

太极实验室 2026 年度“大学生创新实践训练计划”

(引力波数据分析与 AI for Science 方向)

一、简介

中国科学院大学引力波宇宙太极实验室(北京)引力波数据分析与机器学习课题组长期面向全国高校学生开放科研训练机会。现面向全国优秀学生招募参加 **太极实验室 2026 年度“大学生创新实践训练计划”**。本课题组致力于探索引力波天文学、数值模拟与人工智能技术的交叉研究,重点发展新一代 **AI for Science** 方法,用于解决复杂物理系统建模、信号处理与科学数据分析问题。欢迎对 **引力波科学、人工智能算法与科学计算** 充满兴趣的同学加入,在真实科研项目中接受系统训练,并参与国际前沿研究。

本次为 **2026 年度集中招募与培训计划**。此外,本实验室长期欢迎对相关研究方向感兴趣的同学联系咨询,本广告长期有效。

二、申请要求

| 1. 专业背景与申请阶段 | |
|--|--|
| 欢迎但不限于以下专业: <ul style="list-style-type: none">✓ 物理学 / 天文学✓ 计算机科学与技术✓ 软件工程✓ 人工智能✓ 电子信息工程 / 通信工程✓ 自动化 / 控制科学✓ 数学 / 应用数学 / 统计学✓ 量子信息 / 量子计算✓ 数据科学 / 金融工程等 | 申请者可以来自不同学习阶段,包括: <ul style="list-style-type: none">✓ 本科生(不限年级,大二/大三优先)✓ 硕士研究生 / 博士研究生✓ 编程能力突出的高中生 |

我们特别欢迎 **编程能力强、学习主动性高、自驱力强** 的申请者。

| 2. 基础能力 | 3. 参与方式 |
|--|---|
| 申请者应具备: <ul style="list-style-type: none">✓ 良好的数学基础(线性代数、概率统计、微积分等)✓ 熟悉或愿意学习 Python / Linux / 科学计算编程✓ 具备一定的 英文文献阅读能力 | 项目成员需: <ul style="list-style-type: none">✓ 在项目期间持续参与科研训练✓ 每周至少投入半天时间参与远程讨论或学习✓ 寒暑假期间可来中国科学院大学参与科研实践✓ 在项目结束时完成研究报告或项目成果总结 |

三、项目研究方向

本项目围绕 **引力波科学与 AI for Science 交叉研究**, 可能涉及的研究方向包括 (但不限于):

| | |
|--|---|
| <p>1. 引力波科学与物理建模</p> <ul style="list-style-type: none">✓ 引力波数据分析✓ 多信使天文学数据处理✓ 数值相对论与波形建模✓ 偏微分方程数值求解✓ 科学模拟与物理建模 <p>2. 人工智能与科学计算</p> <ul style="list-style-type: none">✓ 机器学习在科学数据中的应用 (AI for Science)✓ 深度学习在时序信号分析中的应用✓ 生成模型 (Flow / Diffusion models)✓ Physics-Informed Neural Networks (PINN)✓ Operator Learning (DeepONet / Neural Operators)✓ 模拟推断 (Simulation-based inference)✓ 概率统计推断与贝叶斯分析 | <p>3. 新型智能算法</p> <ul style="list-style-type: none">✓ 大语言模型 (LLM) 在科研自动化中的应用✓ AI Agents 与自动化科研系统✓ 进化优化算法与启发式搜索✓ AutoML 与科学模型自动设计 <p>4. 科学计算与工程实现</p> <ul style="list-style-type: none">✓ 数字信号处理✓ 高性能计算 (HPC)✓ GPU 并行计算与科学软件开发✓ 科学计算软件工程 <p>5. 前沿交叉方向</p> <ul style="list-style-type: none">✓ 量子计算在科学计算中的潜在应用✓ 复杂系统数据分析✓ 量化分析与科学数据建模 |
|--|---|

- 通过选拔的学生可以根据学生背景、个人兴趣或选拔作业的完成情况, 在指导老师的辅助下匹配研究课题方向。

四、课题组学习资源

课题组入门课程 《引力波数据探索: 编程与分析实战训练营》 已公开发布:

- ✓ Bilibili 课程: <https://space.bilibili.com/76060243/channel/seriesdetail?sid=3896245>
 - ✓ Github 课程资料: <https://github.com/iphysresearch/GWData-Bootcamp>
- 其他推荐学习资源:
- ✓ GW Open Data Workshops: <https://gwosc.org/workshops/>

五、项目支持与发展机会

“科创计划”以项目形式提供远程指导支持, 项目周期一般为 6-12 个月。

表现优秀者将有机会:

- ✓ 参加 **国科大暑期本科生夏令营项目**
- ✓ 参与 **引力波国际学术会议与研讨会**
- ✓ 在本课题组完成 **本科毕业设计 / 研究生阶段学习**
- ✓ 推荐至 **国内外相关研究团队继续深造**
- ✓ 参与科研论文撰写并发表 **SCI 期刊论文**

六、申请方式

请于 2026 年 4 月 1 日前填写报名表: <https://ictp-ap.org/form/113>

并附上申请材料, 包括:

- ✓ 简短个人简历
- ✓ 编程或科研能力证明 (若有, 如 Github 项目等)
- ✓ 选拔题目结果 (若有)

我们将在截止日期后组织遴选与考核, 并通知获得资助的申请人。

七、联系方式

如有任何问题, 欢迎咨询:

| | |
|---|--|
| 贾老师 电话: 010-82642446 邮箱: jiahaihong@ucas.ac.cn | 指导教师: 王赫 邮箱: hewang@ucas.ac.cn |
|---|--|

八、备注

本课题组长期欢迎对 **引力波科学与 AI for Science 研究** 感兴趣的同学联系咨询。即使未赶上本次集中招募, 也欢迎随时邮件联系。

太极实验室科创计划详情:

- <https://github.com/iphysresearch/UndergradResearchLab>
- <https://ictp-ap.org/post/244>

太极实验室 2026 年度“大学生创新实践训练计划”

(引力波数据分析与 AI for Science 方向)

选拔目标:

强自驱动力和自学能力+愿意通过动手编程来解决问题+充分利用搜索引擎快速学习相关概念和知识点, 以及查找相关的代码解决方案

选拔要求与任务说明:

根据个人兴趣和技能偏好, 自主选择题目中的任务进行完成。每个任务对应不同难度等级 (以☑表示), 选拔标准为**累计完成任务难度达到☑☑☑☑ (4 星)**。完成期限为 2 周 (可根据实际进展情况申请适当延长)。

- (一) 基于机器学习的致密双星并合引力波信号搜寻 【☑】
 - (二) 基于机器学习的快速参数推断
 - 子任务一: MCMC+调通并跑通 DINGO 程序 【☑☑☑】
 - 子任务二: LISA 频段下基于 SBI 的 SGWB 参数反演 【☑☑☑】
 - (三) 基于 PycWB 的真实引力波事件信号的时频聚类分析 【☑☑☑】
 - (四) 基于 LISA Analysis Tools 的数据分析技术
 - 子任务一: LISA 背景下实现的 MCMC 贝叶斯推断 【☑☑☑】
 - 子任务二: LISA 背景下实现的 GRry 统计推断 【☑☑☑】
 - (五) LISA Data Challenge / Taiji Data Challenge
 - 子任务一: LISA Data Challenge 数据时频变换及可视化 【☑】
 - 子任务二: Taiji Data Challenge 大质量双黑洞信号参数估计 【☑☑☑】
 - (六) LISA 探测器响应灵敏度 【☑☑】
 - (七) 基于 Docker+GRChombo 的双黑洞数值相对论模拟 【☑】
 - (八) 利用个人电脑或 HPC 编译 Chombo 与 GRChombo 【☑】
 - (九) 利用个人电脑或 HPC 进行双黑洞模拟并提取引力波 【☑☑☑☑】
 - (十) 基于 HDF5 数据的提取、差分处理及回写实现 【☑☑☑☑☑】
 - (十一) 基于 Einstein Toolkit 的双中子星模拟和引力波提取 【☑☑☑☑☑】
 - (十二) 基于 AMSS-NCKU 的相对论三体数值模拟 【☑☑☑☑】
 - (十三) 多尺度 PINN 的实现与二维推广 【☑☑】 ~ 【☑☑☑☑】
 - (十四) 基于新闻文本的股价波动预测与因果关系初探 【☑☑☑☑】
- 涉及数值相对论的题目 (七~十三), 重点考察学生的计算机能力, 尤其是对阅读并理解陌生代码的能力是一种考验。我们对提交结果的精度没有特别严格的要求, 因此学生无需担心个人电脑性能所带来的误差, 相比于结果, 我们更加重视学生学会探索的过程。
 - 我们建议学生在遇到困难时可以寻求 LLM 模型的帮助。此外, 与题目涉及的全部参考资料基本均为英文, 我们认为**英语能力**还有**科学上网**对于科研是必不可少的。
 - 我们准备的所有选拔题目没有“保姆式”地将所有可能的论文或代码链接全部给到学生, 仅提供了必要的关键词和说明细节, 这对学生来说是一种基本研究素养的考验, 也算是**模拟了与你的未来研究生导师交流时最常见到的一种场景: 面对大量陌生的概念、知识和有限的参考资料, 能否“顺藤摸瓜”在 Github/Baidu/Google 等互联网上迅速查找到相关资料和程序代码, 并且“快速吸收”且“为己所用”。**
 - 总之, 请按照你自己的理解, 尽你最大的可能和努力来完成。

(一) 基于机器学习的致密双星并合引力波信号搜寻

- 关键词：信号搜寻、深度学习、PyTorch
- 任务难度：☑□□□□
- 任务描述：

请自学并阅读《引力波数据探索：编程与分析实战训练营》里 Kaggle Data Science Competition (Hackathon) 的数据挑战任务：

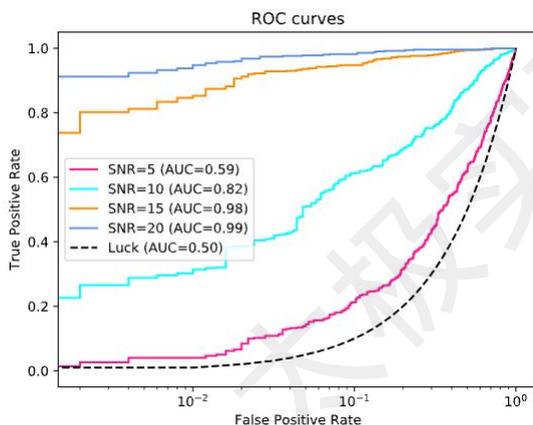
- <https://github.com/iphysresearch/GWData-Bootcamp/?tab=readme-ov-file#kaggle-data-science-competition-hackathon>

- <https://www.kaggle.com/competitions/2023-gwdata-bootcamp>

跑通如下 notebook 链接中的代码：

1. https://github.com/iphysresearch/GWData-Bootcamp/blob/main/2023/deep_learning/baseline/baseline_sugon.ipynb

2. 将跑通的 notebook 上传至自己账号下的 github 仓库 repo 中，报名回复该仓库 repo 链接即为上传作业成功。类似 ROC 结果如下图：



3. 重构并修改源代码：将双黑洞系统修改为双中子星旋近系统的信号搜寻目标。

4. 重构并修改源代码：将 CNN 模型替换为 ResNet (任意版本皆可)模型，绘制双中子星旋近引力波信号搜寻任务的 ROC 图。

5. 将修改的代码和跑通的 notebook 上传至自己账号下的 github 仓库 repo 中，报名回复 repo 链接即为成功。

(二) 基于机器学习的快速参数推断

➤ 关键词: 参数估计, 深度学习, Pytorch, SBI,

子任务一: MCMC+调通并跑通 DINGO 程序

➤ 任务难度:

➤ 任务描述:

1. 先在本地跑通 tutorial-dingo-introduction, 再基于 dingo 程序并且基于 FMPE 模型跑通一个 GW150914 的 toy case 结果, 最终实验结果应该是跑完后反演参数后验分布的 corner 图像。

2. 用 MCMC 或 Bilby 程序给出 GW150914 同样的十五维后验参数分布的 corner 结果, 并和上一步的结果放在同一个图上进行比较。

3. 目标结果图应该接近于: <https://stephengreen.com/research/sbi/> 中的图。

参考链接: <https://github.com/annalena-k/tutorial-dingo-introduction>

参考链接: <https://github.com/dingo-gw/dingo>

子任务二: LISA 频段下基于 SBI 的 SGWB 参数反演

➤ 任务难度:

➤ 任务描述:

1. 基于 arxiv: 2309.07954 文章, 仔细阅读并理解 template-powerlaw 分支下的 SAQQARA 程序, 并且成功在本地跑通示例 notebook 的代码:

https://github.com/PEREGRINE-GW/saqqara/blob/template-powerlaw/saqqara/Example_notebook.ipynb

2. 将训练的模型进一步在指定的 new injection (即 new_obs) 中的物理参数取定为 $\text{amplitude} = -12$, $\text{tilt} = 0$, 同样地给出测试后对应的两套图像。

参考链接: <https://github.com/peregrine-gw/saqqara>

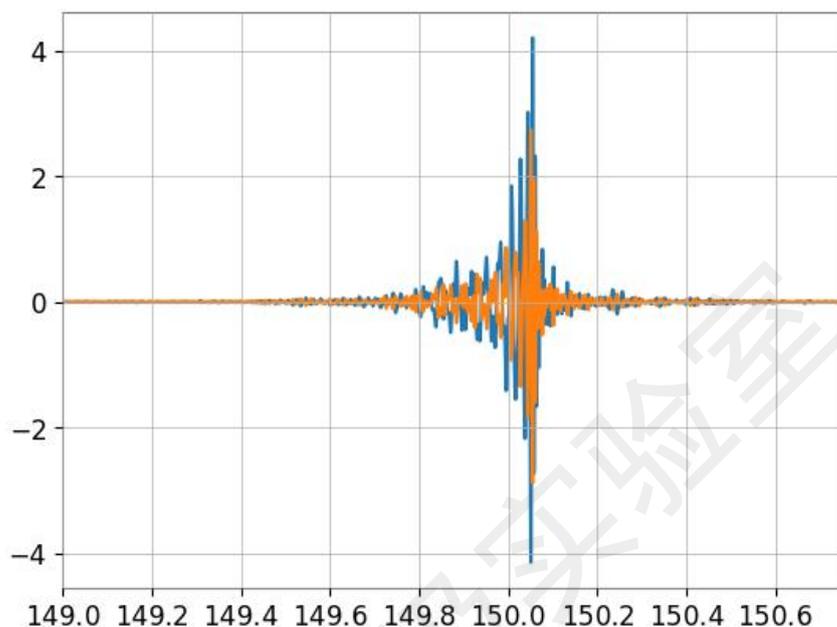
(三) 基于 PycWB 的真实引力波事件信号的时频聚类分析

➤ 关键词：时频分析、科学上网

➤ 任务难度：☑☑☑□□

➤ 任务描述：

1. 请自学并实现 PycWB 的时频分析，完成 GW190521_074359 引力波信号的 PycWB 时频分析和波形重构。波形重构后的结果如下图：



2. (重要) 对 notebook 里数据的时频图的像素进行聚类，并且与 PycWB 所重构的信号进行比较分析，并想办法计算 fitting factor

3. 将跑通的 notebook 上传至自己账号下的 github 仓库 repo 中，报名回复该仓库 repo 链接即为上传作业成功。

PycWB 仓库链接：<https://github.com/PycWB/pycwb>

可以参考的相关文档：

<https://yumeng.xu.docs.ligo.org/pycwb>

<https://pycwb.readthedocs.io/en/latest/>

可以参考的相关文章：<https://arxiv.org/abs/2308.08639>

可以参考的引力波 event list：<https://gwosc.org/eventapi/html/allevnts/>

Hint: 修改下面 notebook 的 GPS 时间即可完成分析

(https://colab.research.google.com/drive/11WK8LPL9sf0Jb1OaM_gYD0tX03FI DvIQ?usp=sharing)

(四) 基于 LISA Analysis Tools 的数据分析技术

子任务一：LISA 背景下实现的 MCMC 贝叶斯推断

➤ 关键词：LISA、Python、TDI、EMRI、参数估计

➤ 任务难度：☑☑☑□□

➤ 任务描述：

1. 按照要求安装并部署环境 LISA Analysis Tools Workshop (LATW)
2. 仔细阅读代码仓库里提供的六个教程代码 notebooks (tutorials/Tutorial1-6.ipynb) ，完成其中所有的代码作业。
3. 参考 Bilby-MCMC 中定义的各类提案分布，在 Eryn 中通过自定义提案分布 (从 Bilby-MCMC 中任选其一即可) 重新完成 Tutorial6.ipynb 的代码任务。
4. 将修改后的 Eryn 源码 (标注修改位置) +完成的 notebooks 与结果分析上传至自己账号下的 github 仓库 repo 中，报名回复 repo 链接即为成功。

参考资料：

LISA Analysis Tools Workshop: <https://indico.physics.auth.gr/event/17/>

Bilby-MCMC 的文档说明：

<https://lscsoft.docs.ligo.org/bilby/bilby-mcmc-guide.html>

LATW 代码仓库地址: <https://github.com/mikekatz04/LATW/>

子任务二：LISA 背景下实现的 GRry 统计推断

关键词：LISA、Python、参数估计，高斯过程

任务难度：☑☑☑□□

作业描述：

1. 按照子任务一中的要求安装并部署环境 LISA Analysis Tools Workshop (LATW)，并且完成其中所有的 Tutorial1-6.ipynb 代码作业
2. 将 Tutorial6.ipynb 代码任务中把完成 MCMC 的 Eryn 程序换成 GPry 程序 (<https://gpry.readthedocs.io/>)，实现同样的后验分布推断的代码任务。
3. 同子任务一的 4。

arXiv:2211.02045 for the core algorithm.

arXiv:2305.19267 for the NORA Nested-Sampling acquisition engine.

arXiv:2503.21871 for LISA inference

(五) LISA Data Challenge / Taiji Data Challenge

子任务一：LISA Data Challenge 数据时频变换及可视化

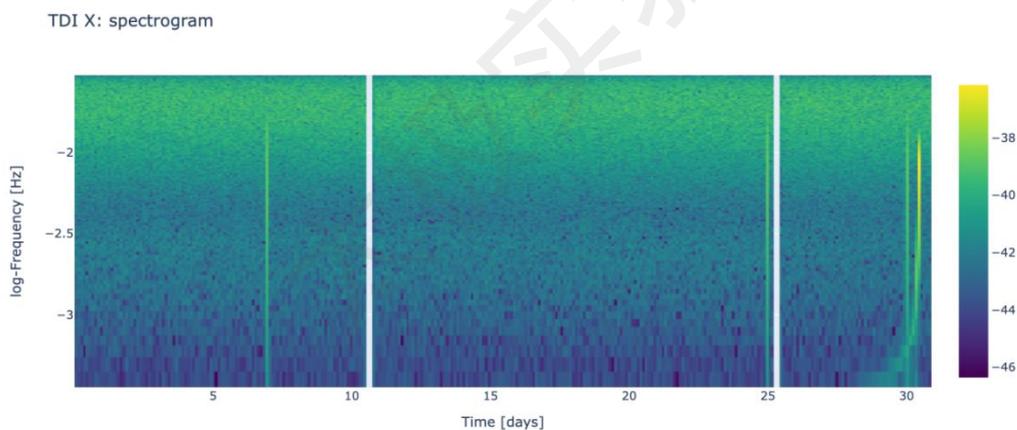
- 关键词：Python、小波变换、时频分析、数据可视化
- 任务难度：☑□□□□
- 任务描述：

1. 从 HDF5 文件中读取 LISA Data Challenge 引力波数据（数据中的 NaN 值可设为 0）；

2. 对数据分别进行(1)Wilson-Daubechies-Meyer (WDM) 小波变换和(2)分数阶傅里叶变换 (FRFT) ；

3. 利用 matplotlib 中的 pcolormesh 等工具对变换结果绘图, 其中(1)WDM 的结果适当调节小波变换参数, 尽量接近下图效果 (不需要完全一致) (2)FRFT 的结果也要足够清晰能看到信号的特征。

4. 将跑通的 Python 上传至自己账号下的 github 仓库 repo 中, 报名回复该仓库 repo 链接即为上传作业成功。



HDF5 下载链接: <https://pan.baidu.com/s/1bfAflgvi9Gfhl3SIgrdUw> 提取码: ucas

通过网盘分享的文件: LDC2_spritz_mbhb1_training_v1.h5

WDM 小波代码链接: [XGI-MSU/WDMWaveletTransforms \(github.com\)](https://github.com/XGI-MSU/WDMWaveletTransforms)

可能用到的函数: transform_wavelet_time 或 transform_wavelet_freq_time

参考资料: LISA Data Challenge-2b 数据说明 [LISA Data Challenge \(in2p3.fr\)](https://in2p3.fr)

FRFT 的变换代码请自行从 github 上寻找合适的 repo 来进行

子任务二：Taiji Data Challenge 大质量双黑洞信号参数估计

- 关键词：Python、大质量双黑洞、贝叶斯统计推断
- 任务难度：☑☑☑□□
- 任务描述：
 1. 根据代码仓库 Triangle-BBH 的 README 指示，安装相应的运行环境；
 2. 运行并理解 Examples 文件夹下，编号为 4 的示例，理解空间引力波探测数据形式、傅立叶变换、大质量双黑洞并合引力波信号、贝叶斯统计推断等基础知识（提示：可借助文献索引或 AI 工具）；
 3. 将用于参数估计的数据时长扩展为 5 天（coalescence_time 前 4 天，coalescence_time 后 1 天），重新运行程序，观察结果有何变化；
 4. 将修改后并跑通的 notebook（包含运行结果图）上传至自己账号下的 github 仓库 repo 中，报名回复该仓库 repo 链接即为上传作业成功。

代码仓库链接：<https://github.com/TriangleDataCenter/Triangle-BBH>

(六) LISA 探测器响应灵敏度

- 关键词: LISA、Python、TDI
- 任务难度:
- 任务描述:

LISA 引力波探测器包含三颗卫星, 假设三颗卫星构成边长为 2.5×10^9 m 的等边三角形, 卫星的位置可视为固定, 按照文献 <https://arxiv.org/pdf/2108.01167> 中 P8 到 P14 页给出的方法 (数值或半解析), 计算 TDI-X2.0 的平均引力波响应 (Average GW Response) 和噪声功率谱密度 (Power Spectral Density), 计算并绘制 LISA TDI-X2.0 通道的灵敏度曲线。

进一步参考文献 <https://arxiv.org/abs/1803.01944>, 加入 LISA 的前景混淆噪声 (confusion noise), 计算并绘制包含前景混淆噪声的 LISA 总灵敏度曲线, 用 python 语言画图。报名时可以提交: 推导过程 (手写或 word 或 latex) + python 代码 (.py 文件或 .ipynb 文件) + 提交运行结果 (灵敏度曲线图)

太极实验室

(七) 基于 Docker+GRChombo 的双黑洞数值相对论模拟

- 关键词: 数值相对论、GRChombo、Docker、科学上网
- 任务难度:
- 任务描述:

基于 Docker 容器化技术与数值相对论软件 GRChombo, 构建等质量双黑洞正面碰撞的数值模拟模型。具体参数配置如下: 黑洞质量均设为 0.5, 初始间距为 20, 初始速度可自行视情况而定, 采用正面碰撞 (head-on collision) 模型。

- 关键提示:
 - ✧ 在 GRChombo 官方网站 (<https://www.grchombo.org/>) 阅读关于 GRChombo 的基础内容。
 - ✧ 参考视频 (https://www.youtube.com/watch?v=6xOZxA_Oz-I) 中关于在 Docker 中运行 GRChombo 的具体方法 (注意: 该方法仅适用于 Intel 芯片)。
 - ✧ 本题的核心任务在于修改 GRChombo/Examples/BinaryBH 目录下 params.txt 文件中的参数。这不仅涉及物理参数 (如黑洞质量、初始间距等), 还涉及到演化中分辨率的合理选取。该任务主要考查学生的自主学习能力与代码阅读能力, 并不严格要求结果的精确性。因此, 学生应根据自身电脑的配置, 合理选取演化精度与演化步长。
 - ✧ GRChombo 每一步演化的结果均保存在 hdf5 格式文件中, 这些文件可以通过 Visit 软件进行查看 (<https://visit-dav.github.io/visit-website/>)。
- 结果提交
 1. 实际运行使用的参数文件 params.txt (请勿提交默认文件)。
 2. 提交/GRChombo/Examples/BinaryBH/data 文件夹中的全部内容。

(八) 利用个人电脑或 HPC 编译 Chombo 与 GRChombo

- 关键词: 环境搭建、Chombo/GRChombo 代码编译
- 任务难度:
- 任务描述:
 1. 搭建 GRChombo 所需的计算环境。
 2. 编译 Chombo 与 GRChombo 中的双黑洞代码。
- 关键提示:
 - 在 GRChombo 官方网站 (<https://www.grchombo.org/>) 阅读关于 GRChombo 的基础内容。
 - 仔细阅读 GRChombo 编译所需库的详细说明：<https://github.com/GRTLCollaboration/GRChombo/wiki>。
 - 本题的核心任务在于修改 Make.defs.local 文件，以适配个人电脑的配置。
- 结果提交
 1. 实际编译过程中使用的 Make.defs.local 文件（请勿提交默认文件）。
 2. 提交 GRChombo/Examples/BinaryBH 目录下的可执行文件（文件名类似于 Main_BinaryBH3d.Linux.64.CC.ftn.OPTHIGH.MPI.OPENMPCC.ex）。

(九) 利用个人电脑或 HPC 进行双黑洞模拟并提取引力波

- 关键词：双黑洞演化、GRChombo、引力波
- 任务难度：☑☑☑☑□
- 任务描述：
 1. 在个人电脑或 HPC 上编译基于 GRChombo 的双黑洞代码。
 2. 构建等质量双黑洞正面碰撞的数值模拟模型。具体参数配置如下：双黑洞质量均设为 0.5，初始间距为 20，初始速度可自行视情况而定，采用正面碰撞 (head-on collision) 模型。
 3. 提取引力波模式 (mode) 为 $l=2, m=0$ 和 $l=2, m=2$ 的信号并作图。
 4. 将整个演化过程制作成视频 (视频中的演化变量可任选)。
- 关键提示：
 - 在 GRChombo 官方网站 (<https://www.grchombo.org/>) 阅读关于 GRChombo 的基础内容。
 - 仔细阅读 GRChombo 编译所需库的详细说明：<https://github.com/GRTLCollaboration/GRChombo/wiki>。
 - 编译 Chombo 以及 GRChombo/Examples/BinaryBH 目录下的代码。
 - 按照要求设置恰当的参数，并运行可执行文件。
 - GRChombo 已内置引力波提取功能，关键提示词：“WeylExtraction”。
 - GRChombo 每一步演化的结果均保存在 hdf5 格式文件中，这些文件可以通过 Visit 软件进行查看 (<https://visit-dav.github.io/visit-website/>)。
- 结果提交
 1. 引力波模式 $l=2, m=0$ 和 $l=2, m=2$ 的信号提取结果图。
 2. 利用 hdf5 生成的演化过程视频。
 3. 提交 GRChombo/Examples/BinaryBH/data 文件夹中的全部内容。

(十) 基于 HDF5 数据的提取、差分处理及回写实现

- 关键词：数据提取、数据处理、数据回写
- 任务难度：☑☑☑☑☑
- 任务描述：
 1. 在个人电脑或 HPC 平台上编译基于 GRChombo 的双黑洞程序。
 2. 构建等质量双黑洞正面碰撞的数值模拟模型。具体参数设置如下：两个黑洞的质量均取 0.5，初始间距设为 20，初始速度可根据实际情况自行选取，采用正面碰撞 (head-on collision) 模型，空间分辨率取 $dx=2$
 3. 提取输出的 HDF5 数据，利用线性差分方法将分辨率由 $dx=2$ 加密到 $dx=1$ ，并将处理后的结果写回 HDF5 文件中，随后在 GRChombo 中继续进行演化。
- 关键提示：
 - 在 GRChombo 官方网站 (<https://www.grchombo.org/>) 阅读关于 GRChombo 的基础内容。
 - 仔细阅读 GRChombo 编译所需库的详细说明：<https://github.com/GRTLCollaboration/GRChombo/wiki>。
 - 编译 Chombo 以及 GRChombo/Examples/BinaryBH 目录下的代码。
 - 按照要求设置恰当的参数，并运行可执行文件。
 - GRChombo 每一步演化的结果均保存在 hdf5 格式文件中，这些文件可以通过 Visit 软件进行查看 (<https://visit-dav.github.io/visit-website/>)。
 - 可使用 Python 中的 h5py 包完成 HDF5 文件的读取与写回。
- 结果提交
 1. 修改前后的 HDF5 文件。
 2. 相关代码，包括：HDF5 数据读入、差分处理（采用线性差分即可）以及结果写回。
 3. 能够证明修改后的 HDF5 文件仍可继续用于演化的可靠证据，例如演化过程视频、约束量输出的 .dat 文件等。

(十一) 基于 Einstein Toolkit 的双中子星模拟和引力波提取

- 关键词: 双中子星演化、Einstein Toolkit、Lorene、引力波
- 任务难度:
- 任务描述:
 1. 在个人电脑或 HPC 上编译 Einstein Toolkit 和 Lorene 的代码。
 2. 构建等质量双中子星绕转的数值模拟模型。具体参数配置如下: 双中子星质量均设为 1.2 倍太阳质量, 初始间距为 45km, 物态方程使用多方模型。使用 Lorene 生成双中子星所需要的初值, 使用 Einstein Toolkit 完成双中子星合并的动态演化过程。
 3. 提取引力波模式 (mode) 为 $l=2, m=0$ 和 $l=2, m=2$ 的信号并作图。
 4. 将整个演化过程制作成视频 (视频中的演化变量可任选)。
- 关键提示:
 - 在 Einstein Toolkit 官方网站 (<https://einsteintoolkit.org>) 阅读关于 Einstein Toolkit 的基础内容。特别地, 在 Einstein Toolkit 网页中给出了双中子星例子的一个描述 (<https://einsteintoolkit.org/gallery/bns/index.html>)
 - 在 Einstein Toolkit 的官方网站下载对应的代码: <https://einsteintoolkit.org/download.html>)。
 - 在 Einstein Toolkit 中, 每一个模块成为 thorn, 整个过程中所需要的所有 thorn 成为 thornlist。请确保代码中包含了编译时所需要的 thornlist 的文件, 以.th 结尾 (例如 `einsteintoolkit.th` 或 `bns.th`)。关于每个 thorn 的介绍, 可参考 Einstein Toolkit 官网的介绍: (<http://einsteintoolkit.org/documentation/ThornGuide.php>)。根据双中子星例子的提示下载双中子星模拟中需要的 thornlist (参考 <https://einsteintoolkit.org/gallery/bns/index.html>), 注意该 thornlist 必须合补充道原来的 thornlist 中
 - 编译 Einstein Toolkit 软件, 安装 visit 和 kuibit 工具进行数据分析 (可以参考 arXiv 论文 Nicholas Choustikov, The Einstein Toolkit: A Student's Guide, arXiv:2011.13314; 官方安装说明: https://docs.einsteintoolkit.org/et-docs/Setup_Tutorial)。
 - 编译 Lorene 代码, 并用其生成所需要的双中子星初值 (关于 Lorene 代码的介绍, 可参考 <https://lorene.obspm.fr/>)。
 - 按照要求设置恰当的参数, 并运行可执行文件以参考 arXiv 论文

Nicholas Choustikov, The Einstein Toolkit: A Student's Guide, arXiv:2011.13314; 以及网络上公开的视频: B 站视频, 数值相对论与 Einstein Toolkit 操作基础培训 2-Einstein Toolkit 的安装与数值相对论基础; 网上公开的 Einstein Toolkit 学习笔记 (<https://www.cnblogs.com/zouastro/p/16415564.html>) 。

■ Einstein Toolkit 已内置引力波提取功能, 可利用 kuibit 工具或 mathematica 工具提取引力波 (kuibit 为官方开发的基于 python 的数据处理工具包, 可参考网页 <https://sbozzolo.github.io/kuibit/>) 。

■ Einstein Toolkit 每一步演化的结果均保存在 hdf5 格式文件中, 这些文件可以通过 Visit 软件进行查看 (<https://visit-dav.github.io/visit-website/>) 。

➤ 结果提交

1. 引力波模式 $l=2, m=0$ 和 $l=2, m=2$ 的信号提取结果图。
2. 利用 hdf5 生成的演化过程视频。

(十二) 基于 AMSS-NCKU 的相对论三体数值模拟

➤ 关键词: 相对论三体模拟、AMSS-NCKU

➤ 任务难度:

➤ 任务描述:

1. 在个人电脑或 HPC 上编译 AMSS-NCKU 代码。

2. 构建等质量的三个黑洞的数值模拟模型。具体参数配置如下: 所有黑洞的质量均设为 $M=1/3$, 初始间距为 $d=10M$, 初始相对静止。

3. 使用 AMSS-NCKU 生成三体所需要的初值 (初值可以按照 Lousto 或曹周键老师的解析结果给定), 并完成三个黑洞合并的动态演化过程 (关于 AMSS-NCKU 的代码和介绍可参考 <https://github.com/xiaoqu0000/NR-amssncku>)。

4. 获取黑洞轨迹并做图, 提取引力波模式 (mode) 为 $l=2, m=0$ 和 $l=2, m=2$ 的信号并作图。

➤ 参考资料:

1. Zhou-jian Cao, Hwei-Jang Yo, and Jui-Ping Yu, A Reinvestigation of Moving Punctured Black Holes with a New Code, Phys.Rev.D 78, 124011 (2008).

2. Chen-Kai Qiao, Yi Zheng, and Zhou-Jian Cao, A User-Friendly Python Interface for the Numerical Relativity Code AMSS-NCKU, Astron.Comput. 55, 101093 (2026).

3. Pablo Galaviz, Bernd Bruegmann, and Zhoujian Cao, Numerical evolution of multiple black holes with accurate initial data, Phys.Rev.D 82, 024005 (2010).

4. Carlos O. Lousto and Yosef Zlochower, Foundations of multiple black hole evolutions, Phys.Rev.D 77, 024034 (2008).

5. Manuela Campanelli, Carlos O. Lousto, and Yosef Zlochower, Close encounters of three black holes, Phys.Rev.D 77, 101501 (2008).

➤ 结果提交

1. 黑洞位置变化轨迹

3. 引力波模式 $l=2, m=0$ 和 $l=2, m=2$ 的信号提取结果图。

(十三) 多尺度 PINN 的实现与二维推广

- 关键词: 数值相对论, PINN, 算子学习, 偏微分方程, 多尺度建模
- 任务描述: 本题目基于多尺度 Physics-Informed Neural Networks (Multi-scale PINN) 方法, 探索神经网络在求解多尺度偏微分方程中的能力。
- 请基于给定参考代码 (1D Poisson 方程), 完成如下任务:

【基础任务: 方法复现】任务难度:

1. 复现一维 Poisson 方程的 PINN 求解结果;
2. 实现并对比以下三种模型:
 - a. 普通神经网络 (NN)
 - b. Fourier Feature 网络 (FF)
 - c. Multi-scale Fourier Feature 网络 (mFF)
3. 在单一频率情形下比较三种方法的表现。

【进阶任务: 多频率分析+方法分析】任务难度:

4. 构造如下多频率源项 (示例): $f(x) = \sin(2\pi x) + 0.5 \sin(10\pi x)$
5. 在该多频率问题下, 比较 NN / FF / mFF 的表现差异;
6. 分析不同模型对低频与高频成分的拟合能力。
7. 分析为什么标准 PINN 难以拟合高频解;
8. 解释 Fourier feature 与 multi-scale embedding 的作用机制;
9. 从实验角度验证 multi-scale embedding 的优势。

【挑战任务: 二维推广】

10. 将问题推广至二维 Poisson 方程: $-\Delta u(x,y) = f(x,y), (x,y) \in \Omega$
11. 构造二维多频率源项 (如不同方向/不同尺度组合);
12. 实现对应的 PINN / FF / mFF 方法;
13. 分析不同方法在二维多尺度问题中的性能差异。

【可选加分】

14. 系统研究不同频率组合 (如频率间隔、强度比例) 对模型表现的影响;
15. 讨论模型在更复杂 PDE (如非线性问题) 中的适用性;
16. 尝试解释该方法与 Neural Tangent Kernel (NTK) 的关系。

- 提交内容包括:
 - 方法说明 (≤ 4 页) + 核心代码 + 数值实验结果与分析
 - 上述内容上传至自己账号下的 github 仓库 repo 中, 报名回复该仓库 repo 链接即为上传作业成功。

- 参考代码:

https://github.com/xgxgnpu/Physics-informed-vibe-coding/blob/main/MultiScalePINN_jax/case2_poisson1d/poisson1d_multiscale_pinn.py

(十四) 基于新闻文本的股价波动预测与因果关系初探

- 关键词: Python, 时间序列分析, 自然语言处理, 情感分析, 回归分析, 因果推断, 机器学习, 知识图谱
- 任务难度:
- 任务描述: 在金融市场中, 新闻信息被认为是影响资产价格的重要因素之一。本题目旨在探索新闻文本与股票价格波动之间的关系。
 1. 从公开数据源获取某一只股票或指数的历史价格数据 (如日频数据);
 2. 收集对应时间段内的相关新闻数据 (可使用公开新闻 API 或爬虫);
 3. 对新闻文本进行处理 (如情感分析、关键词提取或事件分类);
 4. 构造一个简单指标 (例如每日情绪指数或新闻强度);
 5. 分析该指标与股票收益或波动之间的关系;
 6. 尝试回答: 新闻是否对市场具有“预测能力”?
- 鼓励但不强制:
 - 使用统计检验方法 (如 Granger 因果检验)
 - 对比不同文本处理方法 (词典法 vs. LLM 法) 的效果差异;
 - 分析情感信号的衰减效应 (如使用指数衰减模型处理非交易日新闻);
 - 尝试构建简单的“新闻事件-股价反应”知识图谱或关系图;
 - 区分不同来源 (官方公告 vs. 媒体报道)、不同类型 (公司层面 vs. 行业政策) 新闻的影响差异;
 - 讨论“市场有效性”在不同情境下的体现;
 - 识别并分析可能的“伪因果”案例。
- 最终提交内容包括:
 - 方法说明 (不超过 4 页): 清晰阐述数据来源、处理流程、分析模型、主要结果及其经济学/金融学解释, 重点说明对因果性的思考和讨论。
 - 核心代码和可复现结果: 提供完整、注释良好的 Python 代码 (Jupyter Notebook 格式为佳), 以及关键结果的图表和数据。
 - 上述内容上传至自己账号下的 github 仓库 repo 中, 报名回复该仓库 repo 链接即为上传作业成功。

(本题不要求复杂模型实现, 更关注对数据的理解能力、合理建模以及对“相关性”与因果性差异”的认识。)
- Hint 1: 数据来源
 - 股票: Yahoo Finance、Tushare、Akshare
 - 新闻: NewsAPI、GDELT (强烈推荐)、Fred、东方财富、财联社、金十数据等国内新闻
- Hint 2: 最简单可行方案 (baseline): 用词典法做情绪分析; 每天求平均

sentiment; 然后做回归: $\text{return}_t \sim \text{sentiment}_{\{t-1\}}$

- Hint 3: 进阶方向: 滞后效应 (lag 1~5 days) ; 高频 vs 低频; 不同类型新闻 (政策 vs 公司) ; 能识别伪因果; 能讨论 “信息延迟” “市场有效性” “confounding variables”
- 可参考的资料:
 - ✧ 《Advances in Financial Machine Learning》 by Marcos López de Prado
 - ✧ 《Python for Finance》
 - ✧ Tetlock (2007) “Giving Content to Investor Sentiment” (新闻情绪→市场影响的经典工作),
https://www.uts.edu.au/globalassets/sites/default/files/adg_cons2015_tetlock-journal-of-finance-2007.pdf
 - ✧ statsmodels (因果检验) : <https://github.com/statsmodels/statsmodels>
豆瓣电影短评情绪分析 :
https://github.com/iphysresearch/TOP250movie_douban ;
<https://arxiv.org/abs/2411.13965>
 - ✧ 《FNSPID: A Comprehensive Financial News Dataset in Time Series (2024) , 了解大规模新闻-股价对齐数据集的价值与构建 ,
<https://arxiv.org/abs/2402.06698>
 - ✧ Sentiment trading with large language models (2024), 了解 LLM 在金融情感分析中的前沿应用与卓越表现, <https://arxiv.org/abs/2412.19245>
 - ✧ News-Aware Direct Reinforcement Trading for Financial Markets (2025), 了解如何将新闻情感特征直接整合至机器学习/强化学习模型
<https://arxiv.org/abs/2510.19173>
 - ✧ JP Morgan: 融合 NLP 情绪与 LLM 洞察的配对交易策略 (2026), 了解业界如何系统化运用新闻情绪与 LLM 事件分类于量化策略
<https://mp.weixin.qq.com/s/8Q06yZxG1Lgu4JWwsCNk6A>
 - ✧ HuggingFace (NLP) ;
 - ✧ statsmodels (因果检验) ;
 - ✧ 豆瓣电影短评情绪分析:
https://github.com/iphysresearch/TOP250movie_douban
 - ✧ <https://arxiv.org/abs/2411.13965>