ALEPH 93 - 61
PHYSIC 93 - 51
F. Saadi et al.
24.3.1993

March 16, 1993

Measurement of the Partial Width of the Z into b Quark Pairs using a Neural Netwok classifier with Jet shape variables.

F. Saadi, A. Falvard, P. Henrard and P. Perret

University Blaise Pascal-Clermont II/IN2P3

1 Introduction

The aim of this Aleph Note is to describe in a more detailed way the analysis made to measure the Z into $b\bar{b}$ partial width $R^{(b)} = \Gamma_{b\bar{b}}/\Gamma_{had}$ using an Hemisphere Double Tag method with a Neural Network to identify b-jet, which is part of the draft "Measurement of the Partial Width of the Z into $b\bar{b}$ using Event Shape Variables" circulating in the Collaboration.
Our 'official' result will be given for the statistic of 1990 and 1991 since the Heavy Flavour lepton analysis (which is needed to estimate the b-purity in the high $p_\perp$ lepton sample) is not yet available for '92 data.

2 Using a Neural Network to tag b quark events

The Neural Network technique is now widely used in many fields of high energy Physics: to discriminate between quark and gluon jets [1], for the separation of quark flavours in e^+e^- annihilations [2, 3, 4, 5] for the identification of Tau decays [6], for the Higgs [7] and top quark [8] search, for charm physics [9], in the charged particle track reconstruction [10] and also as a first level trigger for futur high energy hadronic colliders [11].

In ref. [5], it has been shown that Multi-layered Neural Networks are very well suited to tag b-quark events at LEP. More precisely, we have shown that multivariate analyses (Neural Network, Linear Discriminant Analysis, ...) can significantly improve the b-tagging compared to single-variable analyses (for instance the Boosted Jet Sphericity Product), and that among the different multivariate approaches, the Neural Network technique which is a non-linear method [1] (i.e. the output given by

[1] This non-linearity is provided by one or several hidden layers.

the Neural Network is not a linear function of the input variables) maps in a better way the complexity of the problem (i.e. the fact that there is not a linear separation in the feature space between the two classes: b and non-b events). This conclusion remains valid when the classification problem becomes almost linear i.e. when one of the variables used as input of the Neural Network is very discriminant. This is the case if we use for instance vertex information from QIPBTAG [12] to tag b-quarks. But even in this limit, Neural Network can help by rejecting the remaining charm background (fig. 1).

3 Choice of the Neural Network

We have chosen a feed forward multi-layered Neural Network trained with backpropagation of the errors. Other models exist such that the Kohonen and the Hopfield models [13] but the multi-layered approach is the best suited to solve classification problems [5].
The number of hidden layers and the number of neurons per layer have been chosen to optimize the separation between b and non-b events. This led to the following structure: one input layer with 9 neurons associeted to 9 physics variables, two hidden layers with 9 and 6 neurons respectively, and one final layer with 1 neuron giving the output of the Network normalized between 0. (target value for non-b events) and 1. (target value for b events). Each neuron of a given layer is connected to all the neurons of the following layer. To each connection is associated a weight W which is determined during a supervised learning phase. Each neuron i performs a weighted sum y_i of the output values x_j from all the neurons j of the previous layer; its output O_i is then computed via the sigmoïd function:

$$O_i = g(y_i) = \frac{1}{1+e^{-y_i/T}} \quad with \quad y_i = \sum_j W_{ij} x_j \tag{1}$$

The coefficient T is usually called "Temperature" and has been fixed to 1. for all the layers. For more technical details, we refer to the ref. [5].

For the training of the network, we have used 7000 $b\bar{b}$, 7000 $c\bar{c}$ and 7000 $u\bar{u}$. $d\bar{d}$. $s\bar{s}$ fully simulated events. They have been obtained by using the standard ALEPH Monte Carlo program HVFL based on DYMU and JETSET 7.3 [14, 19] to simulate the reaction $e^+e^- \rightarrow Z \rightarrow q\bar{q}$. Afterwards, a detailed detector simulation is performed and the events are reconstructed in the same way as it is done for data.
The initial weights of the connections were chosen randomly between [-0.01,0.01]. One event of each class is presented to the Network. We feed forward and backpropagate the errors. The reactualisation of the weights W is done after one exposure of an event of each class by minimizing a cost function E according to the formula [5]:

$$\Delta W(k+1) = \alpha \Delta W(k) + \eta \frac{\partial E}{\partial W}$$

where k is the number of exposures and

$$E = \frac{1}{2} \sum_{events} (\theta - t)^2.$$

θ is the obtained output for a given example and t the target value. The parameter α has been fixed to 0.5 while η decreases linearly from 0.01 to 0.0001 during the learning phase. The training is then stopped when the performance of the Network ceases to improve significantly, i.e., when the function E reaches an asymptotic minimal value. This corresponds roughly to 2 millions exposures. Furthermore, we have checked that a change of the relative fractions of $b\bar{b}$ events from 33% to 50% in the learning set and the order in which we present each class of events do not bias the result of the training. So, the learning of the Neural Network is independent of the value of $R^{(b)}$ which is in the Monte Carlo.
This training procedure has been tested in a validation step with a sample of about 200,000 fully simulated hadronic Z decays. different from those of the learning sample. In particular, we have checked that the Network performances (b-purity versus b-efficiency) are very similar in the two phases.

4 The input variables

4.1 Tagging of b quarks

The large mass of the b quark has several important consequences:
– b quarks lose less energy by gluon bremsstrahlung than light ($udsc$) quarks.
– The fragmentation of b-quarks is harder.
– The track multiplicity of the two most energetic jets is higher for a $b\bar{b}$ than for a light quark events [15].
Thus, the fraction of the beam energy carried by B-hadrons is 70% on average and only 51% for D-hadrons produced in $c\bar{c}$ events, resulting in different topologies for $b\bar{b}$ and light quark events. In particular, $b\bar{b}$ events will appear more spherical in th detector than light quark events and the particles produced in $b\bar{b}$ events will have specific momentum p and transverse momentum $p_\perp$ distributions [2].

4.2 Definition of shape variables

Taking advantage of these characteristics a set of 70 purely kinematical variables has been defined. Two types of variables have been used: variables based on the full hemisphere [3] shape, and variables based on the properties of the most energetic jet

[2]The transverse momentum $p_\perp$ of a track is defined with respect to its jet axis.

[3]The two hemispheres of the events are defined w.r.t. the plane perpendicular to the thrust axis computed by using all the tracks (charged and neutral) of the event.

of each hemisphere. All the variables and axes (thrust and jet axes) used in this analysis are defined by using all 'good' charged tracks, with neutral electromagnetic and hadronic clusters. A 'good' charged track is one that passes through a cylinder of 2 cm radius and 20 cm length around the interaction point, has at least four TPC coordinates, a transverse momentum relative to the beam axis $p_{\perp}^{x,y} \geq 0.25$ GeV, a polar angle between 18° and 162° and a momentum between 0.2 and 100. GeV. The last three cuts are also applied to the neutral tracks. Jets are reconstructed with the scaled-invariant-mass clustering algorithm [16] with the parameter y_{cut} set to $(6.\ GeV/E_{vis})^2$ where E_{vis} is the total energy of each hemisphere as reconstructed by the energy flow algorithm.

Furthermore, some acceptance cuts have been done to restrict the analysis to the region of good resolution and acceptance and to ensure a proper definition of the variables:
– the polar angle of the thrust axis must be in the range $30° \leq \theta_t \leq 150°$,
– the most energetic jet of each hemisphere must have at least 4 'good' energy flow tracks to avoid tails in the $p_{\perp}$ distributions (fig. 2, 3),
– E_{vis} must exceed 5 GeV for both hemispheres,
– only particles with an angle of less than 70^0 with the thrust axis are used to build the input variables.
The last cut is introduced to reduce the correlations between the two hemispheres (called 'Double Tag Correlations' in the following).

4.3 Selection of Input variables

Although there is *a priori* no restriction on the number of variables used as input of the Neural Network, it is obvious that a small number will lead to a more manageable and less time consuming learning. Starting from the original set of 70 variables, we have selected nine variables with the help of a F-test taking into account the discrimination power of each variable and its correlation to the others inside the same hemisphere [17] [4]. The nine selected variables $A(I)_{F-value}$ $(I = 1, 9)$ are the following (for a more detailed description, see ref. [5] and the references given subsequently):
– $A(1)_{869}$ is the boosted sphericity ($\beta_{boost} = 0.965$) of the most energetic jet of the hemisphere.
– $A(2)_{850}$ is the product of the sum of the transverse momenta by the sum of the longitudinal momenta normalized to P_{total}^2 where P_{total} is the sum of the momentum of all the tracks of the hemisphere.
– $A(3)_{463}$ is the sum of the squared transverse momentum of the particles in the jet with respect to the jet axis of the most energetic jet of the hemisphere.
– $A(4)_{459}$ is the longitudinal momentum of the leading particle of the most energetic jet of the hemisphere.

[4] The number of input variables is given by the fact that we only consider variables with a F-value greater than 10% of the F-value associated to the most discriminante variable.

– $A(5)_{272}$ and $A(6)_{212}$ are the transverse momentum of the leading and of the second leading particle of the most energetic jet of the hemisphere with respect to the jet axis.
– $A(7)_{258}$ is the invariant mass of the three most energetic particles of the most energetic jet [2].
– $A(8)_{279}$ and $A(9)_{186}$ are the directed sphericities described in [2].
Higher is the $F - value$, more discriminante is the associated variable. The absolute value of the $F - value$ is not relevant and simply reflects the number of Monte Carlo events used to do this F-test (6000 events of each class in this study).
Fig. 4 shows the shape of the nine variables for b and non-b events and fig. 5. the comparison between data and Monte Carlo. A reasonable agreement is observed between real and simulated data.
The distribution of the resulting Neural Network output R_i computed for each hemisphere i (i=1,2) is shown for simulated b, c and uds quarks in fig. 6, while fig. 7 shows the same distribution for data and Monte Carlo and fig. 8 the distribution of the product $R_1 \times R_2$ for simulated events. The performance of this Neural Network for the b-tagging is illustrated fig. 9 by the curve b-purity versus b- efficiency obtained by applying different cuts on the Neural Network output. Typically, a b-purity of 45% if we cut on one hemisphere $R_{1(2)}$, and of 65% if we cut on the product $R_1 \times R_2$, can be obtained for a b-efficiency of 50%. The comparison with the same curves computed for each variables (fig. 10) shows the interest of the Neural Network technique to combine as well as possible all the informations contained in the input variables.
It has to be noted that (except for A(1) and (A(2)), these nine variables are not the nine most discriminant but have been chosen to reduce the correlations between the two hemispheres since, as we will see in the next section, these correlations cannot be determined from the data and come from the Monte Carlo simulation. So, this Network is in fact a compromise between the best possible discrimination power between $b\bar{b}$ and light quark events and small correlations between the two hemispheres. Another set of input variables will be presented as checks in section 10.

5 Event selection

Selection of hadronic events

The event selection in Scanbook is based on the Heavy Flavour Group selection HE.and.HD.and.HH.and.HM (i.e. events flagged as good for use of ECAL, DEDX, HCAL and muon chambers). This ensures in particular to have good runs for the selection and the identification of our high $p_\perp$ leptons. After the CLAS 16 selection, we obtain 151784 hadronic events for '90 data and 285818 for '91 events. After the acceptance cuts described in the previous section, we finally select a total of 362,246 $q\bar{q}$ candidate events, with 125,175 from the 1990 data and 237071 from the 1991 data set. The main inefficiency of this selection is due to the cut on the polar angle of the thrust axis at 30^0 which removes 14% of the hadronic events. The background

from $\tau^+\tau^-$ in this sample has been estimated using a large sample of Monte Carlo events (about 300,000 simulated $\tau^+\tau^-$) and leads to subtract (376 ± 6) events from the hadronic sample. The error on the number of sutracted events comes from the Monte Carlo statistic and from the present experimental uncertainty on $\Gamma_{\tau^+\tau^-}$). The selection efficiency for $\tau^+\tau^-$ events is 3.12% including the CLAS 16 selection and all the acceptance cuts. Fig. 11 shows the distribution of the Neural Network output obtained for these $\tau^+\tau^-$ events.

Selection of high $p_\perp$ lepton events

The selection and identification of the lepton candidates (ECAL and dE/dX for electrons and HCAL and muon chambers for muons) candidates follows the official procedure of the Heavy flavour lepton working group. We refer to the ref. [19, 20] for more details.
Only leptons with a momentum greater than 3 GeV are considered. Jets are found using the sclaed-invariant-mass clustering algorithm. The $p_\perp$ of the lepton is calculated with respect to its jet axis excluding the lepton. The number of selected lepton candidates is given table 1 for different cuts on the transverse momentum.

Source	$p_\perp \geq 0.$	$p_\perp \geq 1.00$	$p_\perp \geq 1.25$	$p_\perp \geq 1.50$	$p_\perp \geq 1.75$
Electrons	6908 (14587)	2983 (5881)	2296 (4518)	1769 (3431)	1353 (2599)
Muons	11099 (23680)	4355 (9125)	3332 (7003)	2513 (5357)	1917 (4044)

Table 1: Number of selected lepton candidates as function of the applied $p_\perp$ cut (in GeV) for the 1990 data. The numbers in parenthesis correspond to the 1991 data.

6 The Hemisphere Double Tag Method

The measurement of $R^{(b)}$ is usually done by an inclusive tagging of the complete events in which some distributions obtained in the data are compared with the Monte Carlo predictions (for instance, the p and $p_\perp$ distributions of the leptons produced in the semileptonic decays of B-hadrons, the output of an event shape discriminator, etc.). These methods are statistically powerful but rely on the Monte Carlo to predict the shape of these distributions and therefore suffer from systematic errors arising from Monte Carlo uncertainties [21]. We present in this ALEPH Note a Double Tagging method which allows to derive the tagging efficiencies for b and light quarks directly from the data together with the measurement of $R^{(b)}$, eliminating any uncertainties due to b-physics like the b-fragmentation, the B-hadron decay modelisation, Branching ratios, etc. Monte Carlo predictions are only used to determine some small correction coefficients.

This analysis proceeds in three steps:
– We split the $q\bar{q}$ events in two hemispheres according to the plane perpendicular to the thrust axis.
– We calculate the nine variables for each hemisphere and derive a Neural Network output R_i ($i = 1, 2$) which will be used in the following to tag the hemisphere.
– For a given cut on R_i and on the $p_\perp$ of the leptons, we define three classes of events: a class where only one hemisphere is used as a tag ('single tagged events') [5]. a class where both hemispheres are required to satisfy the cut ('double tagged event') and a class of events tagged by a high $p_\perp$ lepton on one side and by a cut on the discriminator on the other side ('single tagged high $p_\perp$ leptons') [6]. From these three samples, we can define the following system of three equations:

$$\left\{ \begin{array}{ll} N^{ST}/2N_{had.} = f_b\epsilon_b & +(1-f_b)\epsilon_{udsc} \\ N^{DT}/N_{had,} = f_b\epsilon_b^2(1+C_b^{DT}) & +(1-f_b)\epsilon_{udsc}^2(1+C_{udsc}^{DT}+c_{udsc}^{DT}) \\ N_{lept.}^{ST}/N_{lept.} = f_b^L\epsilon_b(1+C_b^{ST}) & +(1-f_b^L)\epsilon_{udsc}(1+C_{udsc}^{ST}+c_{udsc}^{ST}) \end{array} \right. \quad (2)$$

– f_b is the fraction of $Z \to b\bar{b}$ events in the hadronic sample. In order to derive the value of $R^{(b)}$ from the measurement of f_b, it is necessary to take into account for the difference of acceptance between $b\bar{b}$ events and all hadronic events. This correction is obtained from the Monte Carlo and is 0.987 ± 0.001, where the error is due to the limited Monte Carlo statistics. This is mainly due to the requirement of at least four energy flow tracks in the most energetic jet of each hemisphere (see fig. 3).
– N^{ST}, N^{DT} and $N_{lept.}^{ST}$ are the number of 'single tagged events', 'double tagged events' and 'single tagged high $p_\perp$ leptons', respectively. $N_{had.}$ and $N_{lept.}$, the number of hadronic events and high $p_\perp$ leptons used for this analysis (for '90 and '91 data, $N_{had.} = 362,246$ and $N_{lept.} = 17,149$ for a $p_\perp$ cut at 1.25 GeV).
– C_b^{ST}, C_{udsc}^{ST}, C_b^{DT}, C_{udsc}^{DT}, c_{udsc}^{DT} and c_{udsc}^{ST} are correction factors which are introduced to take into account possible correlations between the two hemispheres (C^{DT} and C^{ST} coefficients) and the fact that the c quark is heavier than the uds quarks (c^{DT} and c^{ST} coefficients). These correction factors will be described in details in section 8.
– ϵ_b and ϵ_{udsc} are the tagging efficiencies for b and light quarks.
– f_b^L is the b-fraction in the high $p_\perp$ lepton sample.
Then, by solving this system of three equations, it is possible to extract from the data $R^{(b)}$ and the cut efficiencies ϵ_b and ϵ_{udsc} for $b\bar{b}$ and light quark events, provided the hemisphere b-purity f_b^L in the high $p_\perp$ lepton sample is determined by the multi-lepton fit and the correction coefficients are estimated from Monte Carlo.

[5] If the two hemispheres of an event pass the cut, the event is double counted.
[6] In fact, an event contributes in this third sample according to its number of high $p_\perp$ leptons.

7 Determination of f_b^L

Knowledge of the $b\bar{b}$ purity f_b^L in the lepton sample is crucial for the hemisphere double tag method. This purity is determined from a global analysis of single and multi-leptons hadronic events and is defined by the relation:

$$f_b^L = \frac{N_b}{N_b + N_c + N_{uds}} \tag{3}$$

where:

$$N_b = R^{(b)} \times [2 \times f_{b\to l} BR(b \to l) + 2 \times f_{b\to c\to l} BR(b \to c \to l) + f_{b-others}] \times N_Z \tag{4}$$

$$N_c = R^{(c)} \times [2 \times f_{c\to l} BR(c \to l) + f_{c-others}] \times N_Z \tag{5}$$

$$N_{uds} = R^{(uds)} \times f_{uds} \times N_Z \tag{6}$$

f_i is the probability that the $p_\perp$ of a lepton coming from the process i is greater than the applied $p_\perp$ cut.
The details of this study can be found in [19] and a publication is in preparation on this subject; so only a brief description is given here.
This analysis takes advantage of the redundancy of the data to measure simultaneously the partial width of the Z into $c\bar{c}$, the b and c-fragmentation, the semileptonic branching ratios of b and c-hadrons and another determination of the partial width of the Z into $b\bar{b}$.

Two samples of events are considered: events with at least one lepton with $p \geq 3$ GeV (single lepton events) and events with at least two leptons with $p \geq 3$ GeV (multi-lepton events). The first sample is analysed in the $(p, p_\perp)$ plane and the second one in the $(p_{cross}, p_\perp^{min})$ plane with $p_{cross} = p_\perp^1 p_{//}^2 + p_\perp^2 p_{//}^1$ and $p_\perp^{min} = Min(p_\perp^1, p_\perp^2)$.
– The single lepton sample receives three main contributions:

$$BR(b \to l) \times R^{(b)} \tag{7}$$

$$BR(b \to c \to l) \times R^{(b)} \tag{8}$$

$$BR(c \to l) \times R^{(c)} \tag{9}$$

– The multi-lepton sample can be splitted in two samples: same side dileptons, if the angle between the direction of the two leptons is less than 90^0, and opposite side dileptons otherwise.
⋆ The main contribution for the same side dilepton events is

$$BR(b \to l) \times BR(b \to c \to l) \times R^{(b)}. \tag{10}$$

⋆ The opposite side dilepton events are dominated by the following contributions:

$$[BR(b \to l)]^2 \times R^{(b)} \tag{11}$$

$$BR(b \to l) \times BR(\bar{b} \to \bar{c} \to l) \times R^{(b)} \tag{12}$$

$$[BR(c \to l)]^2 \times R^{(c)}. \tag{13}$$

To summarize, opposite side dileptons measure $BR(b \to l)$, same side opposite charge dileptons measure $BR(b \to c \to l)$, the single lepton sample allows to extract $R^{(b)}$ and $R^{(c)}$ and the three samples determine the fragmentation parameters. With the present statistic, we are not able to fit simultaneously the branching ratios $BR(c \to l)$ and $BR(b \to c \to l)$, therefore we have fixed $BR(c \to l)$ to its measured value at low energy: $BR(c \to l) = 0.098 \pm 0.005$.

To determine f_b^L, a 6 parameter fit ($R^{(b)}, R^{(c)}, < X_b >, < X_c >, BR(b \to l)$ and $BR(b \to c \to l)$) of single and dilepton events is performed in the full $(p, p_\perp)$ and $(p_{cross}, p_\perp^{min})$ planes (i.e. no cut on $p_\perp$) and the result is extrapolated in the restricted region $p_\perp \geq 1.25$ GeV/c. Table 2 gives the purities obtained for the '90 and '91 data separately. The value obtained for 1991 is slightly lower than for 1990 mainly because of a higher conversion rate due to the addition of the vertex detector.

	1990	1991	1990+1991
f_b^L (%)	88.49	88.19	88.29

Table 2: b-purity in the high $p_\perp$ lepton sample for '90 and '91 data.

Table 3 shows the various contributions to the total error on f_b^L. The statistical

Source	variation	Δf_b^L (%)
Fit statistics $\sigma(f_b^L)$	$\pm 1\sigma$	± 0.50
b decay model	Altarelli/ISGW	± 0.25
c decay model for $c \to l$	Altarelli/JETSET ($\pm 0.5\sigma$)	± 0.19
c decay model for $b \to c \to l$	Altarelli/JETSET ($\pm 1\sigma$)	± 0.35
elec. Id. efficiency	$\pm 3\%$	± 0.02
muon Id. efficiency	$\pm 3\%$	± 0.03
γ conv.	$\pm 10\%$	± 0.14
$e - misid.$	$\pm 10\%$	± 0.07
$\mu - decay$	$\pm 10\%$	± 0.21
punch-through	$\pm 20\%$	± 0.13
Total syst.		± 0.55
Total stat. + syst.		± 0.74

Table 3: Contributions to the error on the b purity? f_b^L in the global fit of single and dilepton hadronic events.

error $\sigma(f_b^L)$ on f_b^L was estimated in the following way: we first do the 6 parameter

fit; from this fit, we obtain the values $x_i \pm \sigma_i$ for each parameter x_i $(i = 1,6)$. and the correlation matrix ρ_{ij} $(i,j = 1,6)$. Then, we deduce:

$$\sigma^2(f_b^L) = \sum_i \sum_j \frac{\partial f_b^L}{\partial x_i} \frac{\partial f_b^L}{\partial x_j} \rho_{ij} \sigma_i \sigma_j \tag{14}$$

The uncertainty coming from the b-decay modelisation has been estimated by using the Altarelli et al. [22] and the ISGW [23] models to describe the shape of the lepton energy distribution in the b-hadron rest frame for $b \rightarrow l$ transitions, the two models being optimized on the ARGUS and CLEO data. Then, the purity is computed for each model and we take as error one half of the difference between the two results.
A similar method is used to derive the error due to the c-decay modelisation in $c \rightarrow l$ transitions. In this case we compare the result obtained for f_b^L by using the standard JETSET modelisation and the Altarelli et al. predictions optimized on the DELCO data (fig. 12).
To estimate the uncertainty due to the c-decay modelisation in $b \rightarrow c \rightarrow l$ decays, we factorize this transition in $(b \rightarrow c) \times (c \rightarrow l)$. The error due to $c \rightarrow l$ is estimated as previoulsy while for $b \rightarrow c$, we compare the predictions of the standard JETSET modelisation with the Altarelli et al. model optimized on the CLEO data for the $B \rightarrow D^0 X$ and $B \rightarrow D^+ X$ transitions. In fact, the two effects are opposite and would tend to cancel each other since in the $c \rightarrow l$ case, the JETSET prediction leads to a softer distribution than the Altarelli et al. model, while for the $b \rightarrow c$ energy spectrum, the JETSET model gives a harder spectrum. However, we have added linearly the two effects to have a conservative estimate of this error on f_b^L.
Finally, the purity of the lepton sample for $p_\perp \geq 1.25$ GeV is:

$$f_b^L = 0.8829 \pm 0.0050(stat.) \pm 0.0055(syst.). \tag{15}$$

Note that the contribution to $\sigma(f_b^L)(stat.)$ of the parameter $R^{(b)}$ is $\pm 0.21\%$, the main contribution being due to the charm. This induces a 19% correlation between the $\Gamma_{b\bar{b}}/\Gamma_{had}$ values obtained in the global lepton analysis and in this event shape analysis.

8 Determination of the correction coefficients

The determination of the 6 correction coefficients and the confidence that we can have on their values is really the key point of this analysis since they are not up to now measured from the data and therefore can introduce some bias due to the Monte Carlo.

– C_b^{DT} and C_{udsc}^{DT} are correction factors which take into account possible correlations between the two hemispheres due to kinematic constraints (conservation of momentum for instance) and hard gluon radiation by the primary quarks. They are

defined by the following relation:

$$C_i^{DT} = \frac{\epsilon_i^{DT} - \epsilon_i^2}{\epsilon_i^2} \quad with\ i = b,\ udsc. \tag{16}$$

where ϵ^{DT} is the probability that both hemispheres of an event satisfy the cut on the event shape discriminator ($\epsilon^{DT} = N^{DT}/N^{had.}$ for $q\bar{q}$ events).
These two coefficients are determined by using more than one million fully reconstructed hadronic Monte Carlo events [7]. Note that the value for the coefficient C_{udsc}^{DT} is in fact the average of the values obtained for $c\bar{c}$ and uds events independently. We find $C_c^{DT} = 0.026 \pm 0.004$ and $C_{uds}^{DT} = 0.036 \pm 0.002$ leading to $C_{udsc}^{DT} = 0.033 \pm 0.002$; the error is due to the Monte Carlo statistics.

– C_b^{ST} and C_{udsc}^{ST} are correction factors which take into account possible correlations between the two hemispheres due to the presence of an energetic lepton with missing energy carried out by the neutrino produced in the semileptonic decay of D and B-hadrons. They are defined by the following relation:

$$C_i^{ST} = \frac{\epsilon_i^{lepton} - \epsilon_i}{\epsilon_i} \quad with\ i = b,\ udsc. \tag{17}$$

where ϵ^{lepton} is the probability to tag the hemisphere opposite to the high $p_\perp$ lepton ($\epsilon^{lepton} = N_{lept.}^{ST}/N^{lept.}$ for $q\bar{q}$ events).
These two coefficients are expected to be very small especially in the case of uds events for which there is no neutrino produced with the lepton. They have been determined in the same way as C_b^{DT} and C_{udsc}^{DT} by using our large sample of fully reconstructed hadronic Monte Carlo events.

– c_{udsc}^{DT} and c_{udsc}^{ST} are two correction coefficients introduced to account for the fact that, due to the mass of the c quark, $c\bar{c}$ events are less well separated from $b\bar{b}$ events than uds events (fig. 6). This results in a higher tagging efficiency of charm events for a given cut on the Discriminator output. For instance, for a cut at 0.3 on the Neural Network output, Monte Carlo predictions give: $\epsilon_c = 0.4309 \pm 0.0009$ and $\epsilon_{uds} = 0.3670 \pm 0.0005$ leading to $\epsilon_{udsc} = 0.3811 \pm 0.0004$, again the errors are due to the Monte Carlo statistics. So, these two correction coefficients depend on the fractions f_c and f_c^L of $c\bar{c}$ events in the non-b hadronic and leptonic samples respectively and are given by the following equations:

$$c_{light}^{DT} = \frac{(1 - f_c) f_c (\epsilon_c - \epsilon_{uds})^2}{\epsilon_{udsc}^2} \tag{18}$$

$$c_{light}^{ST} = \frac{(f_c^L - f_c)(\epsilon_c - \epsilon_{uds})}{\epsilon_{udsc}} \tag{19}$$

[7] The determination of these correction coefficients has been done by using: 1,050,000 fully simulated $q\bar{q}$ events (210,000 with the '90 geometry and GAL252 + JUL261 and 850,000 with the '91 geometry mainly with GAL242 + JUL258) and about 350,000 $b\bar{b}$ (46,000 with the '90 geometry and GAL252.7 + JUL261.5, 90,000 with the '91 geometry and 210,000 with the '92 geometry). Note that this number of $b\bar{b}$ corresponds to more than 1,5 million additional $Z \rightarrow q\bar{q}$ events.

The charm fractions f_c and f_c^L are determined by the multi-lepton fit. We get: $f_c = 0.21 \pm 0.02$ and $f_c^L = 0.57 \pm 0.07$. The error on f_c is obtained from the errors (stat. + syst.) on $R^{(b)}$ and $R^{(c)}$ in the multi-lepton fit. To obtain the error on f_c^L, we express this quantity as function of f_L^b and of the $c\bar{c}$ purity P_c^L in the high $p_\perp$ lepton sample: $f_c^L = P_C^L/(1 - f_b^L)$ with $P_c^L = 0.0667$. P_C^L is mainly proportional to the product $BR(c \to l) \times R^{(c)}$ which is measured with an accuracy of 3% [19]. We assign an error of 10% on P_c^L to have a conservative estimate of the error on f_c^L.
For a given cut on the Neural Network output, ϵ_c and ϵ_{uds} are then estimated by Monte Carlo and the two coefficients c_{udsc}^{ST} and c_{udsc}^{DT} derived.
The values obtained for the 6 correction coefficients are given in table 4 for a cut at 0.3 on the Neural Network output. and for $p_\perp \geq 1.25$ GeV.

parameter	Monte Carlo estimate
C_b^{DT}	0.001 ± 0.001
C_b^{ST}	-0.001 ± 0.001
C_{udsc}^{DT}	0.033 ± 0.002
c_{udsc}^{DT}	0.005 ± 0.002
C_{udsc}^{ST}	-0.001 ± 0.016
c_{udsc}^{ST}	0.060 ± 0.015

Table 4: Correction coefficients from Monte Carlo estimate for a cut on the event shape discriminator at 0.3 and on the $p_\perp$ of the lepton at 1.25 GeV. The errors are only statistical.

Fig. 13 and 14 shows the values of the correction coefficients as function of several cuts applied on the Neural Network output. The two coefficients associated to $b\bar{b}$ events are stable and always compatible with zero within the errors while the light quark coefficients increase with the applied cut. This is mainly due to the fact that the difference between c and uds hemispheres becomes more important when we cut in the tails of the event shape discriminator. The systematics on these correction factors will be discussed in section 10.

9 Results

Once the b-purity in the high $p_\perp$ lepton sample and the 6 correction coefficients have been estimated for a given cut on the Neural Network ouput and on the $p_\perp$ of the leptons, the system of three equations can be solved and the values of f_b, ϵ_b and ϵ_{udsc} derived. The best cut on the event shape, i.e. the cut which corresponds to the smallest statistical error on f_b is a cut at 0.3 (cf. table 5). To select the high $p_\perp$ lepton sample, we have chosen a cut at 1.25 GeV since this value gives the best overall error on f_b (a lower cut gives a smallest statistical error but this gain is canceled by the

increase of the systematic error of f_b^L, while a harder $p_\perp$ cut leads to a too important statistical error) (table 6).

Table 7 shows the effect of the various correction factors on f_b. Note that the single

Cut on R	f_b (%)	ϵ_b (%)	ϵ_{udsc} (%)
0.21	24.90 ± 0.72	88.72 ± 0.29	56.29 ± 0.36
0.22	24.54 ± 0.67	87.56 ± 0.30	54.02 ± 0.34
0.23	24.29 ± 0.65	86.36 ± 0.31	51.76 ± 0.32
0.24	24.00 ± 0.61	85.05 ± 0.32	49.59 ± 0.31
0.25	23.91 ± 0.59	83.80 ± 0.33	47.45 ± 0.30
0.26	23.54 ± 0.58	82.41 ± 0.34	45.49 ± 0.28
0.27	23.58 ± 0.57	80.89 ± 0.35	43.46 ± 0.27
0.28	23.27 ± 0.56	79.50 ± 0.35	41.63 ± 0.26
0.29	22.98 ± 0.56	77.94 ± 0.36	39.90 ± 0.25
0.30	23.10 ± 0.55	76.32 ± 0.37	38.05 ± 0.24
0.31	22.93 ± 0.55	74.67 ± 0.37	36.39 ± 0.23
0.32	22.75 ± 0.55	73.06 ± 0.38	34.78 ± 0.23
0.33	22.82 ± 0.56	71.21 ± 0.38	33.17 ± 0.22
0.34	23.10 ± 0.57	69.28 ± 0.39	31.15 ± 0.22
0.35	23.12 ± 0.58	67.40 ± 0.40	30.00 ± 0.22
0.36	23.16 ± 0.59	65.58 ± 0.40	28.49 ± 0.21
0.37	23.06 ± 0.60	63.65 ± 0.40	27.19 ± 0.21
0.38	22.89 ± 0.61	61.87 ± 0.41	25.88 ± 0.21
0.39	23.10 ± 0.62	59.88 ± 0.40	24.47 ± 0.20

Table 5: Optimisation of the cut on the Neural Network output R for a given $p_\perp$ cut at 1.25 GeV. The values of f_b are given for '90+'91 data; they are not corrected by the acceptance factor ($\times 0.987$) and by the $\tau^+\tau^-$subtraction (-0.06%) but the 6 correction factors are estimated for each cut and taken into account to solve the system. Note that all the points are correlated.

tag coefficient (C^{ST}) for light quark events is big but its effect on f_b is small because it is multiplied in the third equation of the system by $(1 - f_b^L) = 0.1171$, while the effect of the double tag coefficient (C^{DT}_{udsc}) is quite large.

Table 8 gives the results obtained for $\Gamma_{b\bar{b}}/\Gamma_{had}$ after correcting f_b by the acceptance factor, for different years of data taking. The background from $Z \to \tau^+\tau^-$ is also subtracted and changes the $\Gamma_{b\bar{b}}/\Gamma_{had}$ value by -0.06%. This table shows that the measured value of $\Gamma_{b\bar{b}}/\Gamma_{had}$ is very stable with the year of data taking and that the shape of the Neural Network output (and so the efficiencies) is not affected (fig. 15). The stability of $\Gamma_{b\bar{b}}/\Gamma_{had}$ for several cuts on the Neural Network output with $p_\perp \geq 1.25$ GeV, and for several $p_\perp$ cuts with a Neural Network cut at 0.3 is shown figure 16. The corresponding measured efficiencies ϵ_b and ϵ_{udsc} are displayed fig. 17. Again, the

results are very stable in a wide range of b-efficiency values and of $p_\perp$ cuts showing that the determination of $\Gamma_{b\bar{b}}/\Gamma_{had}$ is independent of the cuts within the errors.

10 Checks and study of the systematics

Test of the method using Monte Carlo events

We have checked on 850,000 $q\bar{q}$ Monte Carlo events treated as if they were real data, that the Double Tag method is able to find the correct fraction of $b\bar{b}$ events, and that the extracted efficiencies obtained by solving the system of three equations are compatible with the true efficiencies. The results are summarized in table 9 and in fig. 18 and show that indeed the method works well on simulated events.

Influence of the Monte Carlo tuning

Although this method does not rely too heavily on the Monte Carlo tuning, we have to take care that the efficiencies ϵ_b and ϵ_{udsc} measured for a given cut on the Neural Network output are not very different between data and Monte Carlo. This is because some correction factors depend on these efficiencies. So, if they do not agree between data and Monte Carlo, the correction factors estimated by Monte Carlo for the data could be wrong. This problem can only occur for light quark events since in the b case, the two coefficients C_b^{DT} and C_b^{ST} are quite stable whatever the efficiency ϵ_b. Figures 19 and 20 show the comparison between real and simulated data for the measured efficiencies (see also table 10). The observed agreement is very good for light quark events giving confidence that we really take the correct coefficients from Monte Carlo. In the b case, the agreement is not at the same level but, because of the stability of C_b^{DT} and C_b^{ST}, this problem is not so crucial.

Another problem coming from the Monte Carlo could be the use of different versions to estimate an average value of the correction factors (210,000 $q\bar{q}$ with the '90 geometry and 850,000 with the '91 geometry). First, we have checked that the two Monte Carlo samples give very similar Neural Network outputs (fig. 21) and that the correction coefficients are compatible within the errors leading to compatible values for f_b (see tables 11 and 12, and figures 22 and 23).

Influence of the Neural Network tuning

Mainly two types of checks have been done. We have verified that the measurement of f_b was not affected by:
– the choice of the learning parameters η and α and the structure of the Neural Network (number of hidden layers, number of nodes per hidden layer),
– the choice of the input variables.
This study was done by repeating the complete analysis with four Neural Networks:
– Option (1): version used to produced our 'official' result. This is a priori the best compromise between small correlations between the two hemispheres and good discrimination power between $b\bar{b}$ and light quark events.

– Option (2): same as (1) but with different values of η and α during the learning: $\alpha = 0.9$ and $\eta = 0.01$.
– Option (3): same as (1) but we have changed the input variables to get the smallest possible correction coefficients even at the expense of a worse discrimination power.
– Option (4): Neural Network with 20 input variables and two hidden layers with respectively 20 and 10 nodes. This option corresponds to the package QNNBTAG which is installed in the UPHY area.
The results obtained with the four options are summarized in table 13 and in fig. 24. They show that the determination of $\Gamma_{b\bar{b}}/\Gamma_{had}$ is not affected by a possible systematic due to the choice of the Neural Network, and that the result is stable in a wide range of values of the correction coefficients C^{DT}_{udsc} and C^{ST}_{udsc} giving confidence that these parameters are well predicted by the simulation.

Looking at the Neural Network output

The main effect of the Double Tag correlations (correlations between the two hemispheres of an event) is to change the hemisphere shape output. This effect is shown fig. 25 and 26 for b and $udsc$ simulated events. Higher are these correlations, more important is the distortion of the distributions. So, it can be interesting to compare more precisely the Neural Network outputs between real and simulated events in order to have a qualitative estimate of these correlations.
Fig. 7 shows that the disagreement between data and Monte Carlo on the hemisphere shape output is more pronounced when a cut is applied on the opposite hemisphere. In other words, the effect of a cut on the opposite hemisphere is more important for data than for Monte Carlo (see fig. 27 and 28). This effect can be due to the fact that the Double Tag correlations are bigger in the data (and therefore underestimated in the system of three equations). But this can be also produced by a disagreement between data and Monte Carlo on the hemisphere shape of $b\bar{b}$ event (since a cut on one hemisphere enriches in $b\bar{b}$ events). This problem is confirmed if we repeat this analysis on a very pure $b\bar{b}$ sample selected with QIPBTAG ($|probevt| \geq 3.5 \times 10^{-3}$ which gives a b-purity of 90%) (fig. 29) and make difficult any conclusion on the C_b^{DT} coefficient.
In order to study C^{DT}_{udsc}, we have selected events both in data and in Monte Carlo by requiring $|probevt| \leq 1.0 \times 10^{-3}$. The light quark purity is, according to QIPBTAG, 96% and 92% after cutting at 0.3 on one hemisphere (fig. 30). Fig. 31 shows the distortion of the hemisphere shapes for data and Monte Carlo. The same effect as in $q\bar{q}$ events (i.e. the fact that the distortion, and therefore the Double Tag correlation, is bigger in the data) is observed.
In conclusion, a determination of the coefficient C^{DT}_{udsc} from the data will be crucial to reduce the systematics of this method and to get a high precision measurement of $\Gamma_{b\bar{b}}/\Gamma_{had}$. The high statistic collected by ALEPH in 1992 and the use of vertex informations could allow in principe to do that [24].

Systematic errors on $\Gamma_{b\bar{b}}/\Gamma_{had}$

The main systematic errors on $\Gamma_{b\bar{b}}/\Gamma_{had}$ arise from the uncertainties in the purity of the high $p_\perp$ lepton tag, f_b^L and in the correction factors C_{udsc}^{ST}, C_b^{DT}, C_{udsc}^{DT} and C_b^{ST}. The error in f_b^L is dominated by the statistics of the fit and by the c- and b-hadron decay modelisations as described in section 7.
For the correction coefficients C^{DT} and C^{ST}, the systematic error contributions are estimated by varying these parameters within their statistical errors in the Monte Carlo. For the two coefficients c_{udsc}^{DT} and c_{udsc}^{ST}, the systematic errors are estimated by varying the charm fractions f_c and f_c^L by $\pm 10\%$ according to the results obtained in section 8 and by changing the charm fragmentation parameter $\epsilon_c(Peter.)$ of the Peterson et al. function from 0.040 to 0.065 [9]. Note that changes in charm fragmentation and fractions produce anticorrelated and nearly cancelling changes in $\Gamma_{b\bar{b}}/\Gamma_{had}$. Therefore the systematic contributions for these corrections are dealt with in a correlated fashion.
For the b correction factors C_b^{DT} and C_b^{ST}, we have changed the bottom fragmentation parameter $\epsilon_b(Peter.)$ from 0.002 to 0.008. This study is not done at the generator level by with the full Monte Carlo events by weighting the Neural Network output. It should be noted that the most rigourous way to do that would be to directly weight the nine input variables and recompute the Neural Network output. But what has been done is certainly more conservative since it assumes that a change of $\epsilon_b(Peter.)$ affects all the variables in the same direction. Fig. 32 shows the influence of the b-fragmentation on the hemisphere Neural Network output for simulated $b\bar{b}$ events and table 14, the values of C_b^{DT} and C_b^{ST} for different values of $\epsilon_b(Peter.)$ (see also fig. 33). By considering the two extreme values of $\epsilon_b(Peter.)$, we obtain a very conservative error on $\Gamma_{b\bar{b}}/\Gamma_{had}$ of $\pm 0.06\%$.
Another systematic can affect the C_b^{ST} parameter. As it has been previously noticed, the coefficient is essentially estimated by comparing the shape of the Neural Network output in an hadronic event (first equation) with the shape of the hemisphere opposite to the high $p_\perp$ lepton (third equation). But in the lepton sample, we do not only consider events with one high $p_\perp$ lepton but also events with two or more leptons and these events enter in the third equation of the system according to their lepton multiplicity. Then, sometimes the hemisphere opposite to a high $p_\perp$ lepton contains also a lepton. Fig. 34 shows a comparison of the shape of the Neural Network output for the hadronic events and in the lepton sample for the hemisphere containing the lepton. The presence of an energetic lepton with missing energy carried out by the neutrino produced a clear distortion of the hemisphere shape. In orfer to define more precisely the origin of C_b^{ST}, it is instructive to look at the shape of the Neural Network output in the lepton sample for events containing only one lepton with p greater than 3 GeV and for dilepton events. These distributions are given fig. 35 and 36 and show that the factor C_b^{ST} is due to two opposite effects: for single lepton event, the presence of missing energy in the hemisphere opposite to the one used for the Neural Network output induces a small positive correlation (fig 35), while in the case of dilepton events the correlation is mainly induced by the presence of missing energy in the hemisphere used for the Neural Network output and is negative (fig. 36). The

result of this study are listed in table 15.
So, the C_b^{ST} coefficient can be badly estimated by Monte Carlo if the ratio of dilepton and single lepton events is in disagreement between data and Monte Carlo. This ratio is in fact directly proportional to the semileptonic branching ratios of b-hadrons (mainly $BR(b \to l) + BR(b \to c \to l)$). To estimate a possible systematic effect on C_b^{ST}, we have changed these branching ratios by $\pm 10\%$ in the simulation and compute the new values of C_b^{ST}. This effect can induce a systematic error of $\pm 0.03\%$ on C_b^{ST} leading to a $\pm 0.03\%$ systematic error on $\Gamma_{b\bar{b}}/\Gamma_{had}$.

The contributions from the difference sources of systematic error are listed in table 16.

11 Conclusion and future of the method

We have measured the $b\bar{b}$ partial width of the Z by using an hemisphere Double Tag method with a Neural Network to preferentially select the $b\bar{b}$ final state. Using the 450,000 hadronic events collected by ALEPH in 1990 and 1991, we obtain:

$$\Gamma_{b\bar{b}}/\Gamma_{had} = 0.2274 \pm 0.0054(stat.) \pm 0.0038(f_b^L) \pm 0.0026(others) \qquad (20)$$

This result is 1.3σ above the highest Standard Model prediction of 0.218.
Assuming two millions of hadronic Z events at the end of 1993, the statistical error will almost reach the 1% accuracy level (in fact ± 0.26). The dominant systematic error comes from the uncertainty on the b-purity in the high $p_\perp$ lepton sample. This error is dominated by the fit statistic and by the b- and c-decay modelitation: $\Delta f_b^L = 0.0050(stat.) \pm 0.0047(model) \pm 0.0029(back.)$. We can therefore expect to divide this error by two for the end of 1993 (i.e. $\pm 0.19\%$ on $\Gamma_{b\bar{b}}/\Gamma_{had}$).
In conclusion, this method can ultimately provide a measurement of $\Gamma_{b\bar{b}}/\Gamma_{had}$ with an accuracy of (1.5-2)%, dominated by the systematic errors arising from the correction coefficients. To go beyond this (1.5-2)% accuracy, it will be necessary to measure at least one of these correction factors from the data in order to decrease the systematic error and/or to introduce some vertex informations as input of the Neural Network. Indeed, a combinaison of event shape and vertex variables is certainly the best solution to reject in the most efficient way the remaining background.

Acknowledgement

We would like to thank A. Bencheikh for useful work on the Neural Network tuning, E. Mannelli for stimulating discussions and J. Proriol for providing us the Fortran code of the Neural Network package and of some input variables.

References

[1] L. Lönnblad et al., Phys. Rev. Lett. **65** (1990) 1321,
L. Lönnblad et al., Nucl. Phys. **B349** (1991) 675,
C. Peterson, LU TP 90-6, Lund (1990),
M. Mjaed and J. Proriol, Phys. Lett. **B217** (1989) 4,
T. Ruan, ALEPH Note 92-184.

[2] L. Bellantoni et al., Nucl. Inst. . Inst. and Meth. **A310** (1991) 618.

[3] C. Bortolotto et al., lUDINE Preprint 91/04/AA - 26 August 1991,
DELPHI Collab., *Classification of the Hadronic Decays of the Z^0 into b and c Quark Pairs using a Neural Network*, CERN-PPE 92-151 (Submitted to Phys. Lett. B).

[4] J. Proriol et al., Workshop on Neural Networks *Neural Network: from biology to High Energy Physics*, editors: O. Benhar et al., Ets. Editrice Pisa (1992).

[5] B. Brandl, et al. Nucl. Inst. and Meth. **A324** (1993) 307.

[6] J. Conway et al., *Particle Identification for Tau Physic in ALEPH Using a Neural Network*, ALEPH Note 92-56,
V. Innocente al., *Identification of Tau Decays Using a Neural Network*, CERN-PPE/92-98, (Submitted to Nucl. Instr. and Meth.).

[7] J. Schwindling, *Search for the standard Higgs boson in the $H^0\nu\bar{\nu}$ channel using a neural network*, ALEPH Note 92-60.

[8] A. Cherubini et al., Z. Phys. **C53** (1992) 139.

[9] A. Bencheikh, University of Clermont-Ferrand PCCF T 9209,
D. Decamp *et. al.*, ALEPH Collab., *Production of Charmed Mesons in Z decays*, Contributed paper to the $XXVI^{th}$ International Conference on High Energy Physics, 6-12 August 1992, Dallas.

[10] G. Stimpfl-Abele, Comp. Phys. Comm. 67 (1991) 183.

[11] G.Athanasiu et al., *Real time pattern recognition with artificial neural networks*, CERN-PPE 92-86.

[12] D. Brown, *QFNDIP a primary vertex finder*, ALEPH Note 92-47,
D. Brown et al., *Tagging b hadrons using track impact parameters*, ALEPH Note 92-135.

[13] J. J. Opfield, *Neural network and physical system with emergent collective computational abilities*, Proceedings of the National Academy of Science, USA, 79 (1982) 2554,

T. Kohonen, *Self-Organization and Associative Memory*, 2nd Ed., (Springer. Berlin, 1987).file

[14] J. E. Campagne and R. Zitoun, Z. Phys. **C43** (1989) 469.
T. Sjöstrand and M. Bengtsson, Comp. Phys. Com. **43** (1987) 367.

[15] The difference between the average charged multiplicity of $b\bar{b}$ events and that of all hadronic Z decays has been measured at SLC:
D. S. Koetke, SLAC-Report 396 (1992).

[16] W. Bartel *et al.*, JADE Collab., Z. Phys. **C33** (1986) 23.
S. Bethke *et al.*, JADE Collab., Phys. Lett. **B213** (1988) 235.

[17] M.S. Srivastava and E.M. Carter, *An introduction to applied multivariate statistics*, North-Holland Amsterdam (1983),
R. A. Fisher, Annals of Eugenics **7** (1936) 179.

[18] The code has been written by J. Proriol. It is part of the library MODULAD (function MLP); Club MODULAD, INRIA, Rocquencourt F-78153 Le Chesnay Cedex.

[19] A. Falvard al., *Production and Semileptonic Decays of Heavy Quarks at the Z*, ALEPH Note 93-30.

[20] D. Abbaneo al., *Lepton and jet definitions for the lepton paper*, ALEPH Note 92-101.

[21] P. Henrard, ALEPH Collab., presented at the 4th Symposium on Heavy Flavour Physics, Orsay, june 1991; editors M. Davier, G. Wormser, Editions Frontieres (1992);
B. Brandl, ALEPH Coll., XXVIIth Rencontres de Moriond on *Electroweak Interactions and Unified Theories*, March 15-22, 1992.

[22] G. Altarelli et al., Nucl. Phys. **B208** (1982) 365.

[23] N. Isgur, D. Scora, B. Grinstein and M.B. Wise, Phys. Rev. **D39** (1989) 799.

[24] E. Mannelli, ALEPH Note 93-28.

$p_\perp$ cut (GeV)	f_b (%)	ϵ_b (%)	ϵ_{udsc} (%)
0.75	23.19 ± 0.46	76.26 ± 0.32	38.02 ± 0.22
1.00	23.23 ± 0.51	76.23 ± 0.34	38.01 ± 0.23
1.25	23.10 ± 0.55	76.32 ± 0.37	38.05 ± 0.24
1.50	23.73 ± 0.62	75.85 ± 0.41	37.88 ± 0.25
1.75	22.79 ± 0.65	76.57 ± 0.46	38.13 ± 0.26

Table 6: Optimisation of the $p_\perp$ cut for a given cut on the Neural Network output at 0.3. Same remarks as for the previous table.

Correction factor	value	Δf_b (%)
C_b^{DT}	0.001	- 0.09
C_b^{ST}	-0.001	- 0.10
$C_{udsc}^{DT} + c_{udsc}^{DT}$	0.038	- 2.59
$C_{udsc}^{ST} + c_{udsc}^{ST}$	0.059	+ 0.38

Table 7: Effect of the correction factors in the solution of the system of three equations; $\Delta f_b = f_b - f_b^0$ where f_b^0 (f_b) is the solution of the system without (with) these correction factors taken into account.

Period	f_b^L (%)	$\Gamma_{b\bar{b}}/\Gamma_{had}$ (%)	ϵ_b (%)	ϵ_{udsc} (%)
1990	88.49	22.74 ± 0.93	77.06 ± 0.63	38.31 ± 0.41
1991	88.19	22.74 ± 0.67	75.98 ± 0.47	37.93 ± 0.30
1990+1991	88.29	22.74 ± 0.54	76.32 ± 0.37	38.05 ± 0.24

Table 8: Determination of $\Gamma_{b\bar{b}}/\Gamma_{had}$ and of the two efficiencies for different years. Note that all the results are obtained by using the same values for the correction coefficients.

Cut on R	f_b (%)	ϵ_b (%) measured	ϵ_b (%) true	ϵ_{udsc} (%) measured	ϵ_{udsc} (%) true
0.20	22.17 ± 0.52	89.59 ± 0.18	89.59 ± 0.03	58.42 ± 0.25	58.46 ± 0.04
0.25	22.05 ± 0.38	83.01 ± 0.22	82.89 ± 0.04	47.73 ± 0.19	47.76 ± 0.04
0.30	22.12 ± 0.36	75.21 ± 0.24	75.06 ± 0.04	38.07 ± 0.15	38.12 ± 0.04
0.35	22.32 ± 0.37	66.40 ± 0.26	66.32 ± 0.05	29.95 ± 0.14	30.11 ± 0.04
0.40	22.42 ± 0.61	57.19 ± 0.26	57.04 ± 0.05	23.27 ± 0.12	23.42 ± 0.03
0.45	22.19 ± 0.45	47.69 ± 0.26	47.24 ± 0.05	17.69 ± 0.12	17.77 ± 0.03
0.50	22.41 ± 0.52	37.48 ± 0.25	37.40 ± 0.05	12.73 ± 0.11	12.83 ± 0.03

Table 9: Test of the Double Tag method on 850,000 simulated hadronic events. The measured values of f_b take into account the 6 correction coefficients but they are not corrected by the accceptance factor. The true b-fraction in this Monte Carlo sample (without acceptance correction) is 22.17%.

Cut on R	ϵ_b (%) M.C.	ϵ_b (%) data	ϵ_{udsc} (%) M.C.	ϵ_{udsc} (%) data
0.20	89.59 ± 0.18	89.88 ± 0.28	58.42 ± 0.25	58.94 ± 0.38
0.25	83.01 ± 0.22	83.80 ± 0.33	47.73 ± 0.19	47.45 ± 0.29
0.30	75.21 ± 0.24	76.32 ± 0.37	38.07 ± 0.15	38.05 ± 0.24
0.35	66.40 ± 0.26	67.40 ± 0.39	29.95 ± 0.14	30.00 ± 0.21
0.40	57.19 ± 0.26	57.99 ± 0.41	23.27 ± 0.12	23.25 ± 0.20
0.45	47.69 ± 0.26	48.13 ± 0.41	17.69 ± 0.12	17.56 ± 0.19
0.50	37.48 ± 0.25	37.77 ± 0.39	12.73 ± 0.11	12.52 ± 0.18

Table 10: Comparison between real and simulated data for the b and light quark efficiencies obtained as solutions of the system of three equations for different cuts on the Neural Network output.

Cut on R	ϵ_b true (%)	C_b^{DT} (%)	C_b^{ST} (%)
0.25	83.06 ± 0.14	0.31 ± 0.37	-0.04 ± 0.54
	$[82.81 \pm 0.05]$	$[-0.02 \pm 0.13]$	$[-0.18 \pm 0.19]$
0.30	75.23 ± 0.15	0.66 ± 0.44	-0.55 ± 0.69
	$[74.99 \pm 0.05]$	$[0.06 \pm 0.15]$	$[-0.21 \pm 0.23]$
0.35	66.58 ± 0.17	0.57 ± 0.49	-1.24 ± 0.85
	$[66.24 \pm 0.06]$	$[0.12 \pm 0.17]$	$[-0.20 \pm 0.29]$
0.40	57.38 ± 0.18	0.67 ± 0.55	-2.10 ± 1.26
	$[56.93 \pm 0.07]$	$[0.28 \pm 0.19]$	$[0.15 \pm 0.35]$
Cut on R	ϵ_{udsc} true (%)	$C_{udsc}^{DT} + c_{udsc}^{DT}$ (%)	$C_{udsc}^{ST} + c_{udsc}^{ST}$ (%)
0.25	47.55 ± 0.10	2.91 ± 0.36	2.47 ± 3.52
	$[49.62 \pm 0.05]$	$[2.89 \pm 0.18]$	$[5.88 \pm 1.68]$
0.30	37.93 ± 0.09	4.02 ± 0.47	1.06 ± 4.36
	$[38.16 \pm 0.05]$	$[3.77 \pm 0.23]$	$[6.86 \pm 2.07]$
0.35	29.99 ± 0.09	3.73 ± 0.63	0.87 ± 5.20
	$[30.13 \pm 0.04]$	$[4.23 \pm 0.37]$	$[6.03 \pm 2.50]$
0.40	23.38 ± 0.08	2.90 ± 0.88	6.30 ± 6.29
	$[23.43 \pm 0.04]$	$[3.90 \pm 0.44]$	$[7.68 \pm 3.55]$

Table 11: Comparison between '90 and '91 Monte Carlo for the true efficiencies and the correction factors. The numbers in brackets are for events simulated with the '91 geometry.

Period	$\Gamma_{b\bar{b}}/\Gamma_{had}$ (%)	ϵ_b (%)	ϵ_{udsc} (%)
1990	$21.79 \pm 0.88 \pm 0.80$	77.49 ± 0.64	38.80 ± 0.40
1991	$23.13 \pm 0.68 \pm 0.37$	76.06 ± 0.45	37.80 ± 0.29

Table 12: Results obtained for '90 and '91 data with the correction factors estimated by Monte Carlo with the '90 and '91 geometries respectively. The first error on $\Gamma_{b\bar{b}}/\Gamma_{had}$ is the statistical error due to the number of real data events and the second one is the error induced by the statistical uncertainty on the correction factors.

N.Net. option	$\Gamma_{b\bar{b}}/\Gamma_{had}$ (%)	C^{DT}_{udsc} (%)	C^{ST}_{udsc} (%)
(1)	22.74 ± 0.54	3.8	5.9
(2)	22.83 ± 0.56	3.8	5.9
(3)	22.72 ± 0.60	-0.14	0.25
(4)	22.88 ± 0.56	8.4	7.3

Table 13: Results obtained on $\Gamma_{b\bar{b}}/\Gamma_{had}$ for different choices of Neural Network. The b-correction factors are not given since they are always compatible with zero for all the options.

$\epsilon_b(Peter.)$	C^{DT}_b (%)	C^{ST}_b (%)
0.008	0.023 ± 0.31	-0.52 ± 0.50
0.006	0.077 ± 0.30	-0.43 ± 0.47
0.0048	0.109 ± 0.30	-0.54 ± 0.48
0.004	0.202 ± 0.30	-0.63 ± 0.50
0.002	0.157 ± 0.36	-0.33 ± 0.65

Table 14: Effect of the b-fragmentation on C^{DT}_b and C^{ST}_b. Note that we have only used one half of the $b\bar{b}$ Monte Carlo events to estimated this effect and that the simulated events used in ALEPH have been generated with $\epsilon_b(Peter.) = 0.006$. Furthermore, the value 0.0048 corresponds to the result obtain by the Heavy Flavour group.

Cut on R	0.25	0.30	0.35	0.40
≥ 1 lepton(s)	-0.32 ± 0.23	-0.36 ± 0.29	-0.18 ± 0.36	-0.10 ± 0.44
1 lepton	-0.48 ± 0.27	-0.55 ± 0.34	-0.44 ± 0.42	-0.51 ± 0.52
≥ 2 leptons	5.91 ± 0.63	7.50 ± 0.84	10.3 ± 1.1	13.5 ± 1.3

Table 15: Comparison of C^{ST}_b (in %) estimated by Monte Carlo for $b\bar{b}$ events with one, at leat one and more than one high $p_\perp$ leptons.

Source	Effect on $\Gamma_{b\bar{b}}/\Gamma_{had}$ (%)
f_b^L	± 0.38
C_b^{DT}	± 0.10
C_{uds}^{DT}	± 0.13
C_b^{ST}	± 0.10
C_{uds}^{ST}	± 0.10
charm fragmentation	± 0.08
f_c	± 0.00
f_c^L	± 0.05
bottom fragmentation	± 0.06
b-semileptonic BR	± 0.03
backgrounds	± 0.01
geometrical affects	± 0.07
acceptance	± 0.02
Total	± 0.46

Table 16: Contributions to the systematic error on $\Gamma_{b\bar{b}}/\Gamma_{had}$.

Figure 1: The b-purity of the remaining sample as function of the b-efficiency of the applied cut on the variable PROBHEMI of qipbtag (triangles) and on the output of a Neural Network including PROBHEMI as one of its input variables (squares).

Figure 2: Track multiplicity (charged and neutral) of the most energetic jet of each hemisphere: comparison between light quark events (solid histogram) and $b\bar{b}$ events (dotted histogram).

Figure 3: Track multiplicity (charged and neutral) of the most energetic jet of each hemisphere: comparison between data (black points) and Monte Carlo (histogram).

Figure 4: Distribution of the nine variables A(I) used as input of the Neural Network, for b-quark (dotted line) and light quark events (solid line). The two contributions are normalized to the same area.

Figure 5: Distribution of the nine variables A(I) used as input of the Neural Network, for data (black points) and simulated events (histogram).

Figure 6: Shape of the Neural Network output $R_{1(2)}$ computed on one hemisphere for different flavour Monte Carlo events: *uds* hemispheres (solid line), c hemispheres (dotted line) and b hemispheres (dashed line).

Figure 7: Neural Network output $R_{1(2)}$ of one hemisphere for data (crosses) and hadronic Monte Carlo events (solid histogram), a) without and b) with a cut at 0.3 on the output of the opposite hemisphere.

Figure 8: Distribution of the product $R_1 \times R_2$ for simulated events (same notations as for fig. 6).

Figure 9: b-purity of the remaining sample as function of the b-efficiency of the applied cut on the hemisphere Neural Network output (squares) and on the product $R_1 \times R_2$ (triangles).

Figure 10: b-purity of the remaining sample as function of the b-efficiency of the applied cut on each of the nine input variables A(I). Note that none of these purities never exceeds 50% while this quantity can reach 70% in the case of the Neural Network output.

Figure 11: Distribution of the Neural Network output for $\tau^+\tau^-$ events (solid histogram) and for $q\bar{q}$ events (dotted histogram). The two contributions are normalized to the same area.

Figure 12: Lepton energy spectum in the c-hadron rest frame for $c \rightarrow l$ transitions: comparison between the Monte Carlo predictions (dotted histogram) and the data from DELCO: a) JETSET prediction, b) model of Altarelli et al.

Figure 13: Values of the correction coefficients $1 + C_b^{DT}$ and $1 + C_b^{ST}$ as function of the cut applied on the hemispshere Neural Network output.

Figure 14: Values of the correction coefficients $1 + C_{udsc}^{DT} + c_{udsc}^{DT}$ and $1 + C_{udsc}^{ST} + c_{udsc}^{ST}$ as function of the cut applied on the hemispshere Neural Network output.

Figure 15: Comparison of the Neural Network output between '90 (stars), '91 (squares) and '92 (triangles) data.

Figure 16: Mesaured values of $\Gamma_{b\bar{b}}/\Gamma_{had}$ in '90+'91: a) for different cuts on the Neural Network output with $p_\perp \geq 1.25$ GeV, b) for different $p_\perp$ cuts and a cut on the Neural Network output at 0.3.

Figure 17: Mesaured values of ϵ_b and ϵ_{udsc} obtained by solving the system of three equations for '90+'91 data for different cuts on the Neural Network output.

Figure 18: Mesaured values of f_b obtained on 850,000 $q\bar{q}$ simulated events treated as if they were real data, for different cuts on the Neural Network output.

Figure 19: Ratio of the b-efficiencies obtained with real and simulated data as function of the b-efficiency measured in the real data.

Figure 20: Ratio of the light quark efficiencies obtained with real and simulated data as function of ϵ_{udsc} measured in the real data.

Figure 21: Neural Network output of one hemisphere for Monte Carlo events simulated with the '90 geometry (histogram) and the '91 geometry (squares): a) without, b) with a cut at 0.3 on the output of the opposite hemisphere.

Figure 22: Comparison of the correction factors $1 + C_b^{DT}$ and $1 + C_b^{ST}$ for b quarks estimated with '90 and '91 simulated events as function of the cut applied on the Neural Network output.

Figure 23: Same as figure 22 for light quark events.

Figure 24: Measured values of $\Gamma_{b\bar{b}}/\Gamma_{had}$ as function of the cut applied on the Neural Network output for different choices of Neural Networks.

Figure 25: Effect of the Double Tag correlations on the shape of the hemisphere output; comparison of the Neural Network output for $b\bar{b}$ simulated events without (solid histogram) and with (black points) a cut at 0.3 on the opposite hemisphere.

Figure 26: Same as figure 25 for light quark events.

Figure 27: Comparison of the Neural Network output of $q\bar{q}$ events with (solid histogram) and without (black points) a cut at 0.3 on the opposite hemisphere: a) for simulated events, b) for data.

Figure 28: Ratio of the bin contents of the Neural Network output obtained with and without a cut on the opposite hemisphere: comparison data (dotted crosses) Monte Carlo (solid crosses).

Figure 29: Comparison data (black points) Monte Carlo (histogram) for a b-enriched sample obtained with qipbtag: a) without and b) with a cut at 0.3 on the opposite hemisphere.

Figure 30: Same as figure 29 for light quark events.

Figure 31: Same as figure 28 for the *udsc*-enriched sample.

Figure 32: Neural Network output on simulated $b\bar{b}$ events for $\epsilon_b(Peter.) = 0.006$ (histogram) and $\epsilon_b(Peter.) = 0.004$ (crosses).

Figure 33: Dependence with $\epsilon_b(Peter.)$ of the correction factors $1 + C_b^{DT}$ and $1 + C_b^{ST}$ for simulated $b\bar{b}$ events as function of the cut applied on the Neural Network output.

Figure 34: Comparison of the Neural Network output for hadronic events (solid histogram) and for the hemisphere containing a high $p_\perp$ lepton (dotted histogram).

Figure 35: Comparison of the Neural Network output for hadronic events (solid histogram) and for the hemisphere opposite to the high $p_\perp$ lepton for single lepton events (black points).

Figure 36: Comparison of the Neural Network output for hadronic events (solid histogram) and for dilepton events (black points).

Figure 1

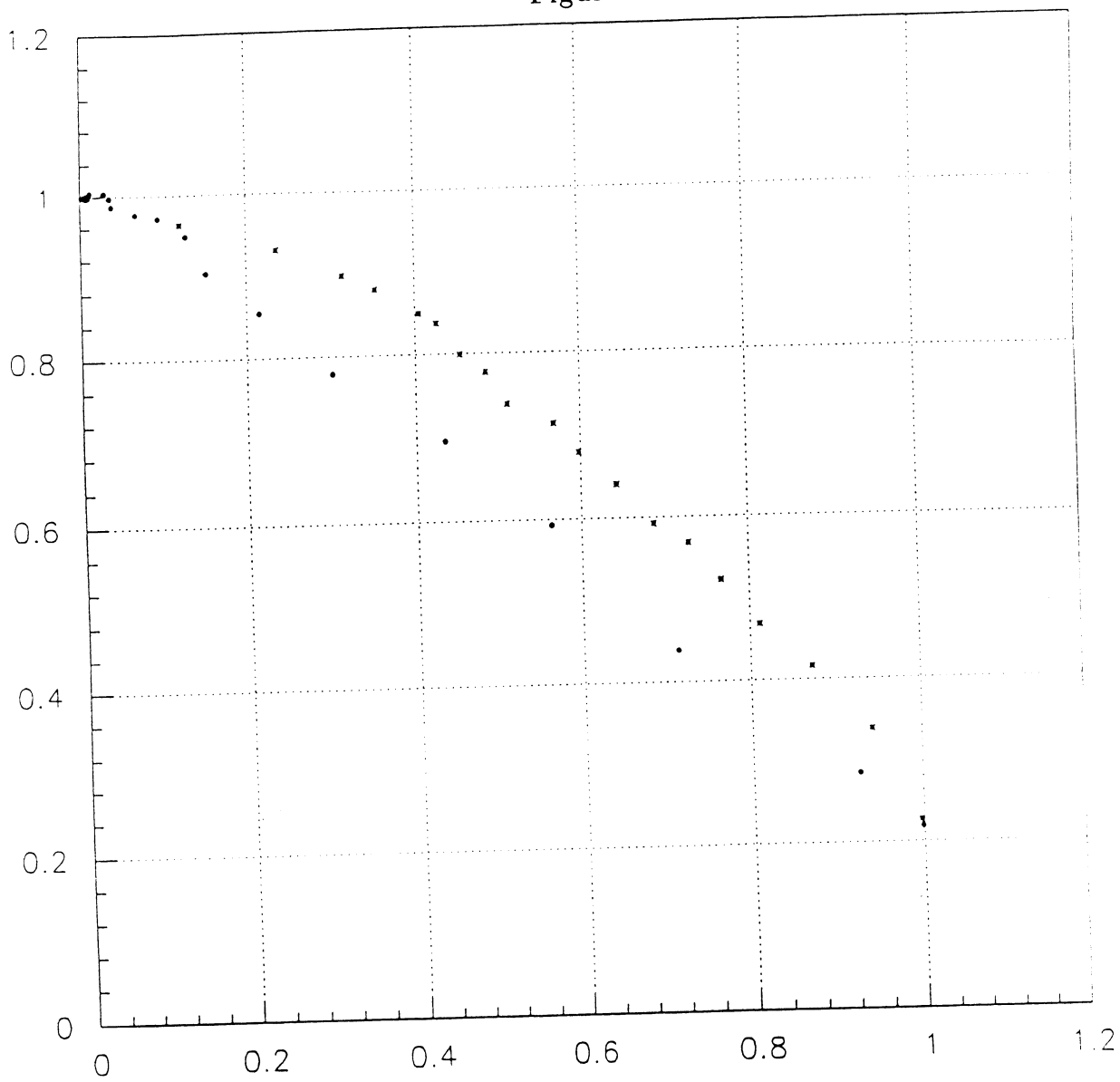


Figure 2

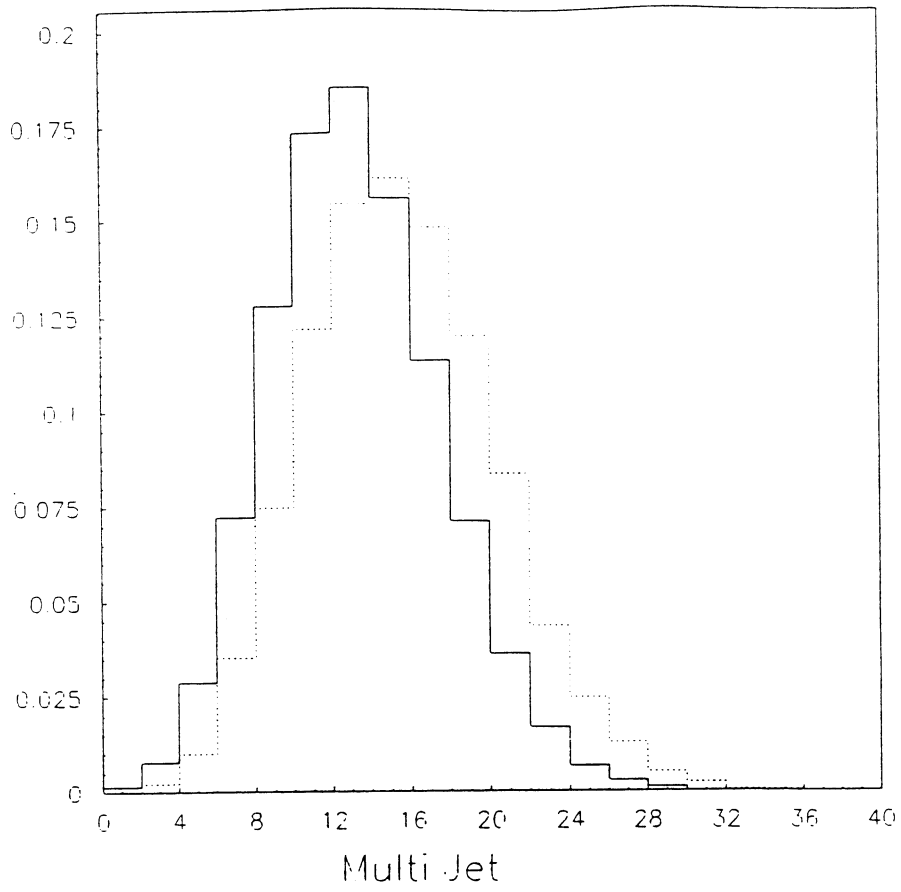


Figure 3

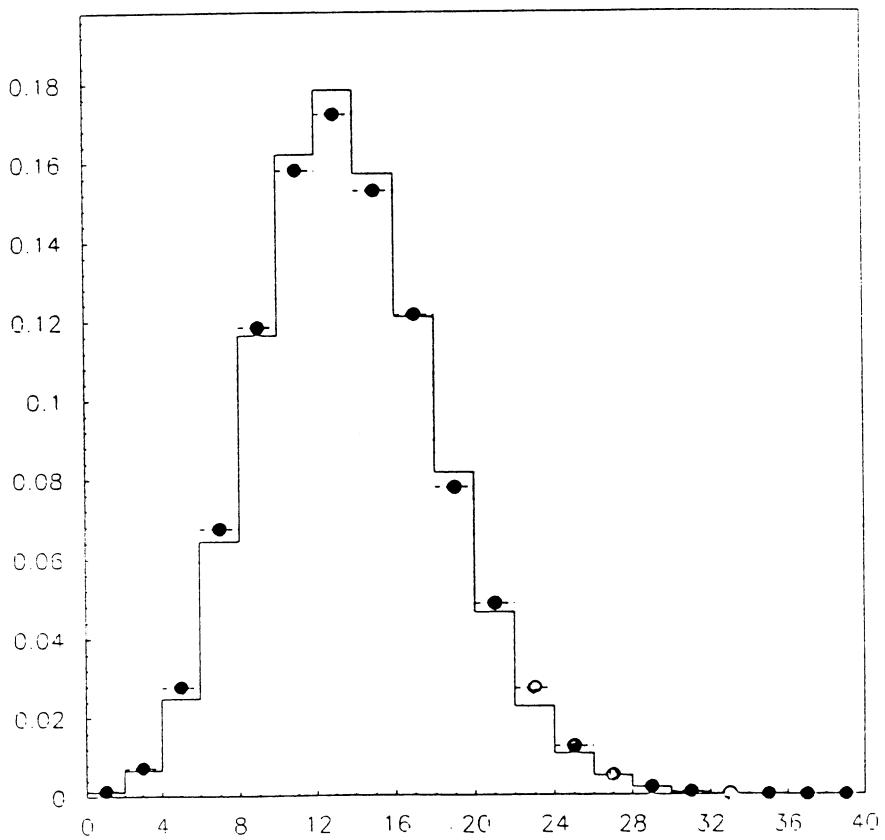


Figure 4

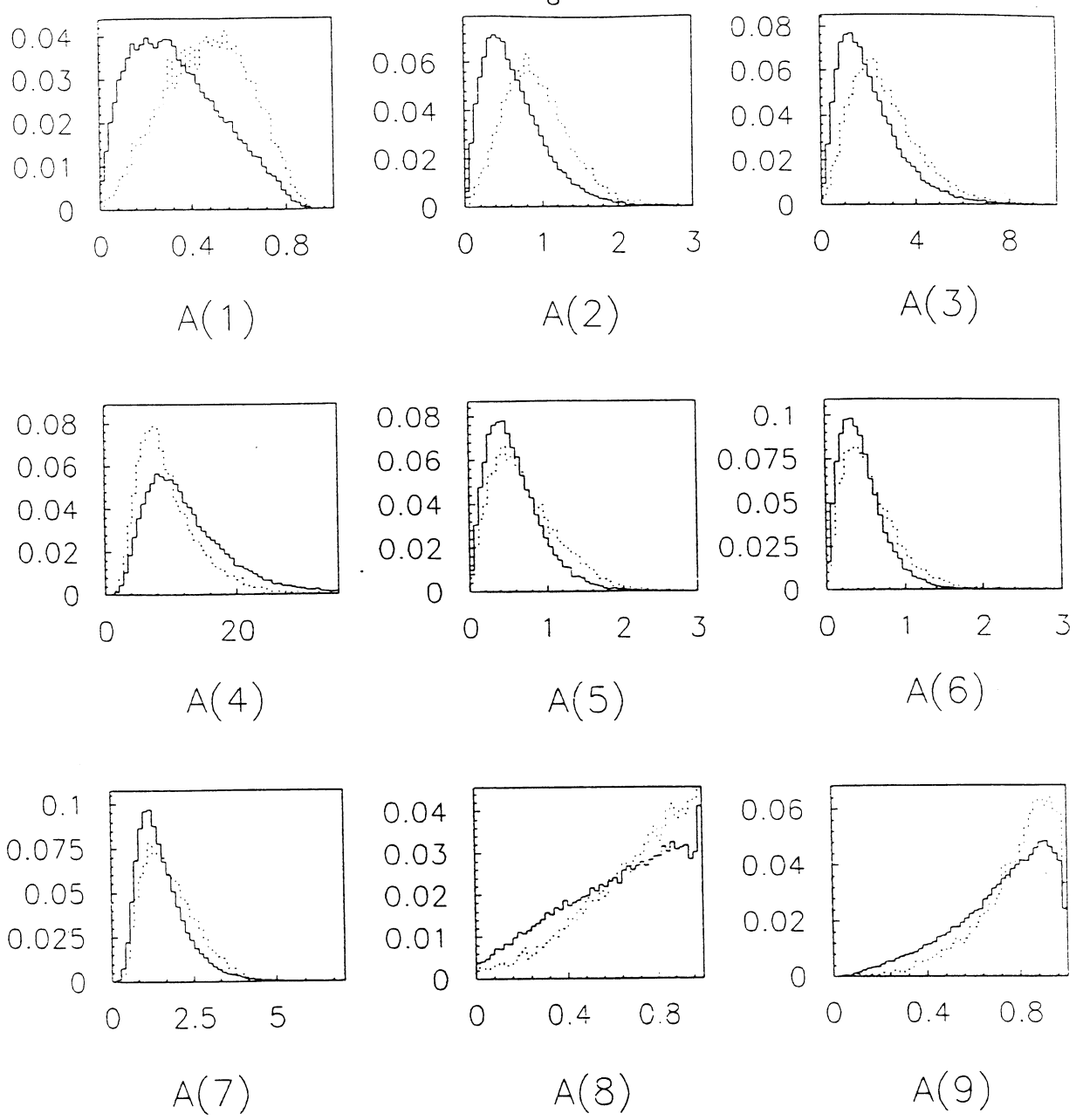

0.04
0.03
0.02
0.01
0
0
0.4
0.8
A(1)
0.06
0.04
0.02
0
0
1
2
3
A(2)
0.08
0.06
0.04
0.02
0
0
4
8
A(3)
0.08
0.06
0.04
0.02
0
0
20
A(4)
0.08
0.06
0.04
0.02
0
0
1
2
3
A(5)
0.1
0.075
0.05
0.025
0
0
1
2
3
A(6)
0.1
0.075
0.05
0.025
0
0
2.5
5
A(7)
0.04
0.03
0.02
0.01
0
0
0.4
0.8
A(8)
0.06
0.04
0.02
0
0
0.4
0.8
A(9)

Figure 5

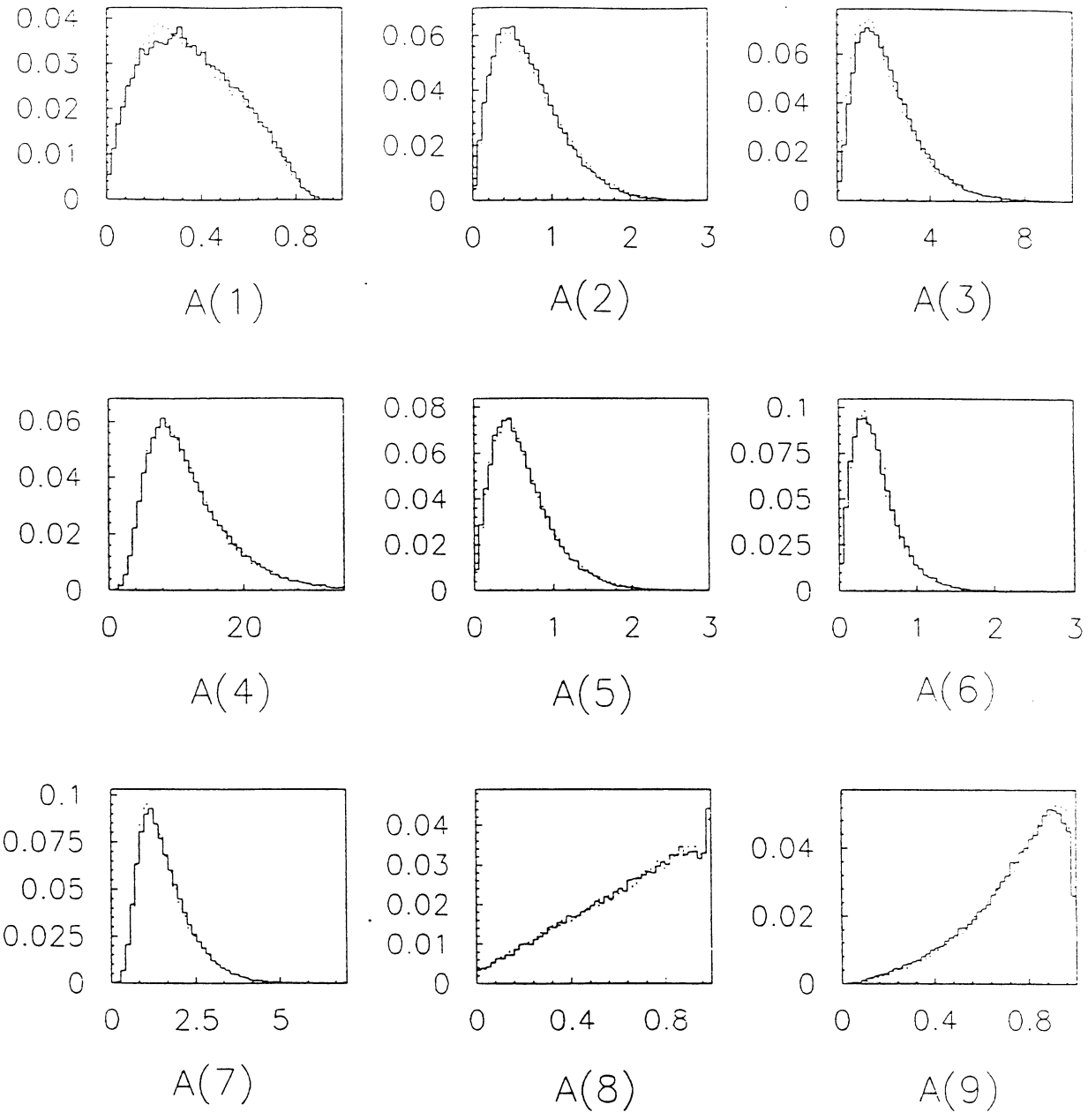

0.04
0.03
0.02
0.01
0
0 0.4 0.8
A(1)
0.06
0.04
0.02
0
0 1 2 3
A(2)
0.06
0.04
0.02
0
0 4 8
A(3)
0.06
0.04
0.02
0
0 20
A(4)
0.08
0.06
0.04
0.02
0
0 1 2 3
A(5)
0.1
0.075
0.05
0.025
0
0 1 2 3
A(6)
0.1
0.075
0.05
0.025
0
0 2.5 5
A(7)
0.04
0.03
0.02
0.01
0
0 0.4 0.8
A(8)
0.04
0.02
0
0 0.4 0.8
A(9)

Figure 6

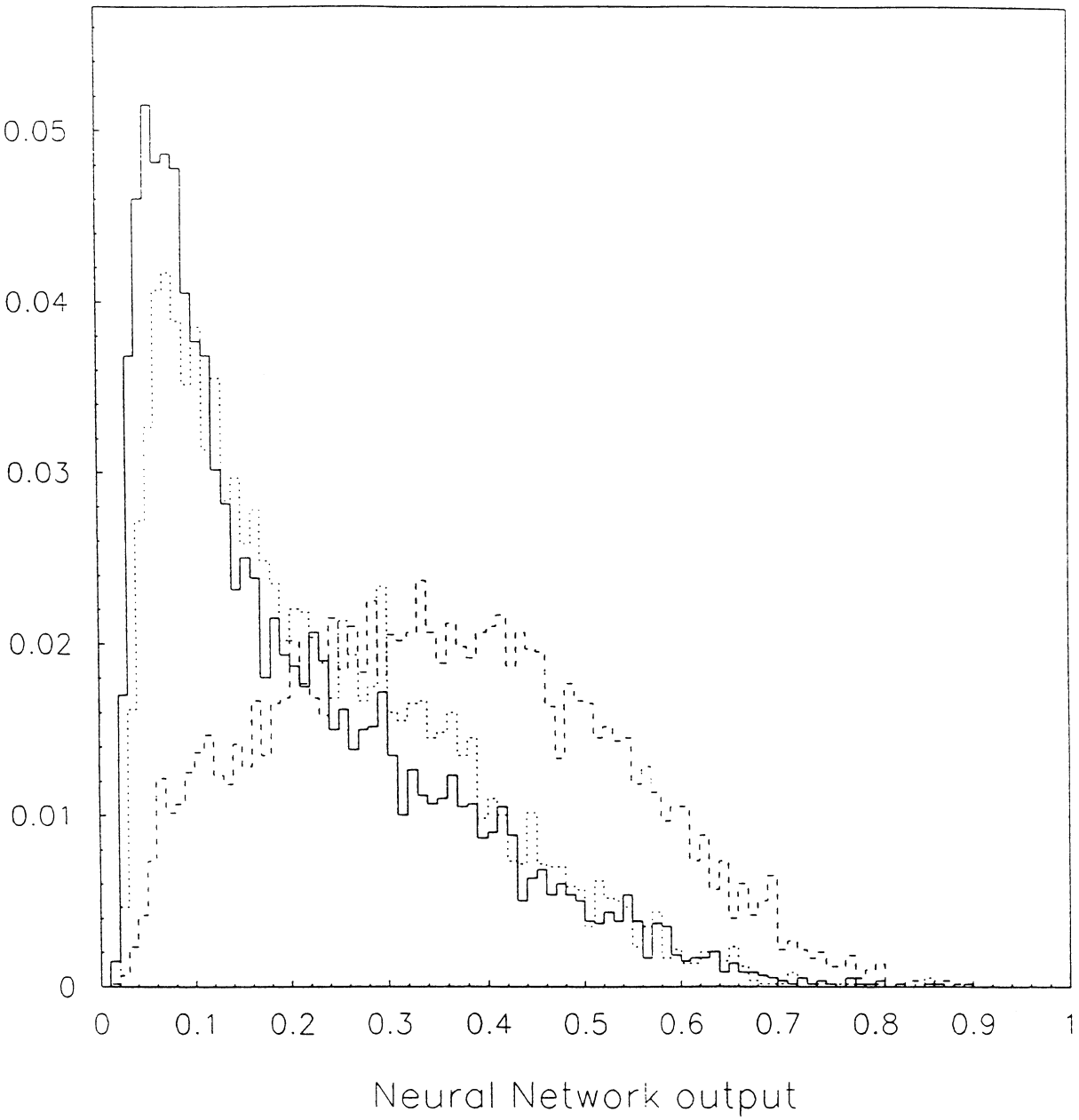

0.05
0.04
0.03
0.02
0.01
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Neural Network output

Figure 7-a

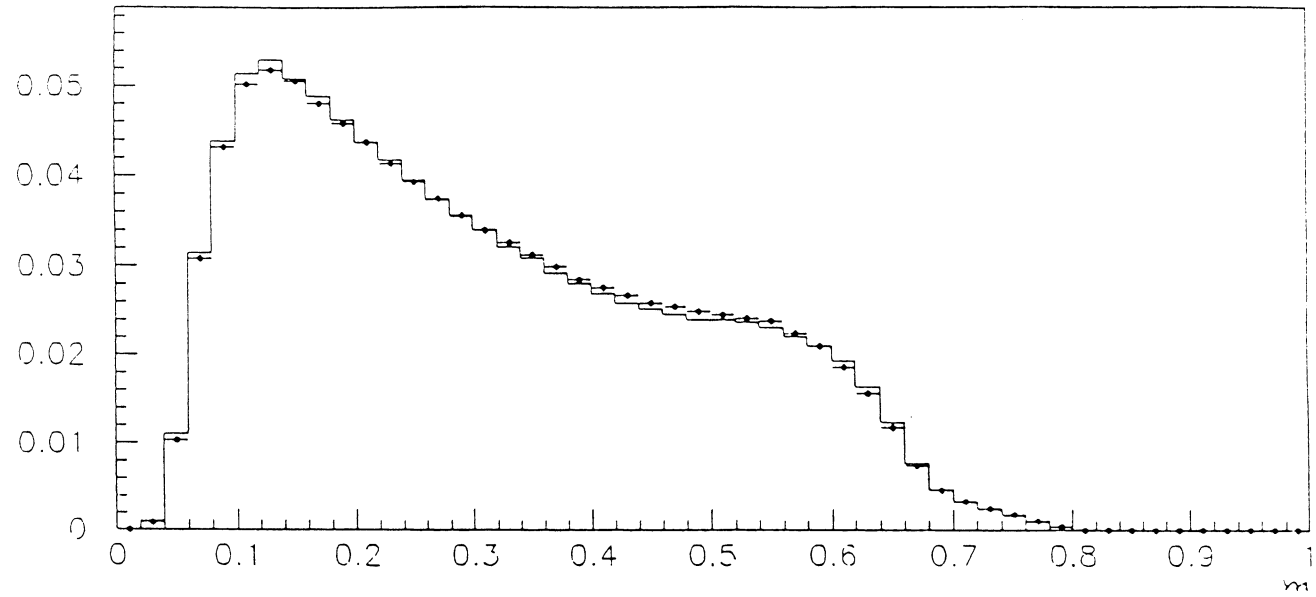


Figure 7-b

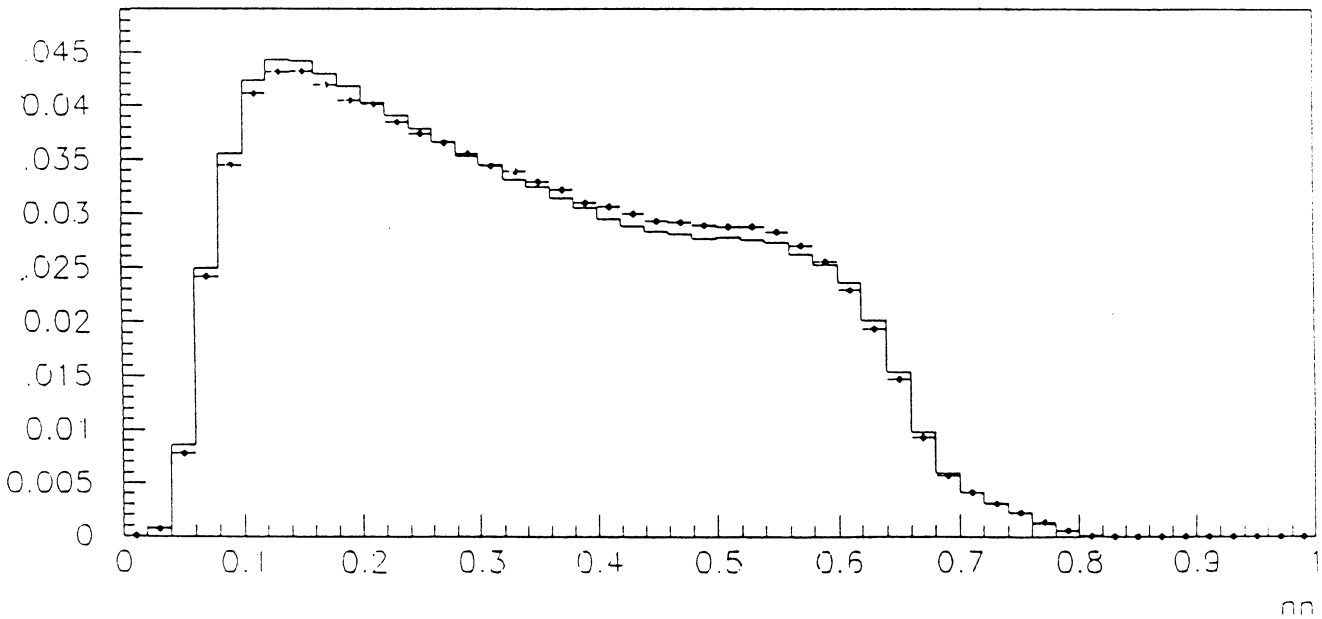


Figure 8

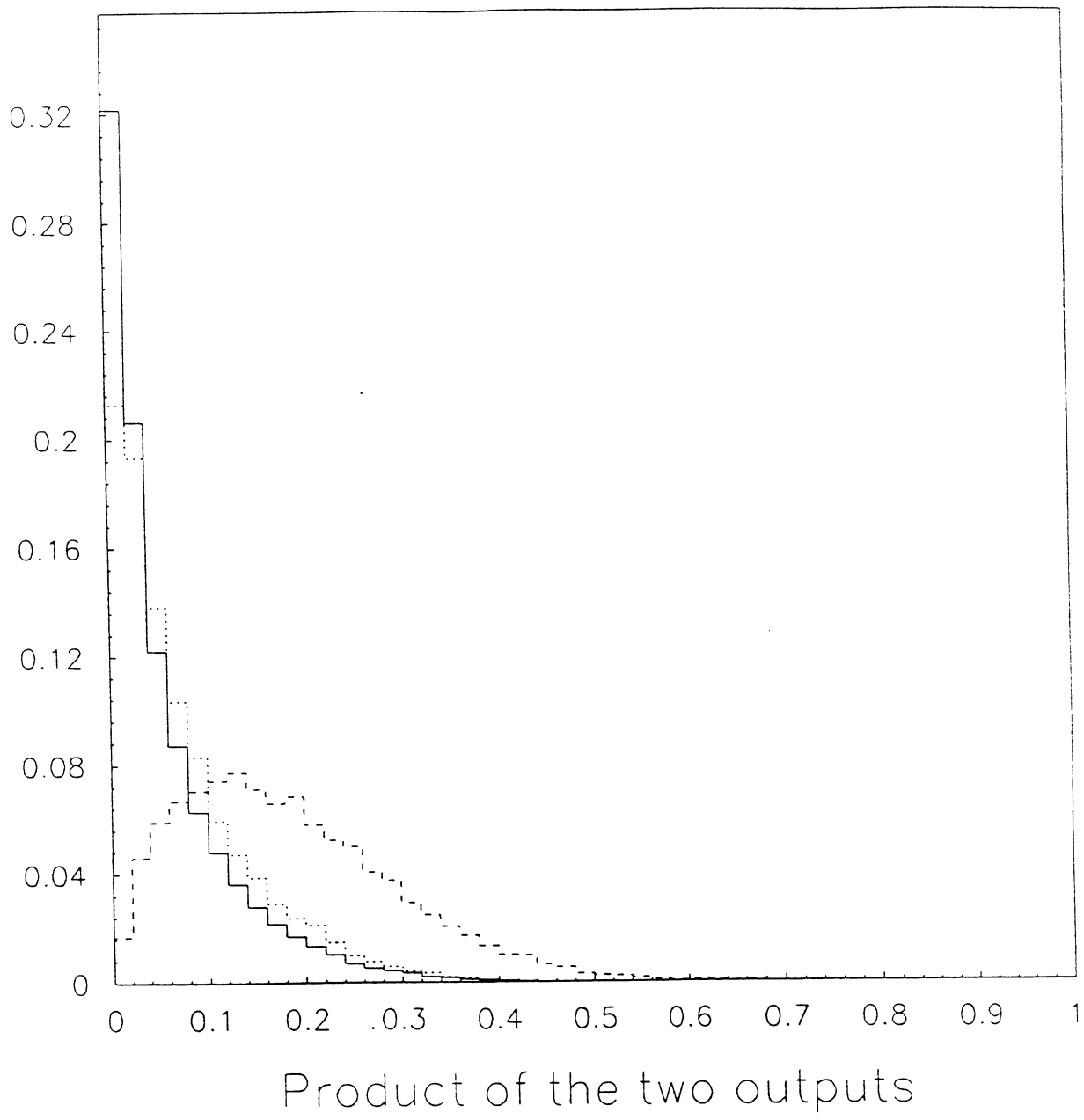

0.32
0.28
0.24
0.2
0.16
0.12
0.08
0.04
0
0
0.1
0.2
.0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Product of the two outputs

Figure 9

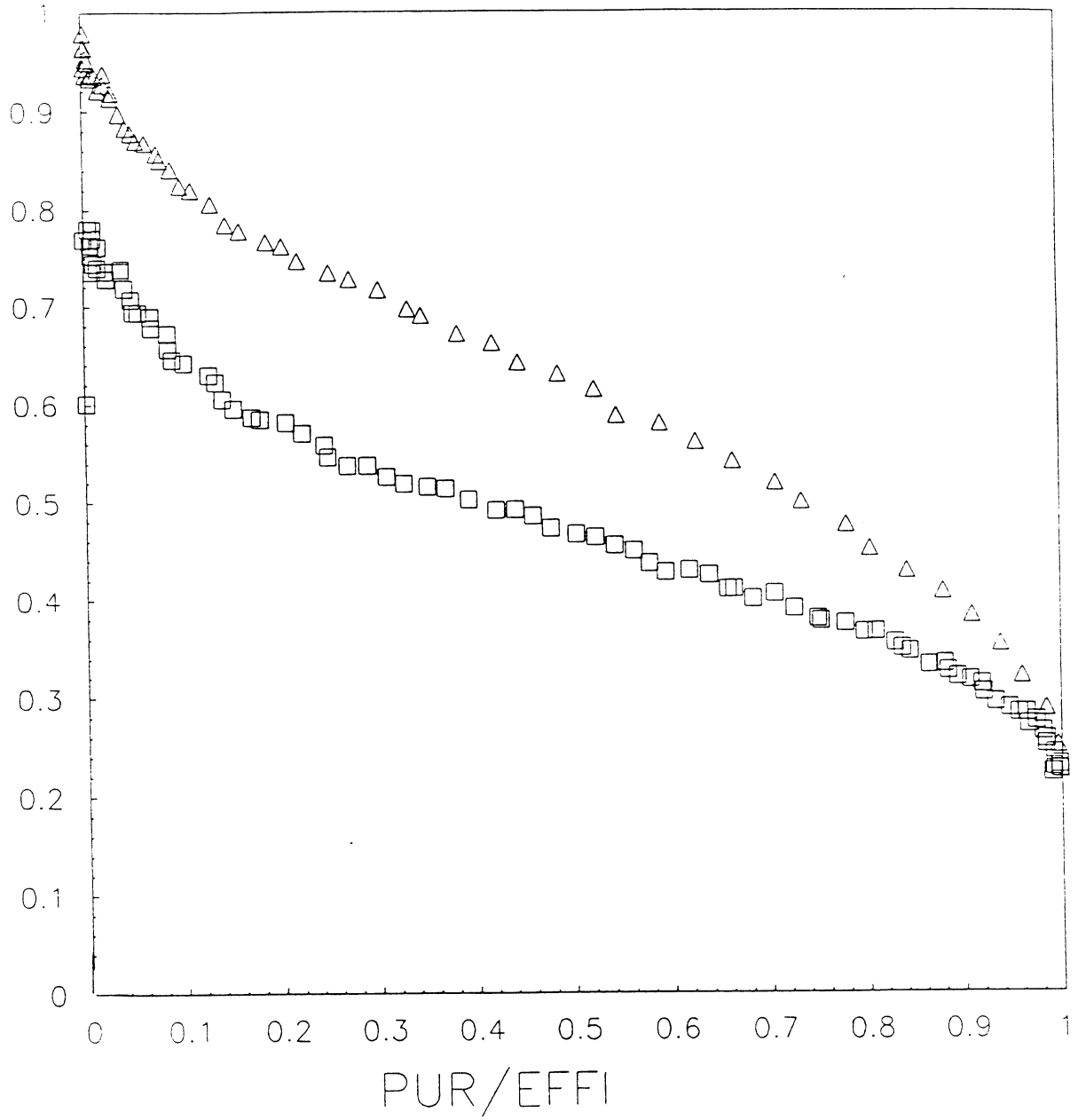


Figure 10

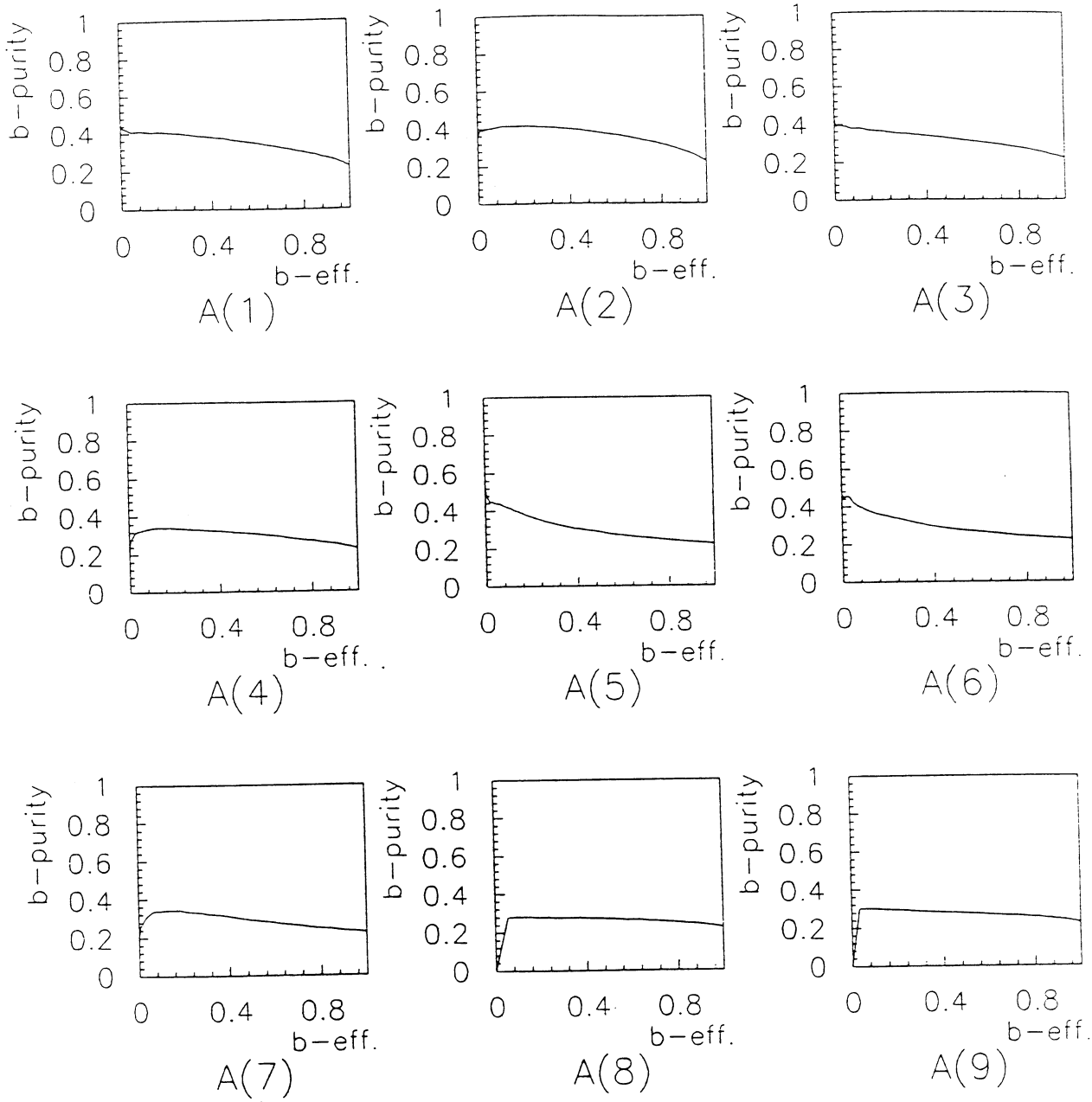


Figure 11

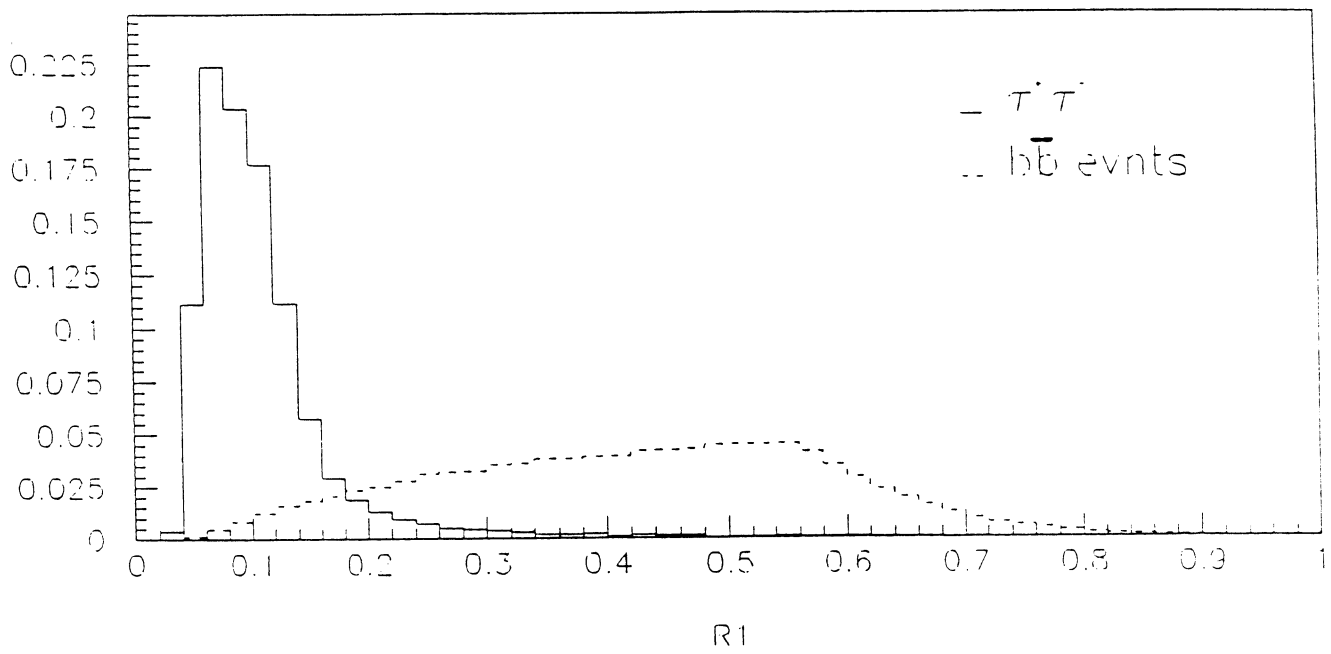

0.225
0.2
0.175
0.15
0.125
0.1
0.075
0.05
0.025
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
R1
τ τ
bb̄ evnts

Figure 12-a

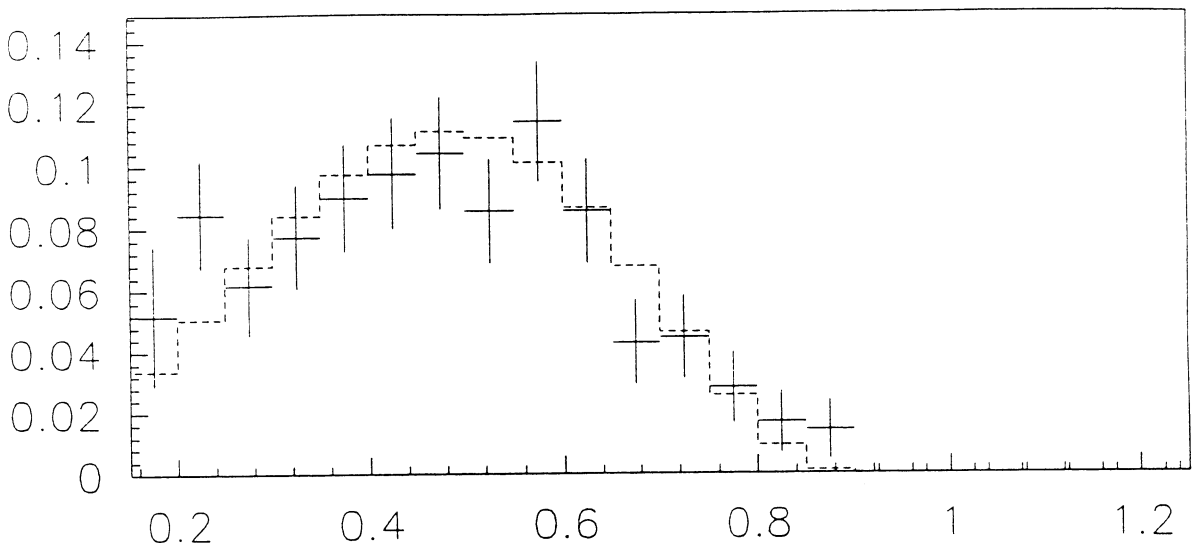


Figure 12-b

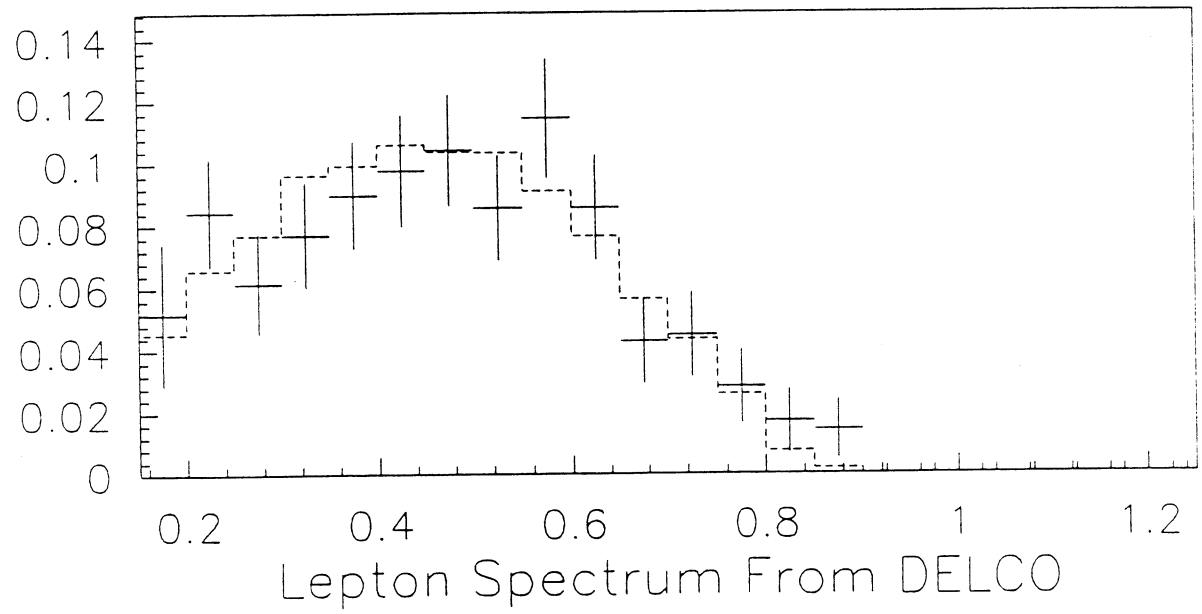


Figure 13

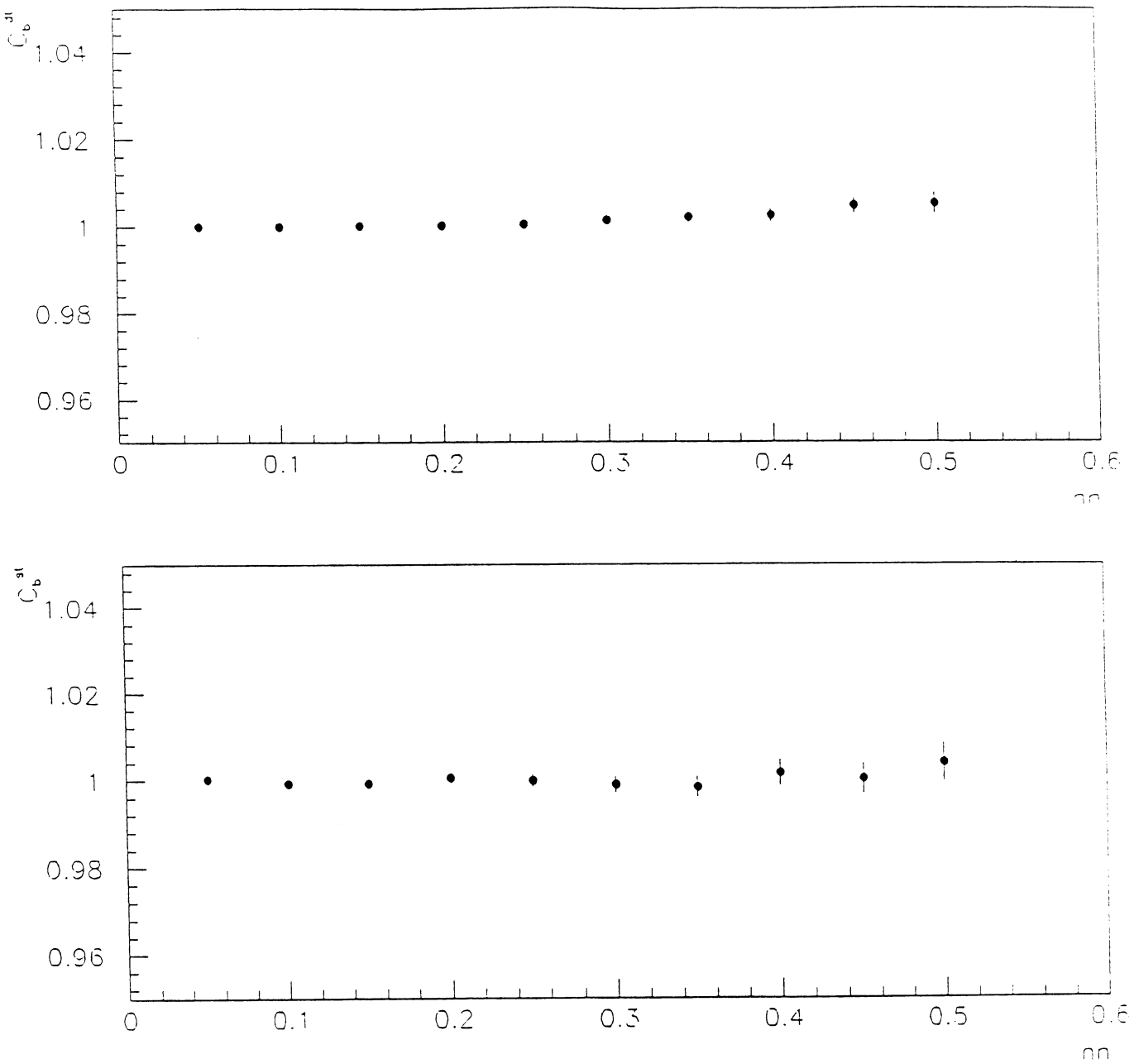


Figure 14

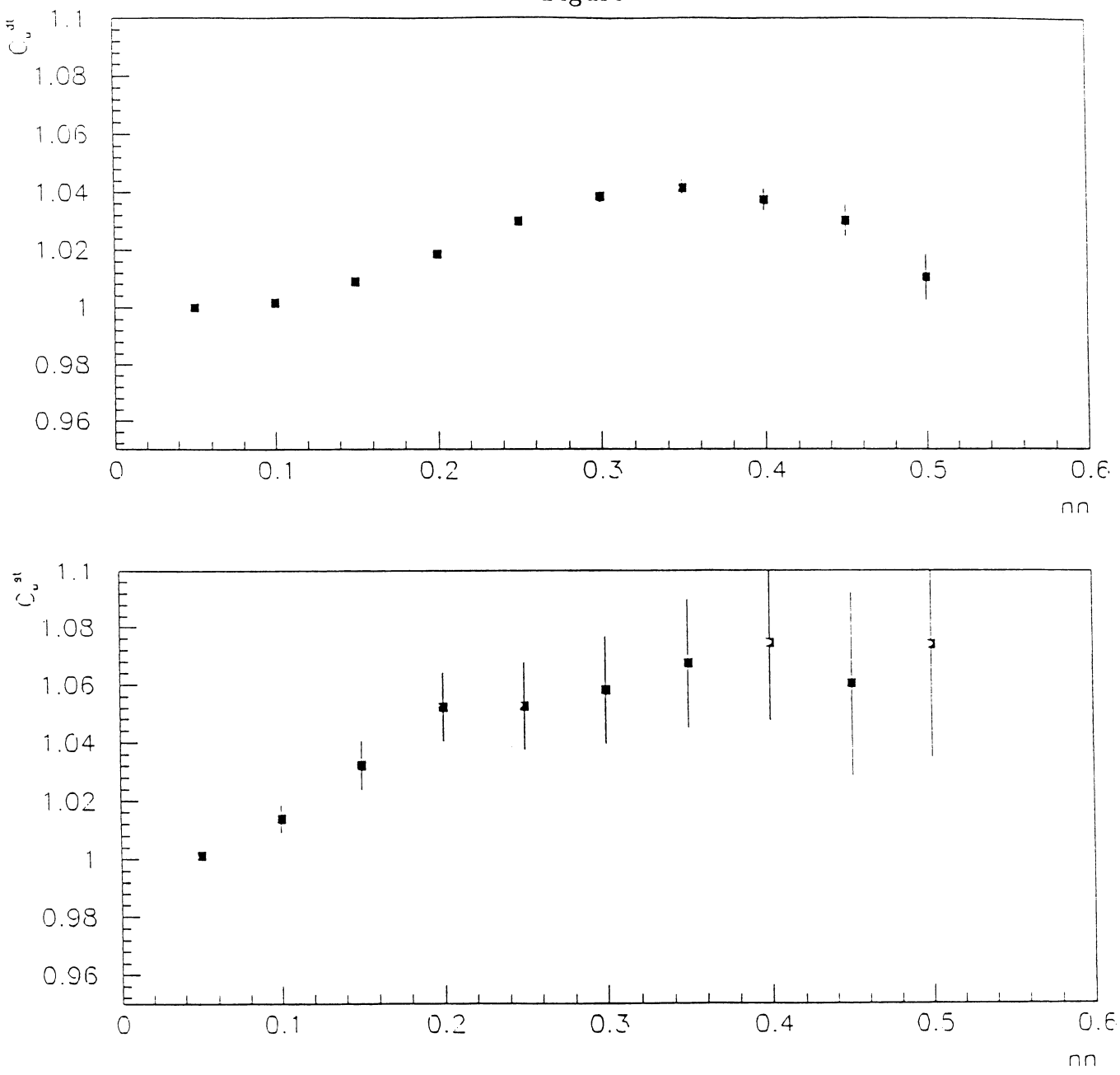


Figure 15

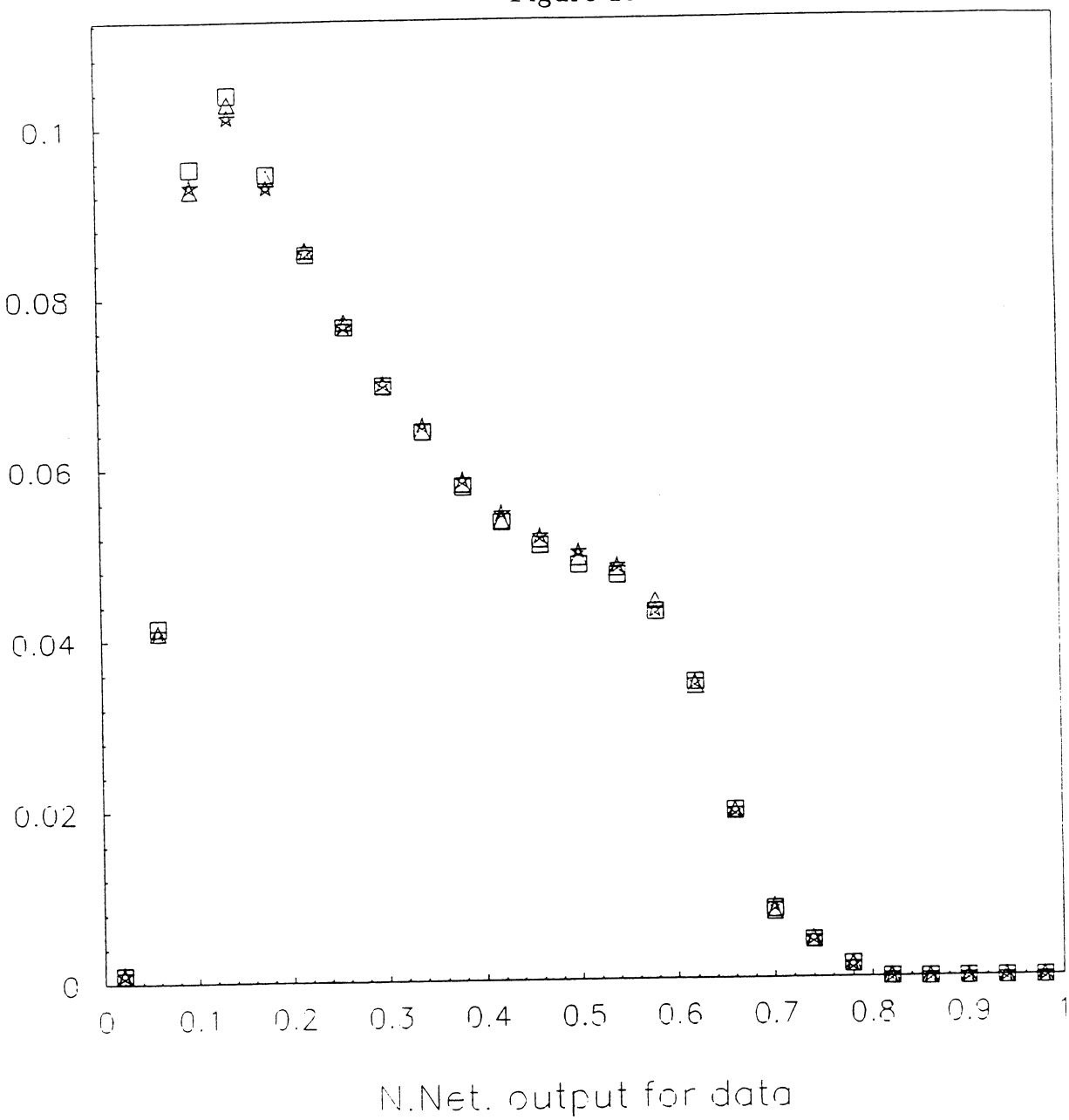

0.1
0.08
0.06
0.04
0.02
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
N.Net. output for data

Figure 16-a

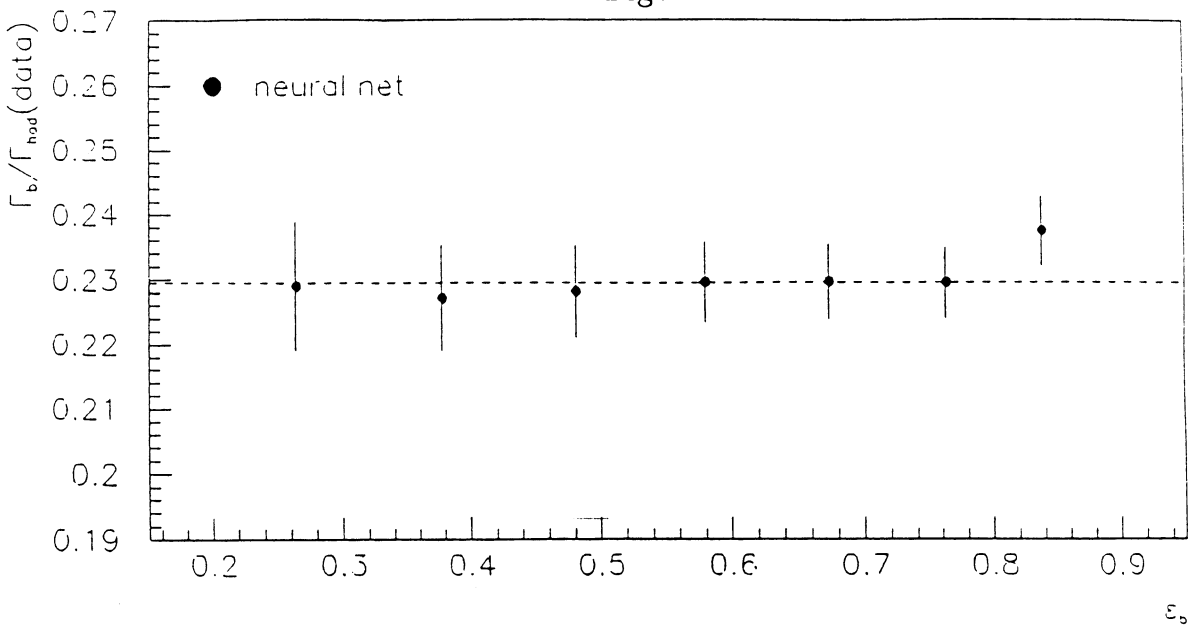


Figure 16-b

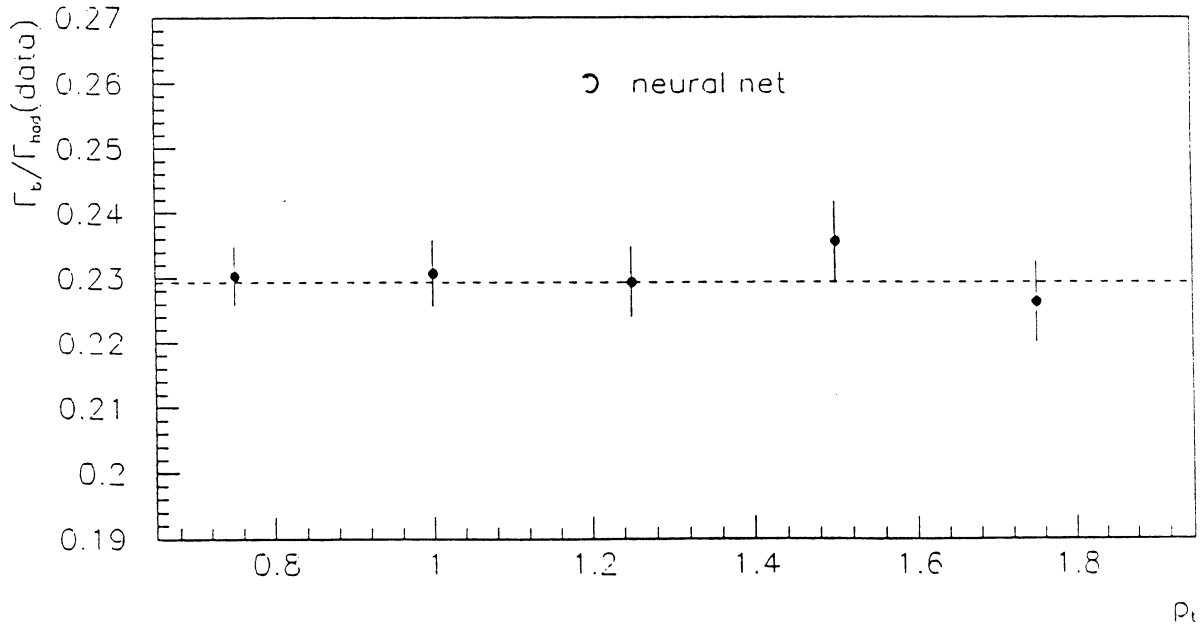


Figure 17

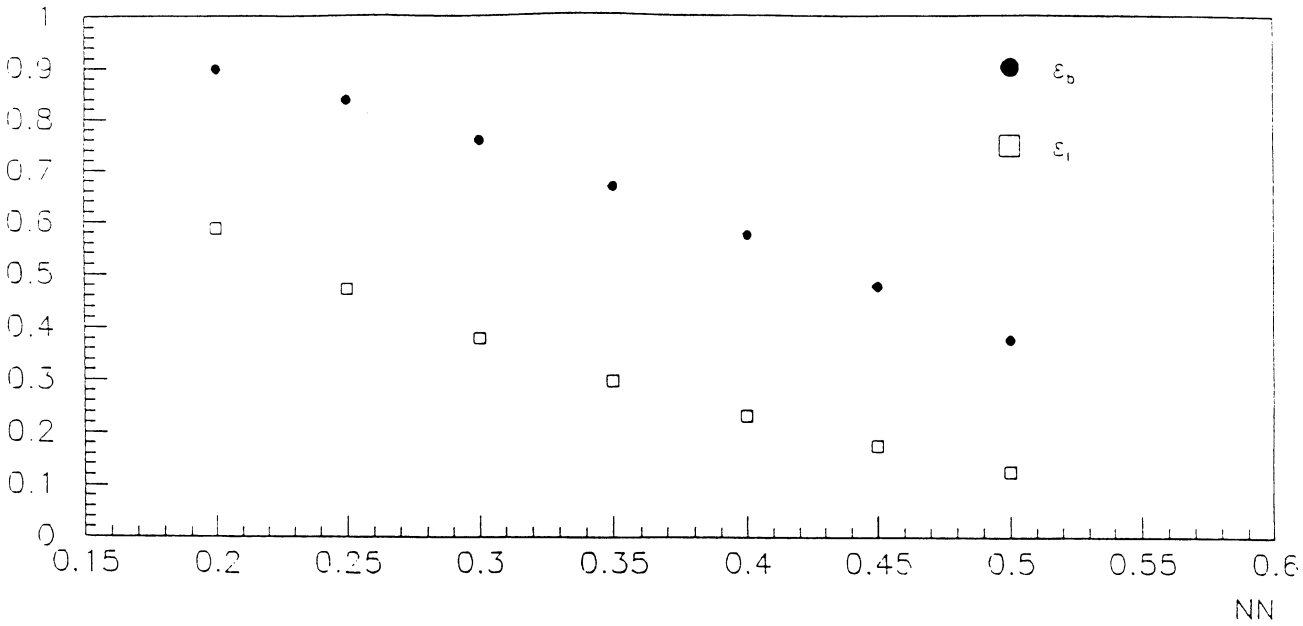


Figure 18

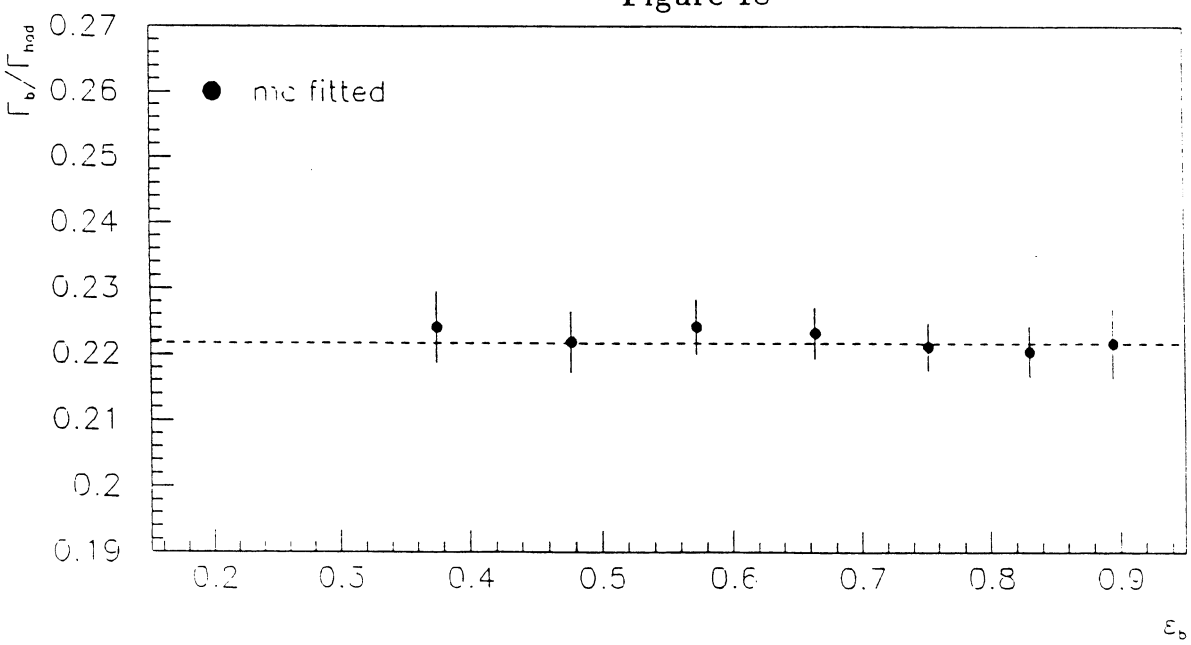


Figure 19

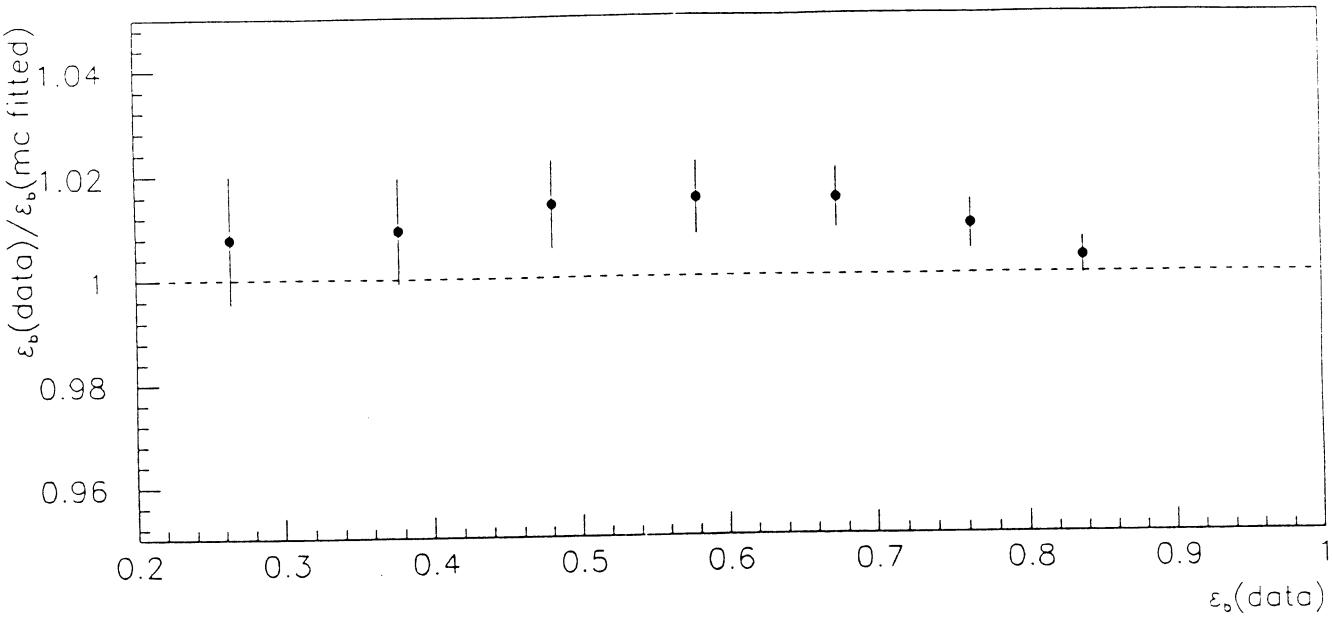


Figure 20

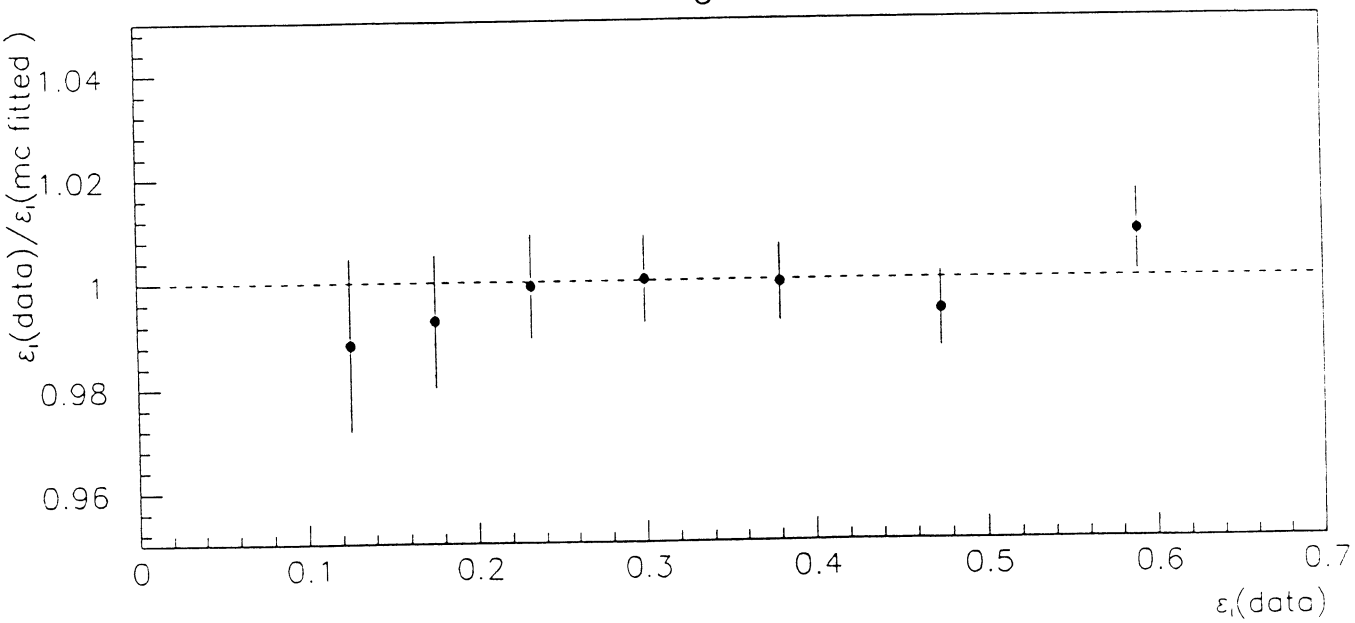


Figure 21-a

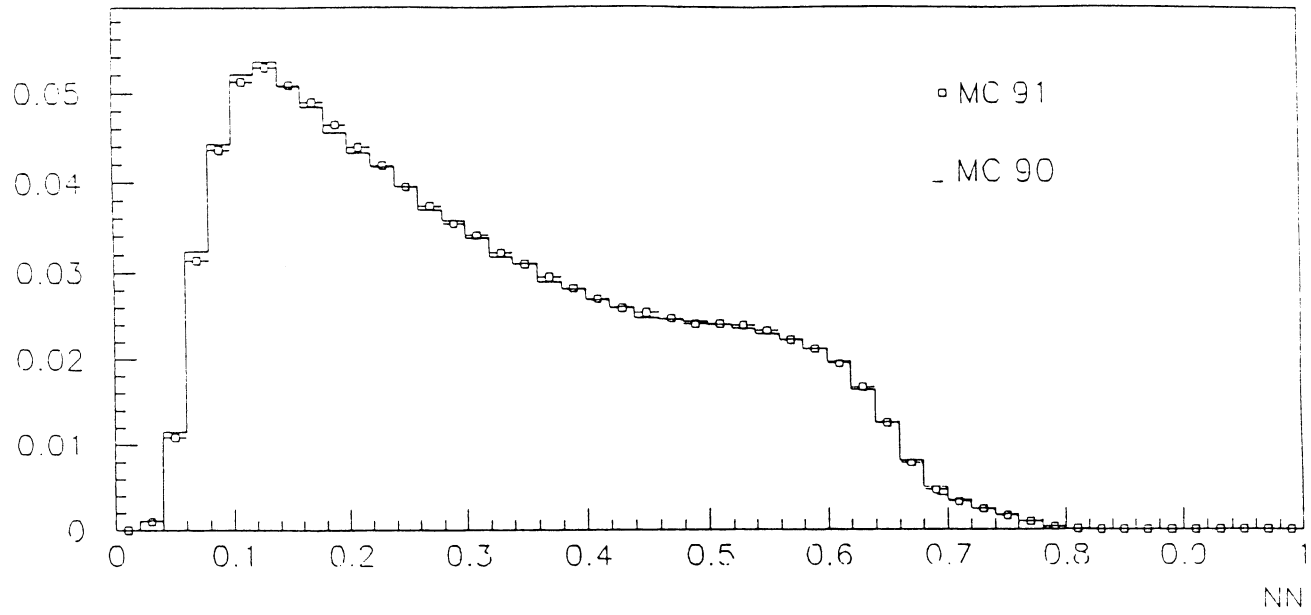


Figure 21-b

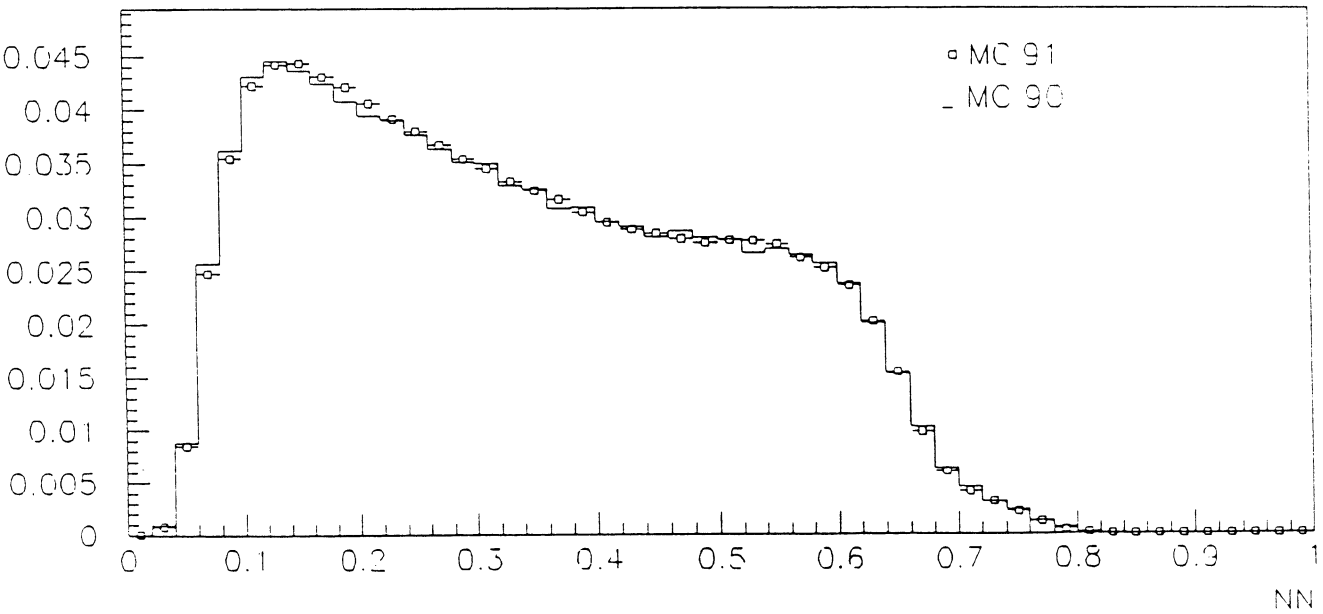


Figure 22

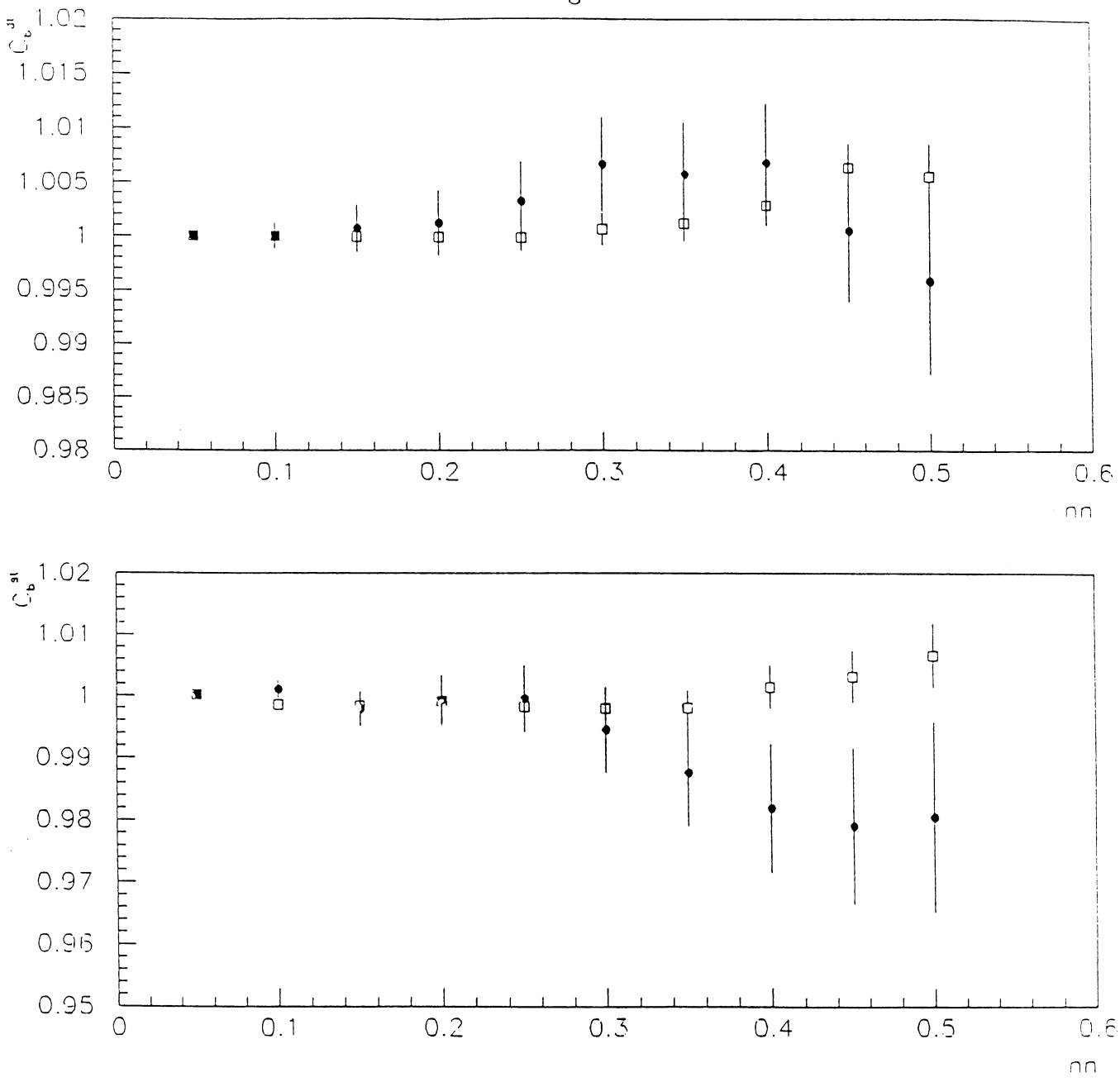
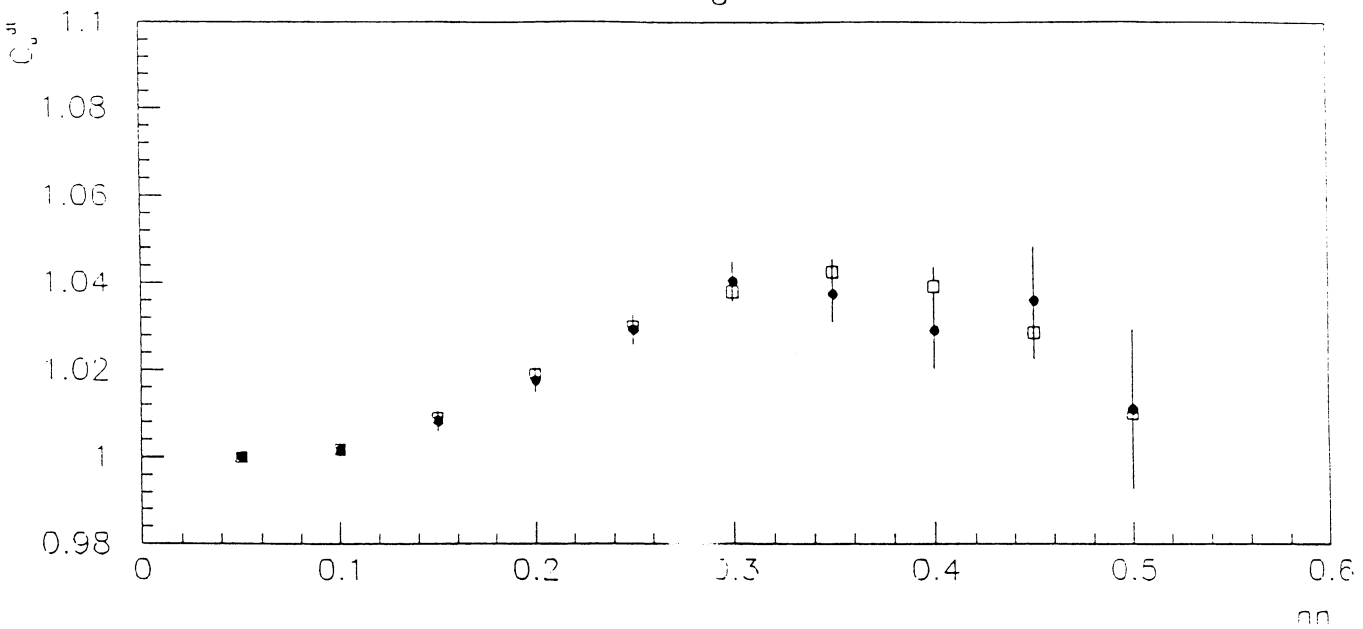

1.02
1.015
1.01
1.005
1
0.995
0.99
0.985
0.98
0
0.1
0.2
0.3
0.4
0.5
0.6
1.02
1.01
1
0.99
0.98
0.97
0.96
0.95
0
0.1
0.2
0.3
0.4
0.5
0.6

Figure 23


1.1
1.08
1.06
1.04
1.02
1
0.98
0
0.1
0.2
0.3
0.4
0.5
0.6

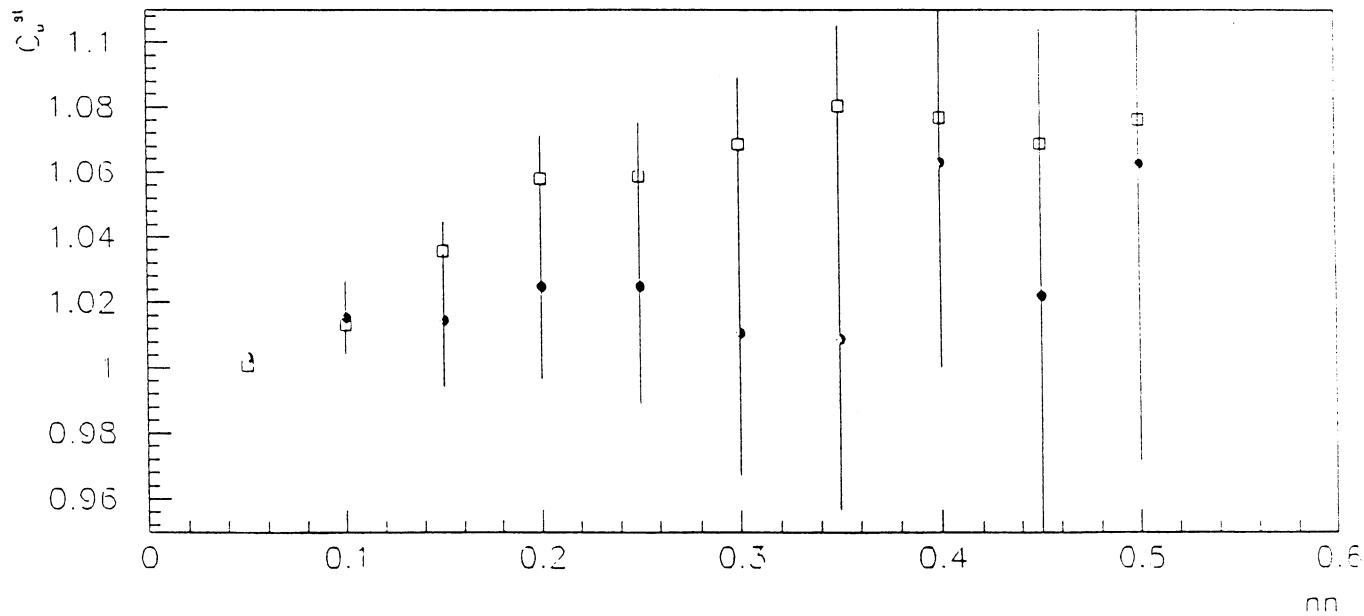

1.1
1.08
1.06
1.04
1.02
1
0.98
0.96
0
0.1
0.2
0.3
0.4
0.5
0.6

Figure 24

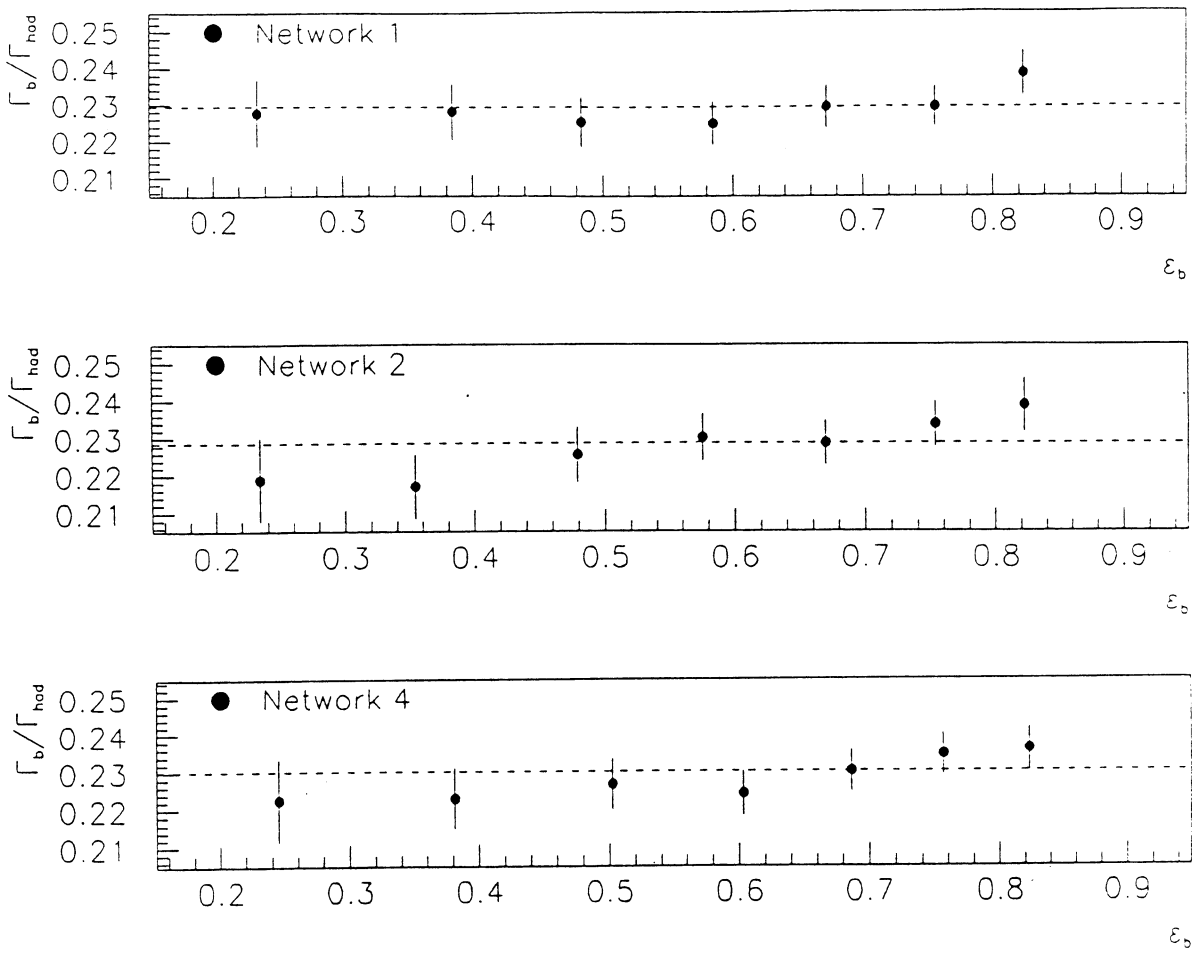

Network 1
Network 2
Network 4
Γ_b/Γ_{had}
0.25
0.24
0.23
0.22
0.21
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
ε_b

Figure 25

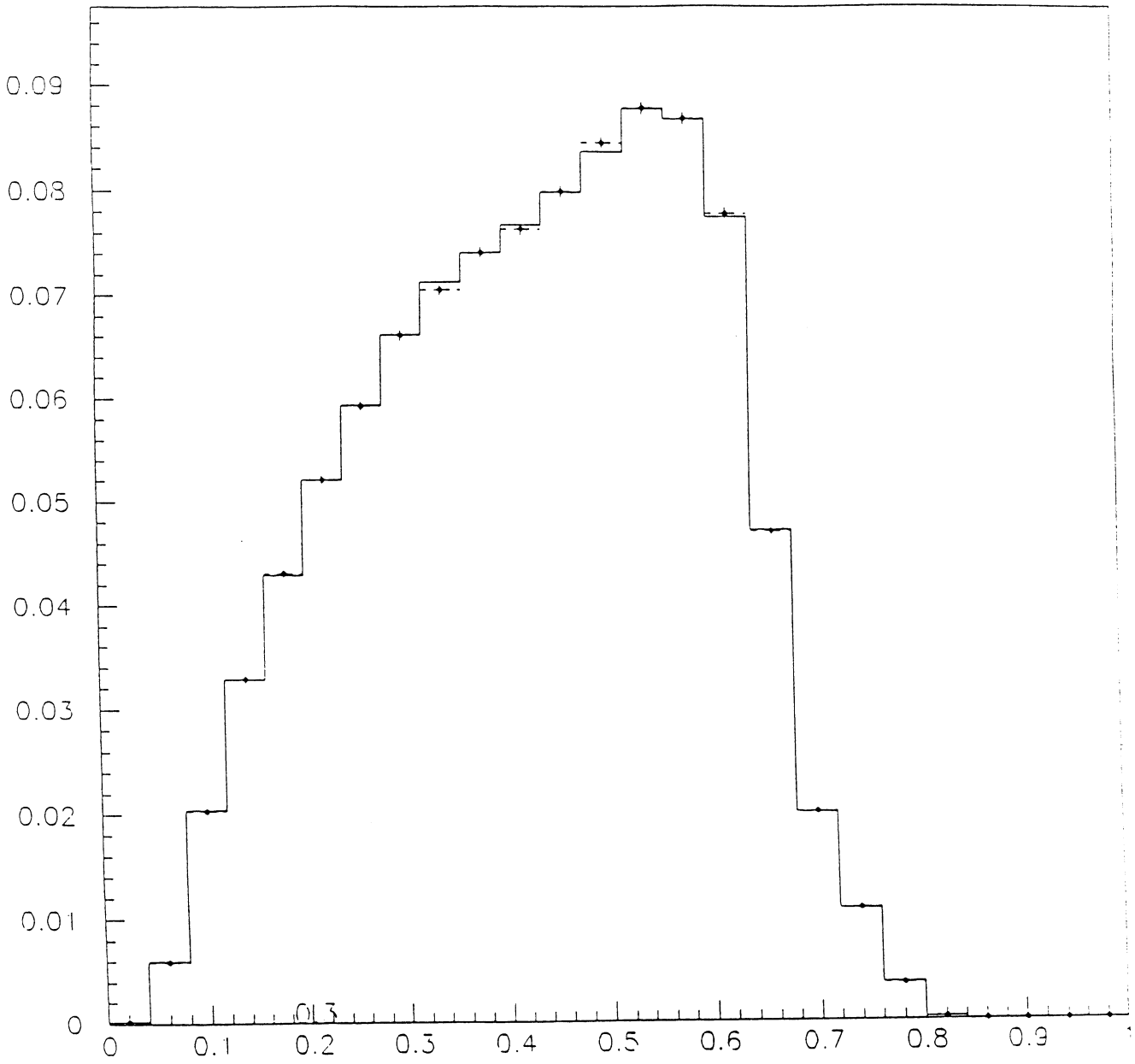


Figure 26

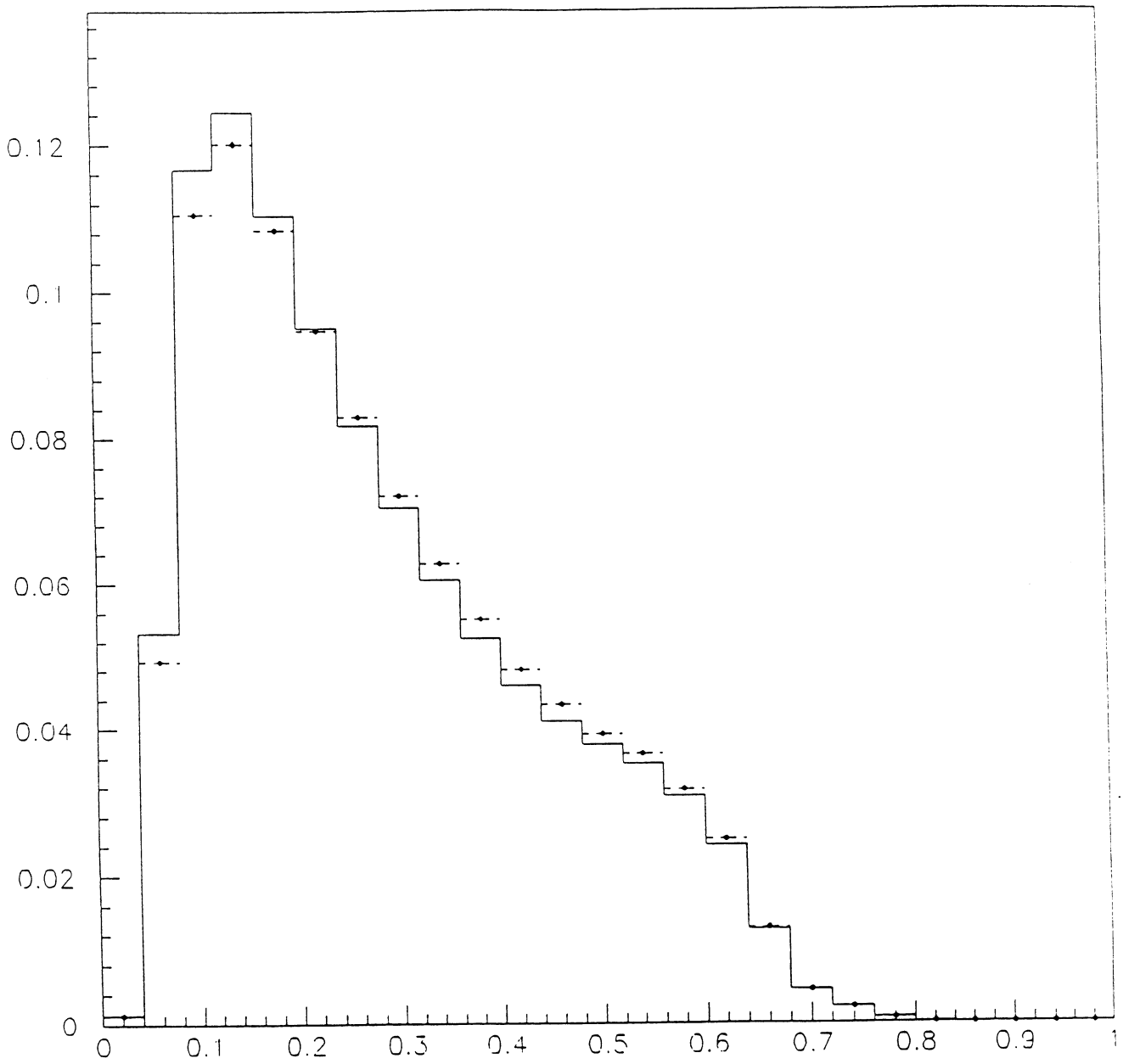


Figure 27-a

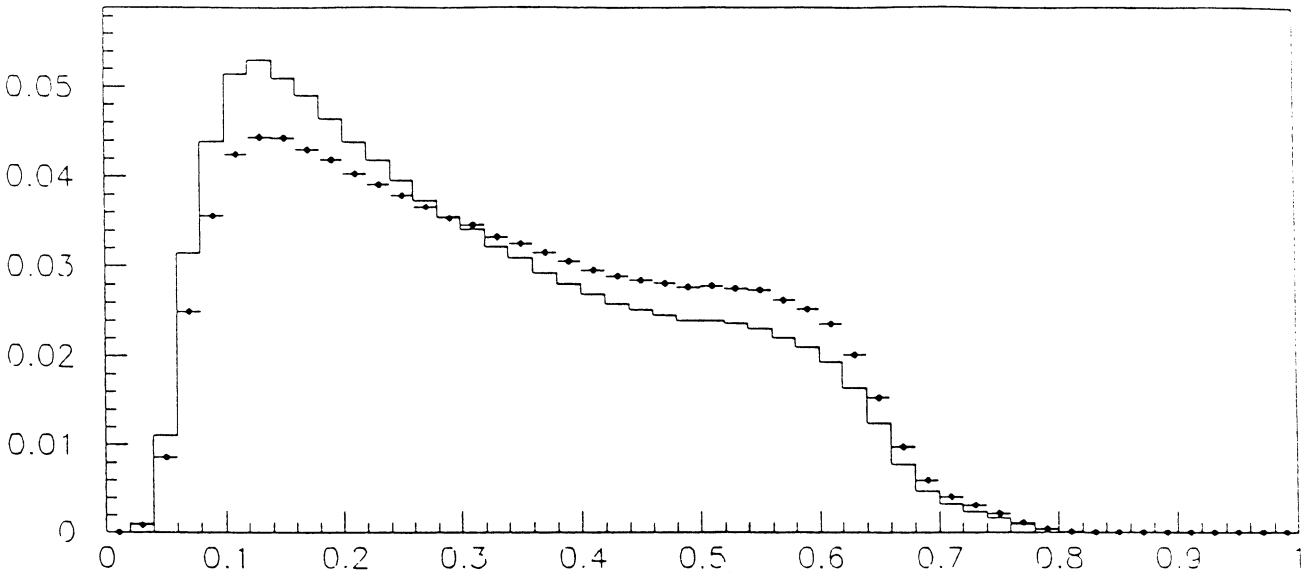


Figure 27-b

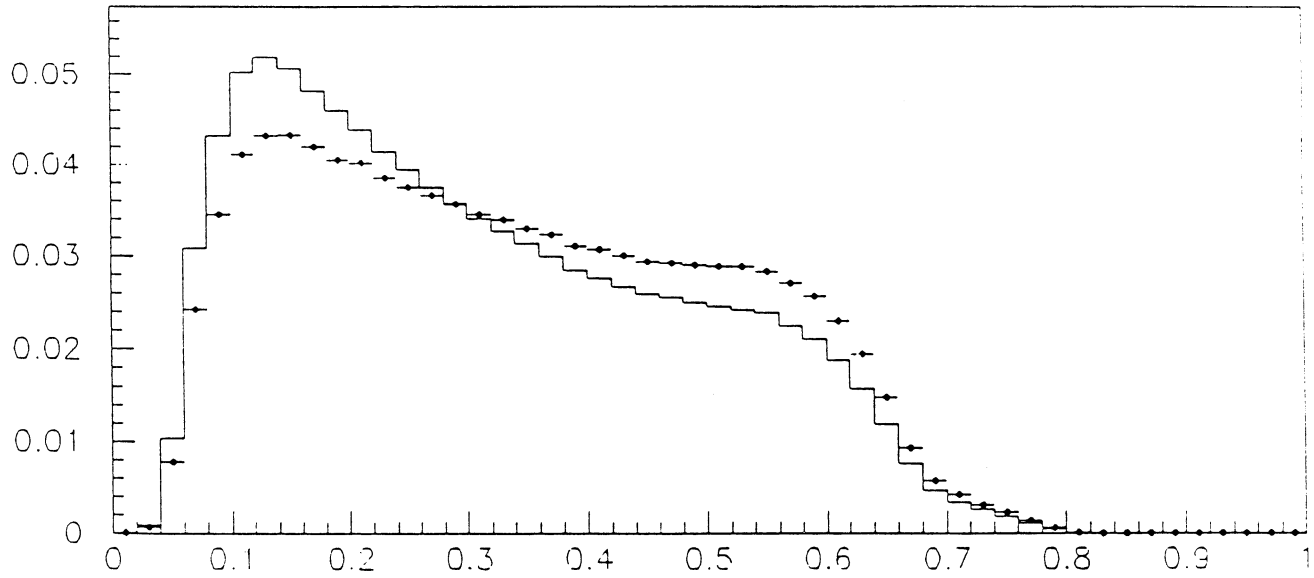


Figure 28

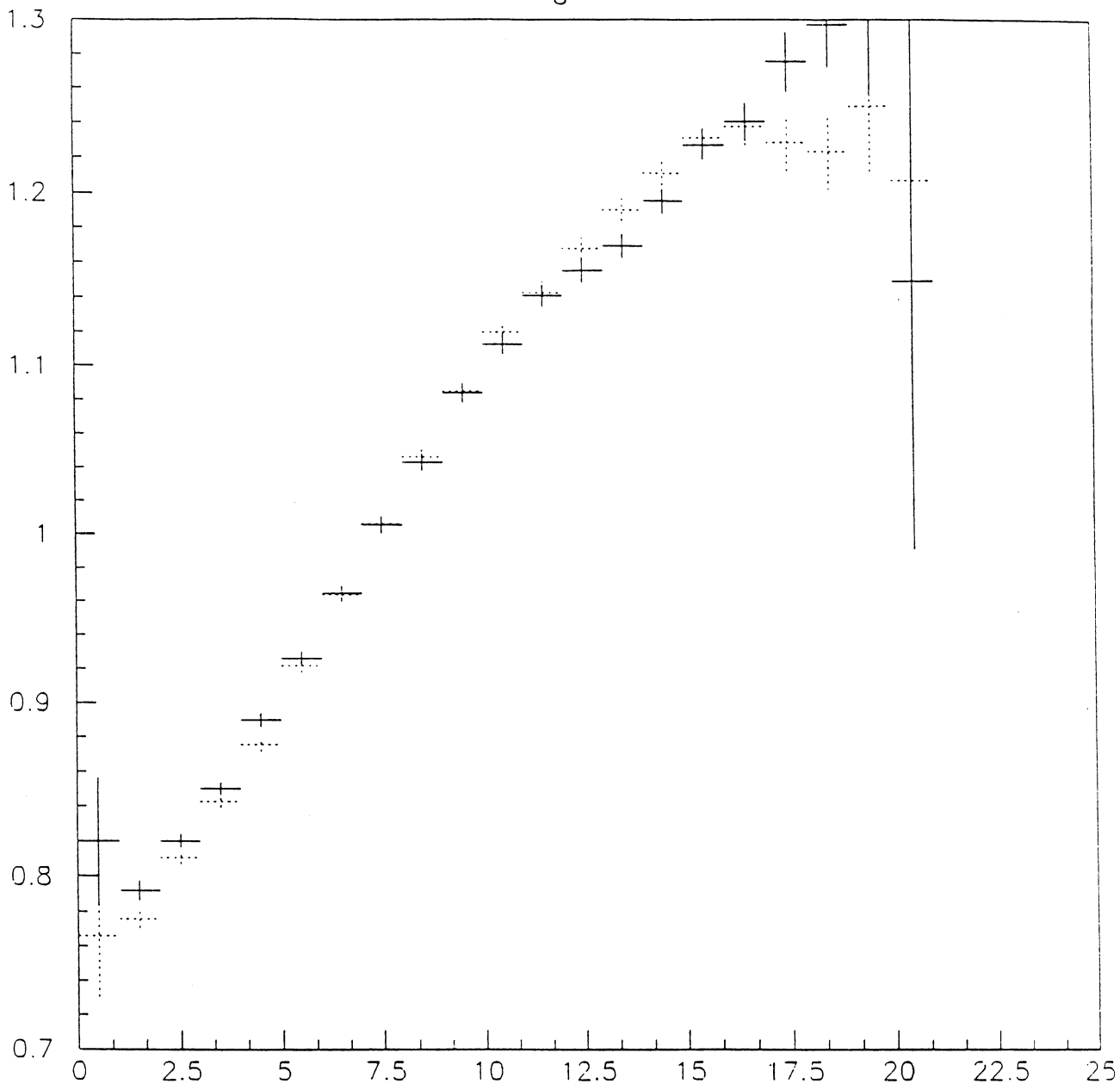


Figure 29-a

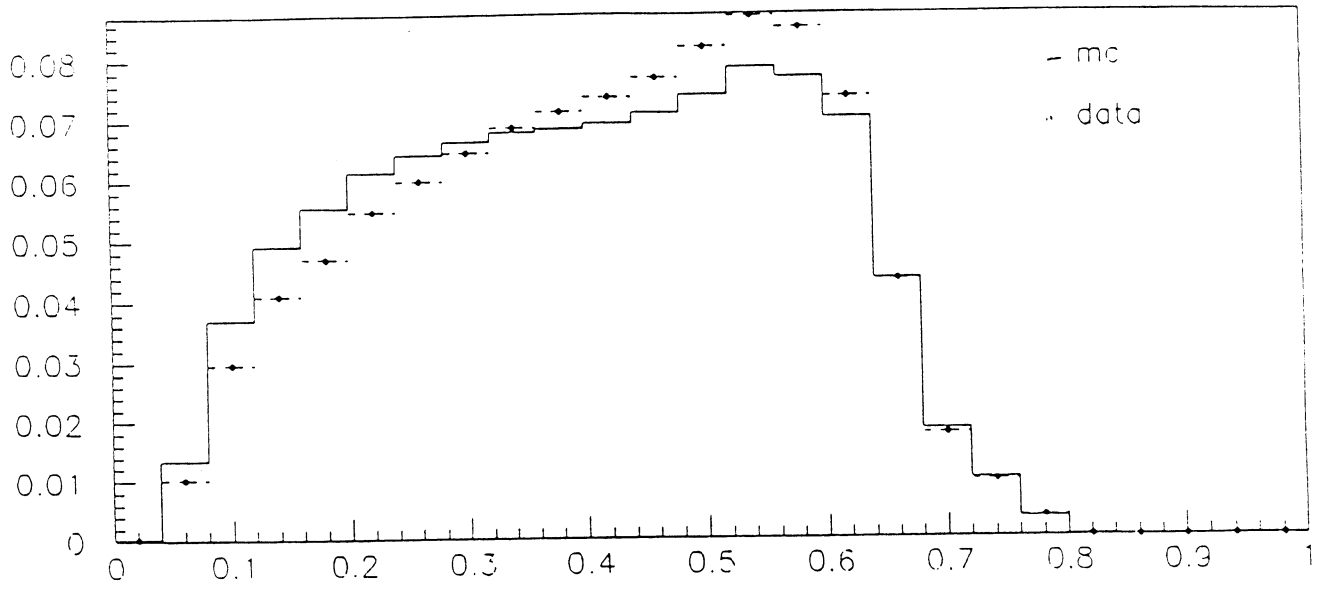


Figure 29-b

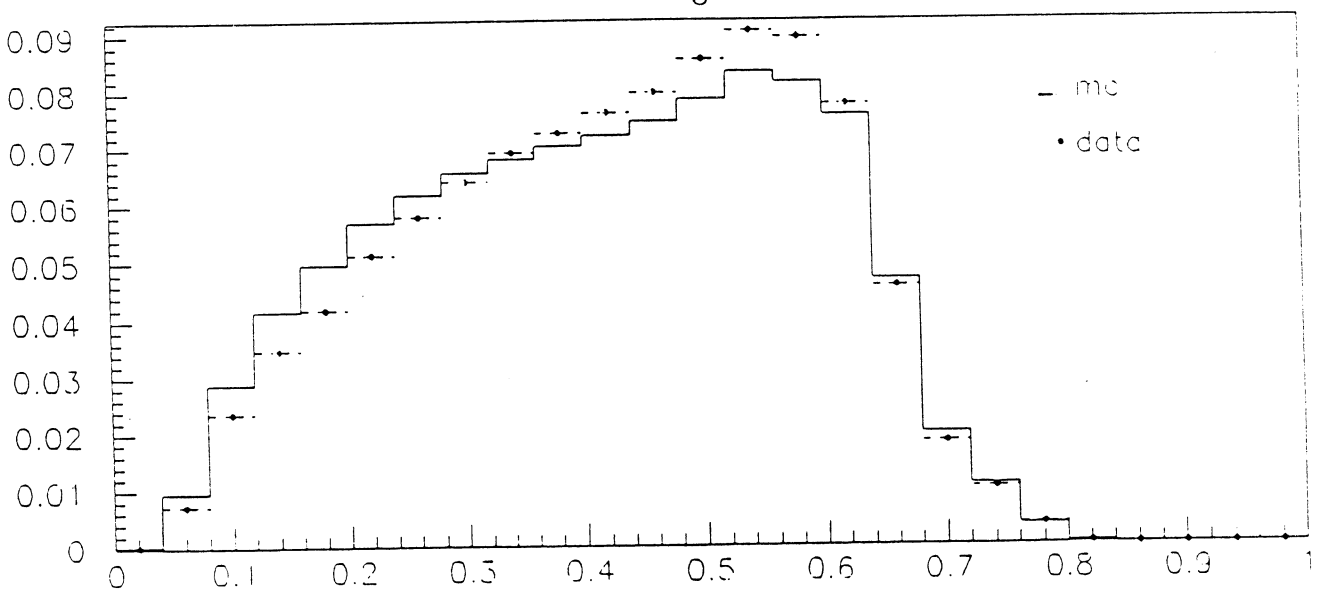


Figure 30-a

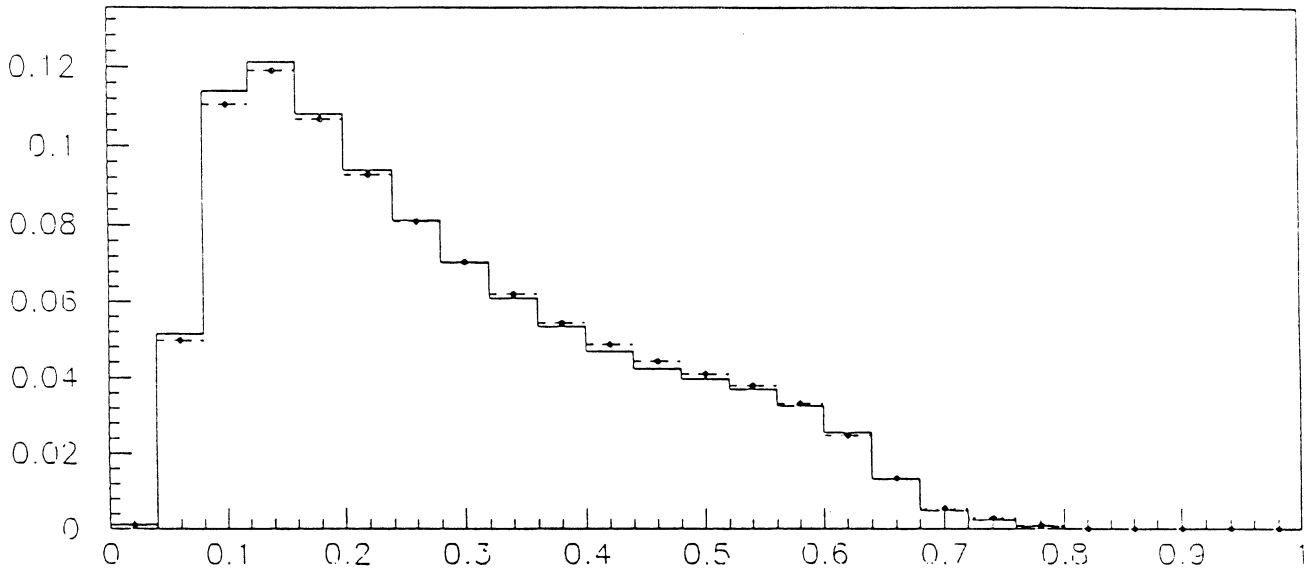


Figure 30-b

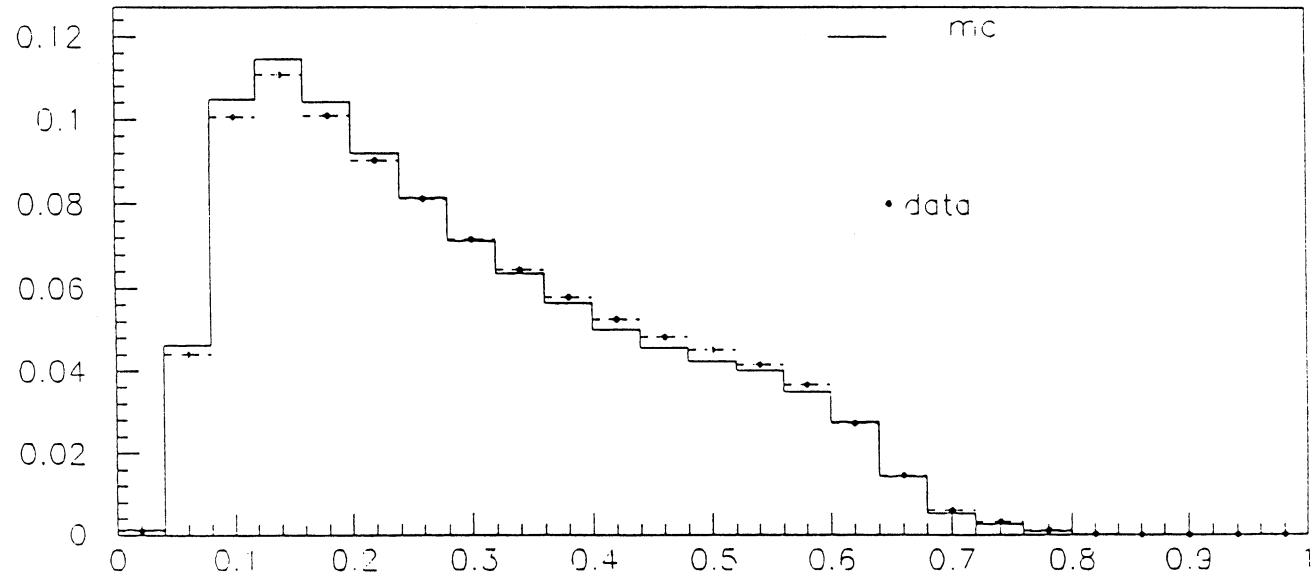


Figure 31

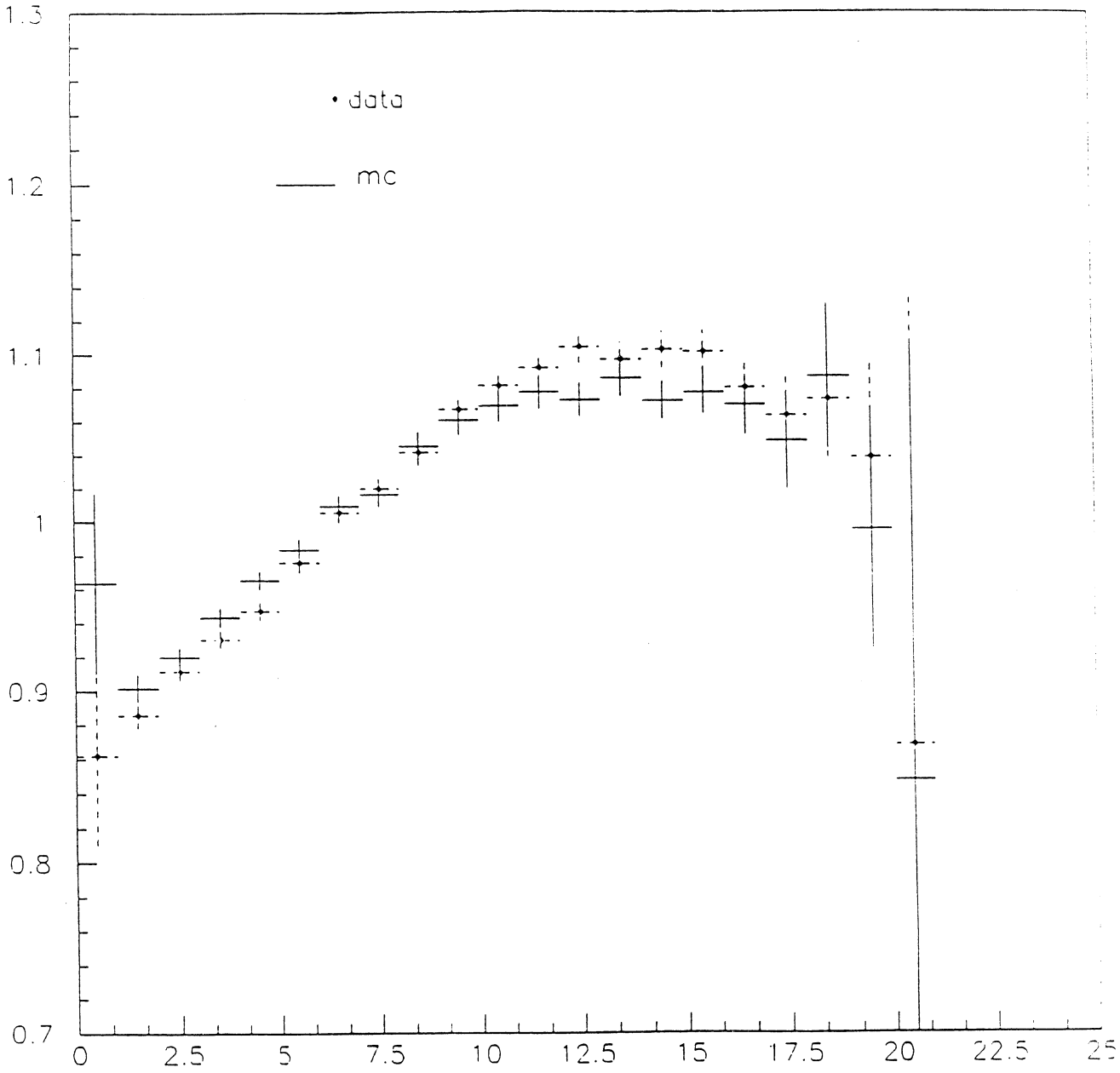


Figure 32

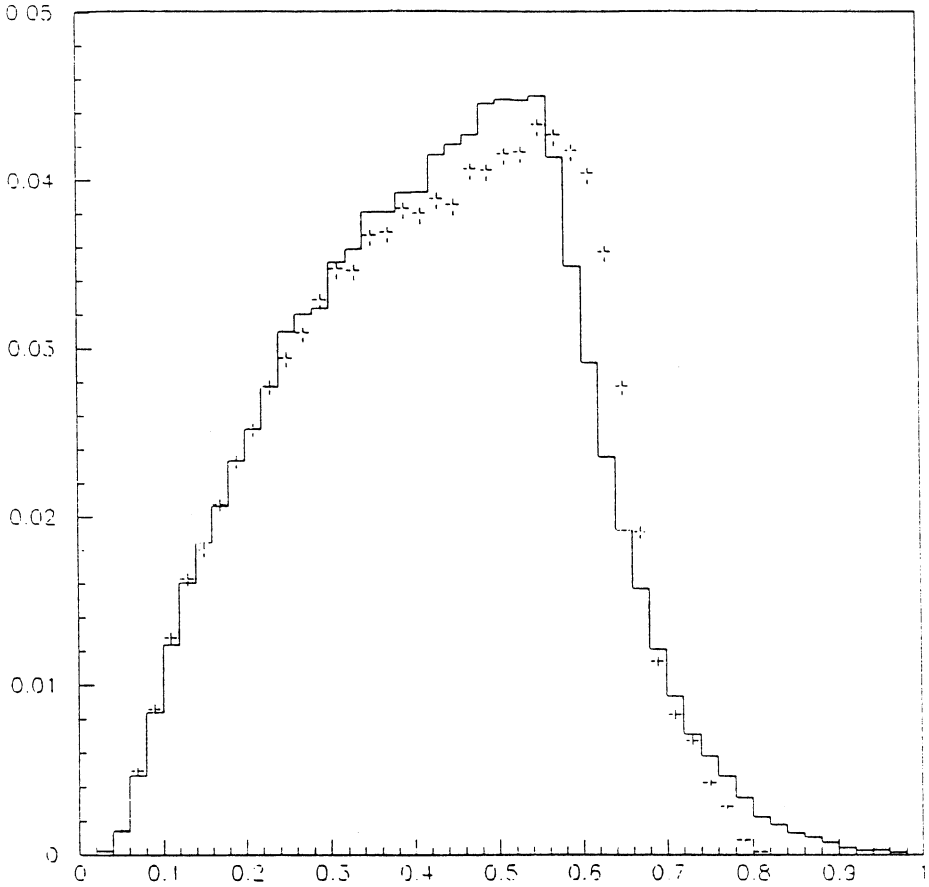
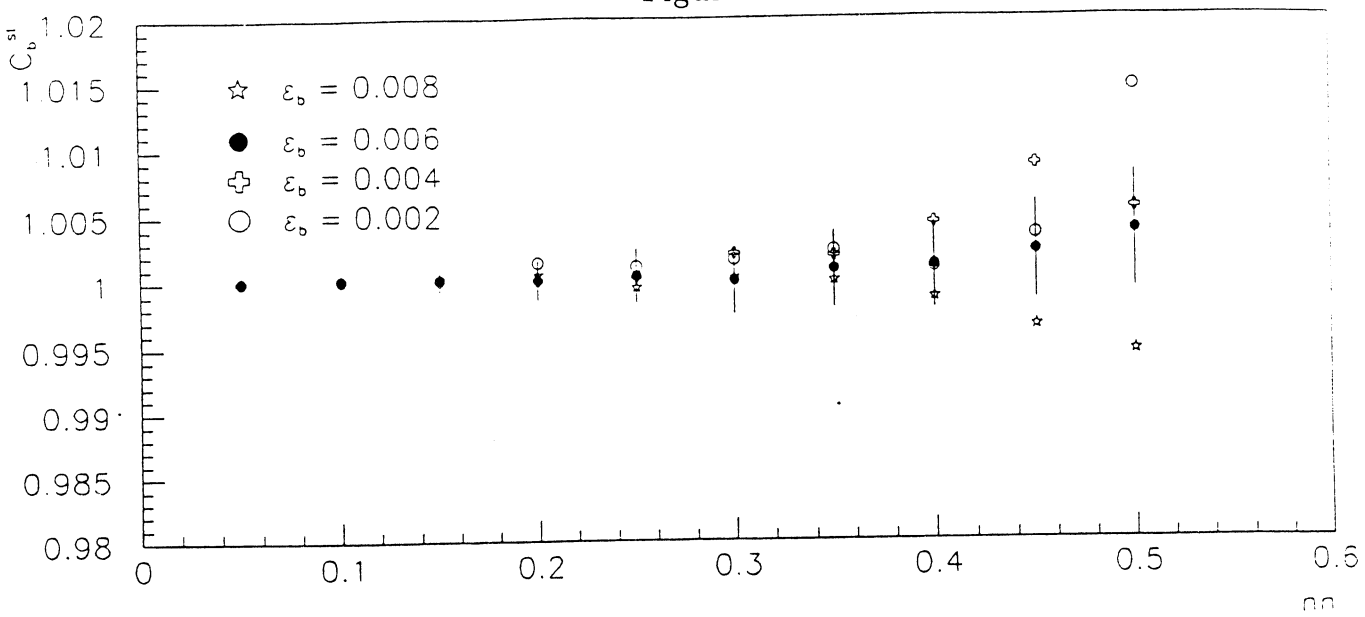


Figure 33


C_b^{st}
1.02
1.015
1.01
1.005
1
0.995
0.99
0.985
0.98
0
0.1
0.2
0.3
0.4
0.5
0.6
ε_b = 0.008
ε_b = 0.006
ε_b = 0.004
ε_b = 0.002

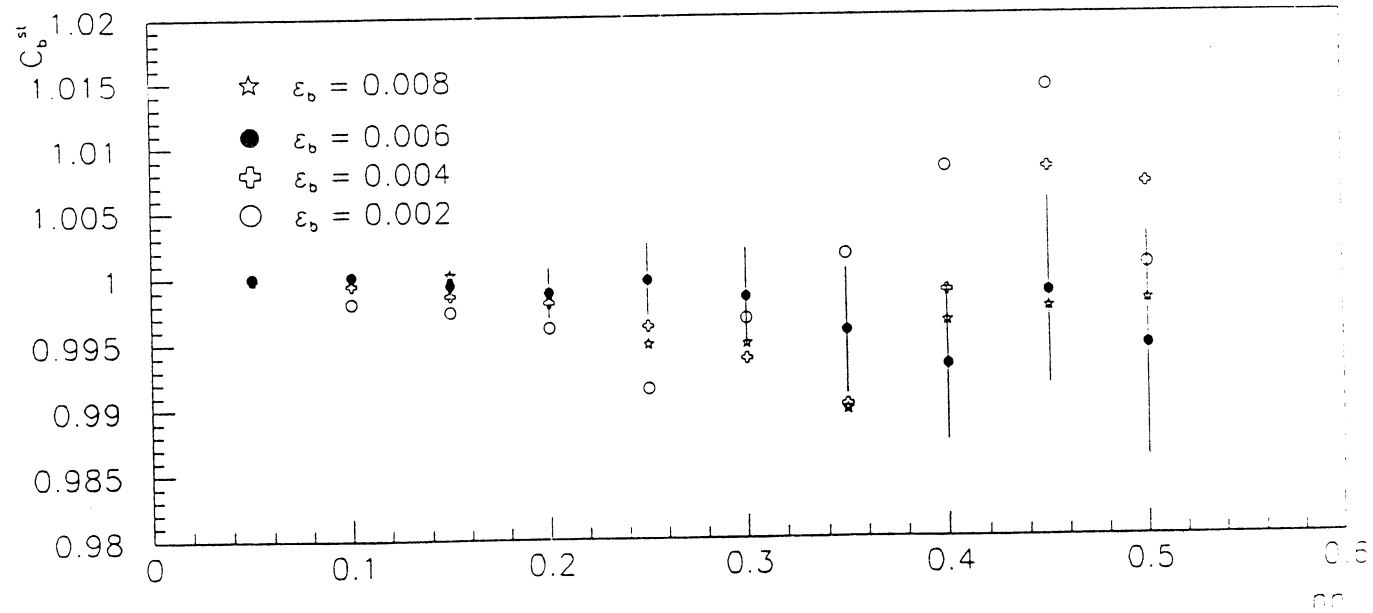

C_b^{st}
1.02
1.015
1.01
1.005
1
0.995
0.99
0.985
0.98
0
0.1
0.2
0.3
0.4
0.5
ε_b = 0.008
ε_b = 0.006
ε_b = 0.004
ε_b = 0.002

Figure 34

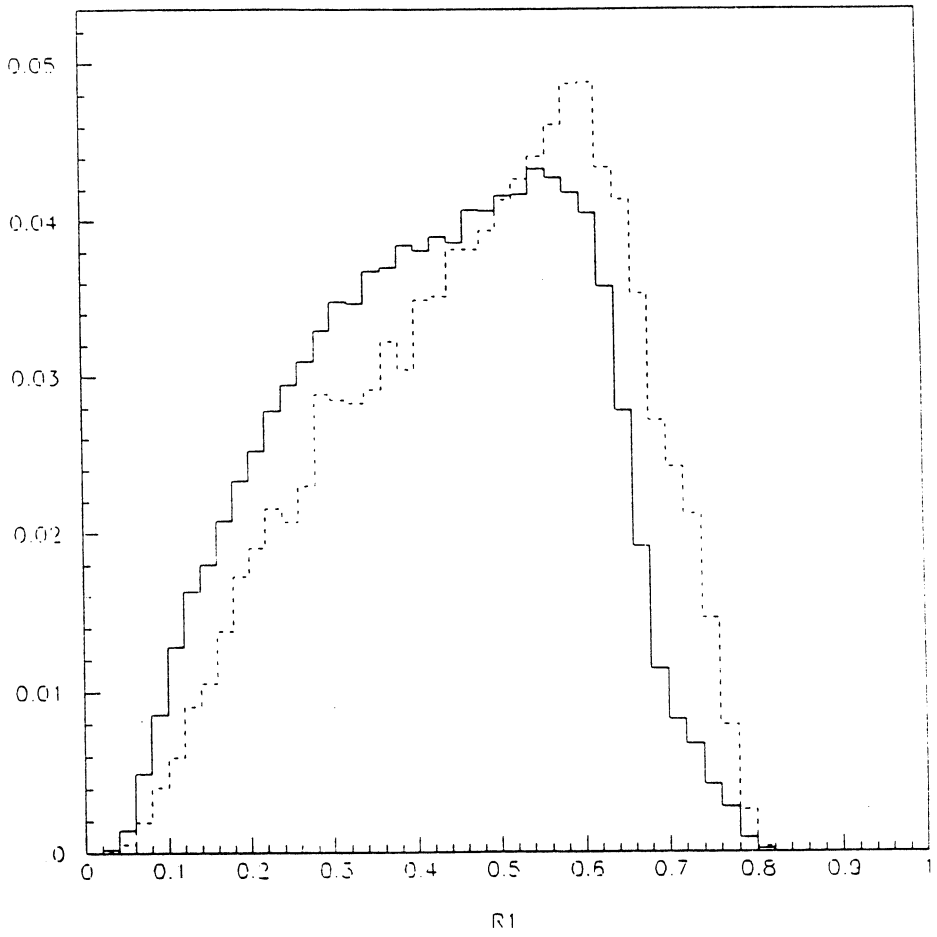


Figure 35

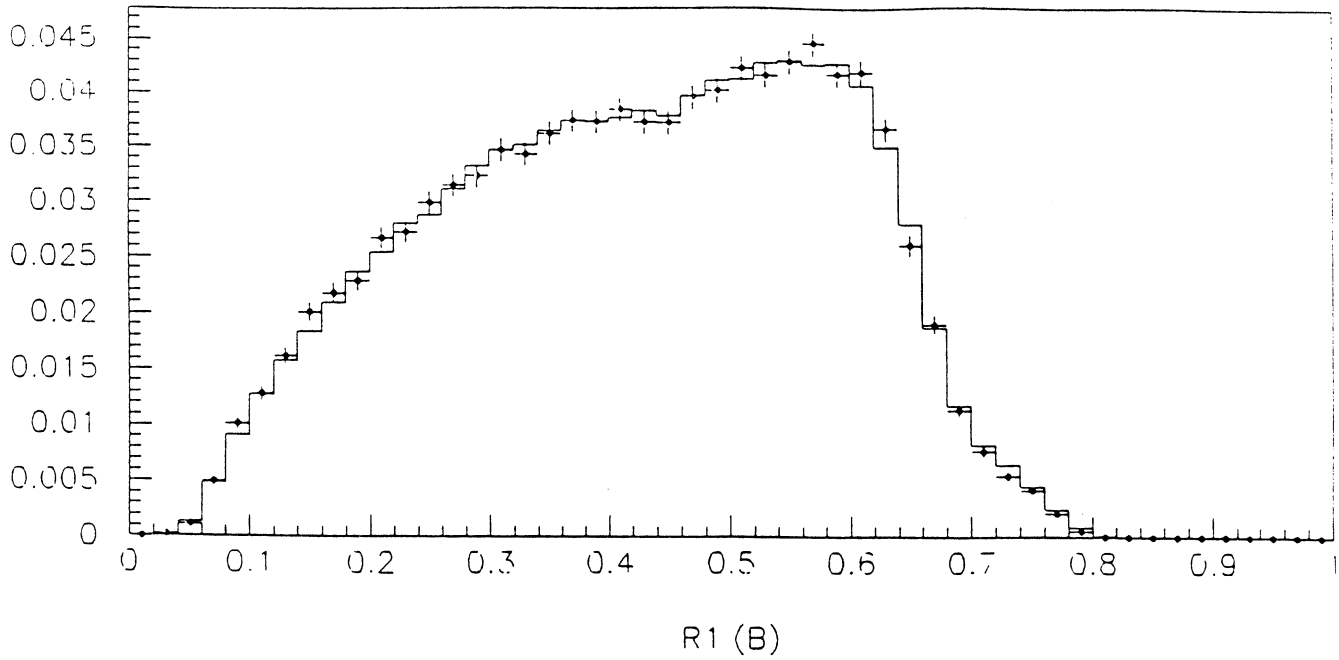


Figure 36