

# 夸克胶子喷注的人工神经网络识别研究<sup>\*</sup>

张昆实<sup>1,2</sup> 刘连寿<sup>1;1)</sup>

1(华中师范大学粒子物理研究所 武汉 430079)

2(长江大学物理科学与技术学院 荆州 434020)

**摘要** 为了将人工神经网络用于高能物理中对喷注的分类识别,用从高能正负电子对撞的蒙特卡罗模拟中得到的不对称三喷注事件中的夸克喷注和胶子喷注的平均多重数、平均横动量和两类喷注所对的夹角的平均值作为输入 BP 神经网络的 3 个特征参量,对 2.5—22.5 GeV 能区的 8 个能量间隔进行等精度的训练.用训练好的神经网络模型对不对称三喷注事件中的夸克喷注和胶子喷注样本进行检验判定,并对混合喷注样本进行分类识别.所得结果表明,有望将人工神经网络用于高能正负电子对撞产生的喷注的分类分析.

**关键词** 人工神经网络 BP 神经网络模型 模式识别 高能电子-正电子碰撞 夸克喷注 胶子喷注

## 1 引言

高能强作用的基本理论量子色动力学(QCD)是量子非阿贝尔规范场论.由于色禁闭,它的基本单元——夸克和胶子不能在实验中被观察到.一旦在实验中产生了夸克和胶子,它们就会演化而最终碎裂为末态强子.当尺度  $Q^2$  大时,演化的动力学可以用微扰 QCD 计算;而当  $Q^2$  小时,演化的动力学是非微扰的,不能解析计算.因此,对夸克和胶子演化、碎裂为强子喷注的过程进行唯象的对比研究非常重要,它有助于加深了解微扰与非微扰 QCD 的动力学性质.所以,对夸克喷注和胶子喷注的差别的研究在国际上很受重视.

研究夸克喷注和胶子喷注的差别的关键是区分这两种喷注. AMY 实验组用角度法则<sup>[1]</sup>,将和喷注间的最大夹角相对的喷注判定为胶子喷注; JADE 实验组用能量法则<sup>[2]</sup>,认为在三喷注事件中,能量最低的喷注主要是胶子喷注.这些方法都是选用某一种参数来识别对象,纯度都不高. OPAL 实验组用 b-tag

方法<sup>[3]</sup>,将能量最高的喷注确定为夸克喷注,再用反 B 标定法从剩下的两个喷注中决定胶子喷注.但这种方法不适用于 u, d, s 夸克发射胶子的情况,而且效率很低.

近十多年来在国际上迅速发展起来的人工神经网络(ANN),是一种模拟人脑神经系统的具有非线性动力学特征的信息处理系统,特别适合于处理需要同时考虑许多因素和条件的,不精确和模糊的信息处理问题<sup>[4]</sup>.现已在包括分类和智能模式识别在内的许多领域中得到了广泛应用<sup>[5]</sup>,并显示出其独特的优越性.对夸克喷注和胶子喷注的区分,实质上是高能物理中的模式识别问题.本文研究将人工神经网络用于对喷注的分类识别,探讨用夸克喷注和胶子喷注的多个特征参量训练人工神经网络,以期得到一个能够以较高的纯度和效率来辨别夸克喷注和胶子喷注的人工神经网络模型.

## 2 人工神经网络与 BP 学习算法

近年来,在智能模式识别中应用得最广泛的是

2004-05-08 收稿

<sup>\*</sup> 国家自然科学基金(10375025),湖北省高等学校优秀中青年科技创新团队计划项目(16)和湖北省教育厅重点项目(2003A002)资助

1)E-mail: liuls@iopp.cnu.edu.cn

具有误差反向传播算法 (back propagation algorithm) 的人工神经网络<sup>[6]</sup>, 简称为 BP 神经网络模型. 具有单隐层的 BP 神经网络如图 1 所示. 它由各有多多个神经元 (又叫节点) 的输入层、隐层和输出层组成. 相邻各神经元之间形成单向全互连接. 连接强度用权值表示. 各神经元有各自的阈值.

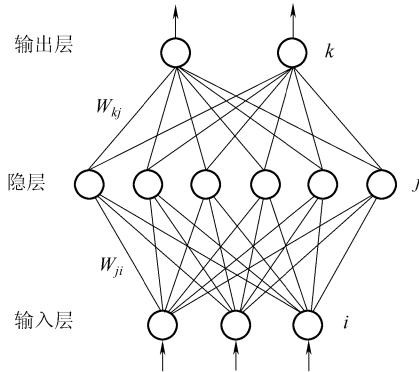


图 1 单隐层前馈式神经网络 BP 模型

要应用 BP 神经网络, 首先必须将它对拟被识别诸对象有代表性的特征参量进行学习. 学习过程由正向传播和反向传播组成. 正向传播是输入信号从输入层经隐层向输出层传播; 反向传播是误差信号从输出层经隐层向输入层传播, 并沿途调整各层间的权值和各神经元的阈值, 以使误差信号不断减小.

BP 学习算法的步骤如下<sup>[7]</sup>:

(1) 网络参数初始化. 给神经网络的各权值  $w_{ji}$ ,  $w_{kj}$  和阈值  $\theta_j$ ,  $\theta_k$  赋予  $[-0.1, 0.1]$  之间的随机小量作为初始值.

(2) 从输入层输入第  $l$  个学习样本信号集  $x_i(l)$  ( $i = 1, 2, \dots, n_1$ ),  $n_1$  为输入层神经元个数.

(3) 计算隐层神经元的输出值

$$h_j = f\left(\sum_{i=1}^{n_1} w_{ji}x_i + \theta_j\right), \quad (j = 1, 2, \dots, n_2), \quad (1)$$

式中  $w_{ji}$  为隐层神经元  $j$  与输入层神经元  $i$  之间的连接权值,  $\theta_j$  为隐层神经元  $j$  的阈值,  $n_2$  为隐层神经元个数,  $f$  为激活函数

$$f(x) = \frac{1}{1 + e^{-\beta x}}. \quad (2)$$

(4) 计算输出层神经元的输出值

$$O_k = f\left(\sum_{j=1}^{n_2} w_{kj}h_j + \theta_k\right), \quad (k = 1, \dots, n_3), \quad (3)$$

式中  $w_{kj}$  为输出层神经元  $k$  与隐层神经元  $j$  之间的连接权值,  $\theta_k$  为输出层神经元  $k$  的阈值.  $n_3$  为输出层神经元个数.

(5) 计算输出层神经元的输出误差

$$\delta_k(l) = [T_k(l) - O_k(l)] \cdot O_k(l) \cdot [1 - O_k(l)], \quad (4)$$

式中  $T_k(l)$  是第  $k$  个输出层神经元对第  $l$  个学习样本的期望输出,  $O_k(l)$  为对应的网络实际输出.

(6) 计算隐层神经元的输出误差

$$\epsilon_j(l) = h_j(1 - h_j) \sum_{k=1}^{n_3} \delta_k(l) w_{kj}(l). \quad (5)$$

(7) 由  $\delta_k$  向误差减小方向调整阈值  $\theta_k$  和权值  $w_{kj}$

$$\theta_k(l+1) = \theta_k(l) + \eta \delta_k(l) + \alpha \Delta \theta_k(l), \quad (6)$$

$$w_{kj}(l+1) = w_{kj}(l) + \eta \delta_k(l) h_j(l) + \alpha \Delta w_{kj}(l). \quad (7)$$

(8) 由  $\epsilon_j$  向误差减小方向调整阈值  $\theta_j$  和权值  $w_{ji}$

$$\theta_j(l+1) = \theta_j(l) + \eta \epsilon_j(l) + \alpha \Delta \theta_j(l), \quad (8)$$

$$w_{ji}(l+1) = w_{ji}(l) + \eta \epsilon_j(l) x_i(l) + \alpha \Delta w_{ji}(l). \quad (9)$$

以上诸式中的  $\eta$  为学习参数.  $l$  为当前学习的次数,  $l+1$  为学习的下一次循环. 含  $\alpha$  的项为动量项, 适当选择动量因子  $\alpha$  可使网络的学习过程加速收敛并且避免振荡发生.

当第一个学习样本经过上述正向和逆向运算后, 网络以调整后的权值和阈值作为新的网络参数接受下一个学习样本, 重复上述学习过程. 每学习一遍, 计算本次网络输出的均方误差并将各次的输出误差累加. 由于网络参数是按使误差梯度下降的方式调整的, 因此 BP 神经网络的训练学习朝误差减小的方向进行. 当学习到一定次数  $N$  以后, 网络的输出误差

$$E = \frac{1}{n_3 N} \sum_{l=1}^N \sum_{k=1}^{n_3} [O_k(l) - T_k(l)]^2 \quad (10)$$

将小于一个预先设定的精度  $\sigma$ , 于是学习结束并输出此时的各权值和阈值. 利用学习结束时更新的网络参数值和网络对不同识别模式的期望输出值, BP 网络模型就可以用来进行模式识别了.

用 FORTRAN 语言编制了上述单隐层前馈式神经网络 BP 模型的程序. 程序流程图如图 2 所示.

### 3 BP 神经网络的训练和对喷注的识别分类

#### 3.1 BP 神经网络的结构

BP 神经网络的输入、输出层中神经元的个数由所研究的实际物理问题的要求来决定. 神经网络作为模式分类识别器, 输入层神经元的个数为模式的输入特征参量的个数; 输出神经元的数目为模式的类别数. 用于识别高能  $e^+e^-$  对撞产生的不对称三喷

注事件中的夸克喷注和胶子喷注的神经网络,选取了这两类喷注的平均多重数,平均横动量和喷注所对夹角的平均值作为 3 个特征参量.因此 BP 网络模型的输入层有 3 个神经元;输出层有 2 个神经元,对应于胶子喷注和夸克(反夸克)喷注两类模式<sup>1)</sup>;隐层神经元的个数取为 6 个,以使神经网络有较高的分类精度.所选用的 BP 网络模型为 3-6-2 型的 3 层网络结构.

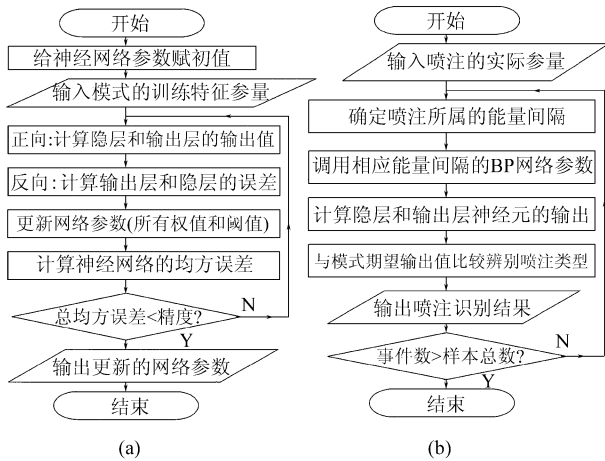


图 2 神经网络 BP 模型的程序流程图

- (a) 用夸克、胶子喷注的特征参量训练神经网络的流程图;
- (b) 用训练好的神经网络辨别夸克、胶子喷注类型的流程图.

### 3.2 BP 神经网络的训练

BP 神经网络的训练是决定其运用成功与否的决定因素.为了得到训练 BP 网络模型的学习参数,用 JETSET 7.4 蒙特-卡洛产生器产生 91.2 GeV 的  $e^+e^-$  碰撞事件.用 Durham<sup>[8]</sup> 法则挑选三喷注事件,将最小夹角所对喷注认同为原始夸克喷注,要求最大夹角和中等夹角之差大于  $30^\circ$  来挑选不对称事件<sup>[9]</sup>.按 Parton-Hadron-Duality 方法,在部分子和强子层次上通过配对<sup>[10]</sup>,分别得到胶子喷注样本和发射了胶子的夸克喷注样本,作为试验 BP 神经网络的检验样本.分别计算夸克喷注和胶子喷注的平均多重数  $\langle n_{jet} \rangle$ , 平均横动量  $\langle p_{t,jet} \rangle$  随喷注能量的变化(如图 3 和图 4 所示),以及这两类喷注所对夹角的平均值  $\langle \theta \rangle$ , 以这 3 个参量作为表征夸克喷注和胶子喷注特征的输入参量.为使神经网络进行有效的训练,同一组输入参量应有大致相同的数量级.为此要对角度参数进行预处理,用实际角度乘以 0.01 作为输入值.参见表 1 中  $\langle \theta \rangle$  的值.

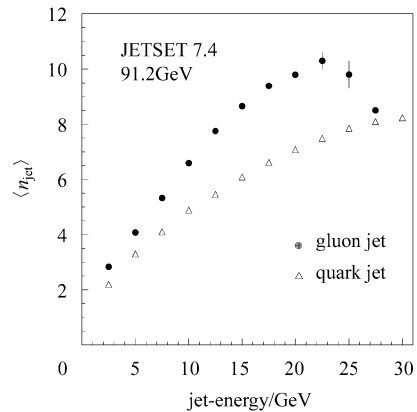


图 3 夸克喷注和胶子喷注的平均多重数随喷注能量的变化

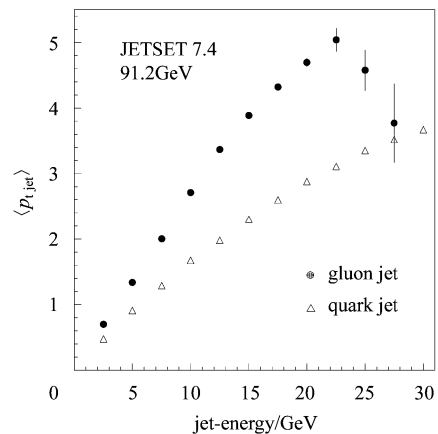


图 4 夸克喷注和胶子喷注的平均横动量随喷注能量的变化

由图 3 和图 4 可见,在不同的能量间隔里,夸克喷注和胶子喷注的  $\langle n_{jet} \rangle$  和  $\langle p_{t,jet} \rangle$  有不同的值.故必须分能量间隔对 BP 神经网络分别进行训练.将 2.5—22.5 GeV 能区均分成 8 个能量间隔.分别以每个能量间隔中夸克喷注和胶子喷注的 3 个特征参量值作为神经网络的输入,取学习参数  $\eta = 0.3$ , 动量因子  $\alpha = 0.3$ , 激活函数中的  $\beta = 2.5$ , 学习精度  $\sigma = 0.005$  来训练神经网络模型.以第 5 个能量间隔 (15.0—17.5 GeV) 为例,夸克喷注和胶子喷注的 3 个输入特征参量以及 BP 网络对应的期望输出和实际输出值见表 1.

当输入样本经 123976 遍 (247952 次) 学习后,网络的均方误差  $E < 0.005$ , 学习结束并输出学习结束时更新的  $n_1 \times n_2 + n_2 \times n_3 = 30$  个权值和  $n_2 + n_3 = 8$  个阈值,共 38 个网络参数.仿照这一过程对 8 个

1) 以下将夸克和反夸克喷注统称为“夸克喷注”

能量间隔分别进行相同精度( $\sigma$ )的学习训练,整个 BP 网络训练结束后,共产生 304 个网络参数.

表 1 在 15.0—17.5 GeV 能量间隔中胶子喷注和夸克喷注的输入特征参量、样本的期望输出和网络的实际输出值

喷注样本	$x_1$	$x_2/\text{GeV}$	$x_3$	预期输出		实际输出	
	$\langle n_{\text{jet}} \rangle$	$\langle p_{\text{t jet}} \rangle$	$\langle \theta \rangle (\times 10^2)^\circ$	$O_1$	$O_2$	$O_1$	$O_2$
胶子喷注	8.651	3.884	1.660	1	0	0.9982	0.0018
夸克喷注	6.088	2.302	1.150	0	1	0.0012	0.9987

### 3.3 BP 神经网络对喷注的分类识别结果

用训练好的 BP 神经网络模型对事先得到的夸克喷注和胶子喷注两个检验样本进行识别判断时,先根据每个喷注的能量,确定它属于 2.5—22.5 GeV 能量范围内的哪个能量间隔,再调用相应能量间隔中的 38 个网络参数对该喷注进行分类判断.经运行,实际分类判断的结果列于表 2 中.两种喷注的纯度都达到 80%,表明以上建立的神经网络有效.

表 2 用训练好的 BP 神经网络识别胶子和夸克喷注检验样本的结果

喷注样本	纯度 (%)	效率 (%)
胶子喷注样本	82.4	65.7
夸克喷注样本	79.6	79.3

在实际应用中,夸克喷注和胶子喷注是混在一起的.我们的任务正是要用神经网络来区分它们.因此,将夸克喷注和胶子喷注合成一个样本来用神经网络分类.同样先确定一个喷注属于 2.5—22.5 GeV 能量范围内的哪个能量间隔,再调用相应能量间隔中的 38 个网络参数对该喷注进行分类判断.结果列在表 3 中.

表 3 用训练好的 BP 神经网络识别混合喷注样本的结果

判断出的夸克喷注的纯度	判断出的胶子喷注的纯度	效率
75.5%	85.6%	70.6%

## 4 结论

本文讨论了用于辨别夸克喷注和胶子喷注的人工神经网络 BP 模型的结构、学习算法和训练过程.用 FORTRAN 语言编写了 3-6-2 型的 BP 神经网络程序.对 JETSET 7.4 蒙特卡罗产生器产生的  $e^+e^-$  对撞,要求最大夹角和中等夹角之差大于  $30^\circ$  来挑选不对称三喷注事件<sup>[9]</sup>.通过部分子和强子层次上的配对,分别得到发射了胶子的夸克喷注样本和胶子

喷注样本,作为试验 BP 神经网络的检验样本.用训练好的 BP 模型对夸克喷注和胶子喷注样本进行了检验判断,并对夸克-胶子喷注混合样本进行了分类.结果表明,将人工神经网络用于高能物理中的模式识别是有效和可行的.

本文的研究表明, BP 网络模型具有以下特征:

(1) 建立模式识别人工神经网络模型的关键是要对其进行有效的训练.一旦选定了夸克和胶子喷注的特征参量,就可以作为学习样本,让 BP 神经网络进行学习建模.因为神经网络具有自组织,自学习,自适应,容错和抗干扰的能力,能够抓住训练样本最本质的特征,所以根据在训练中所获取的知识,训练好后的 BP 模型就可以用来对夸克喷注和胶子喷注进行分类识别.结果客观可靠.

(2) 用已训练好的 BP 神经网络模型来判定喷注类型时,对每个喷注只要按(1)式和(3)式计算网络的实际输出,并与网络的期望输出进行比较就能作出判断.因而计算简单,运算速度快,使用方便,具有实用性.

(3) 作为学习样本的特征参量的选取对神经网络分类模式识别的质量有至关重要的影响.原则上,神经网络的输入参量的数量不受限制.而且,表征两类喷注的特征参量选得越多,不同特征参量之间的独立性越强,训练好后的神经网络进行分类识别的本领就越高.这是人工神经网络在模式识别中优于其他方法的一个显著特点.本文选择两类喷注的  $\langle n_{\text{jet}} \rangle$ ,  $\langle p_{\text{t jet}} \rangle$  和  $\langle \theta \rangle$  作为输入特征参量,仅是一种尝试.运行结果表明,用神经网络对喷注进行分类识别是有效的.但是,分类的纯度只达到 75%—85%,还有待于进一步提高.纯度不够高的原因,可能是由于所用的 3 个输入参量之间存在关联.例如,喷注多重数  $n_{\text{jet}}$  越高,喷注的横动量  $p_{\text{t jet}}$  也就越大.下一步要继续寻找独立性更强的,更多的特征参量,从多个方面来全面刻画两类喷注的本质特征.从而使人工神经网络具有更高的纯度和效率.进一步完善后的人工神经网络模型可望对高能正负电子对撞中产生的喷注进行可靠的、高纯度的分类,为研究夸克和胶子强子化过程的不同动力学特征创造条件.

(4) 在正负电子对撞产生的三喷注事件中,如果 3 个喷注间夹角都相等就称为“对称事件”;如果胶子喷注和发射了胶子的夸克喷注所对的喷注间夹角相等,而和另一个喷注间夹角(未发射胶子的夸克喷注所对的角)不等,就称为 Y 形事件;而如果 3 个喷注间夹角都不相等,就称为“不对称事件”.本文通过要求喷注

之间的最大夹角和中等夹角有大于  $30^\circ$  的差来挑选不对称事件,研究了人工神经网络对这种不对称事件中的夸克喷注和胶子喷注的分类识别.所得到的最大夹角、中等夹角和最小夹角的平均值比

$$\overline{\theta}_{\text{最大}} : \overline{\theta}_{\text{中间}} : \overline{\theta}_{\text{最小}} = 165.3^\circ : 116.8^\circ : 78.1^\circ \approx$$

2:1.5:1

表征了所用的三喷注事件样本的不对称性程度.对于 Y 形事件中的夸克喷注和胶子喷注,适当选用输入特征参量,也能用人工神经网络进行分类识别.这一研究工作正在进行中.

## 参考文献 (References)

- 1 AMY Coll. Kim Y K et al. Phys. Rev. Lett., 1989, **63**:17
- 2 JADE Coll. Bartel W et al. Phys. Lett., 1983, **B123**:460
- 3 OPAL Coll. Alexander G et al. Phys. Lett., 1991, **B265**:462
- 4 HU Shou-Ren et al. Neural Networks Introduction. Changsha: National University of Defense Technology Publishers, 1993 (in Chinese)  
(胡守仁等. 神经网络导论. 长沙:国防科学技术大学出版社, 1993)
- 5 Bishop C M. Neural Networks for Pattern Recognition. Oxford, UK: Oxford University Press, 1995
- 6 Ripsey B D. Pattern Recognition and Neural Networks. Cambridge, UK: Cambridge University Press, 1996
- 7 HAN Jia-Wei. Micheline Kamber, Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, 2002
- 8 Dokshitzer YU L. J. Phys., 1991, **G17**:1537
- 9 ZHANG Kun-Shi, CHEN Gang, YU Mei-Ling et al. HEP & NP, 2002, **26**(11):1110 (in Chinese)  
(张昆实, 陈刚, 喻梅凌等. 高能物理与核物理, 2002, **26**(11):1110)
- 10 YU Mei-Ling, LIU Lian-Shou. Chin. Phys. Lett., 2002, **19**:647

## On the Identification of Quark and Gluon Jets Using Artificial Neural Network Method\*

ZHANG Kun-Shi<sup>1, 2</sup> LIU Lian-Shou<sup>1;1)</sup>

1(Institute of Particle Physics, Huazhong Normal University, Wuhan 430079, China)

2(School of Physics Science and Technology, Yangtze University, Jingzhou 434020, China)

**Abstract** The identification of quark and gluon jets produced in  $e^+e^-$  collisions using the artificial neural network method is addressed. The structure and the learning algorithm of the BP(Back Propagation) neural network model is studied. Three characteristic parameters—the average multiplicity and the average transverse momentum of jets and the average value of the angles opposite to the quark or gluon jets are taken as training parameters and are inputted to the BP network for repeated training. The learning process is ended when the output error of the neural network is less than a pre-set precision ( $\sigma = 0.005$ ). The same training routine is repeated in each of the 8 energy bins ranging from 2.5—22.5 GeV, respectively. The finally updated weights and thresholds of the BP neural network are tested using the quark and gluon jet samples, getting from the non-symmetric three-jet events produced by the Monte Carlo generator JETSET 7.4. Then the pattern recognition of the mixed sample getting from the combination of the quark and gluon jet samples is carried out through applying the trained BP neural network. It turns out that the purities of the identified quark and gluon jets are around 75%—85%, showing that the artificial neural network is effective and practical in jet analysis. It is hopeful to use the further improved BP neural network to study the experimental data of high energy  $e^+e^-$  collisions.

**Key words** artificial neural network, BP network model, pattern recognition, high energy  $e^+e^-$  collisions, quark jet, gluon jet

Received 8 May 2004

\* Supported by National Natural Science Foundation of China (10375025), the Project of Team Research for Excellent Mid-Aged and Young Teachers of Hubei Province(16) and the Major Project of Hubei Provincial Department of Education(2003A002)

1)E-mail: liuls@iopp.ccnu.edu.cn