

文章编号: 1007-4627(2004)01-0012-05

ALICE 光子谱仪光子判别能力模拟研究*

周 静, 孟秋英, 李笑梅, 周书华

(中国原子能科学研究院, 北京 102413)

摘 要: 光子判别能力是 ALICE 光子谱仪 PHOS 的重要性能指标之一. 利用 ALIROOT 软件包对 0.5—100 GeV 能量范围随机分布的 7 种粒子 $n, \bar{n}, p, \bar{p}, \pi^+, K^+$ 和 γ 入射到 PHOS 上时光子判别的效率和纯度进行模拟研究. 在簇射形状分析中, 采用主成分分析方法将 7 参数分析问题约简为 2 参数, 并与 7 参数拓扑分析结果进行了比较. 给出了不同能区中光子判别效率与纯度的关系.

关键词: 光子判别; 主成分分析; ALICE 光子谱仪

中图分类号: O572.212; O572.31 **文献标识码:** A

1 引言

ALICE 是 CERN 的大型强子对撞机 LHC 上的 4 个大型国际合作实验项目之一, 其目标是通过 5.5 TeV/A $^{208}\text{Pb} + ^{208}\text{Pb}$ 等的对撞实验, 研究在高温高密的极端条件下物质的行为, 探索新的物质形态——夸克-胶子等离子体, 了解 QCD 禁闭与手征对称性恢复的机理^[1].

ALICE 实验中, 光子谱仪 (PHOS) 的作用是探测与鉴别实光子, 并以高精度测量光子的四动量, 利用直接光子和轻的中性介子 (主要是 π^0 和 η 介子) 作为探针, 研究感兴趣的物理问题. 在大量产生的多种粒子中鉴别出光子是一个具有挑战性的问题. 特别是强子和光子一样可以发生簇射. 根据簇射形状的不同将光子与强子区分开来, 并在较高的纯度下保持高的判别效率是光子谱仪设计及数据分析中的一项重要课题. 本文利用 ALIROOT 软件包通过模拟计算对 PHOS 的光子判别能力进行研究; 在簇射形状分析中对主成分 (PCA) 分析和拓扑 (TOP) 分析两种方法得到的结果进行了比较.

2 PHOS 的结构

PHOS 由电磁量能器 (EMC) 和带电粒子探测器 (CPV) 组成, 后者用来提供禁阻信号以去除带电粒子^[2].

EMC 分为 5 个模块, 每个模块装配有 64 根垂

直于束流方向的和 56 根沿束流方向的钨酸铅 (PbWO_4) 晶体. 每根晶体长 18 cm, 等价于 20 个辐射长度, 横截面为 22 mm × 22 mm, 相当于 90% 的 Moliere 半径. 模块表面距对撞点 4.6 m. 5 个模块所处的方位角分别为 $\pm 40.7^\circ$, $\pm 20.3^\circ$ 和 0° . PHOS 覆盖的赝快度范围 $|\eta| \leq 0.13$. 钨酸铅晶体受到激发后发出波长为 420 nm 的荧光, 由面积为 5 mm × 5 mm 的雪崩光电二极管 (APD) 探测, 其信号经低噪声前置放大器放大后输送到成形放大器. 电子学线路提供来自每根晶体的两路能量信号、一路时间信号和用作触发的逻辑信号. 两路能量信号有不同放大倍数, 以覆盖 0.5—100 GeV 的能量范围.

每一个量能器模块由一灵敏面积为 144.6 cm × 126.6 cm 的 CPV 模块覆盖. CPV 由多丝正比室组成. 正比室的阳极丝与束流方向平行, 距阴极 7 mm, 每根丝间隔 5.65 mm, 阴极由 56 (沿束流方向) × 128 (垂直于束流方向) 个 2.26 cm × 1.13 cm 的金属片组成. 带电粒子通过探测器时, 相关的阳极丝附近产生雪崩, 在阴极片上诱发的电荷由集成在探测器前面板上的 GASSIPLEX 芯片读出.

3 PHOS 光子判别能力的模拟

PHOS 光子判别能力的模拟分为 PHOS 输出信号模拟、重构运算和粒子鉴别 3 个步骤进行.

收稿日期: 2003-09-09; 修改日期: 2003-12-26

* 基金项目: 国家自然科学基金资助项目 (00121140488)

作者简介: 周 静 (1972-), 女 (汉族), 北京人, 研究实习员, 从事实验核物理研究.

3.1 PHOS 输出信号模拟

首先产生能量在 0.5 – 100 GeV 范围上随机分布的 7 种粒子: γ , π^+ , K^+ , p , \bar{p} , n 和 \bar{n} , 每种 10 000 个, 然后跟踪每一个初始粒子入射到 EMC 的晶体上时在晶体中沉积的能量、该晶体的位置、响应时间、初始粒子种类和在 CPV 中单个阴极片上感应的电荷. 设定 EMC 的阈值为 1 MeV, 对于沉积能量低于阈值的簇射不再继续跟踪. 接着将随机的高斯分布噪声加到所有的探测单元中 (EMC 噪声的平均值为 4 MeV 的等效沉积能, CPV 为 0.01 个电荷单位). 在同一个 CPV 或 EMC 探测单元中, 通过将粒子所沉积的能量相加而得到相应于实验上每一探测单元输出的信号能量. 粒子在 EMC 中沉积的能量大于 10 MeV 及在 CPV 中诱发的电荷大于 0.1 个单位时, 该粒子在 EMC 及 CPV 中产生的信号幅度、相应的探测单元的位置信息以及 EMC 的时间信息数字化后储存在一个数组中, 这样的数组所代表的事件称为 digits.

3.2 重构运算

通过 3 个步骤对 EMC 和 CPV 各单元产生的 digits 进行重构运算: 组成重构点、构建径迹段和构建重构粒子.

(1) 组成重构点

重构点是通过聚类运算将相邻的 digits 组成集团形成的. 任意一个 EMC 的 digits, 其能量在超过最小电离粒子沉积能量的阈值时被认为是一个新的集团的种子. 具有几个局部最大值的集团一般是由相互重叠的簇射造成的, 需将这样的集团分解成不同的重构点, 每个重构点具有一个最大值.

集团的能量 e 取为集团所包含的 digits 的能量之和, 而集团在探测器模块坐标系 (x, z) 中的位置按照用对数权重加权的重心来计算. 此外, 还要对与相互作用顶点出射的粒子方向不平行的晶体做修正. 对数权重

$$w_i = \max \left[0, p + \log \left(\frac{e_i}{e} \right) \right], \quad (1)$$

其中 e_i 是集团中第 i 个 digits 的能量, p 是经验参数.

EMC 中, 集团的特征由 digits 的多重性表示, 而 digits 的多重性是由粒子类型、能量和所选择的

digits 的阈值决定的.

(2) 构建径迹段

PHOS 中的径迹段最多由两个重构点组成: CPV 中一个点, EMC 中一个点. 运算中首先寻找所有彼此之间的距离小于一个给定距离 ($d_{\text{min}} = 10$ cm) 的一对点. 对于一个给定的 EMC 重构点, 相关的径迹段由 EMC 重构点与最接近的 CPV 重构点构成, 或者由 EMC 重构点单独构成. 一个 CPV 的重构点只使用一次.

(3) 构建重构粒子

径迹段作为重构粒子的种子. 它的能量取为 EMC 重构点的能量. 径迹段由两个重构点组成时, 其动量取为通过 CPV 重构点和 EMC 重构点的方向; 径迹段由 EMC 重构点单独组成时, 取为通过相互作用顶点和 EMC 重构点的方向. 粒子的种类根据所测量的表征径迹段的特性参数来定. 光子判别效率定义为初始产生的光子被鉴别为光子的重构粒子数与初始产生的光子数之比, 纯度定义为初始产生的光子被鉴别为光子的重构粒子数与所有被鉴别为光子的重构粒子数之比.

3.3 粒子判别

粒子判别由飞行时间 (TOF)、CPV 和 EMC 簇射形状分析 3 个判别过程完成.

(1) 飞行时间 (TOF)

TOF 定义为粒子由对撞点飞到 PHOS 所经过的时间, 由重构点的时间信息计算. 本工作中快速粒子 (光子、电子) 及慢速粒子 (各种大质量的粒子) 的时间分界选为 0.17×10^{-7} s.

(2) 带电粒子判别

对于由 EMC 和 CPV 二者的重构点组成的径迹段, 将两个重构点中心之间的距离与一个阈值相比较, 该阈值依赖于 EMC 中集团的能量和所需要的粒子判别纯度与效率, 超过这个阈值的径迹段判别为中性粒子 (光子、中子), 低于这一阈值的判别为带电粒子 (电子、带电介子和质子).

(3) 簇射形状分析

依靠 TOF 与 CPV 并不能完全剔除中子及带电粒子. 而光子和强子在 EMC 中都会发生簇射, 簇射形状表征着粒子的种类. 需要通过簇射形状的分析来将光子与强子区分出来. 下面介绍基于簇射形状分析的光子判别方法.

粒子在电磁量能器 EMC 中簇射的形状由下列 7 个参数表征:

横向弥散 d 横向弥散 d 定义为在 EMC 的表面测量的 (x, z) 平面中的横向展宽:

$$d = \frac{\sum_{\text{digits}} w_i [(x_i - x)^2 + (z_i - z)^2]}{\sum_{\text{digits}} w_i},$$

其中 (x_i, z_i) 是对簇射响应的晶体 i 的位置坐标.

椭圆轴 λ_1 和 λ_2 包含簇射的圆锥与 EMC 表面的交面可以用协方差矩阵表示:

$$S = \begin{pmatrix} S_{xx} & S_{xz} \\ S_{xz} & S_{zz} \end{pmatrix},$$

其中

$$S_{xx} = \langle (x - \bar{x})^2 \rangle = \frac{\sum_{\text{digits}} w_i x_i^2}{\sum_{\text{digits}} w_i} - \left(\frac{\sum_{\text{digits}} w_i x_i}{\sum_{\text{digits}} w_i} \right)^2,$$

$$S_{xz} = \langle (x - \bar{x})(z - \bar{z}) \rangle = \frac{\sum_{\text{digits}} w_i x_i z_i}{\sum_{\text{digits}} w_i} - \frac{\sum_{\text{digits}} w_i x_i \times \sum_{\text{digits}} w_i z_i}{(\sum_{\text{digits}} w_i)^2},$$

对 S_{xx} 和 S_{zz} 有相似的定义. 这里, $\langle \rangle$ 表示按对数权重 w_i 求平均, \bar{x} 和 \bar{z} 是集团的重心, 将这个协方差矩阵对角化, 簇射与 EMC 交面的椭圆轴 λ_1 和 λ_2 为协方差矩阵的本征值的平方根.

球度参数 S 球度参数由椭圆轴 λ_1 和 λ_2 确定

$$S = \frac{|\lambda_1 - \lambda_2|}{\lambda_1 + \lambda_2}.$$

核心能量 E_{core} 核心能量取为在最大的 digits 周围给定半径 ($R_{\text{core}} = 3 \text{ cm}$) 内的 digits 的能量之和. 实验表明, 在这一距离之外电磁簇射所沉积的能量小于 2%, 因此由于可能对集团有贡献的本底或簇射的重叠造成核心能量的涨落是较小的.

在单根晶体中能量沉积的最大份额 E_{max} 和 digits 的多重性 M_{digits} E_{max} 反映了簇射的能量分散情况, 而 M_{digits} 为簇射所产生的 digits 的数目.

在簇射形状分析中, 本工作分别采用了主成分分析和拓扑分析两种方法.

拓扑分析需要先对各种粒子做出表征其簇射形状的 7 个参数的谱. 根据不同粒子在这 7 个参数谱

上的差别来选出一套参数范围, 落在范围内的簇射判别为光子. 因为每个参数谱中都有光子与其它粒子重叠的情况, 这 7 个参数中的任何一个都不能单独对粒子做明确的判别, 而且这些参数又是彼此关联的, 必须综合分析和反复调试才可能找到将光子与其它粒子区分开的较好的参数组合. 这种方法计算量大而繁琐.

主成分(PCA)分析是一种多元统计方法, 从自然科学到社会科学的众多领域都得到广泛应用^[3]. 利用主成分分析方法可以减少问题的参数(或维数)而尽量多地保持原有的信息, 便于特征选择.

主成分分析的步骤如下:

(1) 对每一个初始粒子, 求出在 EMC 中簇射的 7 个参数的协方差矩阵.

(2) 计算该协方差矩阵的本征值, 得到一组 7 个统计无关的参数, 即主成分.

(3) 将主成分按由大到小的顺序排列, 舍弃数值小的主成分, 这些主成分对于粒子判别的作用很小. 只采用前两个最大的主成分 p_0 和 p_1 .

(4) 在这两个主成分组成的二维坐标中, 每个粒子的簇射由一个点表示, 点的位置反映着簇射特征, 据此来判别光子.

光子簇射在主成分坐标中的分布用二维高斯函数拟合(见图 1),

$$f(p_0, p_1) = \exp(-R(p_0, p_1)), \quad (2)$$

$$R(p_0, p_1) = \left(\frac{p_0 - x_0}{a} \right)^2 + \left(\frac{p_1 - x_1}{b} \right)^2 + C \frac{(p_0 - x_0)(p_1 - x_1)}{ab},$$

其中参数 a, b, C, x_0 和 x_1 依赖于重构能量. 本工

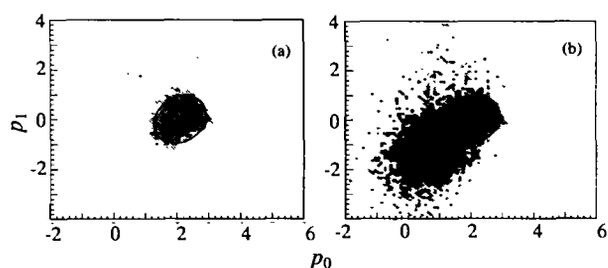


图 1 在两个主成分 (p_0, p_1) 组成的坐标系中光子簇射的分布(a)和 7 种粒子簇射加在一起分布(b)
椭圆是沉积能量为 20—50 GeV 时, 按照参数 $x_0 = 2.1, x_1 = 0, a = 1, b = 0.78, C = 1.13$ 的二维高斯函数表达式(2)画出的; 椭圆内的事件判别为光子, 椭圆外为其它粒子.

作中将粒子能量分为 0.5—10, 10—20, 20—50 和 50—100 GeV 4 个能区. 每一能区中 x_0 , x_1 和 C 值分别是判别为光子的粒子分布的椭圆中心的坐标及椭圆轴的取向. x_0 , x_1 和 C 值选定后, 调节椭圆的

长轴和短轴参数 a 和 b 可选择不同等级的纯度. 在 4 个能量段上选定的这些参数值及相应的光子判别效率与纯度列于表 1. PCA 分析与 TOP 分析结果的比较画在图 2 中.

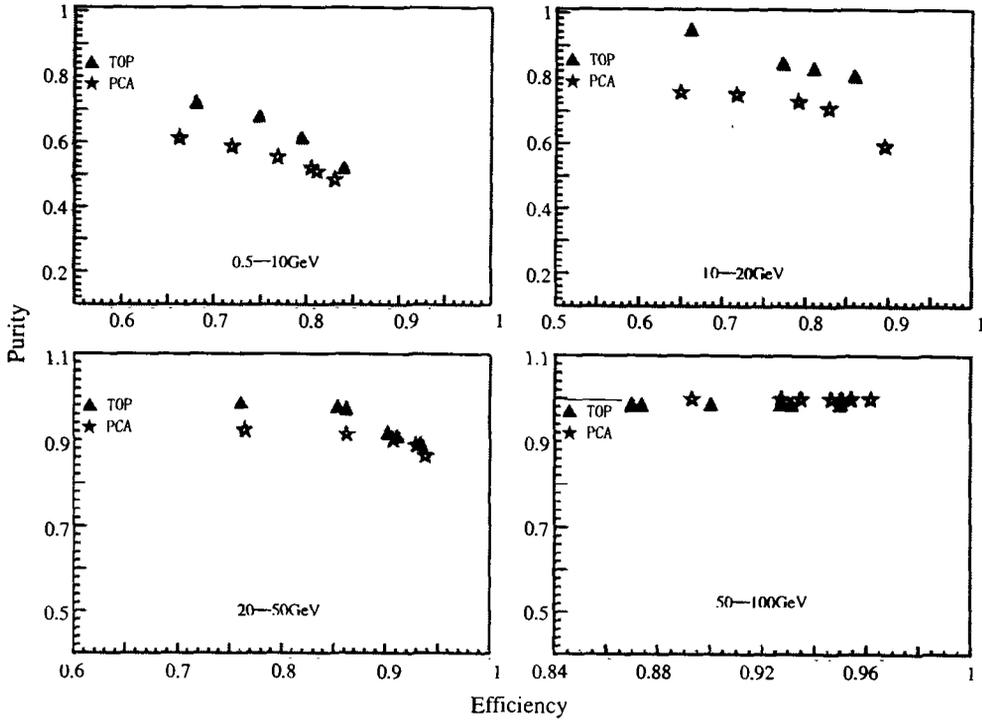


图 2 沉积能在 0.5—10, 10—20, 20—50 和 50—100 GeV 能区中光子判别纯度与效率的关系

表 1 沉积能 E_{dep} 在 4 个能区的椭圆参数

能区	效率	纯度	a	b	x_0	x_1	C
0—10 GeV	0.66	0.61	0.90	0.65	1.30	1.00	1.22
	0.72	0.58	1.00	0.75	1.30	1.00	1.22
	0.77	0.55	1.10	0.85	1.30	1.00	1.22
	0.80	0.52	1.20	0.90	1.30	1.00	1.22
	0.81	0.51	1.25	1.00	1.30	1.00	1.22
	0.83	0.48	1.40	1.05	1.30	1.00	1.22
10—20 GeV	0.65	0.76	1.25	1.00	1.30	1.00	1.22
	0.72	0.75	1.40	1.05	1.30	1.00	1.22
	0.79	0.73	1.45	1.15	1.30	1.00	1.22
	0.83	0.71	1.65	1.35	1.30	1.00	1.22
	0.90	0.59	1.85	1.55	1.30	1.00	1.22
20—50 GeV	0.76	0.92	1.15	0.85	2.05	0.00	1.13
	0.86	0.92	1.30	1.00	2.05	0.00	1.13
	0.91	0.90	1.50	1.15	2.05	0.00	1.13
	0.93	0.89	1.70	1.35	2.05	0.00	1.13
	0.94	0.86	1.90	1.55	2.05	0.00	1.13
	0.94	0.86	1.90	1.55	2.05	0.00	1.13

50—100 GeV	0.89	1.00	1.00	0.95	2.80	-0.80	0.85
	0.93	1.00	1.10	1.05	2.80	-0.80	0.85
	0.94	1.00	1.25	1.20	2.80	-0.80	0.85
	0.95	1.00	1.40	1.25	2.80	-0.80	0.85
	0.95	1.00	1.45	1.35	2.80	-0.80	0.85
	0.95	1.00	1.85	1.75	2.80	-0.80	0.85
	0.96	1.00	2.15	2.05	2.80	-0.80	0.85

从图 2 可以看出, 依靠 TOF, CPV 和 EMC 簇射形状分析, 在 10 GeV 以下的低能区拓扑分析和主成分分析两种方法对于光子的判别都是不能令人满意的. 随着能量增高, 判别效果越来越好. 能量高于 20 GeV 和纯度在 90% 以上时, 效率也可达 90% 以上. PCA 分析与 TOP 分析相比, 在低于 20 GeV 处, 相同效率下, 前者纯度较低, 相差约 10% 左右; 在 20—50 GeV 能区相差小于 10%; 在 50 GeV 以上时, 两者基本相同.

本研究结果显示, 在 PHOS 粒子的簇射形状分析中, 用两个主成分来代替 7 个参数的拓扑分析可简化计算, 两者的结果随沉积能的增加而接近. 在低于 10 GeV 的能区中, 需要对 PHOS 的设计加以改进, 才能得到满意的光子判别效果.

致谢 感谢李霞及刘志毅在计算过程中所给予的帮助.

参 考 文 献:

- [1] ALICE Collaborarion. Technical Proposal[Z]. CERN/LHCC/95-71, 1995.
- [2] ALICE Collaborarion. Photon Spectrometer PHOS [Z]. CERN/LHCC 99-4, ALICE TDR 2, 5, March 1999.
- [3] 李伟明. 多元描述统计方法[M]. 上海: 华东师范大学出版社, 2000, 118—153.

ALICE/PHOS Photon Discrimination Simulation*

ZHOU Jing, MENG Qiu-ying, LI Xiao-mei, ZHOU Shu-hua
(China Institute of Atomic Energy, Beijing 102413, China)

Abstract: The photon discrimination power of the ALICE photon spectrometer is simulated for incident particles n , \bar{n} , p , \bar{p} , π^+ , K^+ and γ with random energy distribution in the range from 0.5 to 100 GeV. The pricipale component analysis method is used to reduce the number of parameters in the shower shape analysis, and the results are compared with that from the seven parameter topology analysis. The efficiency dependence of purity for the photon discrimination is obtained for the deposited energy range 0.5—10, 10—20, 20—50 and 50—100 GeV.

Key words: photon discrimination; pricipal component analysis; ALICE photon spectrometer