自的预期形态(教材级推导, 标注为背景非贡献)。
7. 独立交叉校验: 用重子 Tully – Fisher 关系(同数据库另一投影)复核样本无系统异常。

7 · 新颖性边界

- 本课题不声称发现 RAR(McGaugh et al. 2016)或判定引力理论之争;SPARC 数据为 Lelli et al. 的公开贡献,不计入本项目数据贡献。
- 已有工作:RAR 原文、多篇口径互怼的后续文献(链条需核实)。
- 本项目贡献(主结论): 单一冻结口径的独立重建 + 注入-回收标定的内禀散度估计 + "口径选择 → 结论"敏感性映射表。属"评价维度的转换 + 可复现性贡献"——已有工作各自主张一个口径, 本项目量化口径间的结论距离。
- 价值: 给这场著名争论一份中学生也能重跑的"口径账本";H2 两个方向都是可写的结论。
- 核对获奖清单: 宇宙学/引力相邻获奖为 "Dark Sector Interaction Model / Hubble Tension"(2020 铜, 理论模型拟合)与 "Wet Dark Problem 理论修正"(2025 入围)——均非旋转曲线数据分析; 本赛道无 RAR/SPARC 课题。该谱系显示引力-宇宙学题可入围但奖次偏低, 本题以数据与统计严谨性对冲这一风险。

8 · 决策门槛(go / no-go)

⚠ 第 3 周末:SPARC 下载与字段核实完成(含误差列与质量旗标)。⚠ 第 6 周末: 验收 1 复现达标; 不达标先查单位制(kpc、 $M_{\odot}/L$)与质光比取值。⚠ 第 12 周末: 注入-回收标定判定。阈值: 管线对 0.05 dex 注入散度的回收偏差 ≤ 0.02 dex。若分辨力不足(误差模型吃掉信号),降级: 主结论改为"重子 Tully – Fisher 关系斜率与散度的同口径独立重测"(同数据库、同管线、文献参照充分),RAR 敏感性矩阵降为附录——统一口径检验的主框架保留。⚠ 选择前提: 学生需愿意读 2 – 3 篇英文原文的统计方法节(有指导老师陪读即可); 答辩需能讲清"内禀散度 vs 测量散度"的区别——这恰是本题的教学价值所在。

P40 · 含噪模拟下变分量子本征求解器(VQE)求横场 Ising 链基态：误差随线路深度与噪声率的标度(高风险高上限)

★(高风险) · D 族·理论与标度 · 物理赛道·量子多体/量子信息 · 纯 CPU + Qiskit Aer · 量子力学 + 数值优化

来源说明：本则出自独立撰写的第二批方案。它与前 20 则同样遵循八块结构与硬门槛要求，但撰写时未做英文文献检索，新颖性边界依据的是历届获奖图谱与既有知识，而非当轮查新。因此其「需核实」条目更多，启动前须自行补一轮英文检索。

选择前提(启动即确认)：学生对线性代数(本征值、张量积)有实操能力、对量子力学有认真的自学意愿，且接受两点——其一，本题概念门槛为 20 题最高，英文答辩讲不清楚测量、纠缠与变分原理会整题反噬；其二，保底产出(精确对角化标度研究)必须先于 VQE 完成。满足才启动，否则选 T13。

1 · 研究问题

在 Qiskit Aer 含噪模拟器(纯 CPU)中,VQE 求 $N \leq 10$ 自旋横场 Ising 链基态能量的相对误差，随变分线路深度 p 与去极化噪声率 λ 如何标度？给定 λ 下是否存在最优深度 $p(\lambda)$ (表达能力与噪声积累的竞争), p 随 λ 的变化方向是否符合"噪声惩罚 $\propto$ 门数"的一阶模型？

2 · 研究背景与空白

横场 Ising 链 $H = -\sum Z_i Z_{i+1} - g \sum X_i$ 是量子相变的最小模型，可 Jordan - Wigner 精确求解，小尺寸下直接对角化即得基态真值——是量子算法误差核算的理想靶。VQE(Peruzzo et al. 2014, Nat. Commun.)用参数化量子线路 + 经典优化器变分逼近基态，是含噪中等规模量子(NISQ)时代的旗舰算法；其两大已知病灶——贫瘠高原(barren plateaus,McClean et al. 2018, Nat. Commun.)与噪声下的误差积累——都随深度加剧，而表达能力随深度增强，提示存在最优深度。

"噪声限制 VQE 可达精度"在文献中已被广泛讨论;空白在于：小系统 + 单一噪声通道(去极化)+ 系统扫描 (N, p, λ) 的公开、可复现基准表并不常见(相近基准是否存在，需核实：搜索关键词 "VQE depolarizing noise optimal circuit depth benchmark transverse field Ising")，多数论文只展示个别工况。全部计算在笔记本 CPU 上可行: $N = 10$ 的密度矩阵模拟维度 $4^N \approx 10^6$,Aer 完全支持。

3 · 可检验假设

- H1: 无噪时能量相对误差随 p 单调下降, $p \geq N/2$ 时 $\leq 10^{-3}$ (硬件高效 ansatz, 临界点 $g = 1$)；含噪($\lambda = 10^{-3} - 10^{-2}$)时误差-深度曲线出现极小, $p^*(\lambda)$ 随 λ 增大而减小。
- H2: 极小点附近误差可分解为"表达误差(无噪残差)+ 噪声误差($\approx c \cdot \lambda \cdot$ 门数)"的加和模型，拟合优度 $R^2 \geq 0.9$ ；偏离加和模型的区域与优化器陷入局部极小的种子占比上升同步(可由多种子散布诊断)。

4 · 量化验收标准

1. 方法学校验(硬门槛)：与精确解对照——numpy 精确对角化 $N \leq 10$ 横场 Ising 链，基态能量与 Jordan - Wigner 解析式(有限链开边界的自由费米子谱)逐点一致到 10^{-10} ；无噪 statevector 模式下 VQE 基线在 $N = 4$ 、足够深度时误差 $\leq 10^{-6}$ (优化器收敛验证)。此条不过关，后续全部结论无效。

- 统计口径预先写死：每 (N, p, λ) 配置 ≥ 10 个随机初始化种子，报中位数与四分位距(优化落局部极小使分布偏态，不报均值为主、更不许只报最优种子); $p^*(\lambda)$ 的置信区间由对种子自助法给出;shot 噪声与门噪声分离——主扫描用密度矩阵模式(无 shot 噪声)，另做一组有限 shot(10^4)对照并单独报告。
- 扫描矩阵: $N = 6/8/10 \times p = 1 - 8 \times \lambda = 0/10^{-3}/3 \times 10^{-3}/10^{-2}$ ，全部列表；优化器(COBYLA 或 L-BFGS，选定后冻结)与迭代上限写死。
- 有效参数核查：每配置报告门数、参数量与实际达到的收敛判据。
- 明确声明：结论限于模拟器与所选噪声模型，不外推到真实硬件(外推仅作展望，标注为推测)。
- 可复现性：全部配置 + 种子 + 环境文件开源，单配置 30 分钟内可重跑。

5 · 数据与工具

用途	来源 / 工具
量子模拟	Qiskit + Qiskit Aer(开源,pip；密度矩阵/状态向量模拟器，含去极化噪声模型); $N = 10$ 密度矩阵单次能量评估秒级，单次 VQE(数百次评估)分钟-十分钟级，全矩阵(~ 120 配置 $\times 10$ 种子) $\approx$ 数天 CPU 后台，可行
真值	numpy/scipy 精确对角化($2^{10} = 1024$ 维哈密顿量，毫秒级)+ Jordan–Wigner 解析式自写(推导入报告)
文献对照	Peruzzo 2014(VQE)、McClean 2018(贫瘠高原),仅用于校验与对比，不计入本项目的数据贡献
分析	pandas + scipy 拟合 + matplotlib

6 · 方法路径

- 精确对角化 + Jordan – Wigner 解析对照(验收 1 前半；这一步独立构成保底资产)。
- 无噪 VQE 基线($N = 4$)收敛验证(验收 1 后半)。
- 冻结 ansatz、优化器、噪声模型与扫描矩阵文档。
- 生产扫描(夜间后台，断点续跑)，原始优化轨迹落盘。
- 提取误差-深度曲线与 $p^*(\lambda)$ ，拟合 H2 加和模型。
- 核实工具边界:Aer 去极化模型的门级施加方式(单/双比特门分别的 λ 约定)——读文档 + 单门小实验验证，写入方法节。
- 独立交叉校验：对 $N = 6$ 的 3 个配置换 ansatz(全对易层结构)复跑，确认 $p^*(\lambda)$ 趋势不依赖特定 ansatz。

7 · 新颖性边界

- 本课题不提出新算法、不声称量子优势相关任何结论;VQE、贫瘠高原、噪声限制均为已发表工作，横场 Ising 链解析解为教科书结果,不得声称为本项目发现。
- 已有工作：算法原文、噪声下 VQE 的多篇个案研究(系统基准的覆盖度需核实)。
- 本项目贡献(主结论): (N, p, λ) 三维系统扫描的公开基准表 + "最优深度"的加和模型检验 + 全种子散布报告。属"方法层面的可复现性贡献 + 评价维度转换(个案 $\rightarrow$ 系统标度)"。
- 价值:NISQ 教学与入门研究者缺一张"深度该开多大"的诚实表格;H1 的极小点不存在(噪声下单调劣化)同样是清晰结论。
- 核对获奖清单：本赛道 2020 – 2025 无量子计算/量子信息课题——属未开垦题材，机会与风险并存：评委新鲜感是上限，无先例参照与答辩理解成本是风险；学科归属(物理 vs 计算机赛道)须在报名前对照当年物理学科参赛指南确认(需以官方指南为准)。

8 · 决策门槛(go / no-go)

⚠ 第 6 周末：验收 1 前半(精确对角化 + 解析对照)达标——此为保底资产，不达标说明线性代数基础不足,整题终止改选 T13。⚠ 第 12 周末：无噪 VQE 基线达标。若优化始终不收敛,降级路径一：改用变分经典模拟(矩阵乘积态或直接变分波函数,numpy 实现)研究同一"表达能力-参数量"标度，量子线路改为附录;降级路径二：课题改为"横场 Ising 链的纠缠熵与能隙的有限尺寸标度"(纯精确对角化,N 至 14 仍 CPU 可行)。两条路径都保留"小量子系统 + 数值标度"的主框架。⚠ 第 24 周检查点：扫描完成率 $\geq 50\%$ ，否则砍 $N = 10$ 档(保 6/8)并缩 λ 到 3 档。⚠ 预算裁剪顺序:shot 噪声对照组 $\rightarrow$ 换 ansatz 交叉校验 $\rightarrow$ N 档(最后砍)。

附录 A · 倒排时间表（52 周，通用骨架）

终点是 2027 年 9 月 15 日提交截止（按 2026 赛季日程外推，须以当年官方公告为准）。第 0 周为 2026 年 9 月第一周。每个阶段都有可测的硬门槛，不是任务描述。

周次	阶段	硬门槛（可测，不达标即触发降级）
0—2	选题定稿	从本方案选定 1 条主路线 + 1 条备选路线；确认指导老师；核对当年官方参赛指南的学科研究范围
2—6	环境与方法学校验	跑通该路线验收标准第 1 条的方法学校验。 这是全程最重要的门槛：不过关不得进入下一阶段
6—10	数据/器材到位	数据完成下载并通过完整性检查；或器材到齐且装置重复性达标（实验类 $RSD \leq$ 阈值）
10—14	第一次 go/no-go	按各路线写明的阈值判定。不达标立即切换到该路线的具名降级路径， 不要拖到第 20 周
14—28	主体研究	完成主实验/主计算的全部参数点；中期产出可展示的核心图表 3 张
28—34	第二次 go/no-go	主结论方向已明确（含“差异不显著”这一有效结论）；误差棒已能证明有分辨该差异的能力
34—40	交叉校验	用第二种独立方法重算核心量，两者一致或差异有可解释来源
40—46	中文初稿	全文完成；图表定稿；代码与数据整理为可一键重跑的仓库
46—50	英文稿与查重	完成英文研究报告与查重报告 （总决赛强制英文，且提交材料含查重报告）
50—52	提交与答辩准备	系统提交（9 月 15 日截止，逾期不可修改）；英文 PPT 与答辩问答演练

两个容易被忽略的时间陷阱：

其一，英文化不是最后一周的事。总决赛全程英文（研究报告、PPT、答辩），从第 0 周起就要用英文记录关键术语与方法描述，否则第 46 周会成为瓶颈。

其二，11 月 3—10 日的公示期会把研究报告全网公开。这意味着新颖性问题会暴露在公众视野，查重与文献边界必须在提交前就处理干净。

附录 B · 资源清单（全赛道通用底座）

用途	来源 / 工具	说明
文献检索	Google Scholar、arXiv、PubMed、Semantic Scholar、Connected Papers	必须用英文检索 ；中文检索会系统性漏掉近三年工作
全文获取	arXiv、bioRxiv、PMC、作者主页、通过学校图书馆或指导老师获取	不要使用非法下载渠道，公示期会暴露
代码与数据托管	GitHub（公开仓库）+ Zenodo（存档并获取 DOI）	可复现性是评审加分项，也是自证独立完成的证据
计算环境	Python 3.11+（conda/uv 管理）、Jupyter	环境须导出为 <code>environment.yml</code> 或 <code>requirements.txt</code>
统计与作图	numpy / scipy / statsmodels / matplotlib	图表须含误差棒；配色对色盲友好
实验记录	纸质或电子实验记录本，逐日、带日期、不得回填	评委可能追问原始记录；实验类还需按要求提交实验视频
写作	LaTeX（Overleaf）或 Word；参考文献用 Zotero 管理	英文稿建议请母语者或老师润色，但 内容必须学生本人撰写
版本管理	Git，从第 0 周开始提交	连续的提交历史是“学生本人完成”最有力的旁证

附录 C · 红线清单

以下每一条，触碰即可能导致取消参赛/获奖资格，或在评审中被直接判负。

资格红线 1. 指导老师不得来自公司、培训机构或任何谋利机构。组委会**有理由相信学生受培训机构指导完成研究报告，即保留取消资格的权利**。2. 不得跨校组队；同一学科每人只能报名一次；同一篇论文不得报两个学科。3. 参赛信息须据实填写；报告提交截止后不得更改团队与指导老师信息。4. 曾参加或拟参加其他竞赛、已投稿或发表的，**必须在报名时如实申报**。

学术诚信红线 5. 不得将他人（含指导老师、高校课题组）已完成的工作写成本人贡献。研究报告须明确区分“他人已发表成果”“公开数据库/事件表”与“本项目贡献”。6. 公开数据库、已发表事件表、预训练模型只能作为对照与输入，**不得计入学生的数据或方法贡献**。7. 不得给出未标注为推测的物理/因果外推。讨论部分可以推测，但必须明示。8. 提交材料含查重报告；公示期报告全网公开，抄袭与过度重合会被公开发现。9. AI 使用政策以当年官方参赛规则为准（公开公告中未见明文，须核实）。无论政策如何，研究报告的科学内容必须由学生本人完成；若使用 AI 辅助，按当年规则要求致谢。

方法学红线（不违规，但会被评委问倒） 10. 没有方法学校验的结果不可信——所有路线的验收标准第 1 条都是这个门槛。11. 不平衡分类报 accuracy、时间序列随机划分、幂律只做双对数最小二乘、模拟不做收敛性检查：这四项是评审最常见的追问点，方案里已逐条写死正确口径。12. “差异不显著”是有效结论，但必须有误差棒证明有能力分辨该差异；没有误差棒的“无差异”等于没有结论。

附录 D · 本方案的使用方式

1. 先读赛道导语，判断这个赛道的评委口味是否匹配学生的能力结构。跨赛道选题（例如数学好的学生去做经济金融建模的理论建模族）常常比在本赛道硬拼更有胜算。
2. 按可行性顺序从上往下读，不要从最精彩的那条开始。编号越小越稳。
3. 选定一条主路线后，必须再选一条备选路线，且两条最好落在不同的族——这样第一次 go/no-go 触发时有真正的退路。
4. 第 2-6 周的方法学校验是全案的生死线。这一步跑不通，说明学生的技能栈与该路线不匹配，此时换题的成本远低于第 20 周换题。
5. 本方案的赛制信息取自官方公告（yau-awards.com，2020-2026 全部公示文章），但规则每年会改。报名开放后必须重新核对当年的《参赛规则》与六个学科各自的《参赛指南》，尤其是研究范围、评审标准与 AI 使用政策。