

用神经网络方法鉴别 τ 的强子衰变*

陈国明 陈刚

(中国科学院高能物理研究所 北京 100039)

1994-06-30 收稿

摘 要

用神经网络法鉴别 τ 的单叉强子衰变 $\tau \rightarrow h + n\pi^0 + \nu$, $n = 0, 1, 2, 3$. 并进而测得上述各道的分支比: $\text{Br}(\tau/\text{K}\nu) = (12.18 \pm 0.26 \pm 0.42)\%$, $\text{Br}(\tau/\text{K}\pi^0\nu) = (25.20 \pm 0.35 \pm 0.50)\%$, $\text{Br}(\tau/\text{K}2\pi^0\nu) = (8.88 \pm 0.37 \pm 0.38)\%$, $\text{Br}(\tau/\text{K}3\pi^0\nu) = (1.70 \pm 0.24 \pm 0.39)\%$, 其中第一项误差是统计误差, 第二项是系统误差.

关键词 对撞机, 粒子鉴别, 神经网络, 衰变分支比.

1 引 言

在西欧核子研究中心的 LEP 实验中, 由 Z^0 衰变而得的 τ 具有很高能量 (45 GeV), 由于 Lorentz 增强效应, 最终 τ 的衰变产物都集中在一个很小的锥角内. 对于三个带电强子, τ 衰变末态可以由径迹室的径迹数来判别, 而对于一个带电强子加上 n 个 π^0 ($n = 0, 1, 2, 3$) 的末态则很难判别. 另外, τ 轻子的发现虽然已有 20 年^[1], 对于 $\tau \rightarrow \pi/\text{K} + n\pi^0 + \nu$ ($n = 0, 1, 2, 3$) 的分支比仍旧有争议^[2].

近来神经网络在高能物理中的应用越来越多. 在粒子末态鉴别方面它已被证明是一个强有力的工具. 本文将介绍利用神经网络鉴别 $\tau \rightarrow \pi/\text{K} + n\pi^0 + \nu$ ($n = 0, 1, 2, 3$), 并进而测量各个道的分支比.

文中所用的数据取自 L3 实验. LEP 是个大型正负电子对撞机, L3 探测器安装在其中的一个对撞点上. 我们的数据来自 1992 年物理运行, 质心能量在 Z^0 峰上, 总积分亮度为 21.68 pb^{-1} (扣除了部分探测器工作不正常的数据). L3 探测器由中心径迹室 TEC、BGO 晶体组成的电磁量能器、铀板和多丝正比室夹层而成的强子量能器及 μ 子漂移室构成. 探测器最外层是线圈和轭铁以提供轴向磁场. 详细资料见文献[3]. 在这里需要强调指出的是由 BGO 组成的电磁量能器具有很好的能量和空间分辨率. 在 45 GeV 处能量分辨为 1.34%, ϕ 方向的分辨为 1.31 mrad, θ 方向的分辨为 3.32 mrad. 对于带有一根径迹的喷注, 尽管冲度很大 (> 0.9), 我们仍旧能够将 BGO 中的簇射分解成一个与径迹相联系的带电粒子及 n 个中性粒子 ($n = 0, 1, 2, 3, 4, 5$). 一般来说一个中性粒子是一个

* 国家自然科学基金资助.

光子或 π^0 , 但也可能由光子与 π^0 或 π^0 与 π^0 交叠而成, 或者甚至是噪音或中性强子, 因此简单地区分 $\tau \rightarrow \pi/K + n\pi^0 + \nu$ 还是有困难。这就是为什么要用神经网络来区分的原因。

2 单叉 τ 事例的选择

作为分析的第一步, 选择 1prong + X $\tau\tau$ 事例。这样的事例至少有一头有且只有一根径迹, 这根径迹在 $R-\phi$ 平面上至少有 30 个着火点, 在 z 方向至少有 2 个测量点。应用到同一时期的 $e^+e^- \rightarrow e^+e^-$, $\mu^+\mu^-$ 事例, 可得单径迹测量的效率为 $(95.3 \pm 0.6)\%$, 且在误差范围内, Monte Carlo 模拟结果与其符合。由于强子事例多重数高, 每边的径迹数都远大于 1, 因此被排除在外。为了排除 $e^+e^- \rightarrow e^+e^-$ 事例, 要求电磁量能器总能量小于 65GeV , 而每边的电磁量能器能量小于 40GeV 。为了拒绝 $e^+e^- \rightarrow \mu^+\mu^-$ 事例, 我

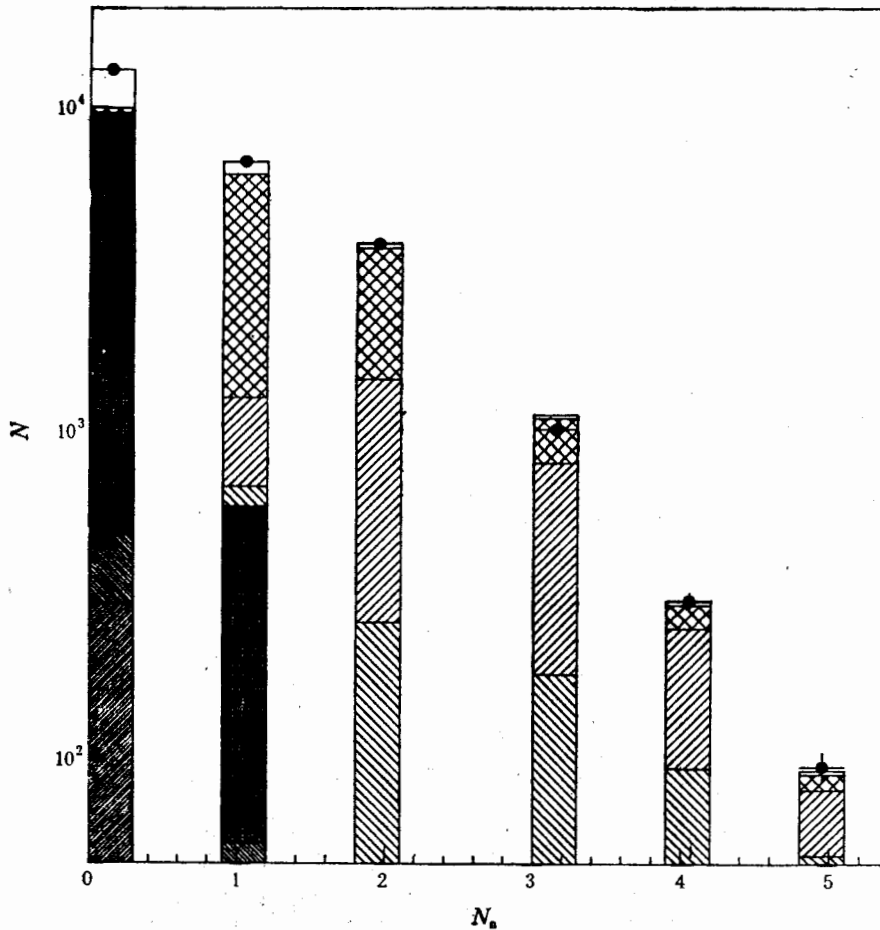


图 1 通过 $\tau\tau$ 事例选择以后, 单个喷注(即一个 τ) 中, 中性粒子数的分布

这里的 Monte Carlo 归一到积分亮度, 图中可见数据与 Monte Carlo 符合。

- 数据
- ▨ MC: $\tau \rightarrow h + 2\pi^0 + \nu$
- ee 本底
- MC: $\tau \rightarrow h + \nu$
- ▧ MC: $\tau \rightarrow h + 3\pi^0 + \nu$
- ▩ $\mu\mu$ 本底
- ▨ MC: $\tau \rightarrow h + \pi^0 + \nu$
- ▨ MC: $\tau \rightarrow e/\mu + \nu\nu$

们去掉两边都是 μ 候选者的事例。一个 μ 候选者在 μ 漂移室有径迹或在量能器中被判定为最小电离粒子。为了去掉双光子事例,要求量能器总能量大于 13GeV 。这样在桶部的选择效率为 67.3% ,纯度为 97.2% ,主要本底来自 e^+e^- 和 $\mu^+\mu^-$, 分别为 1.2% 和 1.5% 。

3 粒子鉴别

通过 $\tau\tau$ 事例选择,得到 24776 个带一根径迹的 τ , 这些 τ 的中性粒子数分布如图 1 所示,图 2 为带电的和中性的电磁量能器的能量分布,图 3 为带电粒子与第一个中性粒子的不变质量分布。在图中,直方图代表 Monte Carlo 模拟,需指出的是这里的归一是绝对归一,即归一到总的积分亮度。对于 $n=0$ (n 为中性粒子的个数) 的情形,主要是 τ 到 e, μ, K, π 的衰变,用简单的 cut 来做衰变模式鉴别。对于 $1 \leq n \leq 4$, 则用神经网络来区分。如果一喷注被选定是某一衰变模式且没有被其它任何拒绝条件所拒绝,则该喷注被认定为这一衰变模式。

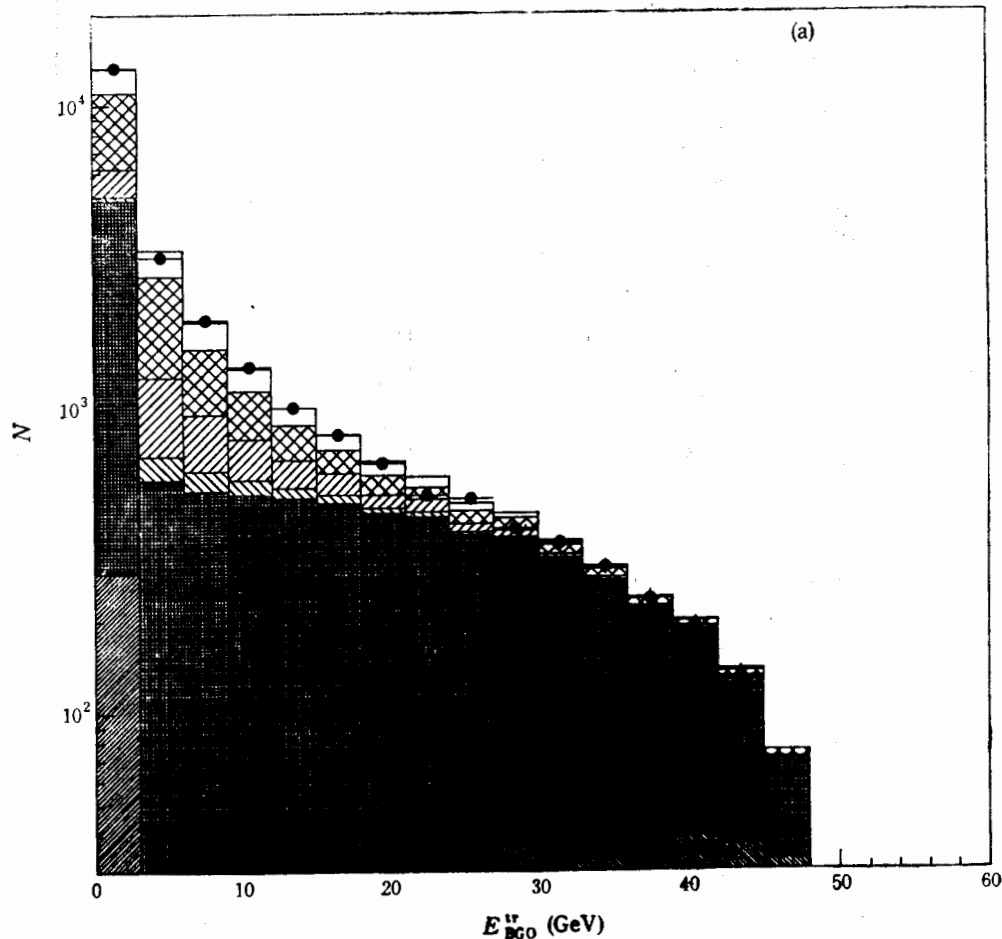


图 2 (a) 带径迹的 BGO 能量的分布

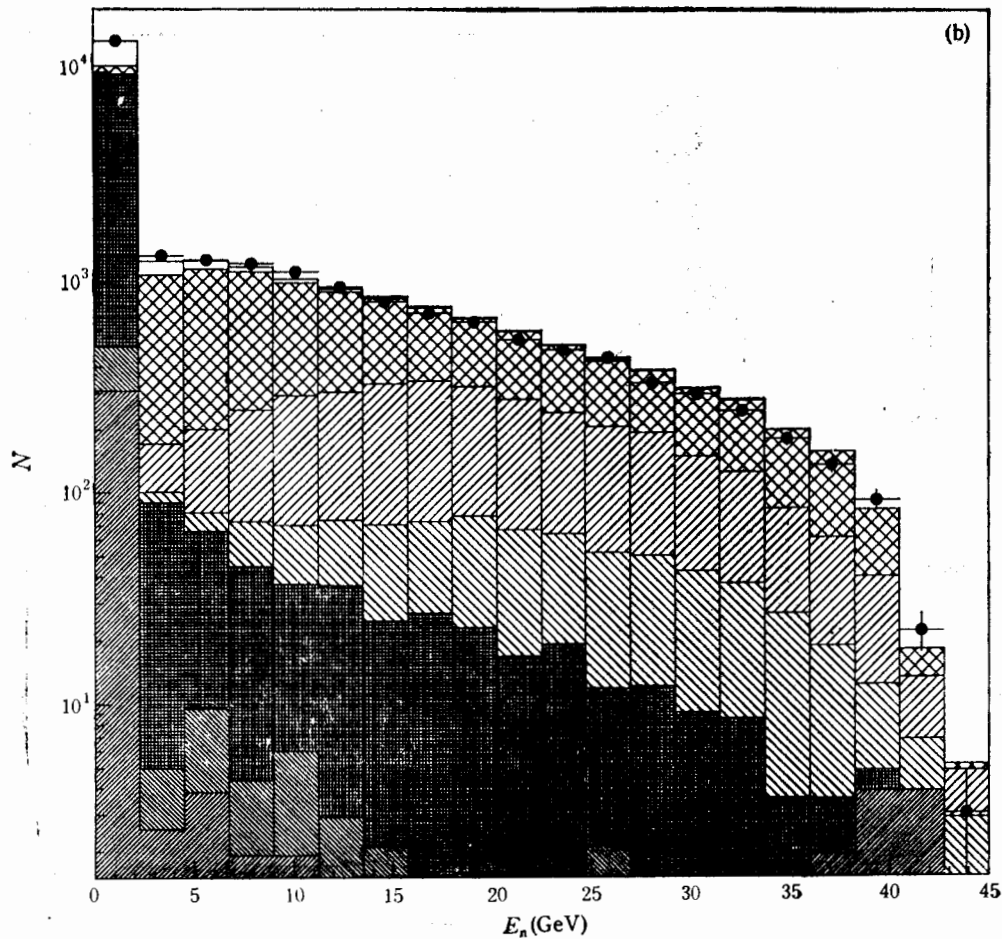


图 2 (b) 中性的 BGO 能量的分布
(图例与图 1 同)

在 $n = 0$ 时感兴趣的道是 K, π , 因此要求在量能器里横向和纵向的能量分布既非电磁型也非最小电离型, 对于 $\pi/K + \pi^0$, 从图 1 可知, 主要分布在 $n = 1, 2$, $n = 0$ 很少, 但为了减少其对 π/K 的本底, 要求电磁量能器的横向能量分布 $E_{\text{BGO}}^T < 0.15$, 其中 $E_{\text{BGO}}^T = \sum E_i^T / E_i$, 而横向则是相对径迹方向而言。

对于 $n = 1, 2, 3, 4$, 分别用一个神经网络来进行衰变模式的鉴别。其输入变量的个数分别是 17, 17, 21, 22。总的来说, 输入变量可分以下四类:

- (1) 各子探测器的输出能量: $P_{\text{TEC}}, E_{\text{ECAL}}, E_{\text{HCAL}}$;
- (2) 有关簇射形状的量: $\Sigma, \Sigma_{25}, \Sigma_1 / \Sigma_{25}, E_{\text{BGO}}^T$, 这里 Σ , 指 9 个最高输出的 BGO 晶体的能量和, 其余类推;
- (3) 中性粒子的能量: $E_i, i = 1, 2, 3, 4$
- (4) 带电粒子和中性粒子, 及中性粒子之间的不变质量。

所用的 Monte Carlo 事例由 KORALZ 3.8^[4] 产生, 然后由 GEANT3^[5] 对 L3 探测器

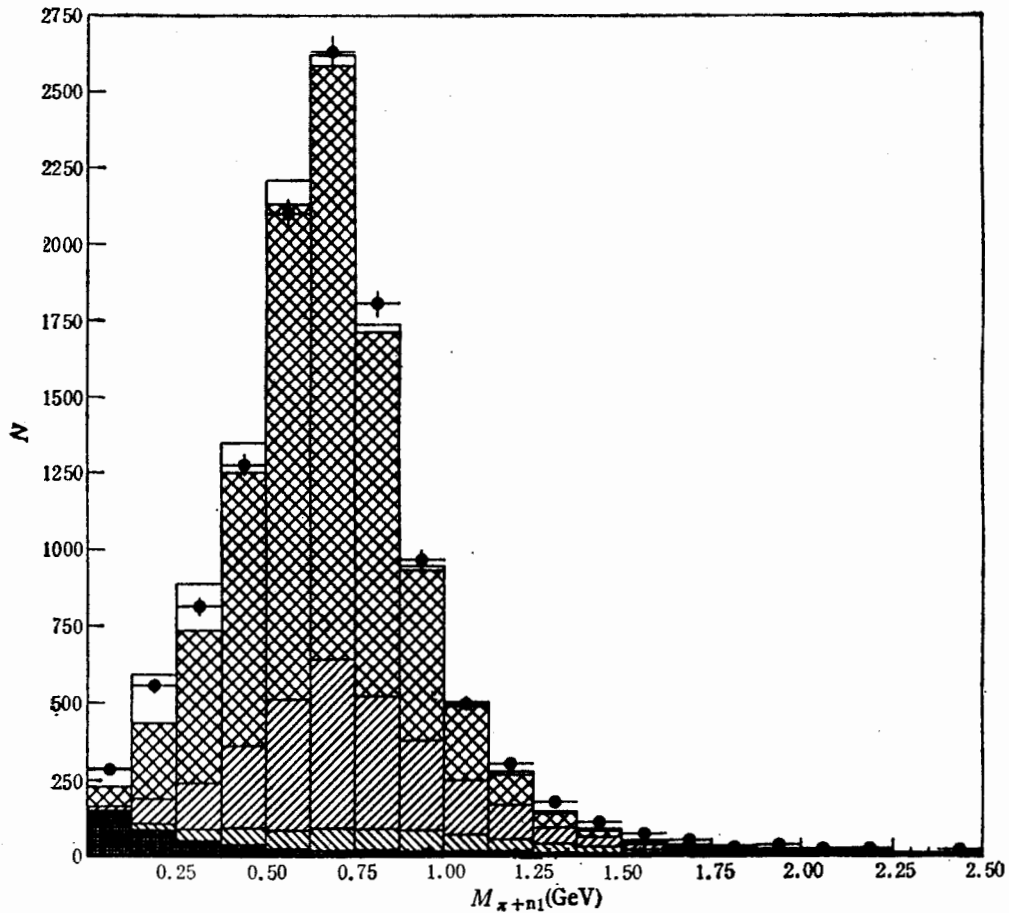


图3 带电粒子与第一个中性粒子的不变质量分布
(图例与图1同)

的响应进行全面模拟。Monte Carlo 不仅很好地描述了能量响应,而且很好描述了簇射的形状,这点从图1—3可以看出。我们用了250,000个这样的 $\tau\tau$ 事例来训练神经网络。

4 神经网络

这一节以 $n=2$ 为例介绍所用的神经网络的结构,对于 $n=1,3,4$ 除输入变量外其它都相同。

使用向后传递模型,并使用由Rumelhart、Hinton、Williams在1986年提出的momentum方法^[6]。网络分成三层:输入层,输出层和中间层。输出层的节点只有一个,输出值分布在0到1之间。输入层和中间层的节点数相同,都为17,其中输入的17个变量为:

- (1) E_{n_1} : 第1个(能量最高的)中性粒子的能量;
- (2) E_n : 所有的中性粒子的能量和;

(3) $E_n/(E_{\text{ECAL}} + E_{\text{HCAL}})$: 其中的 E_{ECAL} , E_{HCAL} , 分别为电磁量能器和强子量能器的能量;

(4) E_{BGO}^T : 电磁量能器的横向能量;

(5) $\Delta\phi_{\text{max}}$: 中性粒子与径迹的最大夹角;

(6) $M_{n1, n2}$: 第一个中性粒子与第二个中性粒子的不变质量;

(7) $M_{\pi, n1}$: 第一个中性粒子与带电粒子的不变质量;

(8) $\sum M$: $M_{n1, n2} + M_{\pi, n1} + M_{\pi, n2}$, 三个不变质量之和;

(9) $\Delta\phi$: BGO 能量重心与带电径迹间的夹角;

(10) χ_{ep}^2 : 表示量能器能量与漂移室动量间的符合程度,

$$\chi_{\text{ep}}^2 = \frac{(E_t - E'_t)^2}{2\Delta E_t^2} + \frac{\left(\frac{1}{p_t} - \frac{1}{E'_t}\right)^2 p_t^4}{2\Delta p_t^2},$$

其中 p_t 为漂移室测出的带电粒子的横动量。这里的横向指垂直于 e^+e^- 方向, E_t 为垂直于 e^+e^- 方向的量能器能量, 而 E'_t 为量能器能量与漂移室动量的加权平均后取横向部分;

(11) Σ_1/Σ_{25} : 被径迹击中的 BGO 晶体能量 Σ_1 与周围 (包括本身) 25 块晶体的能量 Σ_{25} 之比

(12) Σ_9/Σ_{25} ;

(13) E_{n1}^* : 第一个中性粒子中具有最高输出的晶体的能量;

(14) E_{n2}^* : 第二个中性粒子中具有最高输出的晶体的能量;

(15) E_{n2}/E_{n1} ;

(16) M_{n1}^{**} : 假设第一个中性粒子是由两个光子形成时两个光子的不变质量;

(17) M_{n2}^{**} : 假设第二个中性粒子是由两个光子形成时两个光子的不变质量。

在训练时, 滞留因子 $\alpha = 0.07$, 训练强度参数 η 的初值 0.03, 在训练过程中不停下降, 最小值为 0.001。训练以后为了确定效率和本底, 使用另外独立的 110,000 Monte Carlo $\tau\tau$ 事例对训练结果进行检验。图 4 为神经网络的输出值 Y_p 的分布。其中阴影部分为本底, 通过优化, 我们确定: 当 $Y_p > 0.6$ 时, 为 $\pi/K + \pi^0$, 当 $0.22 < Y_p < 0.6$ 时, 认为有 $\pi/K + \pi^0$ 的可能性, 为减小该道对 $\pi/K + 2\pi^0$ 的本底的贡献而拒绝之。

以上定义 $\pi/K + \pi^0$ 为讯号, $\pi/K + 2\pi^0$, $\pi/K + 3\pi^0$ 为本底进行训练, 还定义了以下的讯号与本底进行训练:

(1) 以 π/K 为讯号, 以 $\pi/K + \pi^0$, $\pi/K + 2\pi^0$, $\pi/K + 3\pi^0$ 为本底进行训练。这是 $n = 2$ 的情形, 从图 1 知 π/K 很少, 在这里鉴别 π/K 本底很高, 得不偿失, 训练的目的在于排除 π/K 对 $\pi/K + \pi^0$ 的本底贡献。

(2) 以 $\pi/K + 3\pi^0$ 为讯号, 以 $\pi/K + 2\pi^0$, $\pi/K + \pi^0$, π/K 为本底。图 5 为其输出值的分布。

通过训练首先选择 $\pi/K + 3\pi^0$, 然后排除剩下的且有一定可能是 $\pi/K + 3\pi^0$ 的事例。由于有一定可能性的 $\pi/K + \pi^0$, 或 $\pi/K + 2\pi^0$ 的事例都被排除了, 最后剩下的就是 $\pi/K + 2\pi^0$ 。

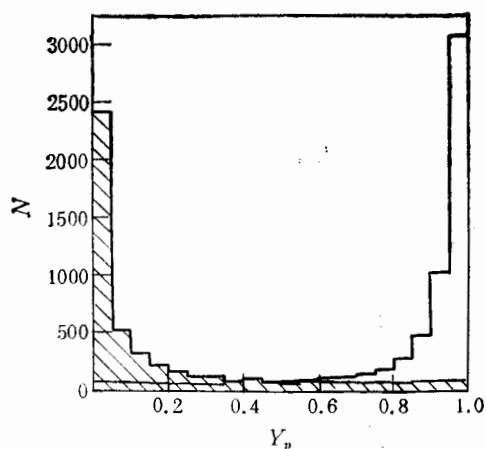


图4 讯号为 $h + \pi^0$ 时的神经网络输出值阴影部分为本底

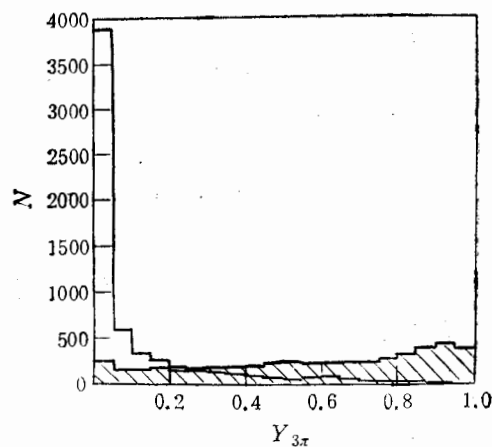


图5 讯号为 $h + 3\pi^0$ 时的神经网络输出值

5 结 果

通过上述粒子鉴别,在所用的数据中选定 π/K , $\pi/K + \pi^0$, $\pi/K + 2\pi^0$, $\pi/K + 3\pi^0$ 的事例数分别为 3109, 6613, 1060, 293. 利用 Monte Carlo 得到效率矩阵如表 1 所示. 选择后的 $\pi/K + \pi^0$, $\pi/K + 2\pi^0$, $\pi/K + 3\pi^0$ 的不变质量分布见图 6.

表 1 粒子鉴别的效率矩阵 ε_{ij} (单位: %)

来 源	选 择			
	h	$h + \pi^0$	$h + 2\pi^0$	$h + 3\pi^0$
e	1.18	0.93	0.24	0
μ	0.80	0.61	0.02	0
h	69.05	3.96	0.36	0.04
$h + \pi^0$	3.64	78.64	1.49	0.18
$h + 2\pi^0$	0.37	12.33	33.13	4.79
$h + 3\pi^0$	0.15	2.81	12.38	28.58
3prong	0.08	0.21	0.06	0
ee	0.03	0.004	0	0
$\mu\mu$	0.26	0.04	0	0
$h\eta\pi^0$	0	2.61	7.48	34.70
$h4\pi^0$	0	1.42	6.03	40.78

为了确定这四个道的分支比,令

$$N_i^{\text{exp}} = N_\tau \cdot \sum_{j=1}^4 \text{Br}_j \varepsilon_{ij} + \sum_{k=1}^2 N_k^{\text{bg}} \varepsilon_k^i \quad (1)$$

其中, $i=1,2,3,4$, 指 π/K , $\pi/K + \pi^0$, $\pi/K + 2\pi^0$, $\pi/K + 3\pi^0$, j 代表 τ 的各个衰变道, ε_{ij} 代表从 j 道鉴别成 i 道的效率. (1) 式中第二项代表 ee , $\mu\mu$ 事例对 i 道

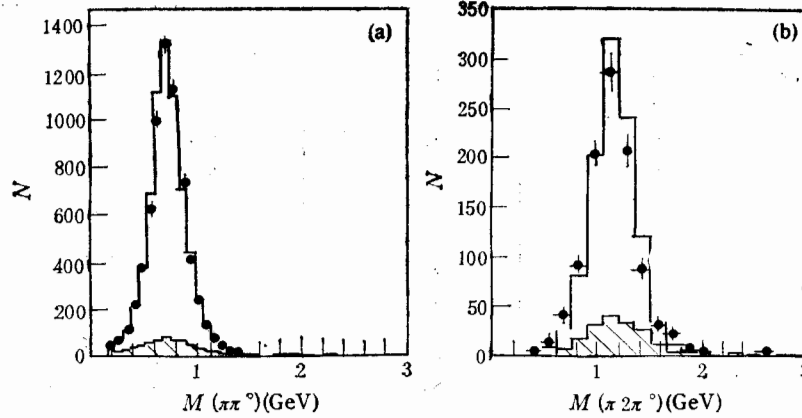
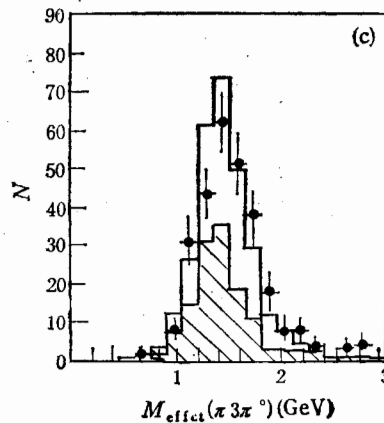


图 6(a) 被鉴别成 $h + \pi^0$ 的喷注的不变质量

● 数据 □ MC: $\tau \rightarrow h + \pi^0 + \nu$
 ▨ MC: 本底

(b) 被鉴别成 $h + 2\pi^0$ 的喷注的不变质量

● 数据 □ MC: $\tau \rightarrow h + 2\pi^0 + \nu$
 ▨ MC: 本底



(c) 被鉴别成 $h + 3\pi^0$ 的喷注的不变质量

● 数据 □ MC: $\tau \rightarrow h + 3\pi^0 + \nu$ ▨ MC: 本底

的贡献。对比 N_i^{exp} 和对应道的测量值, 利用最大似然法拟合可得分支比:

$$Br(\pi/K\nu) = (12.18 \pm 0.26 \pm 0.42)\%;$$

$$Br(\pi/K\pi^0\nu) = (25.20 \pm 0.35 \pm 0.50)\%;$$

$$Br(\pi/K2\pi^0\nu) = (8.88 \pm 0.37 \pm 0.38)\%;$$

$$Br(\pi/K3\pi^0\nu) = (1.70 \pm 0.24 \pm 0.39)\%,$$

其中第一项误差是统计误差, 第二项是系统误差, 系统误差的来源见表 2。需要指出的是, 在拟合时未加任何限制条件。我们仔细研究了神经网络输入鉴别所造成的系统误差。首先, 改变每个网络输出值 cut 的 10%, 得到网络输出对系统误差的贡献, 而网络输入对系统误差的贡献则比较复杂, 这是因为网络个数多, 每个网络的变量数大, 且这些变量并非相互独立。为此, 首先依次将每个变量置于该量的平均值以观察该量对分支比的影响

表2 系统误差分析(单位: %)

	h	$h + \pi^0$	$h + 2\pi^0$	$h + 3\pi^0$
中性粒子能量阈值	0.032	0.065	0.065	0.095
MC 统计性	0.18	0.26	0.18	0.15
$\tau\tau$ 事例选择	0.14	0.20	0.11	0.17
τ 总数	0.10	0.22	0.08	0.02
神经网络的输出	0.04	0.13	0.22	0.18
神经网络的输入	0.29	0.28	0.22	0.20
简单 cut	0.16	0	0	0
$h + 4\pi^0\nu$	0	0	0.04	0.15
$h + \eta\pi^0\nu$	0	0	0.02	0.04
总和	0.42	0.50	0.38	0.39

力。每当一个量置常数时,其它量仍保持原来的分布。而这些操作只对 Monte Carlo 数据,实验数据保持不变。这样得到独立的,且对网络有最大影响力的量为: P_{TEC} 、 E_{ECAL} 、 E_{HCAL} 、 E_{BGO}^T 、最大的两个中性粒子能量、中性粒子之间及中性粒子和带电粒子之间的夹角。为了确定以上各量对系统误差的贡献,在 Monte Carlo 模拟中将每个量的输入值作一些晃动(包括位移,模糊,伸缩),找到与实验数据符合得最好的晃动,与此晃动对应的分支比与标准值之差即为该量对系统误差的贡献。如果不论如何晃动, Monte Carlo 与实验数据的符合都比标准的差, χ^2 只能上升,则以 $\Delta\chi^2 = 1$ 时的晃动来估计。最后将以上各量对系统误差贡献的平方和的开方作为网络输入的系统误差。 τ 总数的不确定性主要来自于总积分亮度的误差。由于蒙特卡罗事例数的限制造成鉴别效率的不确定性对误差的贡献也相当大。在估计 $\tau\tau$ 选择的贡献时,将所有 cut 平移 10% (保守估计,大于 1σ),再将所造成的分支比的变化平方相加。在重建中性粒子时为了去掉噪音,加了 0.5GeV 的能量阈值,由此造成的系统误差几乎可以忽略。 $h + 4\pi^0$, $h + \eta\pi^0$ 的贡献则来自于这两道分支比的不确定性。

6 讨 论

表 3 列出了 LEP 和其他一些实验中 τ 的强子衰变分支比的最新结果^[7]。

从表中可见所有实验的 $\tau \rightarrow h\nu$ 的分支比结果都相符合, $\tau \rightarrow h\pi^0\nu$ 的分支比也都相符合,而我们的结果则最接近 LEP 的平均值, $\tau \rightarrow h2\pi^0\nu$ 的分支比结果与 CLEO II 很接近, $\tau \rightarrow h3\pi^0\nu$ 的分支比与 ARGUS 和 ALEPH 92 年的结果相符合。

根据我们的结果以及 L3 的其它 1prong 分支比可得到^[7]:

$$\sum Br = (84.2 \pm 1.2)\% \text{ (遍举)},$$

$$B1 = (85.6 \pm 0.67)\% \text{ (单举)}.$$

即 1prong 疑难已被解决,遍举分支比和单举分支比相符合。

上述结果表明神经网络用于粒子衰变末态的鉴别是一种十分有效的工具。特别是当衰变粒子末态分布在很小的空间,而一般常规方法已很难进行鉴别时,神经网络显示出很大的优越性。

表 3 到 94 年 3 月为止 τ 的强子衰变分支比的最新测量结果 (单位:

实验组	$h\nu$	$h\pi^0\nu$	$h2\pi^0\nu$	$h3\pi^0\nu$
L3	12.18 ± 0.49	25.20 ± 0.60	8.88 ± 0.53	1.70 ± 0.46
ALEPH 92	13.32 ± 0.53	25.02 ± 1.07	10.53 ± 1.05	1.53 ± 0.61
ALEPH 94	12.07 ± 0.34	25.04 ± 0.55		
DELPHI	11.84 ± 0.56	23.57 ± 0.92		
OPAL	12.20 ± 0.5	26.25 ± 0.63		
LEP 平均	12.07 ± 0.22	25.22 ± 0.32		
ARGUS			9.7 ± 0.58	1.7 ± 0.36
CLEO II		25.87 ± 0.44	$8.92 \pm 0.47^*$	$1.06 \pm 0.15^*$
PDG 92	12.7 ± 0.4	24.4 ± 0.6	10.3 ± 0.9	2.7 ± 0.9

注: 这里的 h 指 π 或 K , 数据来自于 94 年 3 月在法国召开的 Mariond Conference.

本文作者感谢麻省理工学院 (MIT) 的 A. Kuning 博士, 王怡芳博士, 苏黎士联邦工学院 (ETHZ) 的 M. Pohl 博士的帮助, 并感谢陈和生教授和金炳年教授对我们的工作的支持.

参 考 文 献

- [1] M. Perl et al., *Phys Rev. Lett.*, **35**(1975)1489.
- [2] W. Marciano, Second Workshop on τ physics, 1992.
- [3] B. Adeva et al., *L3 Nucl. Instr and Meth.*, **A289**(1990)35.
- [4] S. Jadach, B. Ward, Z. Was, *Comp Phys. Comm.* **66**(1991) 276.
- [5] R. Brun, et al., GEANT 3 Users' Guide, CERN DD/EE/84-1.
- [6] P. Wasserman, *Neural Computing: Theory and Practice* Van Nostrand Reinhold, N. Y, 1989.
- [7] Mariond Conference, France, March, 1994.

Identifying Tau Hadronic Decay with Neural Network

Chen Guoming Chen Gang

(The Institute of High Energy Physics, The Chinese Academy of Sciences, Beijing 100039)

Received 30 June 1994

Abstract

The identification of tau one prong hadronic decay using neural network is presented. Based on the identification, we measured the branching ratios: $Br(\pi/K\nu) = (12.18 \pm 0.26 \pm 0.42)\%$, $Br(\pi/K\pi^0\nu) = (25.20 \pm 0.35 \pm 0.50)\%$, $Br(\pi/K2\pi^0\nu) = (8.88 \pm 0.37 \pm 0.38)\%$, $Br(\pi/K3\pi^0\nu) = (1.70 \pm 0.24 \pm 0.39)\%$.

Key words collider, particle identification, neurae net work, decay branching ratio.