KEK Preprint 96-9
April 1996
H

Measurement of the forward-backward asymmetry in $e^+e^- \to b\bar{b}$ and the b-quark branching ratio to muons at TRISTAN using neural networks

AMY Collaboration

National Laboratory for High Energy Physics, 1996

KEK Reports are available from:

Technical Information & Library
National Laboratory for High Energy Physics
1-1 Oho, Tsukuba-shi
Ibaraki-ken, 305
JAPAN

Phone: 0298-64-5136
Telex: 3652-534 (Domestic)
(0)3652-534 (International)
Fax: 0298-64-4604
Cable: KEK OHO
E-mail: Library@kekvax.kek.jp (Internet Address)

Measurement of the forward-backward asymmetry in $e^+e^- \to b\bar{b}$ and the b-quark branching ratio to muons at TRISTAN using neural networks

AMY Collaboration

K.Ueno [a], S.Kanda [a], S.L.Olsen [a], P.Kirk [b], T.J.Wang [c], A.Abashian [d], K.Gotow [d], L.Piilonen[d], S.K.Choi [e], C.Rosenfeld [f], L.Y.Zheng [f], Winston Ko [g], R.E.Breedon [g], R.L.Lander [g], J.Rowe [g], K.Abe [h], Y.Fujii [h], S.Igarashi [h], Y.Kurihara [h], M.H.Lee [h], F.Liu [h], A.Maki [h], T.Nozaki [h], T.Omori [h], H.Sagawa [h], Y.Sakai[h], T.Sasaki [h], Y.Sugimoto [h], Y.Takaiwa [h], S.Terada [h], F.Kajino [i], T.Aso [j], K.Miyano [j], H.Miyata [j], N.Nakajima [j], K.Ohkubo [j], M.Sato [j], M.Shirai [j], N.Takashimizu [j], Y.Yamashita [k], S.Schnetzer [l], S.Behari [m], H.Fujimoto [m], S.Kobayashi [m], A.Murakami [m], M.Yang [m], S.K.Sahu [n], J.S.Kang [o], D.Y.Kim [o], S.S.Myung [o], H.S.Ahn [p], S.K.Kim [p], S.Matsumoto [q]

[a] *University of Hawaii, Honolulu, HI 96822, USA*
[b] *Louisiana State University, Baton Rouge, LA 70803, USA*
[c] *Institute of High Energy Physics, Beijing 100039, China*
[d] *Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA*
[e] *Gyeongsang National University, Chinju 660-701, South Korea*
[f] *University of South Carolina, Columbia, SC 29208, USA*
[g] *University of California, Davis, CA 95616, USA*
[h] *KEK, National Laboratory for High Energy Physics, Ibaraki 305, Japan*
[i] *Konan University, Kobe 658, Japan*
[j] *Niigata University, Niigata 950-21, Japan*
[k] *Nihon Dental College, Niigata 951, Japan*
[l] *Rutgers University, Piscataway, NJ 08854, USA*
[m] *Saga University, Saga 840, Japan*
[n] *National Taiwan University, Taipei 10764, Taiwan*
[o] *Korea University, Seoul 136-701, South Korea*
[p] *Seoul National University, Seoul 151-742, South Korea*
[q] *Chuo University, Tokyo 112, Japan*

The forward-backward asymmetry in $e^+e^- \to b\bar{b}$ at $\sqrt{s}$ =57.9 GeV and the b-quark branching ratio to muons have been measured using neural networks. Unlike previous methods for measuring the $b\bar{b}$ forward-backward asymmetry where the estimated background from c-quark decays and other sources are subtracted, here events are categorized as either $b\bar{b}$ or non $b\bar{b}$ events by neural networks based on event-by-event characteristics. The determined asymmetry is $-0.429 \pm 0.044(sta) \pm 0.047(sys)$ and is consistent with the prediction of the standard model. The measured $B\bar{B}$ mixing parameter is $0.136 \pm 0.037(sta) \pm 0.040(sys) \pm 0.002(model)$ and the measured b-quark branching ratio to muons is $0.122 \pm 0.006(sta) \pm 0.007(sys)$.

1. Introduction

The most widely used method of obtaining the $b\bar{b}$ forward-backward asymmetry is to enrich the $b\bar{b}$ event sample by a high-p_t lepton cut and to subtract the c-quark and other background contributions. LEP, PEP, PETRA, and TRISTAN have all used this method[1,2]. Because of the large number of events, LEP experiments can afford to make the p_t cut high and lessen the ambiguity due to the background subtraction. However TRISTAN experiments suffer from poor statistics if a high-p_t cut is used.

At the TRISTAN energy, because the $b\bar{b}$ asymmetry is nearly maximal, it can, in principle, be measured relatively precisely. In previous results, all three TRISTAN experiments favor a value of the asymmetry that is closer to the value with $B\bar{B}$ meson mixing not included, $A_b = -0.586$, than the standard model prediction with $B\bar{B}$ meson mixing, $A_b = -0.440$. However, the effect of $B\bar{B}$ meson mixing is clearly seen at LEP and hadron colliders[3]. Since the first publication on $b\bar{b}$ asymmetry by AMY[2], the number of collected events has increased by more than a factor of ten. With the corresponding reduced statistical error, it is desirable to use an improved method for obtaining the $b\bar{b}$ asymmetry with smaller systematic errors.

In the present study, $b\bar{b}$ events are identified by neural networks[4]. In this analysis, either the b or $\bar{b}$ quark is required to decay semileptonically to a muon. The input to the neural networks consists of information on both the muon and the hadrons produced in the event. As usual, the hadron related inputs include quantities such as the boosted sphericity and the invariant masses of certain combinations of particles. The muon related inputs include quantities such as the total and the transverse momentum of the muon. The muon related information is very effective in separating the $b\bar{b}$ events from the background events.

This method identifies events based on the event topologies and does not rely on estimation of the backgrounds. As a result, the difficulty of the non neural network approach associated with background subtraction is absent here.

2. Identification of the $b\bar{b}$ events

This analysis used all the data taken in the AMY experiment and corresponds to an integrated luminosity of 338 pb^{-1}. For selection of hadronic events with leptons, only the muon identification analysis was used since electron identification in hadronic events has a much larger background. The AMY detector and the hadron event selection criteria are described elsewhere [5]. The muon identification system consisted of four layers of drift chambers and an array of scintillation counters placed outside of 9.8 nuclear absorption lengths of steel. A charged track was identified as a muon if the track in the central drift chamber extrapolated to a hit in the muon chambers within a 2 σ radius determined by multiple scattering. The efficiency was 99.5 % for muons with momentum greater than 1.9 GeV/c and with $\cos\theta_\mu < 0.7$. Details of the muon detection system and the muon identification analysis are described elsewhere [6].

For this analysis, hadronic events were selected by the process described in reference [8]. The number of hadronic events with an identified muon is 1625 with an average center-of-mass energy of 57.9 GeV. The particles in each of these events were divided into two hemispheres, H_1 and H_2, with H_2 defined as the hemisphere containing the muon with the largest transverse momentum with respect to the thrust axis. At least four particles were required to be present in each hemisphere. The angle θ was defined as the angle of the thrust axis of the particles in the H_2 hemisphere if the associated muon were negative and as the angle of the thrust axis of the particles

in the H_1 hemisphere if the muon were positive. If there were more than one muon in a hemisphere, the muon with the largest momentum transverse to the beam axis was used. The muon momentum was not included in the determination of θ.

Monte Carlo hadronic events were generated by the LUND7.3 parton shower model[7] with default parameter settings and simulated by the standard AMY simulator. The $B\bar{B}$ mixing was not included in the event generation. The passage of particles through the muon detection system was simulated by the GHEISHA[8] program. The generated $x = E_{hadron}/E_{beam}$ distributions for c and b hadrons can be parametrized by the Peterson fragmentation function[9],

$$D(x, \epsilon_q) = \frac{N}{x[1 - 1/x - \epsilon_q/(1-x)]^2}, \quad (1)$$

where $\epsilon_c = 0.18$ and $\epsilon_b = 0.006$.

Two neural networks were used in a cascaded way as shown schematically in Fig. 1. The first network was trained using information from hemisphere H_1 as follows:

- $E_v/\sqrt{s}$ where $\frac{1}{2}E_v$ is the visible energy observed in the hemisphere.
- $\sum p_t^2$ where p_t is the momentum transverse to the thrust axis.
- The invariant mass of the four highest momentum particles.
- The sphericity of the four highest momentum particles.
- The invariant masses of each three particle combination of the four highest momentum particles.
- The sphericities of each three particle combination of the four highest momentum particles.

- $\int_{0.0}^{4.2} EEC\delta(y_{ij}-y)dy$, $\int_{0.0}^{1.0} EEC\delta(y_{ij}-y)dy$, $\int_{1.0}^{2.2} EEC\delta(y_{ij}-y)dy$, $\int_{2.2}^{3.2} EEC\delta(y_{ij}-y)dy$ where $EEC = \sum E_i E_j / E_v^2$.
- muon p_t^μ if there was a muon in H_1.

The summation is over particles in the hemisphere H_1. The integration is over rapidity with $y_{ij} = -ln(\theta_{ij}/2)/2$ where θ_{ij} is the angle between two particles of energies E_i and E_j. Most of these quantities are known to be effective in separating $b\bar{b}$ events from light-quark events as described in the papers given in [10]. The EEC distributions are shown in Fig. 2.

After the first network was trained, its output was fed into the second network together with the total p_t^2 and the following muon-related information from hemisphere H_2:

- Muon p_t^μ.
- Muon momentum p_o^μ.
- $M_1 \equiv \int_{0.8}^{4.2} y^2 EMC\delta(y_{i\mu}-y)dy / \int_{0.8}^{4.2} EMC\delta(y_{i\mu}-y)dy$, where $EMC = \sum E_\mu / E_v$.
- $M_2 \equiv \int_{0.8}^{4.2} yEEC\delta(y_{i\mu}-y)dy / \int_{0.8}^{4.2} EEC\delta(y_{i\mu}-y)dy$, where $EEC = \sum E_i E_\mu / E_v^2$.
- $M_3 \equiv \int_{0.8}^{4.2} y^2 EEC\delta(y_{i\mu}-y)dy / \int_{0.8}^{4.2} EEC\delta(y_{i\mu}-y)dy$.

Here $y_{i\mu}$ is the rapidity between particle i and the muon.

Hadronic events with an associated muon were classified into three categories. Category A consisted of all non $b\bar{b}$ events that yielded a muon signal in the detector including $c\bar{c}$ events. Category B consisted of $b\bar{b}$ events with a direct $b \to \mu$ decay. Category C consisted of $b\bar{b}$ events with indirect muon production such as the cascade decays $b \to c \to \mu$ and $b \to \tau \to \mu$. A complement A^* of category A was defined as a composite of categories B and C. Complements B^* and C^* were defined similarly.

Because of the limited number of Monte Carlo events with associated muons, this cascading of the two networks was used so that the first network operating on the hadron related information could be trained with a larger sample of events. The first network was trained using 30,000 hadronic samples with an equal number of $b\bar{b}$ and non-$b\bar{b}$ events. For each hadronic event, two samples corresponding to the two hemispheres were used for training. The second network was trained using 2,400 hadronic events with an associated muon with equal amounts of the three categories A, B, and C. This network had three outputs corresponding to the likelihood that the input event was category A, B or C with an output of 1 corresponding to maximum likelihood. For each event, the sum of the three outputs was normalized to 1. The results of the training are shown in Fig. 3 where each point corresponds to an event for which the distances from the three sides of the triangle are equal to the three outputs of the networks.

When the output of the first network is greater than 0.5 the event is identified as a $b\bar{b}$ event. Using only this first network, the identification efficiency for non $b\bar{b}$ and $b\bar{b}$ events is 64 % and 62 %, respectively. Note that this performance is obtained using the information from only one hemisphere of a hadronic event. With the second network, the efficiency for identifying events with the muon $p_t^\mu \geq 0.6$ GeV/c as B and B^* is 84% and 86%, respectively, where each event is identified to be in category B or B^* when the normalized output $O_B/(O_A + O_B + O_C)$ is greater or less than 0.24, respectively. This identification capability is sufficient because requiring a muon in the event already enriches $b\bar{b}$ events. Using the muon-related information in the second network, greatly improves the identification of $b\bar{b}$ events.

3. Determination of the $b\bar{b}$ cross-section

Once the neural networks were trained, the outputs of the networks were treated as quantities associated with the event. The definition of B^* and B as defined above gave about the same number of events identified as B^* and B for the data and the Monte Carlo at $p_t^\mu > 0.4$ GeV/c. The efficiencies e_{B^*} and e_B for identifying events as B^* or B were calculated as functions of $\cos\theta$.

Since the efficiencies are not 100%, the numbers N_B and N_{B^*} identified as categories B and B^* are not the true numbers of B and B^* events; n_B and n_{B^*}. The following relations were inverted to solve for n_B and n_{B^*}:

$$N_B = (1 - e_{B^*}) \times n_{B^*} + e_B \times n_B. \tag{2}$$

$$N_{B^*} = e_{B^*} \times n_{B^*} + (1 - e_B) \times n_B, \tag{3}$$

The calculation was done on a bin-by-bin basis in $\cos\theta$.

The cross-section of the process $e^+e^- \to b\bar{b} \to \mu\nu X$ was obtained using the same procedure as in [6]. The angular distribution was divided by the luminosity and the muon detection efficiency and was multiplied by a factor which is a pure Born term cross-section divided by the full detector simulation. The cross-section which is shown in Fig. 4 was fit to the formula:

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4s} R_b B(b \to \mu\nu X)(1 + \cos^2\theta + \frac{8}{3} A_b \cos\theta), \tag{4}$$

where α is the fine structure constant, R_b is the total $b\bar{b}$ cross-section in units of

the lowest order $e^+e^- \to \mu\mu$ cross-section, B is the branching ratio for b-quark decays to muons, and A_b and R_bB are free parameters. The fitted values are $A_b = -0.429 \pm 0.044$ and $R_bB = 0.0682 \pm 0.0032$.

Table 1 shows a summary of the systematic errors for the measurements of the asymmetry A_b and the cross-section product R_bB. The uncertainty in the luminosity was 1.8%. A study of the systematic errors was made as follows: Quantities in the LUND7.3 Monte Carlo were varied from default values by 30%. The parameters ϵ_c, ϵ_b of the Peterson fragmentation function were varied by $\pm 50\%$. This changes the momentum spectra of particles produced from the c and b quark fragmentation. The amount of the background due to the u, d, s, c-quarks was varied by $\pm 30\%$. The asymmetry of the c and b-quarks was varied by $\pm 30\%$. The ratio of the direct to the cascade decay processes was varied by $\pm 30\%$. The cut on output O_B of the second network was varied by ± 0.04. The muon track matching criteria for the muon detection system was changed from 2σ to 3σ. The muon p_t^μ cut was varied from 0.0 to 0.6 GeV/c.

The values of the $b\bar{b}$ asymmetry of the three previous TRISTAN measurements are shown in Fig. 5. Although the systematic errors are large, they are consistent with each other. If the three measurements are combined, the asymmetry is -0.57 ± 0.10 and is closer to the value without $B\bar{B}$ meson mixing but is also consistent with mixing. The present new AMY measurement also shown in Fig. 5 has improved statistical and systematic errors and is consistent only with the standard model with $B\bar{B}$ meson mixing included.

We also fit the center of mass energy dependence of the asymmetry to obtain the mixing parameter χ by the formula:

$$A_b = A_{Born}(1 - 2\chi), \quad (5)$$

where A_{Born} is the $b\bar{b}$ asymmetry obtained from the Born term without $B\overline{B}$ mixing. Here we assumed χ to be constant with energy. All available measurements at TRISTAN, PEP, PETRA and LEP were used in the fit. We used values of 91.2 GeV, 2.49 GeV and 0.232 for the mass of Z, the width of Z and the weak mixing angle, respectively[11]. The obtained mixing parameter is $0.130 \pm 0.028(sta + sys) \pm 0.020(model)$. The error labelled as model is obtained by varying the mass of Z, the width of Z and the weak mixing angle by 0.007 GeV, 0.007 GeV and 0.002, respectively. The error is comparable with other methods[3]. The curves in Fig. 5 are for the Born term prediction and for the fit. Using only our data, the value of χ obtained is $0.136 \pm 0.037(sta) \pm 0.040(sys) \pm 0.002(model)$.

The standard model predicts $R_b = 0.560$ with an uncertainty of 0.002. Using the measurement described above for $R_b B$, we calculate the branching ratio of b quarks to muons to be $0.122 \pm 0.006(sta) \pm 0.007(sys)$ where model errors are neglected. Conversely, if we use the LEP average[11] of 0.110 for this branching ratio, R_b is calculated to be as $0.620 \pm 0.029(sta) \pm 0.035(sys)$.

4. Conclusion

The $b\bar{b}$ forward-backward asymmetry was measured by a new method using neural networks. The input to the neural networks included muon related information as well as the conventional hadron related information. The muon information is effective for separating $b\bar{b}$ events from the background. The determined value of the asymmetry is $A_b = -0.429 \pm 0.044(sta) \pm 0.047(sys)$. It is consistent with the

prediction of the standard model. In addition, the determined value of the mixing parameter χ is $0.136 \pm 0.037(sta) \pm 0.040(sys) \pm 0.002(model)$ and the branching ratio of b quarks to muons is $0.122 \pm 0.006(sta) \pm 0.007(sys)$.

Acknowledgement

We thank the TRISTAN staff for the excellent operation of the storage ring. We would like to thank Prof. L.Lonnblad for providing us the neural network program. This work has been supported by the Japan Ministry of Education, Science and Culture and by the Japanese Society for the Promotion of Science, the US Department of Energy and National Science Foundation, the Korean Science and Engineering Foundation and Ministry of Education, and the Academia Sinica of the People's Republic of China.

Reference

[1] DELPHI collab., P.Abreu et al., Phys. Lett. B276 (1992) 536. JADE collab., W.Bartel et al., Phys. Lett. B146 (1984) 437. TASSO collab., W.Braunschweig et al., Z. Phys. C48 (1990) 433. TOPAZ collab., K.Nagai et al., Phys. Lett. B278 (1992) 506. VENUS collab., M.Shirakata et al., Phys. Lett. B278 (1992) 499.

[2] AMY collab., H.Sagawa et al., Phys. Rev. Lett. 63 (1989) 2341.

[3] ALEPH collab., D.Buskulic et al., Phys. Lett. B284 (1992) 177. L3 collab., B.Adeva et al., Phys. Lett. B288 (1992) 395. DELPHI collab., P.Abreu et al., Phys. Lett. B301 (1992) 145. OPAL collab., R.Akers et al., Z. Phys. C60 (1993) 199. UA1 collab., C.Albajar et al., Phys. Lett. B262 (1991) 171. CDF

collab., F.Abe et al., Phys. Rev. Lett. 67 (1991) 3351.

[4] L.Lonnblad et al., Nucl. Phys. B349 (1991) 675. L.Lonnblad et al., Comput. Phys. Commun. 70 (1992) 167. C.Peterson, LU-TP 90-6.

[5] AMY collab., T.Kumita et al., Phys. Rev. D42 (1990) 1339.

[6] AMY collab., F.Liu et al., Phys. Rev. D49 (1994) 4339.

[7] T.Sjostrand et al., CERN-TH-6488-92.

[8] H.Fesefeld, Nucl. Instrum. Methods A263 (1987) 114.

[9] C.Peterson et al., Phys .Rev. D27 (1983) 107.

[10] L.Bellantoni et al., Nucl. Instrum. Methods A310 (1991) 618. C.Bortolotto et al., Nucl. Instrum. Methods A306 (1991) 459. P.Abreu et al., Phys. Lett. B295 (1992) 383. G.Marchesini et al., Nucl. Phys. B394 (1993) 541.

[11] Particle Data Group, Phys. Rev. D50 (1994).

Figure caption

Fig. 1. The neural networks were used in a cascaded way. The first network used mostly hadron-related information as input. The second network used the output of the first network in addition to the muon related information as input.

Fig. 2.(a) The energy-energy correlation EEC as a function of rapidity y. (b): The third muon-related quantity M_1. (c): The fourth muon-related quantity M_2. (d): The fifth muon-related quantity M_3. M_1, M_2 and M_3 are defined in the text. The solid and dotted lines are for the events B^* and B, respectively.

Fig. 3.(a) The distribution of the three outputs O_i of the second network for the Monte Carlo category A shown as a Dalitz plot. See text. (b): The same as (a) but for the Monte Carlo category B. (c): The same as (a) but for the Monte Carlo category C. (d): The same as (a) but for the data.

Fig. 4. The cross-section of the process $e^+e^- \to b\bar{b} \to \mu\nu X$ for $p_t^\mu \geq 0.4$ GeV/c. The curve is a fit to the data. The asymmetry A_b is -0.429 ± 0.044.

Fig. 5. The present measurement of the asymmetry A_b together with other experiments. The statistical and systematic errors are added in quadrature. The two curves are the Born term prediction without mixing(broken line) and the fit to the data(solid line) with mixing parameter χ. See the text.

Table 1: Systematic error

Error sources	$1000 \times \Delta A_b$	$10000 \times \Delta R_b B$
$\epsilon_c \times 0.5 - 1.5$	6	2
$\epsilon_b \times 0.5 - 1.5$	2	1
$M_{u.d.s} \times 0.7 - 1.3$	1	0
$M_c \times 0.7 - 1.3$	7	6
$M_b \times 0.7 - 1.3$	6	4
$A_c^m \times 0.7 - 1.3$	4	1
$A_b^m \times 0.7 - 1.3$	0	0
$M(direct)/M(cascade) \times 0.7 - 1.3$	7	5
cuts on network output	10	11
μ selection	23	27
p_t cut $0.0 - 0.6$GeV/c	28	21
Monte Carlo statistics	25	4
luminosity		12
Total	47	39

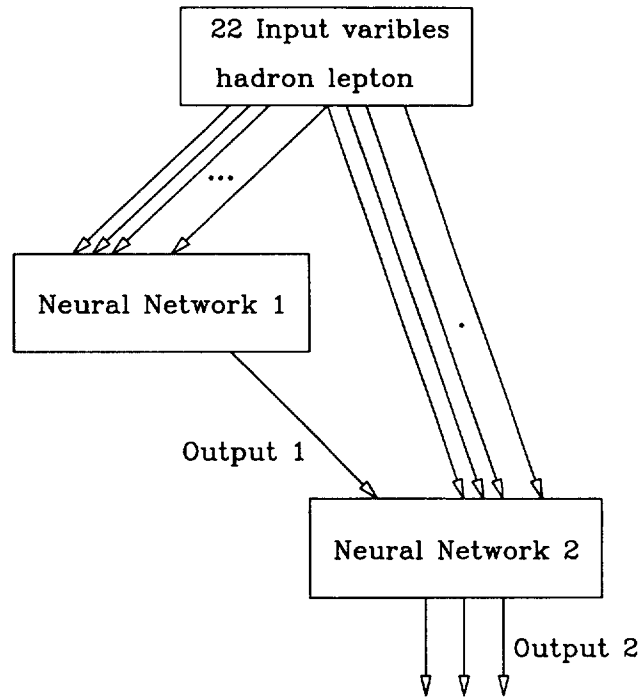


Fig. 1

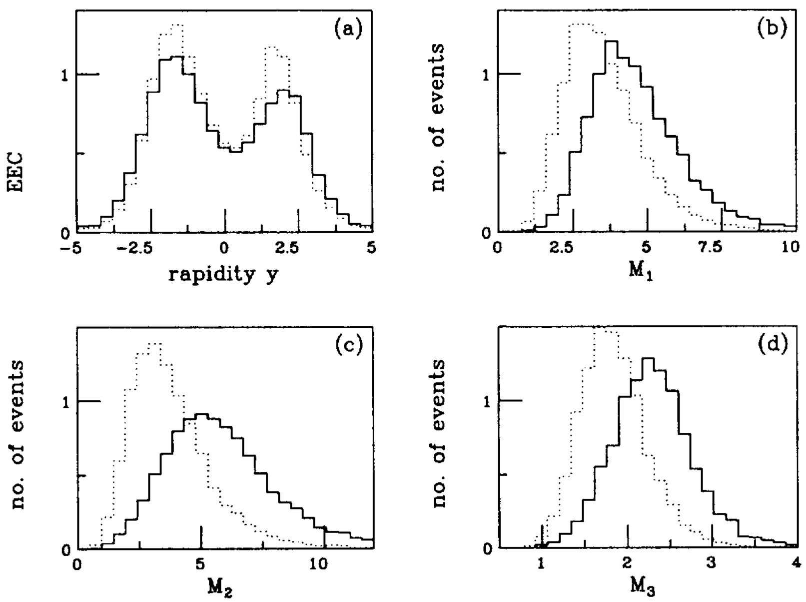

(a)
EEC
rapidity y
(b)
no. of events
M1
(c)
no. of events
M2
(d)
no. of events
M3

Fig. 2

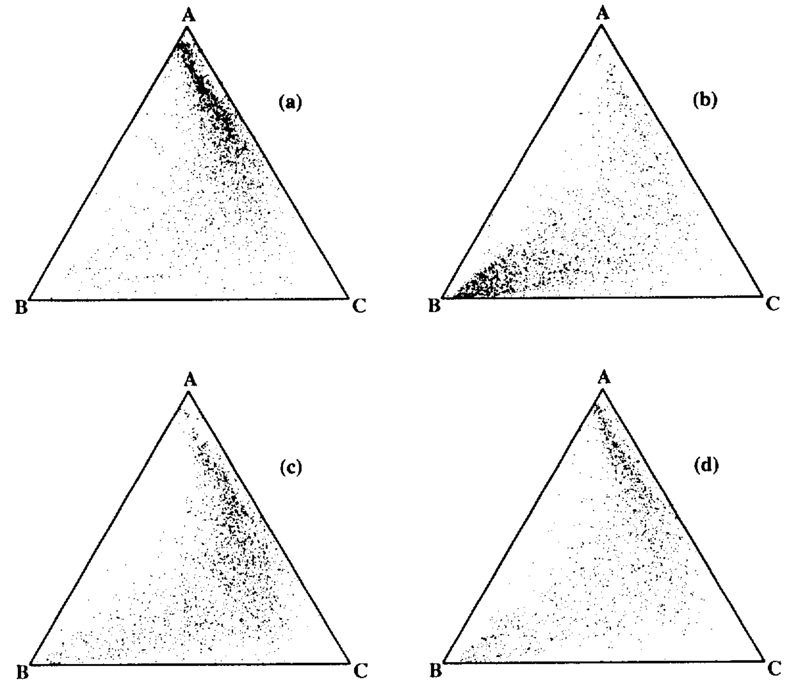
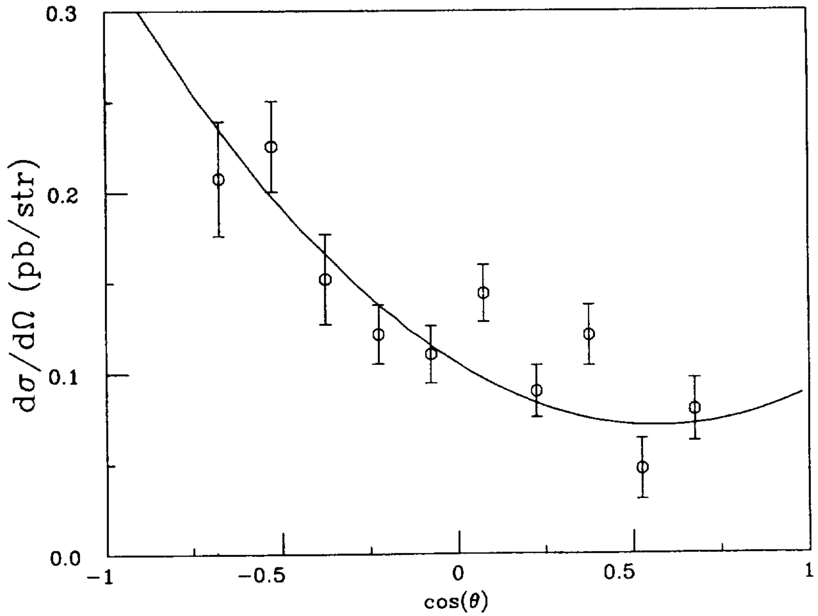
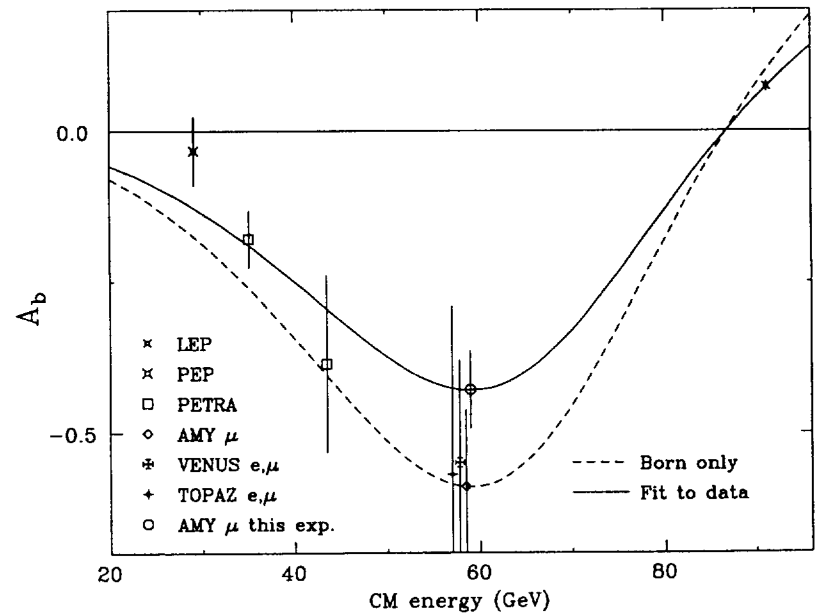

A
B
C
(a)
(b)
(c)
(d)

Fig. 3



0.3
0.2
0.1
0.0
dσ/dΩ (pb/str)
−1
−0.5
0
0.5
1
cos(θ)

Fig. 4



0.0
−0.5
A_b
LEP
PEP
PETRA
AMY μ
VENUS e,μ
TOPAZ e,μ
AMY μ this exp.
Born only
Fit to data
20
40
60
80
CM energy (GeV)

Fig. 5