

Codex-b/beta 实验上的数据处理流程 与径迹重建

报告人：任博涵

中山大学
2026 年超级陶粲装置研讨会

汇报主线

1 物理动机与实验定位

2 数据处理流程

3 径迹与顶点重建

4 总结

核心任务

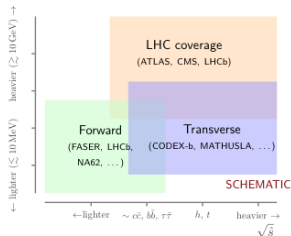
将近似无磁场、低占有率环境下 RPC 条带读出的通道编号与时间戳，构建为三维 Hits、径迹和可用于 LLP 搜索的位移多径迹顶点候选

CODEX-b: 用低本底衰变体积寻找 LLP

LLP 信号的可观测特征

长寿命粒子可能飞出主探测器后才衰变, 信号形态可以概括为:

- 从 LHCb 相互作用点 IP8 方向飞来, 但在远离 IP8 处产生二级顶点
- 衰变体积内出现多条来源于同一顶点的径迹
- 衰变点之前没有对应的入射粒子信号



CODEX-b 与 FASER 等在几何位置和寿命覆盖上互补, 图取自 Fig. 1 [1]

CODEX-b: 在 IP8 横向构建低本底衰变体积

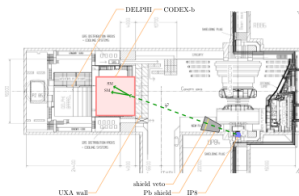


FIG. 2: Top view layout of the LHCb experimental cavern

LHCb IP8 与 CODEX-b 位置示意，取自综合报告
Fig. 2 [1]

CODEX-b 的核心取舍

不追求强磁场动量测量，而是利用**屏蔽、veto、时间、指向性和顶点重建**把普通碰撞本底压低到可控水平

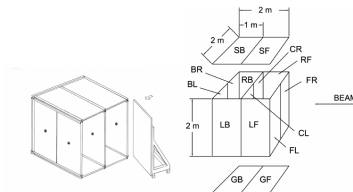
基准设计

- 距 IP8 约 25 m 的 10^3 m^3 空气体积
- RPC 面测量空间点、时间与径迹方向
- LHCb 事件信息提供产生端标签

近零本底策略

- UXA 区域混凝土墙提供被动屏蔽
- 主动 veto 抑制 μ 子及屏蔽层级粒子本底
- 时间、IP8 来源相容性与位移顶点进一步筛选

CODEX- β : 验证 RPC 硬件、DAQ 与软件设计



CODEX- β 布局, 取自 Technical Design Report [2]

示范器规模

- $2 \times 2 \times 2 \text{ m}^3$ 立方体
- 14 个 RPC 模块, 共 42 个 singlet
- 模块由 3 个 singlet 构成 triplet, 提供冗余空间和时间测量

物理与软件任务

- 1 实测 UX85A 区域的 μ 子与中性本底
- 2 验证 RPC 硬件响应、LHCb DAQ、软件触发和时间同步的集成
- 3 用 K_L^0 衰变标定时序与顶点重建
- 4 闭合 simulation、digitization、reconstruction 与 analysis

从 RPC 测量到 LLP 候选：重建对象与物理约束

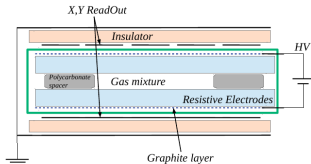


FIG. 3. Schematic of the structure of a singlet RPC, taken from [13].

RPC singlet 结构，取自 Technical Design Report [2]

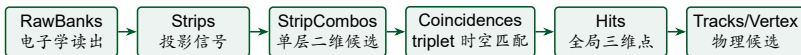
单次穿越的测量层次

- 正交 ϕ/η strips 给出二维位置
- 三个 singlet 的时空符合形成 module 级测量
- 多个 module 级 Hits 形成近似直线径迹
- 多条径迹的共同起点形成位移顶点

算法约束来自物理拓扑

近似无磁场使局部径迹可按直线处理；正式重建仍需保留 surface ID、时间、测量协方差，才能评价效率、假率与顶点分辨率

COBRA 数据流：I/O 调度与物理对象构建相互分离



软件角色	主要职责	物理含义
COBRAConverter	读取、依次调用processor、输出	不定义重建物理
COBRAEvent	保存同一事件的逐级数据输出	支持回放与交叉检查
reco processors	解码、聚类、时空匹配、几何映射	构建可解释的测量对象
tracking processor	pattern recognition、fit、vertex	尚处于原型阶段

设计原则

每一级均保留可写入 ROOT 的中间量，使效率损失或空间偏差能够定位到“电子学解释、组合选择、几何转换、径迹拟合”中的具体环节。

StripDecoder: 电子学读出的 strip 语义恢复

解码目标

StripDecoder 是 RawBanks 之后的第一层重建：把电子学记录整理成带有探测器语义的一维 strip 测量。

- 展开读出通道与时间戳
- 恢复 singlet、本地 strip 和投影方向
- 为 cluster、combo 和 coincidence 提供统一输入

标准输出

`Strips`: $[dct, singlet, strip, time, type]$ 。
其中 *type* 标记正交方向，*time* 保存该 strip 的读出时间。

实现路径

对每条 RawBank，代码将两路时间列展开为 strip 候选；再用 offset 把通道号拼成 singlet 内索引：

$$\begin{aligned}s &= \text{channel} + \text{offset}, \\ \text{singlet} &= \lfloor s/96 \rfloor, \\ \text{local} &= s \bmod 96.\end{aligned}$$

最后按本地通道区间确定投影类型，并写出 `Strips`。

StripComboBuilder: 一维电荷簇与二维投影组合

每个投影独立聚类

对 ϕ 与 η 两组 **strip** 分别处理。按 $(dct, singlet, strip, time)$ 排序, 相邻信号满足

$$\begin{cases} strip_{i+1} = strip_i + 1, \\ |t_{i+1} - t_i| \leq n_t \end{cases}$$

时归入同一 **cluster**, 并以

$$\bar{s} = \text{round} \left(\frac{1}{N} \sum_i s_i \right),$$

$$\bar{t} = \frac{1}{N} \sum_i t_i$$

表示该投影上的条带簇。

同一 singlet 内二维配对

只在相同 $(dct, singlet)$ 内组合 ϕ 簇与 η 簇, 并要求两组时间相容:

$$|\bar{t}_\phi - \bar{t}_\eta| < n_t.$$

输出单个 RPC singlet 上的二维候选:

$$[dct, singlet, \phi, \eta, t_\phi, t_\eta].$$

本层职责

`StripComboBuilder` 负责把单投影 **strip** 信号整理为单层二维候选; **triplet** 层间筛选与全局 Hit 生成由后续处理器完成。

StripCoincidenceBuilder: 利用 triplet 的时空冗余抑制组合本底

兼容性定义

对每个 singlet 候选定义

$$t_i = \frac{t_{\phi,i} + t_{\eta,i}}{2}.$$

新成员必须满足

$$|\phi_i - \bar{\phi}| + |\eta_i - \bar{\eta}| \leq n_s,$$
$$|t_i - \bar{t}| \leq n_t.$$

输出同时保存

$$\bar{\phi}, \bar{\eta}, \bar{t}, \Delta t, N.$$

- 1 构造 singlet 0-1 doubles
- 2 尝试加入 singlet 2 形成 triples
- 3 剩余候选可形成 1-2 或 0-2 doubles
- 4 删除共享同一 StripCombo 的重复 coincidence, 优先保留 triple 候选

后续验证项

后续需在带 truth 关联的样本上扫描 n_s, n_t , 量化 coincidence 保留效率、随机/错误组合率, 以及 double/triple 候选组成。

HitsGenerator: 从离散 strip 坐标进入探测器全局坐标

局部测量模型

singlet 尺寸与 strip 数定义 pitch:

$$\begin{aligned}p_\eta &= (y_{\max} - y_{\min})/32, \\p_\phi &= (z_{\max} - z_{\min})/64, \\y &= y_{\min} + (\eta + 0.5)p_\eta, \\z &= z_{\min} + (\phi + 0.5)p_\phi.\end{aligned}$$

层级坐标变换

$$\mathbf{x}_g = R_D[R_M(R_S\mathbf{x}_l + \mathbf{o}_S) + \mathbf{o}_M] + \mathbf{o}_D.$$

- YAML 保存 14 个 module 的位姿、singlet 偏移与有效尺寸
- coincidence 模式使用 $(\bar{\phi}, \bar{\eta})$, 在中间 singlet 参考面生成一个 module 级 Hit
- 输出

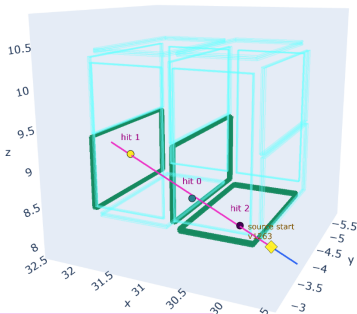
$$[x, y, z, \bar{t}, \Delta t]$$

数据压缩

三个 singlet 被压缩成一个 Hit, 适合快速重建。若要利用层内斜率和非均匀分辨率, 应保留三个测量面及其协方差。

完整模拟重建示例

- incoming parent particle
- RawBank truth-source segment
- ◆ truth process vertices
- Hits
- As Hit signal sources



RawBank truth source, not fitted track
src 3692: pid=-211, RawBanks=30, modules=CL:6, BL:8, LF:16
Hit straight-line residual max: 0.053 m
Truth process: parent [3690] -> source 3691 -> child []

event 6

- Geant4 模拟提供 truth 径迹
- RawBanks: 30
- Strips: 31
- StripCombos: 8
- Coincidences: 3
- Hits: 3

closure check

三个 Hits 到 SVD 拟合直线的残差:

5.27, 1.96, 3.31 cm.

Hits 数据特点

从重建链得到的 Hits

- Hits 已是全局坐标点：
 $[x, y, z, \bar{t}, \Delta t]$
- event 6 中
30 $\rightarrow$ 31 $\rightarrow$ 8 $\rightarrow$ 3 $\rightarrow$ 3:
RawBanks 到 Hits 逐级压缩
- CODEX- β 的小规模组合搜索可作为 CODEX-b 径迹 seeding 原型
- 当前 Hits 压缩了 triplet 信息，未显式保留完整 surface ID

对算法的约束

- 先解决 Hit 归属，再评价直线拟合质量
- 同一探测面或互斥类别中的 Hits 不应进入同一条 track
- 时间信息可帮助限制组合和抑制非同一起穿越

径迹重建原型：从 Hits 到候选 Tracks



输入与分组

- **Hits**: 已重建的空间与时间点
- **互斥 box**: 标记不可放入同一 track 的 Hits
- **回溯枚举**: 在互斥约束下生成事件级分组

评分与输出

- **SVD 评分**: 用直线拟合方差比较候选
- **候选 Tracks**: 选择较合理的 Hit 归属解释
- **Vertex fit**: 作为后续位移顶点重建输入

本节定位

当前径迹部分是可解释的算法原型：用 Hits 特点引出互斥约束，再通过回溯枚举列出可行事件解释并进行候选评分。

互斥 box 构建：把探测器约束转化为组约束

box 的含义

一个 box 是一组彼此不能同时属于同一条 track 的 Hits:

- 同一 track 在每个 box 中最多选一个 Hit
- box 可由时间、module、singlet、surface ID 等信息定义
- 先建立互斥关系，再进入候选枚举

box 作用

- 限制共享 Hit，降低组合爆炸
- 把探测器层级约束提前到 pattern recognition
- 控制候选组合空间，并影响效率和假组合率
- 为后续 ambiguity solving 留下接口

当前 toy 实现

代码暂时用相同时间构造 box:

$box(h_i) = box(h_j)$, 当 $t_i = t_j$ 。正式实现可替换为 module/surface 信息。

组合含义：一条 track 等价于从多个 box 中各取至多一个 Hit，再用候选评分比较不同事件解释。

回溯枚举：在互斥约束下选择事件解释

枚举流程

- 1 按 box 顺序遍历所有 Hits
- 2 对当前 Hit，尝试放入已有 track 组
- 3 若该 track 已使用当前 box，则禁止放入
- 4 也可以新建一个 track 组
- 5 若满足剪枝条件，则剪枝

候选评分

对每个 track 候选，用 SVD 估计主方向 $\hat{\mathbf{u}}$ ，并令 $\mathbf{r}_i = \mathbf{x}_i - \mathbf{c}$ ：

$$d_i^2 = \|\mathbf{r}_i\|^2 - (\mathbf{r}_i \cdot \hat{\mathbf{u}})^2$$

当前代码以各 track 的 $\text{Var}(d_i^2)$ 总和选择较优分组。

核心约束

回溯枚举依赖两个核心约束：

- box 互斥：同一 track 中，每个 box 最多贡献一个 Hit
- 可完成性：若剩余 box 不足以补成有效 track，则提前剪枝

总结

- 1 CODEX-b 以屏蔽后的横向衰变体积补充 LHC 主探测器与前向实验，核心信号是**低本底环境中的多径迹位移顶点**
- 2 CODEX- β 不仅是 RPC 硬件示范器，也承担本底测量、LHCb DAQ 集成、SM 长寿命粒子标定及软件闭环验证
- 3 COBRA 已建立可逐级检查输出的数据链：

RawBanks $\rightarrow$ Strips $\rightarrow$ StripCombos $\rightarrow$ Coincidences $\rightarrow$ Hits.

重建链路同时满足读出约定、聚类连续性和探测器几何

- 4 径迹重建部分已形成可解释的算法原型：先分析 Hits 的低 occupancy 与时间特征，再用互斥 box 约束共享 Hit，通过回溯枚举和 SVD 评分得到候选 Tracks，并为后续多径迹 vertex fit 提供输入

目标：把电子学信号转化为可量化效率与本底的
LLP 候选

参考资料 I

- [1] CODEX-b Collaboration. “CODEX-b: Opening New Windows to the Long-Lived Particle Frontier at the LHC”. In: (2025). arXiv: 2505.05952 [hep-ex].
- [2] CODEX-b Collaboration. “Technical Design Report for the CODEX-beta Demonstrator”. In: (2025). arXiv: 2406.12880 [physics.ins-det].
- [3] CODEX-b Collaboration. “Expression of Interest for the CODEX-b Detector”. In: *European Physical Journal C* 80 (2020), p. 1177. DOI: 10.1140/epjc/s10052-020-08711-3.
- [4] Matt Dobbs et al. *HepMC 2 User Manual*. Version 2.06. 2010.
- [5] COBRA developers. *COBRA Reconstruction and Event-I/O Framework*. Local development repository. 2026.

谢谢!