

项目说明文档

喷注聚类问题的量子计算优化

参赛队伍 apmcmlz2500206fj

2025 年 12 月

目录

1	项目概述	2
2	数据文件	2
2.1	输入数据	2
2.2	配置文件	2
2.3	结果文件	2
3	Python 脚本	3
3.1	数据生成	3
3.2	QUBO 模型构建	3
3.3	基准求解器	3
3.4	开悟 SDK 求解器	4
3.5	分析和可视化	4
4	输出目录	5
4.1	结果和图表	5
5	外部依赖	5
5.1	开悟 SDK	5
6	工作流程摘要	6
6.1	数据生成	6
6.2	基准求解	6
6.3	量子硬件求解	6
6.4	分析和可视化	6
7	文件统计	6
8	关键技术	7

1 项目概述

本项目将量子计算技术应用于高能物理中的喷注聚类问题。我们将问题建模为 QUBO（二次无约束二元优化）模型，并使用经典算法和开悟 SDK 量子求解器进行求解。解决方案在真实量子硬件（玻色量子 CPQC-550）上进行了验证。

2 数据文件

2.1 输入数据

- `events_data.npz`

用途：包含 50 个模拟粒子碰撞事件的压缩 NumPy 二进制文件。

内容：粒子属性 (p_T 、 η 、 ϕ) 和预计算的 ΔR 距离矩阵。

- `events_csv/events_table.csv`

用途：CSV 格式的粒子事件数据（682 行，50 个事件）。

字段：`event_id`（事件编号）、`particle_id`（粒子编号）、`pT`（横向动量）、`eta`（赝快度）、`phi`（方位角）。

- `events_csv/generation_params.csv`

用途：数据生成参数，用于确保可重现性。

内容：随机种子（2025）、粒子数范围（12-15）、 p_T 范围、 η 范围。

2.2 配置文件

- `real_event0_lam200_mu30.json`
`real_event0_lam200_mu100_tmax100.json`
`real_event0_lam300_mu60_tmax100.json`
`real_event0_lam300_mu60_tmax100_oh1.json`
`real_event1_lam300_mu60_tmax100_oh1.json`
`real_event2_lam300_mu60_tmax10_oh1.json`

用途：QUBO 模型参数和量子求解器结果的 JSON 配置文件。

参数： λ （约束惩罚系数）、 μ （平衡系数）、解向量。

2.3 结果文件

- `results_baseline_sa.csv`

用途：经典模拟退火求解器在所有 50 个事件上的结果。

指标：最优能量、平均能量、能量标准差、求解时间、可行性。

- `results_baseline_random.csv`

用途：随机搜索基准方法在所有 50 个事件上的结果。

- `results_kaiwu_sa.csv`

用途：开悟 SDK 模拟退火求解器在所有 50 个事件上的结果。

- `results_baseline.csv`

用途：K-means 聚类基准方法的结果。

- `kaiwu_runs/*.pkl`

用途：包含开悟求解器详细输出的 Python pickle 对象。

内容：完整求解历史、能量轨迹、求解器参数。

3 Python 脚本

3.1 数据生成

- `generate_data.py`

用途：生成用于测试的合成粒子碰撞事件。

输出：`events_data.npz` 和 `events_csv/`目录。

主要函数：

- `generate_toy_events()`: 创建随机粒子事件
- `deltaR_matrix()`: 计算 ΔR 距离矩阵
- `save_events_npz()`: 保存为 NumPy 格式
- `save_events_csv()`: 保存为 CSV 格式

3.2 QUBO 模型构建

- `qubo_model.py`

用途：基础 QUBO 模型构建（已弃用，请使用 `real_qubo_model.py`）。

函数：`build_qubo_matrix_k2()`: 为 K=2 聚类构建 QUBO 矩阵。

- `real_qubo_model.py`

用途：生产环境的 QUBO 模型，包含约束处理和平衡项。

主要函数：

- `build_qubo_matrix_k2()`: 构建带约束的 QUBO 矩阵
- 参数： λ （约束）、 μ （平衡）、 w （距离权重）

3.3 基准求解器

- `baseline_random.py`

用途：随机搜索基准求解器。

方法：随机采样可行解（50,000 个样本）。

输出：`results_baseline_random.csv`。

- `baseline_sa.py`

用途：经典模拟退火求解器。

方法：自定义 SA 实现，带自适应温度调度。

参数: 10,000 步, 每个事件运行 10 次。

输出: results_baseline_sa.csv。

- baseline_kmeans.py

用途: K-means 聚类基准 (非 QUBO 方法)。

方法: 在 (η, ϕ) 空间中聚类粒子。

输出: results_baseline.csv。

3.4 开悟 SDK 求解器

- baseline_kaiwu_sa.py

用途: 开悟 SDK 模拟退火求解器 (模拟模式)。

方法: 使用开悟的量子启发式 SA 求解器。

参数: 10,000 次扫描, 每个事件读取 10 次。

输出: results_kaiwu_sa.csv。

- kaiwu_config.py

用途: 开悟 SDK 身份验证配置模块。

内容: API 凭据、平台端点。

- run_real_event0_cpqc550.py

run_real_event1_cpqc550.py

run_real_event2_cpqc550.py

用途: 向真实量子硬件 (CPQC-550) 提交任务的脚本。

方法: 连接到玻色量子云平台。

输出: 包含量子求解器结果的 JSON 文件。

3.5 分析和可视化

- analyze_results.py

用途: 对所有事件的求解器性能进行统计分析。

生成: 汇总统计、性能比较。

- compare_solvers.py

用途: 不同求解器的多维比较。

指标: 能量质量、求解时间、可行性、稳定性。

- plot_energy_vs_time_sa_vs_kaiwu.py

用途: 生成 SA 和开悟求解器的能量-时间对比图。

输出: energy_vs_time_sa_vs_kaiwu.png/pdf。

- plot_jet_efficiency_hist.py

用途: 绘制喷注聚类效率直方图。

- `make_paper_plots.py`

用途：生成所有出版质量图形的主脚本。

输出：文中图表/目录中的所有图表（PNG 和 PDF 格式）。

4 输出目录

4.1 结果和图表

- 文中图表/

用途：包含所有出版图表和表格的目录。

内容：

- 能量比较图（直方图、箱线图、散点图）
- 求解时间直方图
- 帕累托前沿（能量 vs 时间）
- 方法说明图
- 结果比较表（CSV 格式）

- `kaiwu_runs/`

用途：存储开悟求解器输出文件的目录（.pkl 格式）。

命名：WuYueCup_real_event{idx}_N{N}_lam{lam}_mu{mu}_..._results.pkl。

- 真机使用截图/

用途：玻色量子云平台的截图。

内容：

- 任务列表.png: 任务列表截图
- 用量总览.png: 用量总览截图

5 外部依赖

5.1 开悟 SDK

- `kaiwu-1.3.0-cp310-none-any.whl`

用途：适用于 Python 3.10 的开悟 SDK wheel 包。

安装：`pip install kaiwu-1.3.0-cp310-none-any.whl`。

- `kaiwu-pytorch-plugin/`

用途：开悟 PyTorch 插件源代码和示例。

内容：文档、示例 notebook、测试脚本。

6 工作流程摘要

6.1 数据生成

1. 运行 `generate_data.py` → 创建 `events_data.npz` 和 `events_csv/`

6.2 基准求解

1. 运行 `baseline_random.py` → `results_baseline_random.csv`
2. 运行 `baseline_sa.py` → `results_baseline_sa.csv`
3. 运行 `baseline_kmeans.py` → `results_baseline.csv`
4. 运行 `baseline_kaiwu_sa.py` → `results_kaiwu_sa.csv`

6.3 量子硬件求解

1. 配置 `kaiwu_config.py` 中的凭据
2. 运行 `run_real_event0_cpqc550.py` → JSON 输出
3. 运行 `run_real_event1_cpqc550.py` → JSON 输出
4. 运行 `run_real_event2_cpqc550.py` → JSON 输出

6.4 分析和可视化

1. 运行 `analyze_results.py` → 统计摘要
2. 运行 `compare_solvers.py` → 比较指标
3. 运行 `make_paper_plots.py` → 文中图表/中的所有图形

7 文件统计

类别	数量
Python 脚本 (.py)	13
数据文件 (NPZ, CSV)	3
配置文件 (JSON)	6
结果文件 (CSV)	4
求解器输出 (PKL)	5
图表 (PNG, PDF)	31
表格 (CSV)	3

表 1: 项目文件统计

8 关键技术

- 编程语言： Python 3.10+
- 核心库： NumPy、Pandas、Matplotlib
- 优化方法： QUBO、模拟退火、K-means
- 量子平台： 玻色量子 CPQC-550（550 量子比特）
- SDK： 开悟 SDK v1.3.0