

BES III μ 子鉴别器 μ, π 鉴别的研究^{*}

尤郑昀^{1;1)} 冒亚军¹ 张家文²

1(北京大学物理学院技术物理系 北京 100871)

2(中国科学院高能物理研究所 北京 100039)

摘要 用事件产生软件 PYTHIA 模拟了 τ 粲能区正负电子对撞的各反应道, 得到各反应道产生末态 μ 子的截面和动量分布. 用模拟软件 Geant3 对 BES III 的 μ 子鉴别器进行了模拟. 对动量为 0.5 GeV/c 以上的 μ 子和 π 介子的模拟数据, 用 BP 神经网络方法和 Fisher 判别方法进行 μ, π 鉴别. BP 神经网络具有非线性的模式识别能力, 其 μ, π 鉴别效果显著优于线性的 Fisher 方法.

关键词 BES III PYTHIA Geant3 粒子鉴别 神经网络方法

1 引言

μ 子是高能物理实验中的重要粒子. 在北京谱仪(BES)实验中, μ 子不仅是 $J/\psi, \psi', \tau, D$ 的直接衰变产物, 还间接来源于其他一些末态粒子(如 K, π)的衰变. 此外, 在一些稀有反应道(如 $D^0 \rightarrow \mu^+ \mu^-$, $D_s^+ \rightarrow \mu^+ \nu_\mu$)的分析中, μ, π 鉴别效果的好坏直接影响到物理结果的精度. 因此, 在建设中的第三代北京谱仪(BES III)上, μ 子的测量与分辨具有重要的作用. μ 子鉴别器位于 BES III 探测器的最外层, 它的主要功能是测量反应末态中的 μ 子, 并把 μ 子与其他带电粒子(尤其是 π 介子)区分开来. 运用 Geant3 软件, 对 BES III 的 μ 子鉴别器进行了模拟, 对 μ 子和 π 介子的模拟重建结果分别用 Fisher 判别方法和 BP 神经网络方法进行了鉴别, 从而对 BES III μ 子鉴别器的性能有更深入的了解.

2 物理事件模拟

在进行探测器模拟之前, 有必要先进行物理模

拟, 得到进入 μ 子鉴别器的 μ 子信息. 用事件产生软件 PYTHIA5.7 进行了 BES III 上 e^+e^- 对撞的模拟. 在不同的质心系能量下, 通过模拟得到各反应道^[1]的末态粒子中含 μ 子事件的截面 σ (表 1) 和所产生 μ 子的动量分布(图 1). 在一定时间内, 各反应道产生的末态粒子中含 μ 子的事例数

$$N = \sigma \times L \times T, \quad (1)$$

其中 L 为加速器的束流亮度, T 为运行时间, σ 为各反应道中末态粒子含 μ 子事件的截面. 按照 BPEC II $0.5 \times 10^{33} \text{ cm}^{-2} \cdot \text{s}^{-1}$ 的平均亮度, 每年实际运行时间 $1 \times 10^7 \text{ s}$, 可得出一年中各反应道所产生 μ 子的累计数目(表 1).

从表 1 和图 1 可以看出, BES III 上产生的 μ 子主要来自 $J/\psi, \psi'$ 等共振态的衰变和双 μ 过程, 分布在高动量区 ($> 1.5 \text{ GeV}/c$); 而来自 τ, D 和 D_s 衰变的 μ 子, 则主要分布在低动量区 ($0.1-1.2 \text{ GeV}/c$), 数量较少. 但 τ, D 和 D_s 物理是 BES III 上研究的重要物理内容, 因此 μ 子鉴别器探测 μ 子的最低截止动量以及如何做好低动量区的 μ, π 鉴别, 是在研究 μ 子鉴别器的性能时比较关注的.

2004-01-13 收稿

* 国家自然科学基金(10375002)资助

1) E-mail: youzy@hep.pku.cn

表 1 各反应道的末态粒子中含 μ 子事件的截面, 动量分布和运行 $1a(10^7s)$ 积累的 μ 子数目

反应道	质心系能 量/GeV	截面 /nb	一年累计 μ 子数目	动量分布 /(GeV/c)
$e^+e^- \rightarrow J/\psi, \psi \rightarrow \mu^+\mu^-$	3.097	202	2×10^9	~ 1.5
$e^+e^- \rightarrow \psi', \psi' \rightarrow \mu^+\mu^-$	3.686	4.9	49×10^6	~ 1.8
$e^+e^- \rightarrow \mu^+\mu^-$	4.03	5.8	58×10^6	~ 2.0
$e^+e^- \rightarrow \tau^+\tau^-, \tau \rightarrow \mu + \text{any}$	3.67	0.65	3.6×10^6	0.1 ~ 1.2 (图 1(a))
$e^+e^- \rightarrow D\bar{D}, D \rightarrow \mu + \text{any}$	3.77	0.076	0.41×10^6	0.1 ~ 0.9 (图 1(b))
$e^+e^- \rightarrow D_s\bar{D}_s, D_s \rightarrow \mu + \text{any}$	4.14	0.049	0.26×10^6	0.1 ~ 1.0 (图 1(c))

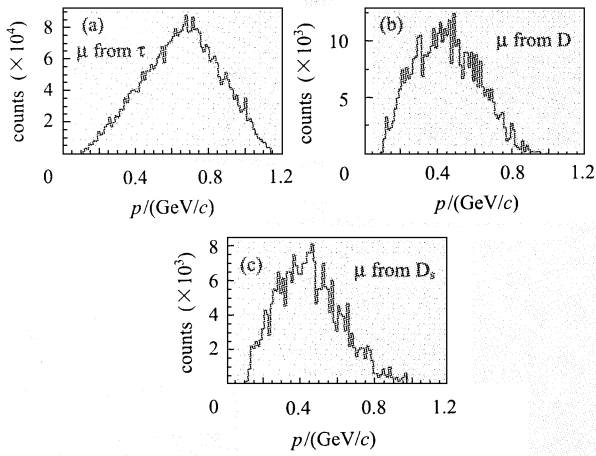


图 1 各反应道所产生 μ 子的动量分布
(a) 来自 τ 衰变的 μ 子; (b) 来自 D 衰变的 μ 子; (c) 来自 D_s 衰变的 μ 子. 各曲线下的积分面积均已归一至 $1a(10^7s)$ 内产生的 μ 子数目.

3 探测器模拟

BES III 探测器主要包括以下几个部分: 1) 主漂移室, 用于测量事例中带电粒子的动量, 方向和能量损失; 2) 电磁量能器, 用于测量电子和 γ 光子的能量和位置信息; 3) 飞行时间计数器, 通过其所测量的飞行时间信息, 辨别粒子的种类; 4) 场强为 1.0T 的超导螺线管磁铁; 5) μ 子鉴别器, 主要功能是测量 μ 子, 并与其他带电粒子 (尤其是 π) 区分开来.

用 Geant3.21 软件包, 对 BES III 探测器进行了模拟, 几何结构如图 2 所示. 其中对 μ 子鉴别器和电磁量能器做了具体几何结构的描述, 内部的束流管、主漂移室和飞行时间计数器的总物质量, 按 2cm 厚的

铁计算, 超导线圈的物质量按 3cm 厚的铁计算.

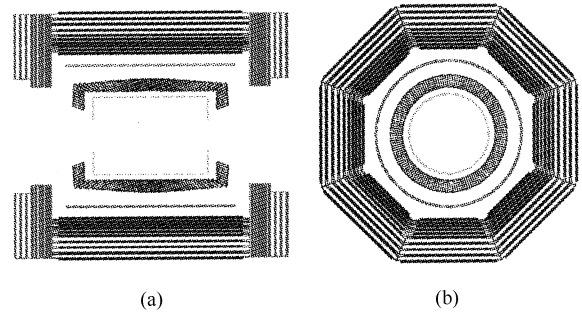


图 2 用 Geant3 模拟的 BES III 结构图
(a) 是 $yz(x=0)$ 截面; (b) 是 $xy(z=0)$ 截面.

μ 子鉴别器分桶部和端盖两部分, 为吸收体铁层和探测器 RPC (Resistive Plate Counter) 的夹层结构, 按八边形排列. 具体结构见 BES III 的设计报告^[2]. 产生模拟数据时, 一个粒子 (μ 或 π) 以一定动量和方向从对撞点出射, 由 Geant3 给出粒子在 μ 子鉴别器各层 RPC 读出条上的响应信号, 从而得到一个单粒子模拟事例.

4 径迹重建和判别变量的选取

由于模拟产生的是单粒子事例, 对于 μ 子鉴别器响应信号的输出, 结合几何结构, 可以简单的用直线拟合的方法重建出入射粒子的单根径迹. 在这个初步的研究中, 不考虑 RPC 的噪声和效率, 以及超导线圈外的反向磁场对径迹偏转的影响.

对重建出的径迹样本, 选择对于 μ, π 鉴别具有物理意义的 5 个变量^[3], 作为判别变量, 见表 2.

表 2 判别变量及其物理意义

变量	定义	对于判别的物理意义
p	粒子动量	动量本身不作为判别变量, 但其他变量将依赖动量而变化
Depth	粒子在 μ 子鉴别器铁层中的穿透深度	与铁作用时, μ 子比 π 介子有更强的穿透能力
Max_Hits	各层 RPC 上被击中读出条数目的最大值	与铁作用时, π 介子有较大的几率产生 Shower
Max_Hits_Layer	Max_Hits 所在 RPC 的层	π 介子产生 Shower 时, 其所在的层
Total_Hits	μ 子鉴别器各层 RPC 上击中数目的总和	综合了穿透能力与 Shower 的因素

由于动量本身并不能作为判别变量,而其他判别变量却依赖动量的变化而变化,因此训练样本必须具有均匀的动量谱.而仅利用 μ 子鉴别器的信息,难以对 $p < 0.5\text{GeV}/c$ 的 μ 子和 π 介子进行有效鉴别(见第 6 节),因此目前只对 $p > 0.5\text{GeV}/c$ 的 μ 子和 π 介子进行鉴别.

在样本的选取上,以两组无关的 μ, π 单粒子模拟事件,分别做为训练样本和检验样本.单粒子从对撞点出射,方向均匀分布在整个 4π 立体角内,动量在 $0.5\text{--}2.1\text{GeV}/c$ 内均匀分布.对于在 μ 子鉴别器上有输出信号的单粒子事件重建出粒子径迹,并去掉探测器边界处的径迹,得到 μ, π 样本,再使样本的动量谱均匀化,从而得到训练样本.

我们用一组含 μ 子和 π 介子各 15000 个事例的样本做为 Fisher 方法和神经网络方法的训练样本,用另一组含有 23088 个 μ 子事例和 17520 个 π 介子事例的样本做为检验样本,来进行 μ, π 鉴别.

5 Fisher 判别分析方法

先用 Fisher 判别分析方法进行 μ, π 鉴别.假设样本有 n 个判别变量,那么可以寻找这 n 个变量的线性组合,将样本从 n 维特征空间投影到一条直线上.在一般情况下,总可以找到一个最好方向,使样本投影到这个方向的直线上是最容易区分开的.这称为线性判别分析. Fisher 方法就是解决如何找到最好的投影方向和实现投影变换^[4].

假设有两类样本 C_i 和 C_j , 输入变量矢量的平均值为 g_i 和 g_j , 协方差矩阵为 S_1 和 S_2 , 定义样本类内离散度矩阵 $T = S_1 + S_2$. 则最好的投影方向 $\omega = T^{-1}(g_i - g_j)$. 对于一个输入矢量 x_0 , Fisher 线性判别式为 $F(x_0) = \omega^t x_0$. 对 $F(x_0)$ 取一个分割值 Y_{cut} . 则有

$$x_0 \in \begin{cases} C_i, & \text{if } F(x_0) \leq Y_{cut}; \\ C_j, & \text{if } F(x_0) > Y_{cut}; \end{cases} \quad (2)$$

将 Fisher 方法应用于 μ 子和 π 介子各 15000 个单粒子模拟事例的训练样本,以表 2 中的 5 个变量为判别变量矢量 $x_0 = (p, \text{Depth}, \text{Total_Hits}, \text{Max_Hits}, \text{Max_Hits_Layer})^t$, 求出 Fisher 判别式

$$Y = F(x_0) = \omega^t x_0, \quad (3)$$

最好投影方向 $\omega = (-1.66, 0.091, 0.172, -0.422, -0.315)^t$, 用 $F(x_0)$ 对检验样本进行判别,得到如图 3 所示的 μ, π 投影分布. 为保证所有动量区间的 μ 子效率均大于 90%, 取 $Y_{cut} = 0.3$, 则对于

检验样本, μ 的总效率为 96.3%, π 的总误判率为 17.1% (表 3).

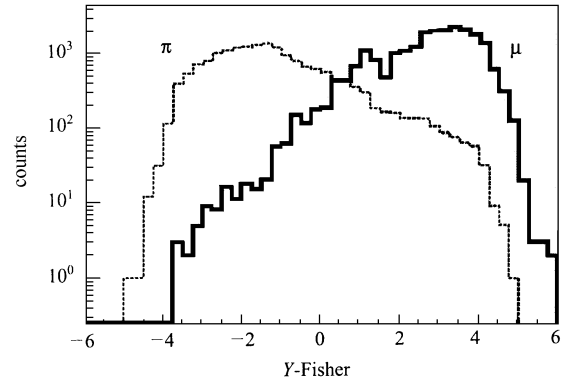


图 3 Fisher 方法对 μ, π 检验样本的输出

表 3 Fisher 方法和 BP 神经网络方法对 μ, π 检验样本的判别效果

	μ 的效率 (%)	π 的误判率 (%)
Fisher	96.3	17.1
BP 神经网络	96.6	9.4

6 神经网络方法

神经网络方法近年来被广泛应用于高能物理中的粒子鉴别,在线触发和径迹重建等方面^[4-7]. 由于具有优越的非线性判别能力,神经网络方法的模式识别能力优于传统的线性判别分析方法. 我们用一种当前应用最为广泛的 BP (Back Propagation) 神经网络来进行 μ, π 鉴别. 有关 BP 神经网络的具体介绍见文献[8].

构造一个三层的前向网络,经过多次试验与优化后,将网络结构定为 5-12-1,输入层为表 2 中的 5 个判别变量,隐藏层节点数取为 12,输出层为 1. 用与 Fisher 方法中同样的一组 μ 子和 π 介子训练样本,通过误差反向传播和梯度下降法,调节网络的连接权阈因子,使 μ 子样本的输出趋于 1, π 介子样本的输出趋于 0.

其他的训练参数优化为,第一层权阈因子的学习速率 $\eta_1 = 0.01$,第二层权阈因子的学习速率 $\eta_2 = 0.1$,动量项系数 $\alpha = 0.5$. Sigmoid 变换函数为

$$f(x) = \frac{1}{1 + e^{-x}}, \quad (4)$$

初始化时,各权阈因子在 $[-0.1, 0.1]$ 内随机赋

值,以一组 μ 子和 π 介子事例(各 15000 个)为一个世代循环训练,直到训练样本的误差方差总和趋于收敛.在训练时一般取循环上限为 1000 个世代,即训练 1.5×10^7 次.由于最优化问题中局域极小值的存在,需要通过多次试验对比,才能得到最佳的训练结果,使误差方差总和最小.

训练结束后,将所有权因子固定,用与 Fisher 方法中相同的检验样本进行检验.检验结果如图 4 所示.设定网络输出的截断值 $Y_{cut} = 0.5$,得到 BP 神经网络对 μ, π 的鉴别效果,见表 3.

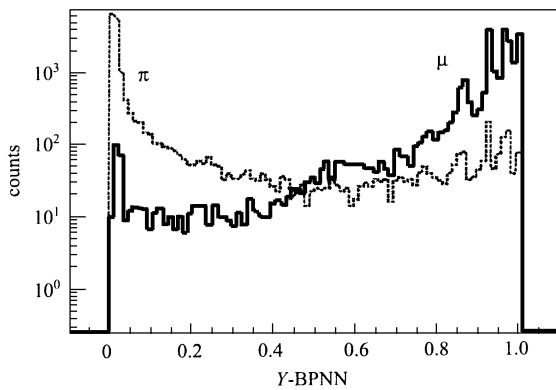


图 4 BP 神经网络方法对 μ, π 检验样本的输出

在保证所有动量区间的 μ 子效率均大于 90% (图 5) 的情况下,从表 3 可以看出, BP 神经网络和 Fisher 方法对 μ 子的总效率都在 96% 左右;而对 π 介子的总误判率 BP 神经网络则低至 9.4%,和 Fisher 方法的 17.1% 相比,降低了 7.7%.

实际上,对判别方法的研究已经证明^[8],在判别分析问题中神经网络方法通常比线性判别函数方法(如 Fisher 方法)更有效.具体到 BES III 的 μ, π 鉴别问题中,在 5 个输入变量所构成的五维特征空间中, μ, π 两类样本的分界面是一个复杂的超曲面,并且由于探测器本身结构的限制,这两类样本在分界面处还存在着一定程度的重叠. Fisher 方法试图以一个最佳超平面来划分两类样本,必然不如以超曲面为分界面的划分效果好.而 BP 神经网络方法通过对输入变量进行非线性变换(见 4 式),能够形成一个较好的超曲面作为 μ, π 样本的分界面,因此得到了比 Fisher 方法更好的判别结果.

图 5 是 BP 神经网络方法对 μ 的效率和对 π 的误判率随动量的分布.随着动量的下降, μ 的效率下降,而 π 的误判率逐渐上升.总的来说判别效果随动量下降而变差.为此,分析了 π 检验样本的径迹信

息,发现由于 π 介子的衰变($\pi^+ \rightarrow \mu^+ \nu_\mu$), π 介子检验样本中约有 6% 在进入 μ 子鉴别器之前已经衰变成 μ 子,而且动量越低,衰变的比率也越大(图 6).若仅仅依靠 μ 子鉴别器的信息,对于这些实际上已经成为 μ 子的“ π 介子”,当然是无法分辨的.由于这部分 π 介子的衰变,导致 μ, π 样本在特征空间内重叠,而且动量越低,重叠的比例越大.因此在保持一定的 μ 子效率的情况下,对 π 介子的误判率存在一个下限.这个下限将随着动量的降低而升高.

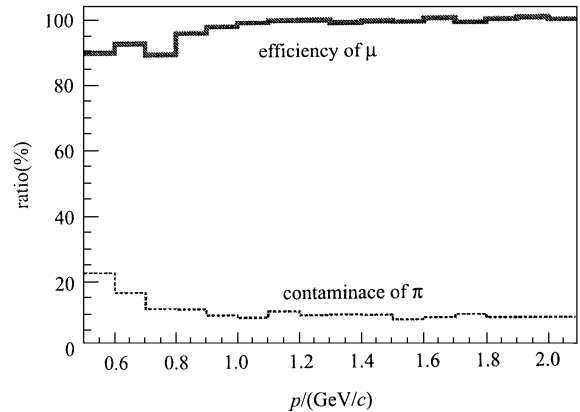


图 5 BP 神经网络方法对 μ 子的判别效率和 π 介子的误判率随动量的分布

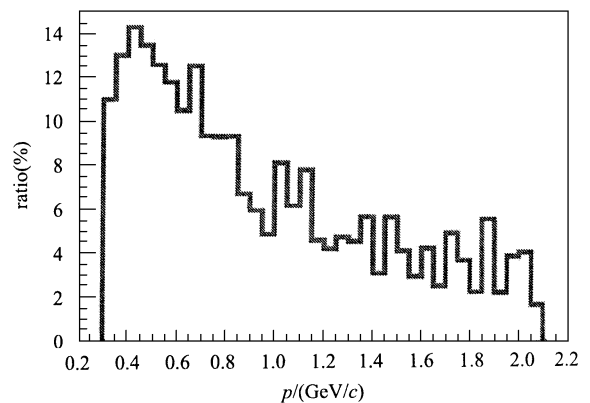


图 6 进入 μ 子鉴别器之前, π 介子衰变成 μ 子的比率随动量的分布

对动量为 $0.5 \text{ GeV}/c$ 以下的 μ 子和 π 介子也进行了模拟.观察表明,当动量下降到 $0.35 \text{ GeV}/c$ 时,只有 10% 的 μ 子能够进入 μ 子鉴别器.因此 μ 子鉴别器对 μ 子的截止动量应当在 $0.35 \text{ GeV}/c$ 左右. $0.35-0.5 \text{ GeV}/c$ 动量区的 μ, π 鉴别,存在着以下两个困难,一是该动量区的 π 介子在进入 μ 子鉴别器之前衰变成 μ 子的比率很大,达到了 14% 左右(图

6);二是该动量区的 μ 子和 π 介子一般只能在第 1—3 层 RPC 上产生信号,并且 π 介子产生簇射的几率减小,从而使主要判别变量失效.因此小于 0.5 GeV/c 动量区的 μ, π 鉴别,需要加入内部探测器的信息,才能达到较好的鉴别效果.

7 总结

本文针对建设中的第三代北京谱仪(BES III) μ

子鉴别器的设计方案,通过探测器模拟,研究了 BP 神经网络方法在 μ, π 鉴别中的应用,对比 Fisher 判别方法,得到了较好的 μ, π 鉴别效果.待 BES III 离线软件中的模拟和重建部分完成后,使用完整而精确的模拟数据,优化重建径迹,并综合 BES III 探测器的全部信息进行判别,相信能够进一步提高 BES III μ, π 鉴别的效果和可信度.

参考文献 (References)

- 1 YUAN Chang-Zheng et al. HEP & NP, 2002, **26**(12): 1—8 (in Chinese)
(苑长征等.高能物理与核物理, 2002, **26**(12): 1—8)
- 2 LI Wei-Guo. Overview of BEPC II / BES III Project, Talk Given on BES III International Review, Sep. 16—28, 2002, Beijing
- 3 MAO Ya-Jun et al. Atomic Energy Science and Technology, 1998, **32**(2): 146—151 (in Chinese)
- 4 Mjahed M. Nucl. Instrum. and Methods, 2000, **A449**: 602—608
- 5 Horn D. Nucl. Instrum. and Methods, 1997, **A389**: 381—387
- 6 Eide A et al. Nucl. Instrum. and Methods, 1996, **A368**: 855—858
- 7 ZHANG Zi-Ping et al. HEP&NP, 1997, **21**(4): 297—303 (in Chinese)
(张子平等.高能物理与核物理, 1997, **21**(4): 297—303)
- 8 Hagan M T et al. Neural Network Design. U.S.A.: PWS Publishing Company, 1996, 197—257

Research on Identification of μ, π with BES III μ Identifier*

YOU Zheng-Yun^{1;1)} MAO Ya-Jun¹ ZHANG Jia-Wen²

1(Department of Technical Physics, School of Physics, Peking University, Beijing 100871, China)

2(Institute of High Energy Physics, CAS, Beijing 100039, China)

Abstract The final states created by e^+e^- annihilation at τ -charm energy region are simulated with PYTHIA5.7, and the resultant inclusive muon production cross sections and momentum distributions are obtained. The muon identifier of BES III detector is fully simulated with Geant3.21. For the simulated events containing muons and pions with momentum greater than 0.5 GeV/c, the BP neural network and Fisher discriminant method are used for muon/pion identification. BP neural network performs better than Fisher discriminant method. With the same muon efficiency of 96%, the misidentification of pion by BP neural network is 9.4%, in comparison with that of 17.1% by Fisher method.

Key words BES III, PYTHIA, Geant3, particle identification, neural network

Received 13 January 2004

* Supported by National Natural Science Foundation of China(10375002)

1)E-mail: youzy@hep.pku.cn