ALEPH 91-85
PHYSIC 91-78
B. Brandl *et al.*
18.6.1991

Event Shape Variables and Multivariate Analyses: An Approach to b Quark Production and Fragmentation

B. Brandl**), A. Falvard*), P. Henrard*), J. Jousset*), J. Proriol*)

*) – Univ. Blaise Pascal-Clermont II- IN2P3

**) – IHEP Heidelberg

1 Introduction

Measuring the $Z \rightarrow b\bar{b}$ partial width with an accuracy of the order of a few percent is one of the interesting goals to be pursued at LEP. A lot of effort was spent to analyze Heavy Flavour production through the semileptonic b and c decays. At the present time, a statistical precision of 5% on $\Gamma(Z \rightarrow b\bar{b})$ can be achieved. Detailed systematics generated by our bad knowledge of strong interaction were essentially not studied but should be in the future and will constitute the main limitation of this approach. In this note, we will describe another way to access the problem where, from the beginning, statistical error is negligible and where we are directly in touch to the systematics.

To each Z hadronic event is associated in this study a number which is the answer of a specific discriminator to which are submitted a given sample of variables computed for that event. These variables are chosen to have some characterisation capabilities of $b\bar{b}$ events. The discriminator can be a Neural Network, a Linear Discriminator, a Canonical Discriminator, etc.. This can be as usual simply a number like the p_t of a lepton. In the present analysis all discriminators use after a learning phase 5 to 10 variables. This is a typical problem of Multivariate Analysis.

Anyway, Multivariate Analysis is not commonly used for such a task and more specifically for a precision measurement. A large amount of time was spent to convince ourselves that these methods were used reliably. A first attempt is described in Ref. [1] for events containing a lepton. For the present study, several methods were also used and it was shown that they give very consistent results. A large number of variables were defined and used. Some of them are more related to the global shape of the event; others characterize

the structure of a b quark jet with respect to the structure of a lighter quark jet. This large preparation period essentially led us to maintain two independent analysis chains. One using the Neural Net method working on global variables (the numerical methods is new and the variables are classic); the other use a Canonical Discriminant method working on "jet" variables (the method is more classic but variables are generally more performant). While different, these two approaches provide important cross checks and give very similar results which will be described in the forthcoming sections.

Starting with 160,000 hadronic Z events, a statistical accuracy of $\sim$ 0.5 % and no systematics is expected on $\Gamma(Z \rightarrow b\bar{b})$ for a perfect discriminator working on all events. With the present analysis, we obtain a statistical error of 2% due for one half to MC. But clearly the problem is not statistics. Comparing the variables with their expected Monte Carlo shape indicates small discrepancies for some of these variables. It is also shown that the fine tune of Monte Carlo parameters from the QCD studies tends to compensate this effect and a deeper study will be started when sufficient Monte Carlo statistics with the best tuning will be available. In a first step we used two directions to look at the systematic effects:

- We know how to treat rigorously our uncertainty on heavy quark fragmentation with fully reconstructed MC events. This fragmentation is expected to generate the largest systematic error on the $\Gamma(Z \rightarrow b\bar{b})$. The idea is also that, if we are sensitive to b fragmentation our data can give a measurement of the ϵ_b parameter as usually done with leptonic events. So the study is self-contained and don't require any input from other sources. The other important advantage is to provide a cross-check of the result obtained with leptons. Other uncertainties are more difficult to treat very reliably but good order of magnitude can be determined.

- The Multivariate Analyses rely on the simulation of Z hadronic decays. Using this method on leptonic events provides an interesting cross check of the method. Comparing the b purity of leptonic events with different p_t cuts indicates how the analyses are reliable. Certainly this approach contains several important potentialities.

2 The Variables

The first objective of these analyses was the tagging of b quark events without using particle identification in order to provide an alternative method to the usual tagging with high p_t leptons. Two groups (Clermont-Ferrand, Heidelberg) have defined independently sets of variables which discriminate between b and $udsc$ events and are based on charged tracks only. Though the definition of the discrimination power of a single variable as well as the sets of variables have not been identical (one approach favoured global variables, the other jet variables) the results of the analyses were consistent. Finally a common set of 27 variables was fixed as the base of the multivariate analyses.

Below the variables are listed and a detailed definition is given where references are not available. The references are not complete but sufficient for the understanding of the variables. The shape of the different variables are shown in fig. 1-27 for data and Monte Carlo events. The small discrepancies between data and Monte Carlo are discussed subsequently.

2.1 Global variables

- Boosted hemisphere sphericity product ($HSPROD$). We use $\beta = .96$ [2].
- Sphericity (SPH), aplanarity (APL), thrust (THR), oblateness (OBL), sum of P_t in event plane ($PTIN$), sum of P_t out of the event plane ($PTOUT$) [3].
- Fox-Wolfram-Moments 1-5 normalized to the 0^{th} moment ($FWM1-5$) [4].
- Transverse mass (YMT), missing transverse momentum ($YPTNU$) [5].
- Momentum of leading particle normalized to the sum of the absolute momenta ($PMAXE$).

2.2 Jet variables

In the following:

PGT = sum of the absolute value of the track momenta

EGT = sum of the energies of the tracks

p_t = momentum transverse to jet axis

p_l = momentum longitudinal to jet axis

For the jet finding the Jade Algorithm $QJMMCL$ is used with $Y_{cut} = (6.0/EGT)^2$.

- Sum of the momenta of the leading particle of each jet normalized to PGT ($PMAXJ$).
- Sum of the jet masses ($SMJETS$).
- The total transverse momentum of a jet P_t^{jet} is defined as the sum of the p_t (absolute value) of each particle in the jet. As variables we use the sum over the jets and the maximum of P_t^{jet} normalized to PGT ($PTJETS$, $PTMAXJ$).
- Same as above but p_l instead of p_t ($PLJETS$, $PLMAXJ$).
- Same as above but $p_t p_l$ instead of p_t and normalized to PGT^2 ($PTLJET$, $PTLMJE$).
- Energy, mass, sum of the squared transverse momenta of particles in the jet and charge of the most energetic jet ($EJET$, $MJET$, SPT, $QJET$) [1].

Another variable with a high discrimination power is the number of charged tracks of an event (NGT). NGT was up to now only used in the Neural Net analysis but will not be used anymore in the future to avoid systematic effects due to not expected track losses. Moreover, some of our systematic checks rely on events not processed with *Galeph* but only with *Kingal* and NGT is not easy to handle on generator level.

3 The Neural Net Analysis

We use a Feed Forward Multi Layer Perceptron (MLP) with Back Propagation of the errors. Let us consider a 2 layers perceptron. In fig. 28 we have 3 neurons for input and 2 neurons for output. We introduce the inputs ξ_i^μ in the neural net. We assume the existence of the weights Wji between the input and the output. The outputs O_i^μ are computed following the feed forward formulas:

$$g(x) = (1+e^{-x})^{-1} \;\; ; \;\; h_i^\mu = \sum_k W_{ik}\xi_k^\mu \;\; ; \;\; O_i^\mu = g(h_i^\mu)$$

We know the class μ of the input, we are waiting for the S_i^μ values for the input. This value is different from O_i^μ. We use this difference to compute a correction to the weights W_{ji} by back propagation of the error. The general cost is given by the formula:

$$E = \frac{1}{2}\sum_{i,\mu}(S_i^\mu - O_i^\mu)^2$$

if we choose:

$$\Delta W_{ji} = -\eta\frac{dE}{dW - ji}$$

the E value decreases during the learning phase. Then:

$$\Delta W_{ji} = \eta\sum_\mu \delta_i^\mu g(h_i^\mu)$$

For the output layer:

$$\delta_\mu^\mu = g'(h_i^\mu)(S_i^\mu - O_i^\mu)$$

If we have introduced hidden layers:

$$\delta_i^\mu = g'(h_i^\mu)\sum_i W_{ij}\delta_i^\mu$$

where δ_i^μ is the value for the preceding layer. The actualization of the weights is generally done by the formula

$$\Delta W_{ij}(t+1) = \alpha\Delta W_{ij}(t) + \eta\sum_\mu \delta_i^\mu g(h_i^\mu)$$

The parameter α is of the order of 0.5 and the η parameter is about $[0.01 - 0.5]$.

3.1 The learning

The initialization of the weights is done by small random numbers. We present an event to the MLP, we feed forward, we back propagate the error and compute ΔW_{ij}. We reactualize the W_{ij} value when we have presented *one event of each class.*

The learning set of events is composed of 10000 events of each class. We present the learning set several times to get a convergence of the learning. The E value lakes then a limit value.

In our case we have 3 classes of Monte-Carlo events: b, c, uds. The data set for the learning is composed of 10,000 b events, 10,000 c events, 10,000 uds fully simulated events. This data set was presented randomly 150 times for the learning. The parameters of the actualization are $\alpha =$ 0.5 and $\eta =$ 0.1. The selection of variables has given 9: $HSPROD$, APL, $FWM3$, $FWM4$, $FWM5$, YMT, $MJET$, SPT and NGT. We have chosen the MLP with:

- 9 input neurons in the 1^{st} layer
- 12 hidden neurons in the 2^{nd} layer
- 9 hidden neurons in the 3^{rd} layer
- 3 output neurons in the 4^{th} layer

The output neuron 1 is affected to the b class. We expect an output 1 for the b events and 0 for the other events. The output neuron 2 is affected to the c class and the output

neuron 3 is affected to the *uds* class. When the learning is finished, we can get the average value of the output for the events of the same class as the output class and for the different ones. We get the following values (Table 1). We see in this table that the behaviour of the c and *uds* are almost the same.

3.2 Test of the Neural Net on Monte Carlo events

Using another sample of Monte Carlo events, we look at the output of the Neural Net working on these events. Fig. 29 shows the output of the Net in the b class for b events and non-b events. To get a sample of increasing purity in b events, we can do a cut on the output value of the b filter. We have defined for each cut the values:

$$\text{Purity} = \frac{\text{number of true } b \text{ events above the cut}}{\text{number of classified } b \text{ events above the cut}}$$

$$\text{Efficiency} = \frac{\text{number of true } b \text{ events above the cut}}{\text{number of original } b \text{ events}}$$

Fig. 30 gives the curve purity versus efficiency for the Neural Net analysis.

3.3 Measurement of $\Gamma(b\bar{b})$

Then the Neural Net is run on data events. The comparison of the output in the b class is compared to the Monte carlo distribution assuming the Standard Model for quark flavour fractions. A fairly good agreement is obtained except for a small discrepancy for small values of the output. The comparison of data and Monte Carlo is done with generated parameters for fragmentation, $\epsilon_b = 6 \times 10^{-3}$ and $\epsilon_c = 20 \times 10^{-3}$. This is certainly not the optimum and we looked at the variation of the χ^2 of the data to Monte Carlo fit by introducting ϵ influence by a weighting method. We determined the best ϵ_b value corresponding to the minimal χ^2 by fitting the higher part of the neural output distribution (output $>$ 0.5). This leads to:

$$\epsilon_b = .75 \, ^{+.32}_{-.22} \, \%$$

This value is in good agreement with other determination from leptons. The reason why we determine ϵ_b in this enriched b region is that our result is unsensitive to the *charm* fragmentation. For this ϵ_b value and error the $b\bar{b}$ partial width is:

$$\Gamma_b = (22.2 \pm 0.4(stat.) \pm 0.8(b - frag.)) \, \%$$

In the statistical error one half is due to the data statistics the remaining part being from the limited MC statistics. The charm fragmentation also influences the partial $b\bar{b}$ width. A conservative error due to charm fragmentation uncertainty is $\pm.5\%$. Then we see that the systematic error is strongly dominant in this analysis. Certainly the error due to inacurrate heavy-quark fragmentation is the dominant part of this error. Anyway the Monte Carlo simulation of the hadronic Z events by Parton-Shower models is not perfect even if it seems to be the best present approach to describe the reality. That is the reason which pushed us to compare our measurement of the $b\bar{b}$ fraction in a leptonic sample to that of the standard lepton analysis with different lepton p and p_t cuts. This will be presented in a forthcoming section and shows the good reliability of the present determination.

4 The Canonical Discriminant Analysis

Canonical discriminant analysis (CD) is related to principal component analysis and canonical correlation and not to the usual linear discriminant analysis first introduced by Fisher in 1936 [6]. CD has to the best of our knowledge not yet been used in High Energy Physics and was discussed for the first time in [2].

4.1 The selection of the input variables

Given the variables defined in Section 2 Stepwise Selection (StS) [7] is used to define the variables to be used in the discrimination model. The discriminatory power of a single variable or a model is defined by a F-statistic based on Wilks' Λ [8]. StS begins with no variables in the model. At each step the model is examined. If the variable in the model that contributes least to the discriminatory power of the model fails to meet the criterion to stay (defined by a certain significance level of the F-test), then that variable is removed. Otherwise, the variable not in the model that contributes most to the discriminatory power of the model is entered. When all variables in the model meet the criterion to stay and none of the other variables meets the criterion to enter, the StS process stops.

In this analysis a sample of 10,000 *uds*, 10,000 *c* and 10,000 *b* fully simulated Monte Carlo events was used to define the input variables for the discrimination model. StS selected the variables: $HSPROD$, $FWM1$, $FWM3$, $FWM4$, $SMJETS$, $PTJETS$, $PTMAXJ$, $PLJETS$, $EJET$ and SPT. The number of ten variables in the model was forced by choosing the appropiate significance level. A model with ten variables is easy to handle and higher dimensional models do not improve the discriminatory power significantly. The F-values for all variables are given in (Table 2).

4.2 The discriminator

CD derives a linear combination of the variables in the discrimination model that has the highest possible multiple correlation with the groups (*uds*, *c*, *b*). This maximal multiple correlation is called the ***first canonical correlation***. The coefficients of the linear combination are the ***canonical coefficients*** or ***canonical weights***. The variable defined by the linear combination is the ***discriminator***.

The first canonical correlation is at least as large as the multiple correlation between the groups and any of the original variables. If the original variables have high within-group correlation, the first canonical correlation can be large even if all the multiple correlations are small. In other words, the ***discriminator*** can show substantial differences among the classes, even if none of the original variables do. Canonical discriminant analysis is equivalent to canonical correlation analysis between the quantitative variables and a set of dummy variables coded from the class variable.

Again a sample of 10,000 *uds*, 10,000 *c* and 10,000 *b* fully simulated Monte Carlo events was used to find the ***discriminator*** that was afterwards standardized to range from [0-1].

Approximately 150,000 fully simulated Monte Carlo events have been used to test the ***discriminator***. Fig. 31 shows the shape of the ***discriminator*** for data and Monte Carlo events. A significant statistical seperation between *b* quark and *udsc* quark events

is observed while the difference between c quark and uds quark events is marginal. Fig. 32 shows purity versus efficiency for CD defined like in section 3.2.

4.3 Measurement of $\Gamma(b\bar{b})$

Fitting the *discriminator* shapes for b and $udsc$ events obtained from a large sample of Monte Carlo events to the 1990 data sample a measurement of the $Z \rightarrow b\bar{b}$ partial width with a small statistical error is possible. To study the influence of heavy quark fragmentation a weighting method similar to the one applied in the D^* analysis is used. Taking in account current results from the D^* and prompt lepton analyses the following values for ϵ_b and ϵ_c are reasonable:

$$\epsilon_b = .6 \, ^{+.2}_{-.2} \, \%$$

$$\epsilon_c = 4 \, ^{+2}_{-1} \, \%$$

For this values and errors the $b\bar{b}$ partial width is:

$$\Gamma_b = (20.9 \pm 0.34(stat.) \, ^{+1.0}_{-.92}(c, bfrag.)) \, \%$$

The statistical error includes the uncertainties due to the limited Monte Carlo statistics. In future the influence of ϵ_b and ϵ_c will be studied in the same way as described in section 3.3.

Though the error from inaccurate heavy-quark-fragmentation is expected to be the dominant systematic error more effort must be spend to understand the influence of light quark fragmentation and strong interaction. The small discrepancies between data and Monte Carlo present in the event shape variables seem to have their origin in a not well tuned Monte Carlo generator. Unfortunately at present only a small number of Monte Carlo events with the best tuning is available. Fig. 33 compares the $THRUST$ distribution of c events for the old tuning , up to now used for this analysis, and a new, but not yet perfect, tuning.

5 Analysis of a leptonic sample with the Multivariate Analyses

The objective of this work is to measure the purity of the leptonic sample extracted from hadronic Z decays and to compare it with the expected value from standard lepton analysis. This has to be considered as a check of the Multivariate Analyses.

The study was performed using the full $q\bar{q}$ statistics (158,566 events). Leptons were selected in the common Heavy Flavour way. The events were processed through the Neural Network with the same learning as that used for the full $q\bar{q}$ sample. The fraction of $b\bar{b}$ events was fitted and results are summarized in Table 3. The Neural Net result is given for the same values of fragmentation parameters than in the Monte-Carlo generation ($\epsilon_b = 6 \times 10^{-3}$; $\epsilon_c = 40 \times 10^{-3}$).

The same check has been done using the Canonical Discriminator. They are summarized in Table 4. This study was extended in a non trivial direction. This consists is analyzing with the Neural Net a sample of events selected by a p_t cut on the lepton. Tables 5 and 6 summarize the results for two different cuts on the lepton momentum 3

Gev/c and 5 *GeV/c*, respectively. The comparison of Monte Carlo and data purities does not indicate any clear deviation. It seems that the simulation gives a good description of real events. Then at the level of the present statistics and systematics, the lepton sample cannot help to improve the Multivariate Analysis approaches.

6 Conclusion

We have studied b quark production and fragmentation using two independent Multivariate Analysis methods based on global and jet variables.

From the Neural Net analysis we obtain:

$$\epsilon_b = .75 \, ^{+.32}_{-.22} \, \%$$

$$\Gamma_b = (22.2 \pm 0.4(stat.) \pm 0.8(b - frag.) \pm 0.5(c - frag.)) \, \%$$

and from the Canonical Discriminant analysis:

$$\Gamma_b = (20.9 \pm 0.34(stat.) \, ^{+1.0}_{-.92}(c-, b - frag.)) \, \%.$$

These results are preliminary because the estimation of the systematic errors is not yet complete.

The errors from inaccurate heavy-quark-fragmentation quoted above are expected to be the dominant systematic errors but the influence of light quark fragmentation and strong interaction has to be understood. Moreover, because this analysis is sensitive to the tuning of the used Monte Carlo generator, it has to proven that no artificial bias is introduced by assuming it the tuning algorithm the Standard Model value 21.7% for Γ_b.

References

[1] J. Proriol *et al.* - Aleph Note Physic 90-058 - Event recognition: Multivariate analysis method to tag b quark events in ALEPH.

J. Jousset *et al.* - Aleph Note Physic 90-066 - Comparaison of three pattern recognition methods for the tagging of b quark events in ALEPH: Linear discriminant analysis, Classification tree, Neural Networks.

[2] A. Putzer *et al.* - Aleph Note 90-161

[3] ALEPH Coll. - D. Decamp *et al.* - Phys. Lett. **B255** (1991) 623

ALEPH Coll. - D. Decamp *et al.* - Phys. Lett. **B234** (1990) 209

[4] G. C. Fox, S. Wolfram - Phys. Rev. Lett. **41** (1978) 1581

[5] P. Henrard PC/CT RI 88-08

[6] R. A. Fisher - Annals of Eugenics **7** (1936) 179

[7] W. R. Klecka - *Discriminant Analysis* - Sage University Paper Series on Quantitative Application in the Social Sciences, Series No. 07-019, Berverly Hills: Sage Publications

[8] C. R. Rao - *Linear Statistical Interference and Its Applications* - (1965) New York: Wiley

Figure captions

Fig. 1-27 Shape of the various global and jet variables for data and Monte Carlo events

Fig. 28 Description of an elementary Neural Net

Fig. 29 Output of the Neural Net for b and $non-b$ events

Fig. 30 Purity versus efficiency for the Neural Net analysis

Fig. 31 a) Shape of the *discriminator* for data and Monte Carlo events b) Shape of the *discriminator* for b and $udsc$ Monte Carlo events

Fig. 32 Purity versus efficiency for the Canonical Discriminant analysis

Fig. 33 $THRUST$ distribution of c events for two different Monte Carlo tunings

	b class	c class	uds class
output b events	0.449		
output non b	0.263		
output c events		0.358	
output non c		0.318	
output uds events			0.381
output non uds			0.323

Table 1: Neural net output for events in the various classes

variable	F-value
$HSPROD$	1772.7
SPH	25.0
APL	125.4
$THRUST$	1.0
OBL	8.8
$PTIN$	20.9
$PTOUT$	287.2
$FWM1$	222.5
$FWM2$	4.2
$FWM3$	218.6
$FWM4$	57.6
$FWM5$	306.3
$PMAXE$	536.1
YMT	287.7
$YPTNU$	2.6
$PMAXJ$	738.8
$SMJETS$	434.9
$PTJETS$	586.9
$PTMAXJ$	221.1
$PLJETS$	48.32
$PLMAXJ$	90.3
$PTLJET$	256.5
$PTLMJE$	116.7
$EJET$	43.5
$MJET$	335.1
SPT	783.3
$QJET$	12.8

Table 2: F-values for the various variables

Lepton Nb.	M.C. Predic.	Neural Net	$\Gamma(b\bar{b})$
> 0	57.8±0.7	58.3±1.2	22.7±0.5
1	55.0±0.7	55.2±1.3	22.5±0.5
2	81.7±2.9	85.2±2.8	22.7±0.9

Table 3: Leptonic events through the neural Net

Lepton Nb.	N_{light}/N_b (M.C.)	Can. Dis. Fit
1	.82	0.81±0.02
2	.25	0.22±0.03

Table 4: Leptonic events through the Canonical Discriminator

p_t cut	M.C.	Neural Net Fit
0.	57.8±0.6	58.4±1.2
0.3	64.8±0.7	64.8±1.4
0.6	79.4±1.0	81.2±2.0
0.9	89.5±1.5	92.0±2.4
1.2	94.2±2.1	95.7±2.6
1.5	95.8±3.2	93.6±2.7

Table 5: Analysis of the Leptonic sample ($p > 3\ GeV/c$)

p_t cut	M.C.	Neural Net Fit
0.	61.8±0.8	63.0±1.4
0.3	67.7±0.9	67.7±1.6
0.6	80.8±1.2	81.6±2.1
0.9	90.7±1.7	94.0±2.5
1.2	95.0±2.3	97.6±2.5
1.5	96.4±3.3	95.4±2.6

Table 6: Analysis of the Leptonic sample ($p > 5\ GeV/c$)

Figure 1

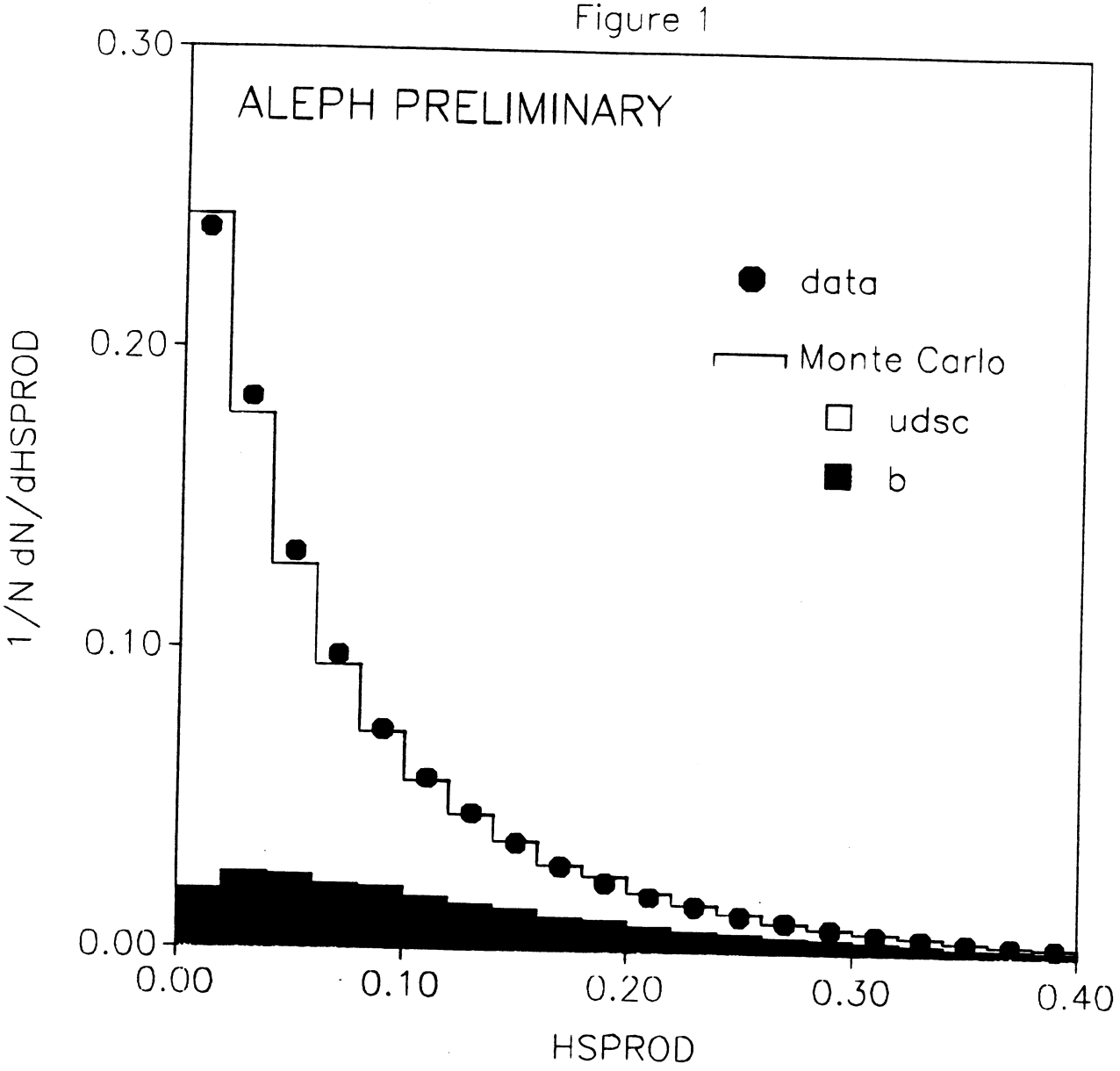

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
1/N dN/dHSPROD
0.30
0.20
0.10
0.00
0.00
0.10
0.20
0.30
0.40
HSPROD

Figure 2

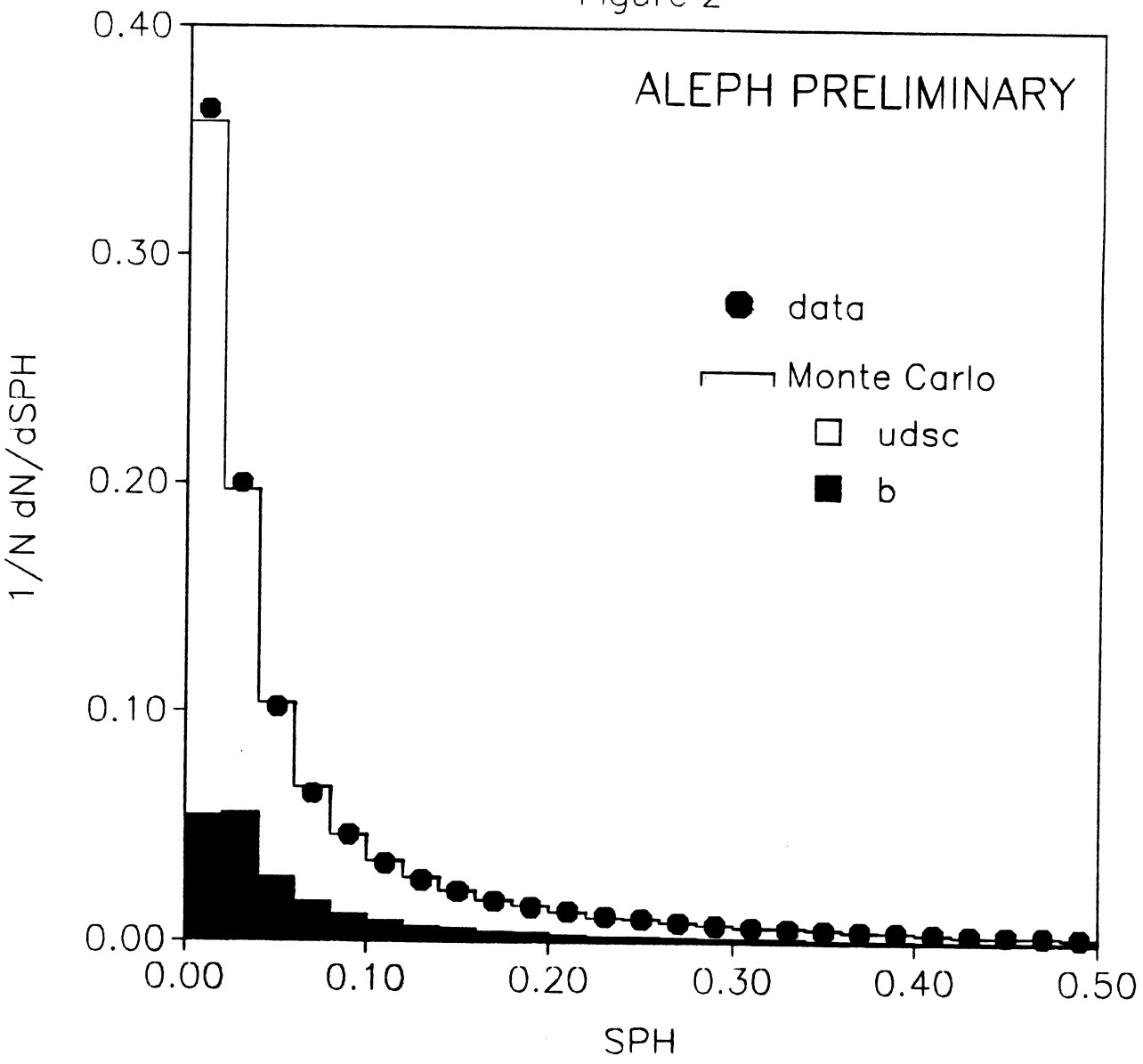

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
1/N dN/dSPH
SPH
0.00
0.10
0.20
0.30
0.40
0.50

Figure 3

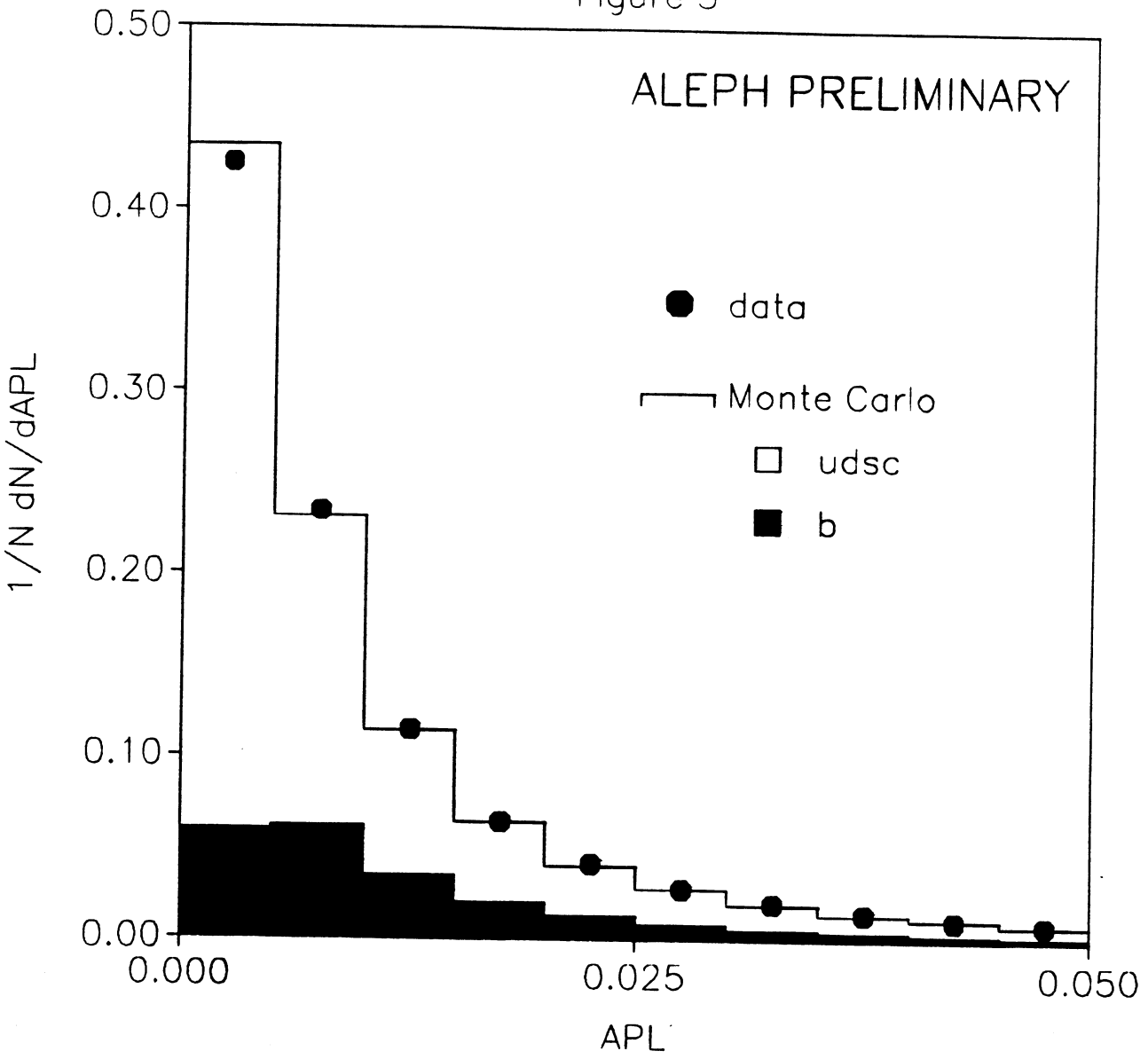

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.50
0.40
0.30
0.20
0.10
0.00
1/N dN/dAPL
0.000
0.025
0.050
APL

Figure 4

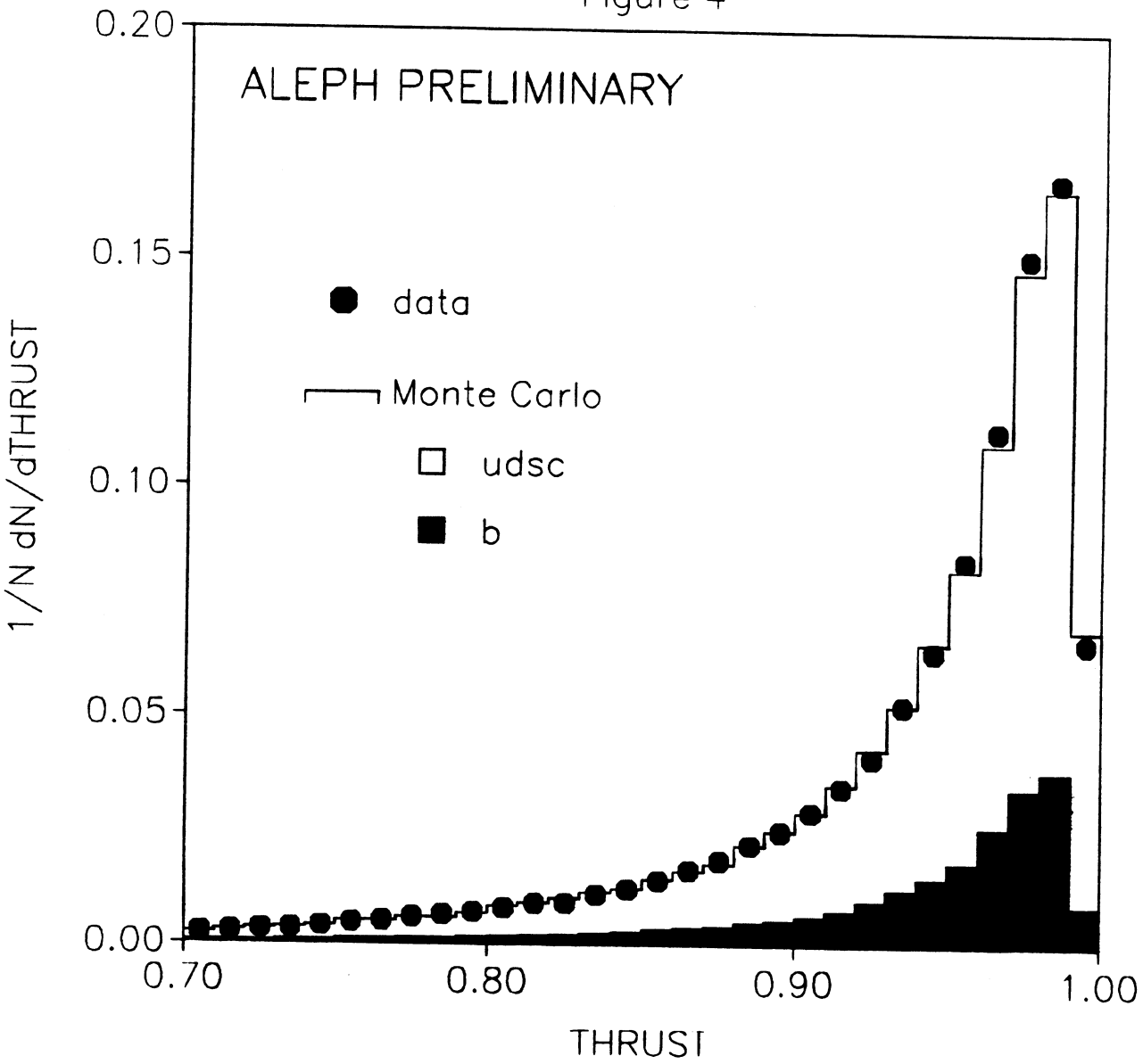

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
1/N dN/dTHRUST
THRUST
0.00
0.05
0.10
0.15
0.20
0.70
0.80
0.90
1.00

Figure 5

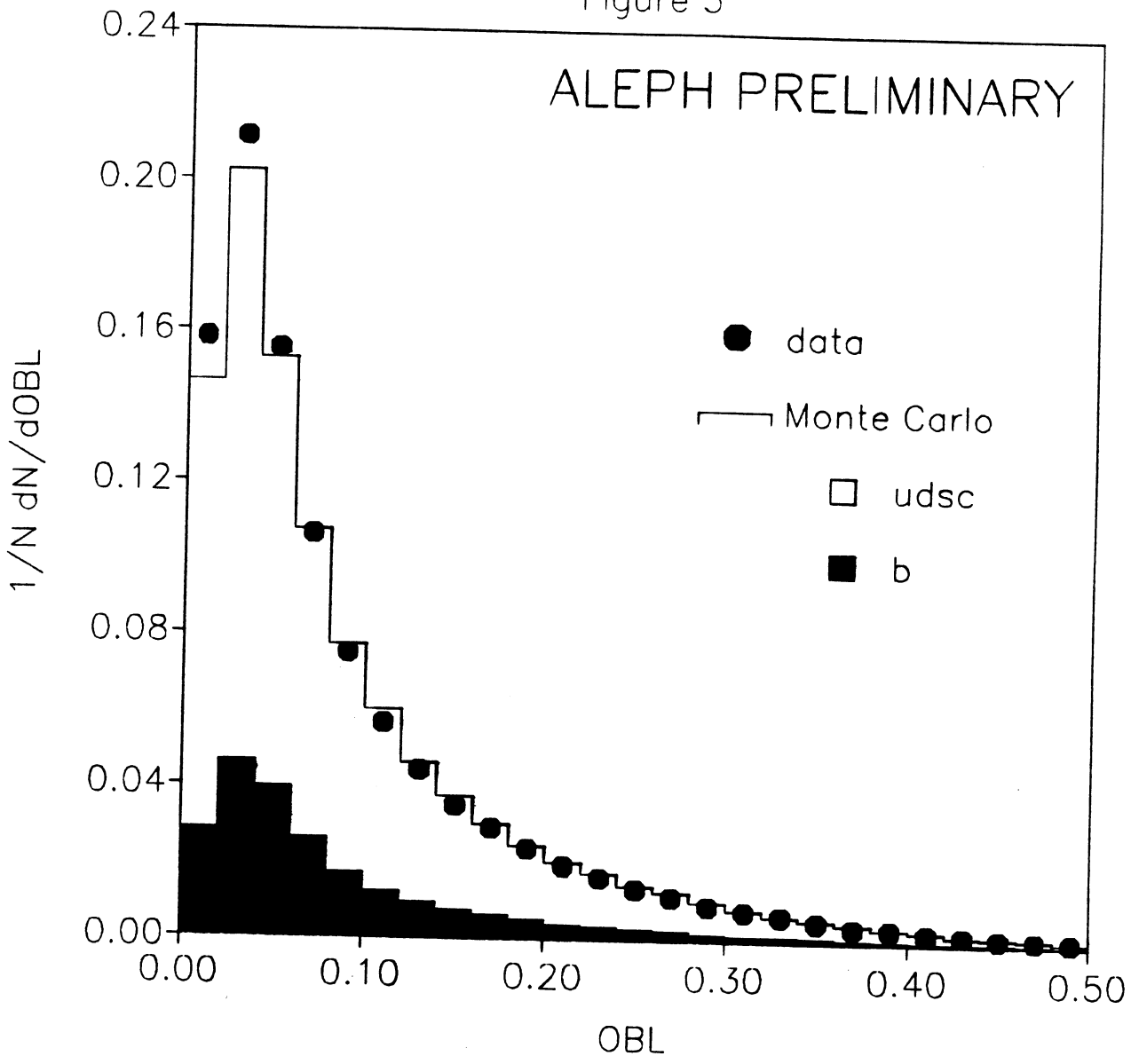

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
1/N dN/dOBL
OBL
0.24
0.20
0.16
0.12
0.08
0.04
0.00
0.00
0.10
0.20
0.30
0.40
0.50

Figure 6

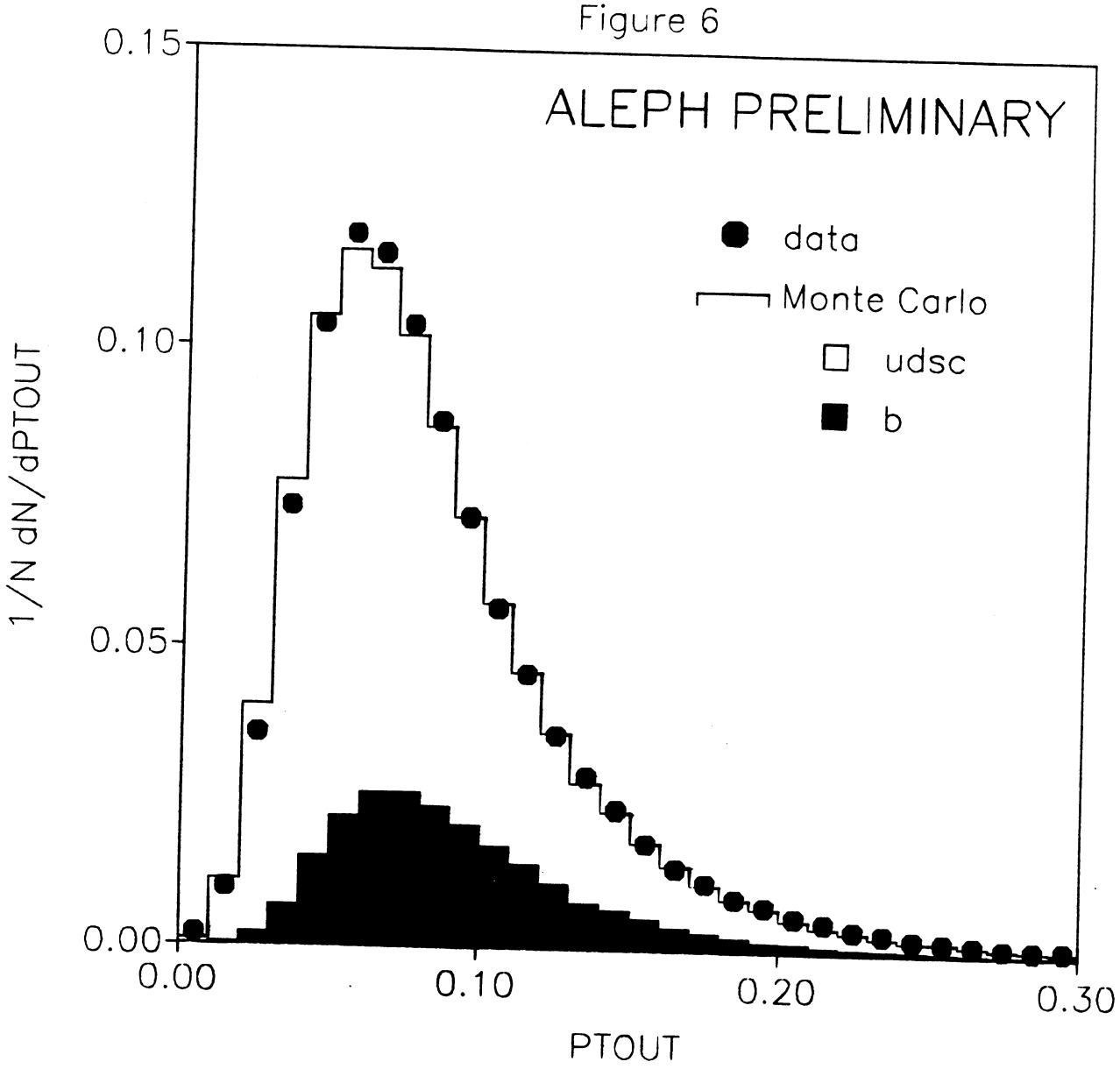

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.15
0.10
0.05
0.00
1/N dN/dPTOUT
0.00
0.10
0.20
0.30
PTOUT

Figure 7

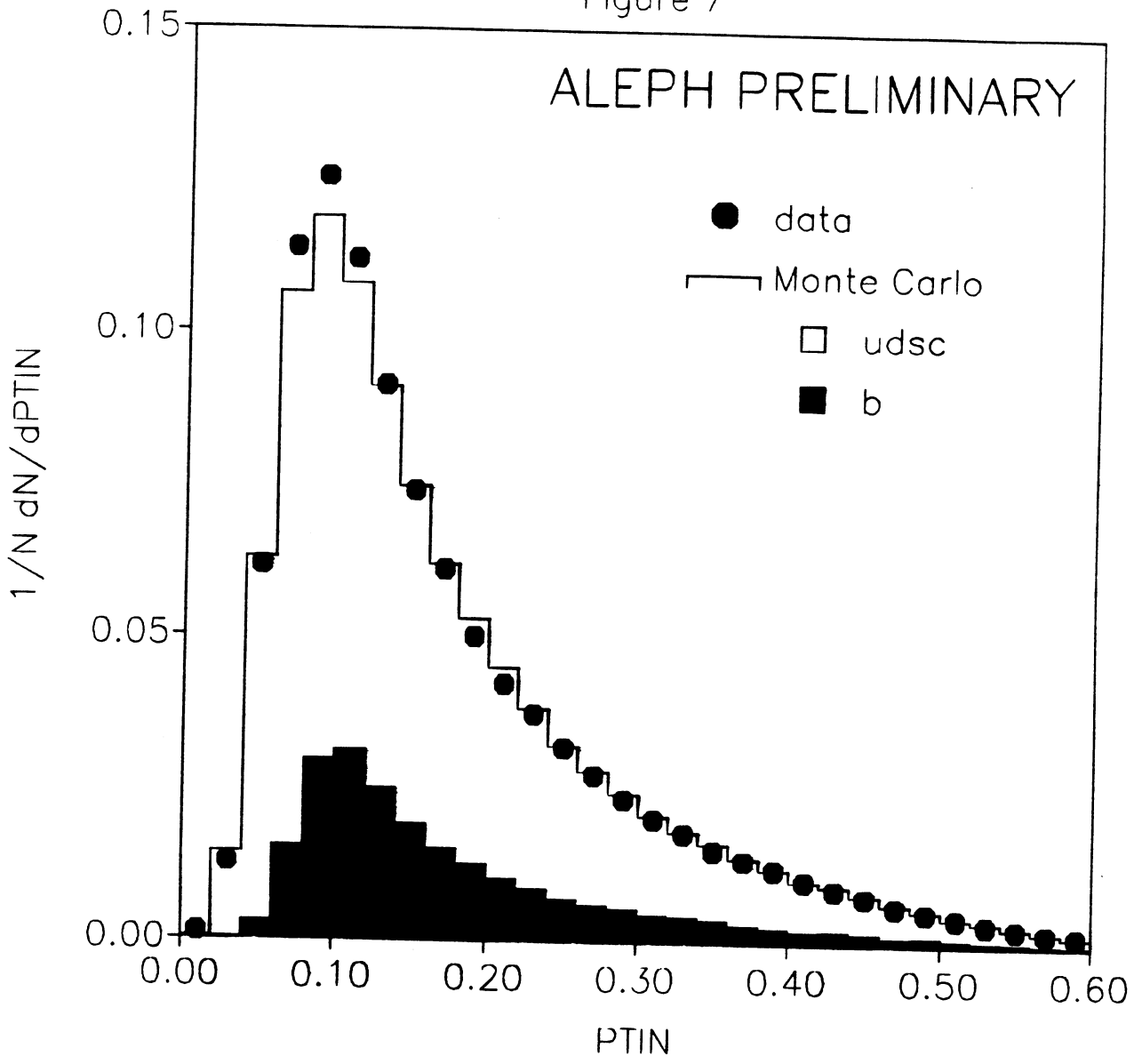

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.15
0.10
0.05
0.00
1/N dN/dPTIN
0.00
0.10
0.20
0.30
0.40
0.50
0.60
PTIN

Figure 8

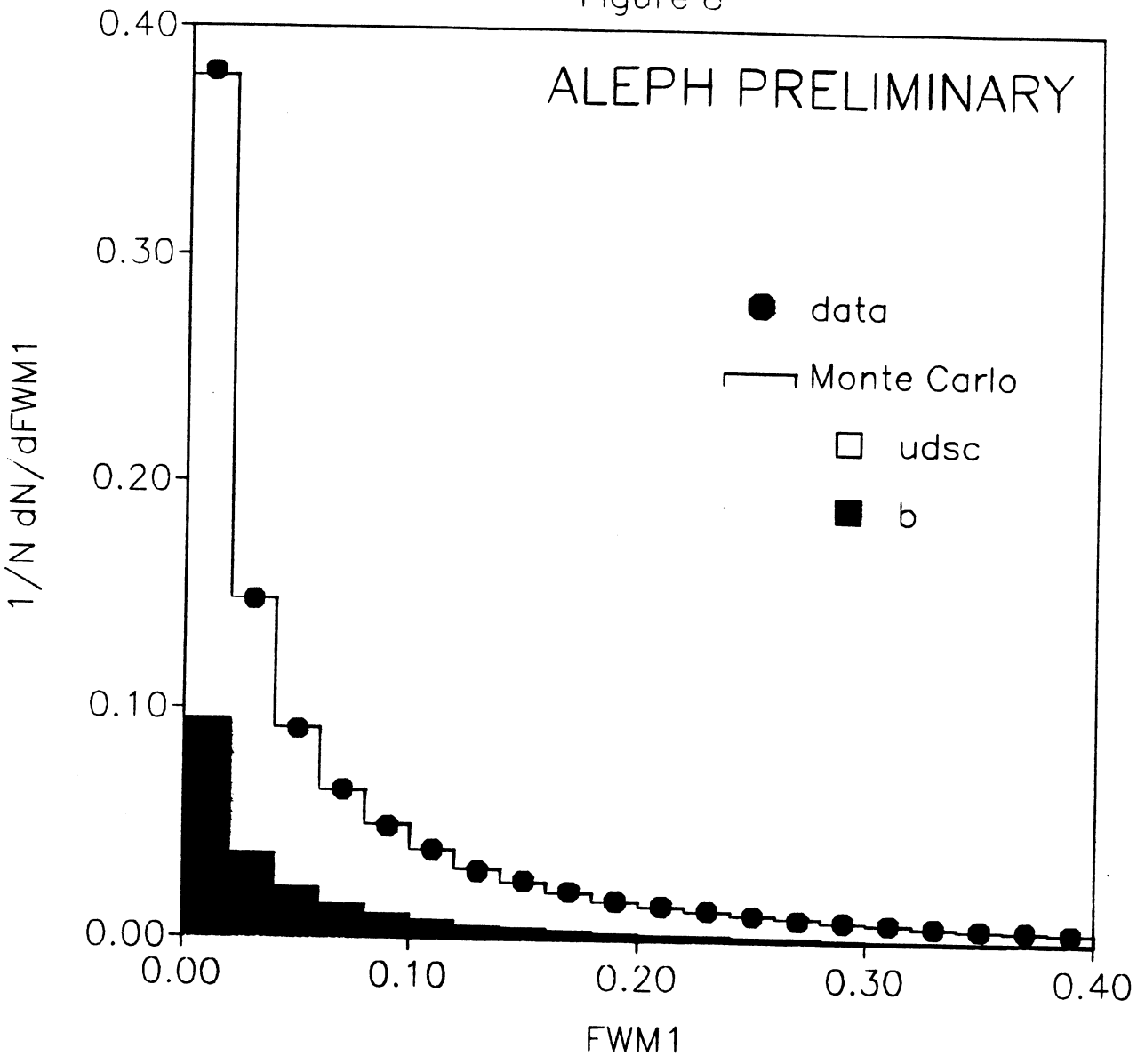

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
1/N dN/dFWM1
FWM1
0.00
0.10
0.20
0.30
0.40

Figure 9

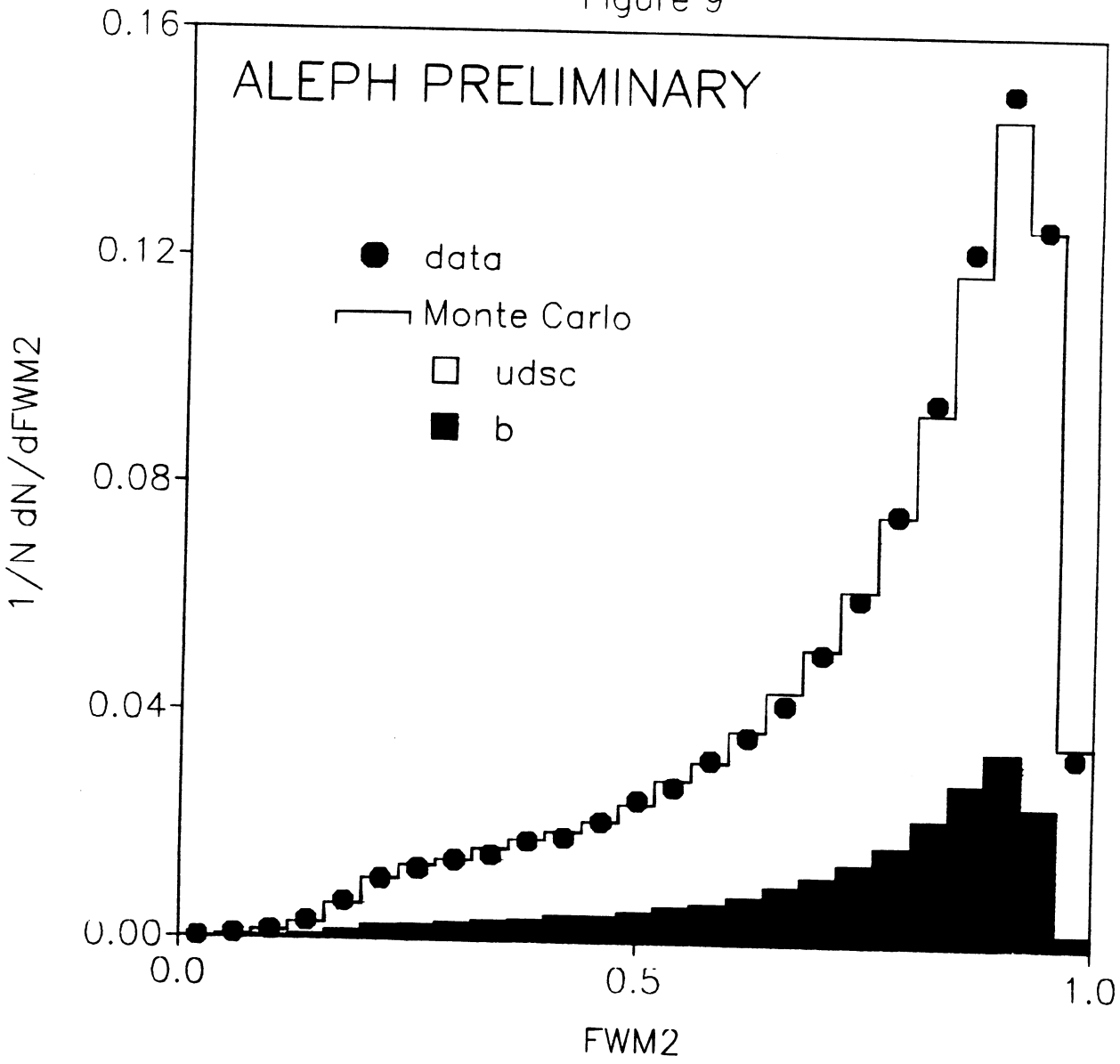

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.16
0.12
0.08
0.04
0.00
0.0
0.5
1.0
1/N dN/dFWM2
FWM2

Figure 10

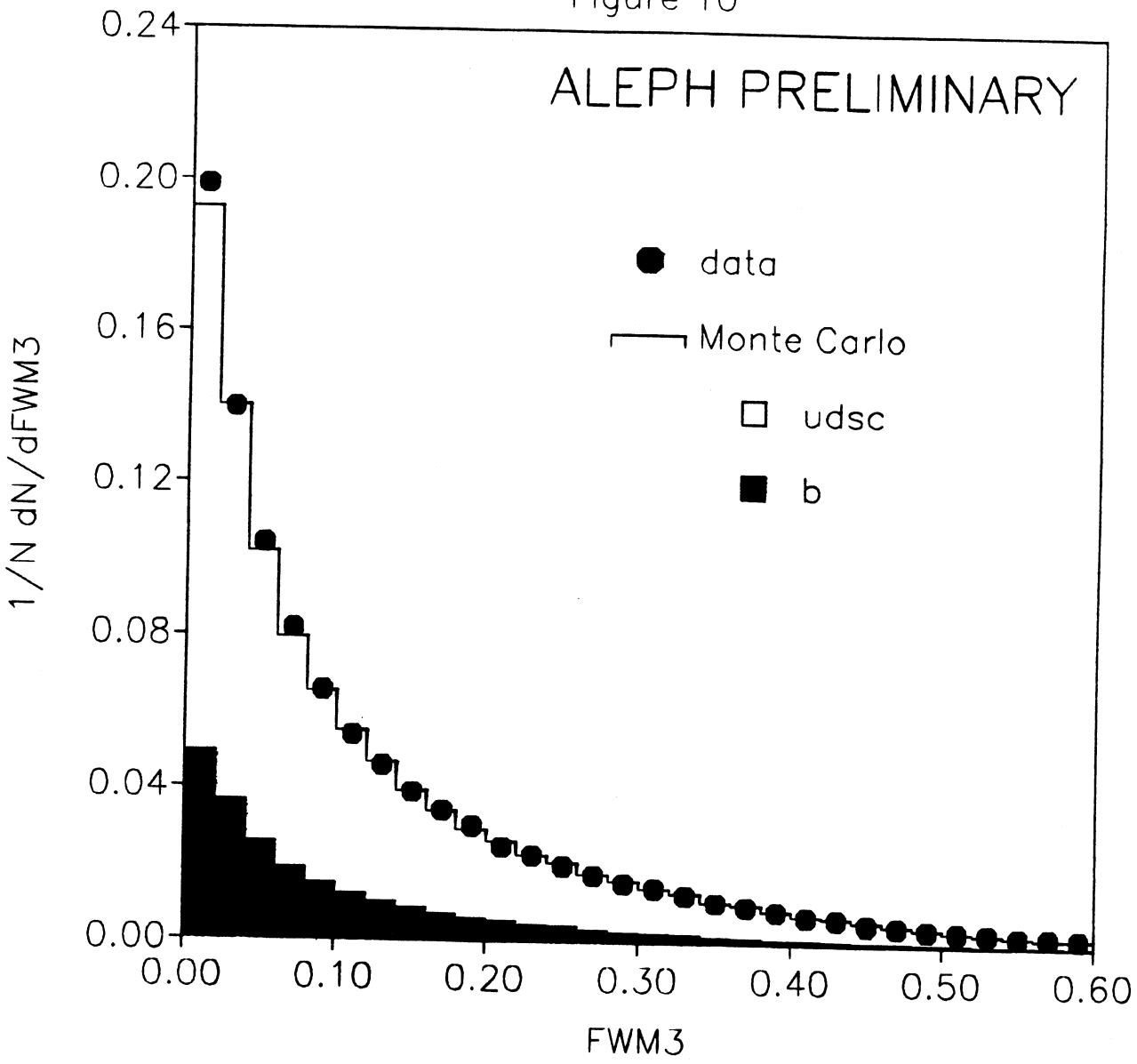

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
1/N dN/dFWM3
FWM3
0.24
0.20
0.16
0.12
0.08
0.04
0.00
0.00
0.10
0.20
0.30
0.40
0.50
0.60

Figure 11

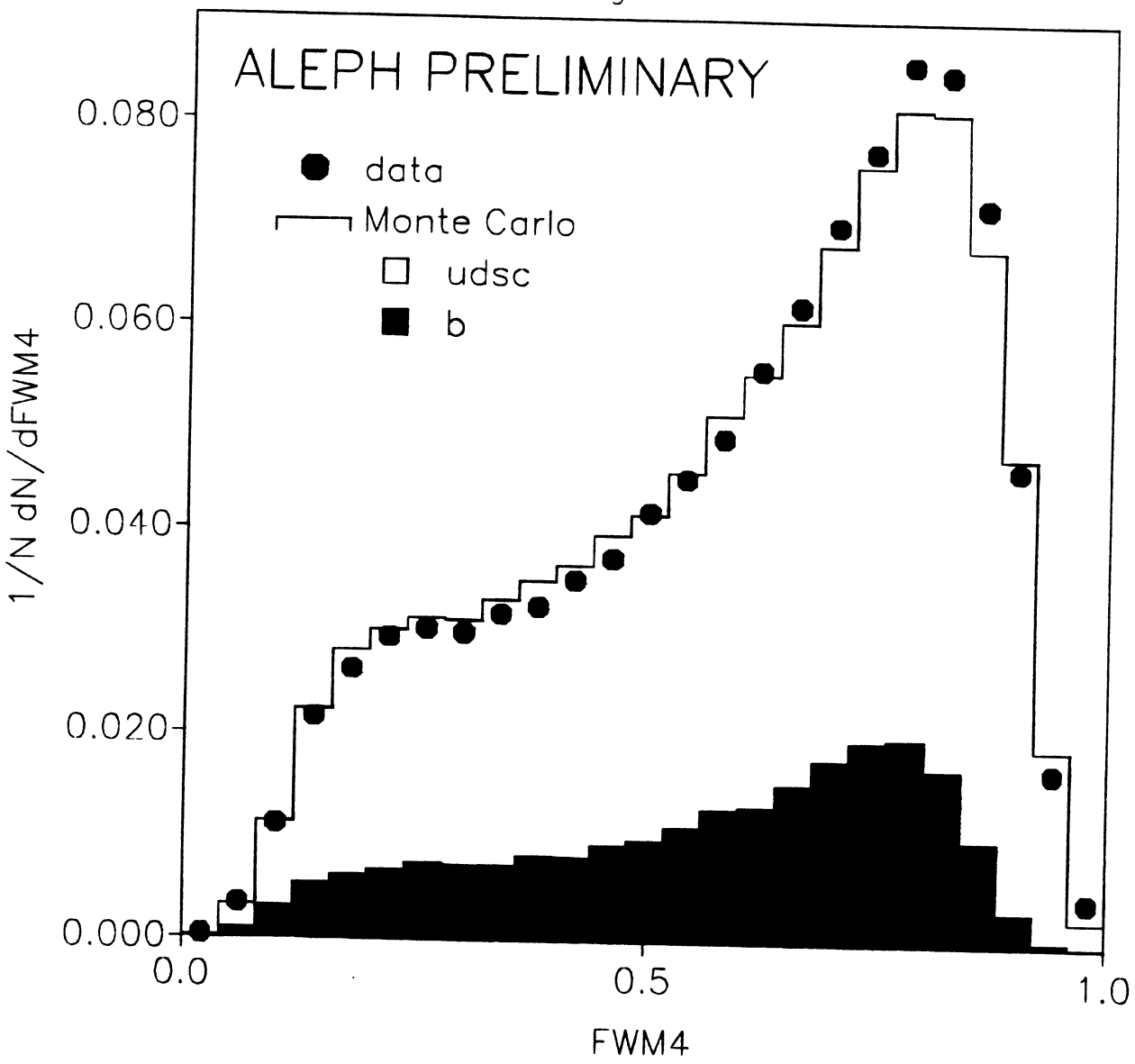

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.080
0.060
0.040
0.020
0.000
1/N dN/dFWM4
0.0
0.5
1.0
FWM4

Figure 12

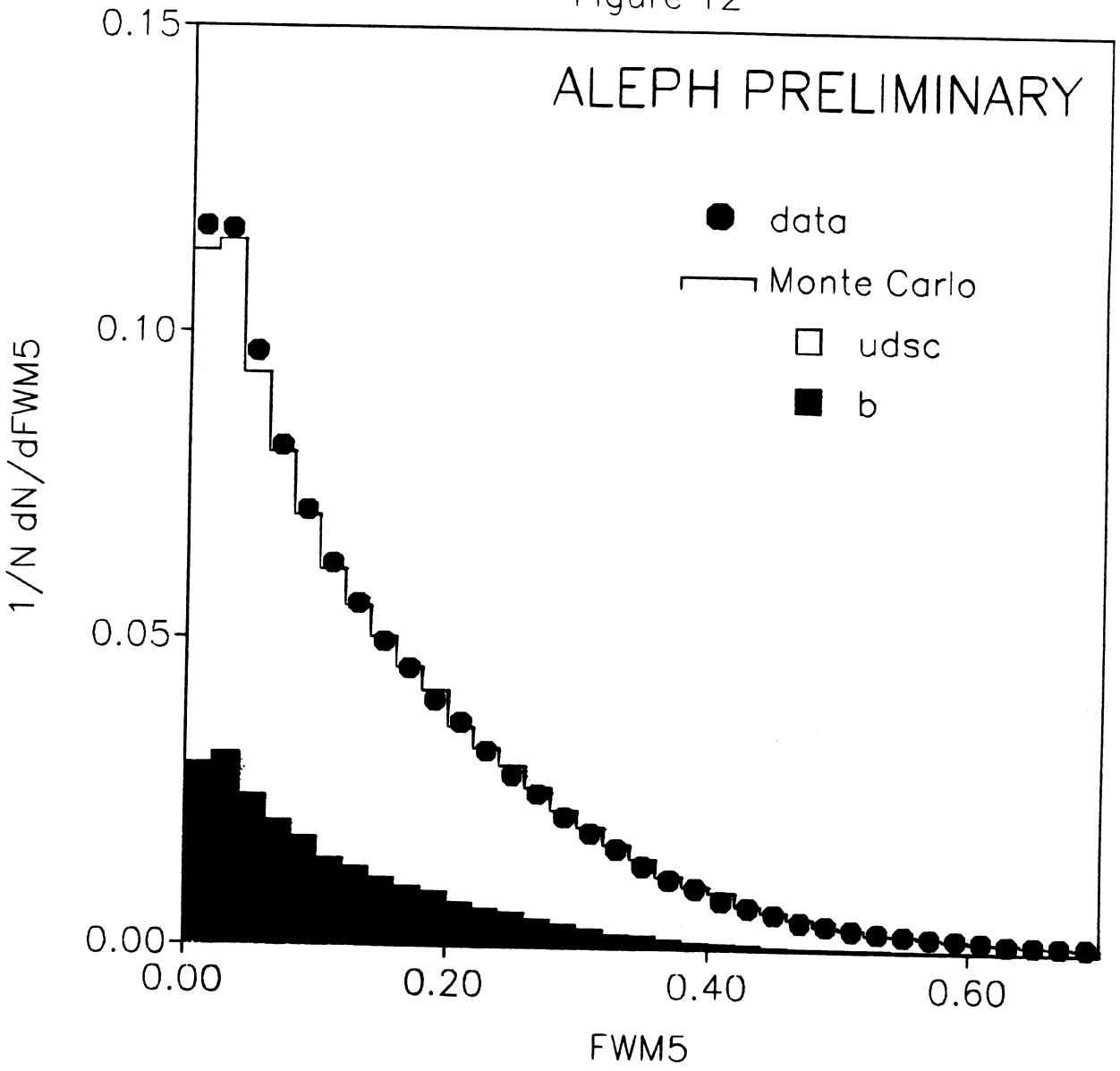

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.15
0.10
0.05
0.00
0.00
0.20
0.40
0.60
1/N dN/dFWM5
FWM5

Figure 13

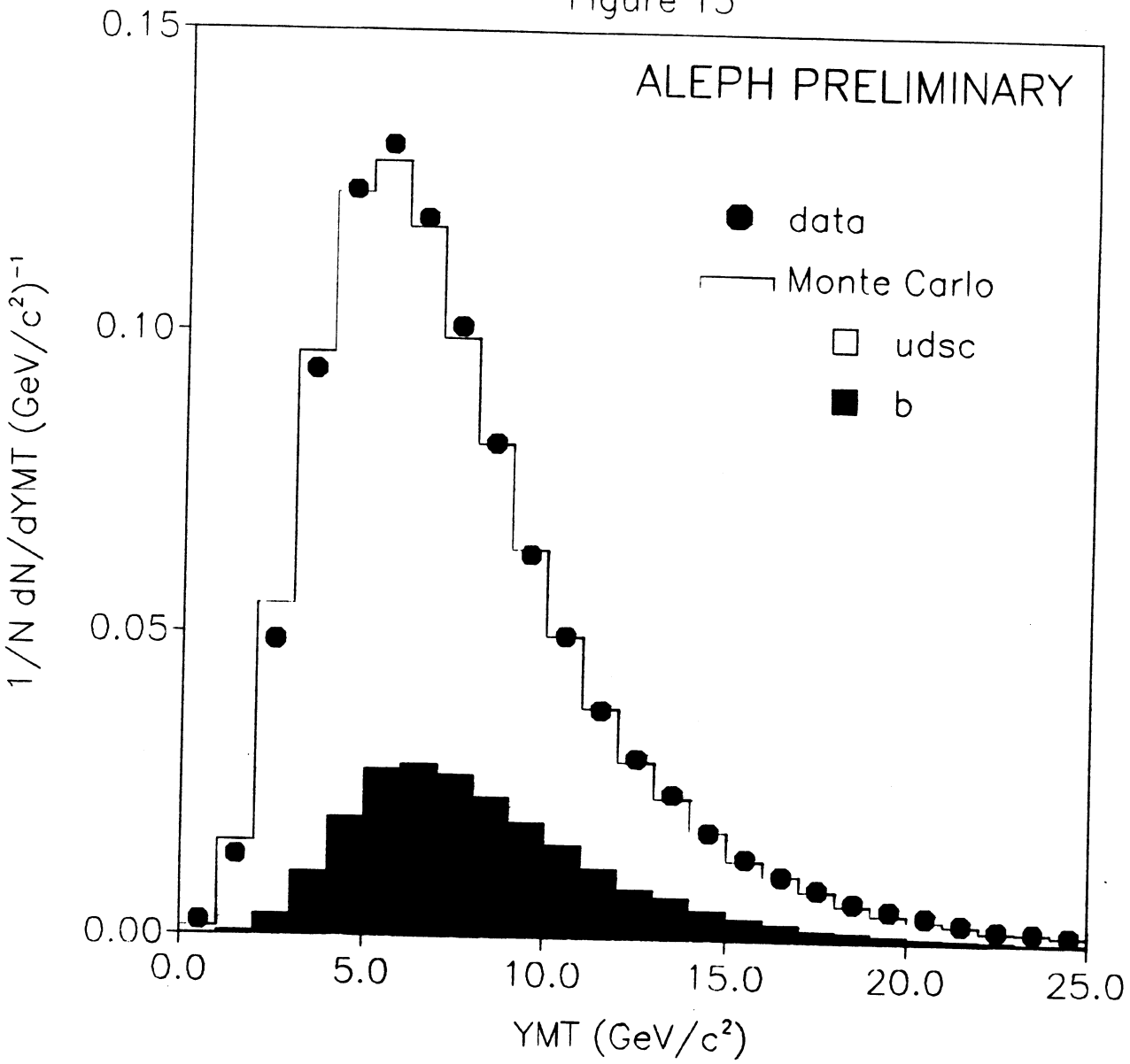

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.15
0.10
0.05
0.00
0.0
5.0
10.0
15.0
20.0
25.0
1/N dN/dYMT (GeV/c²)⁻¹
YMT (GeV/c²)

Figure 14

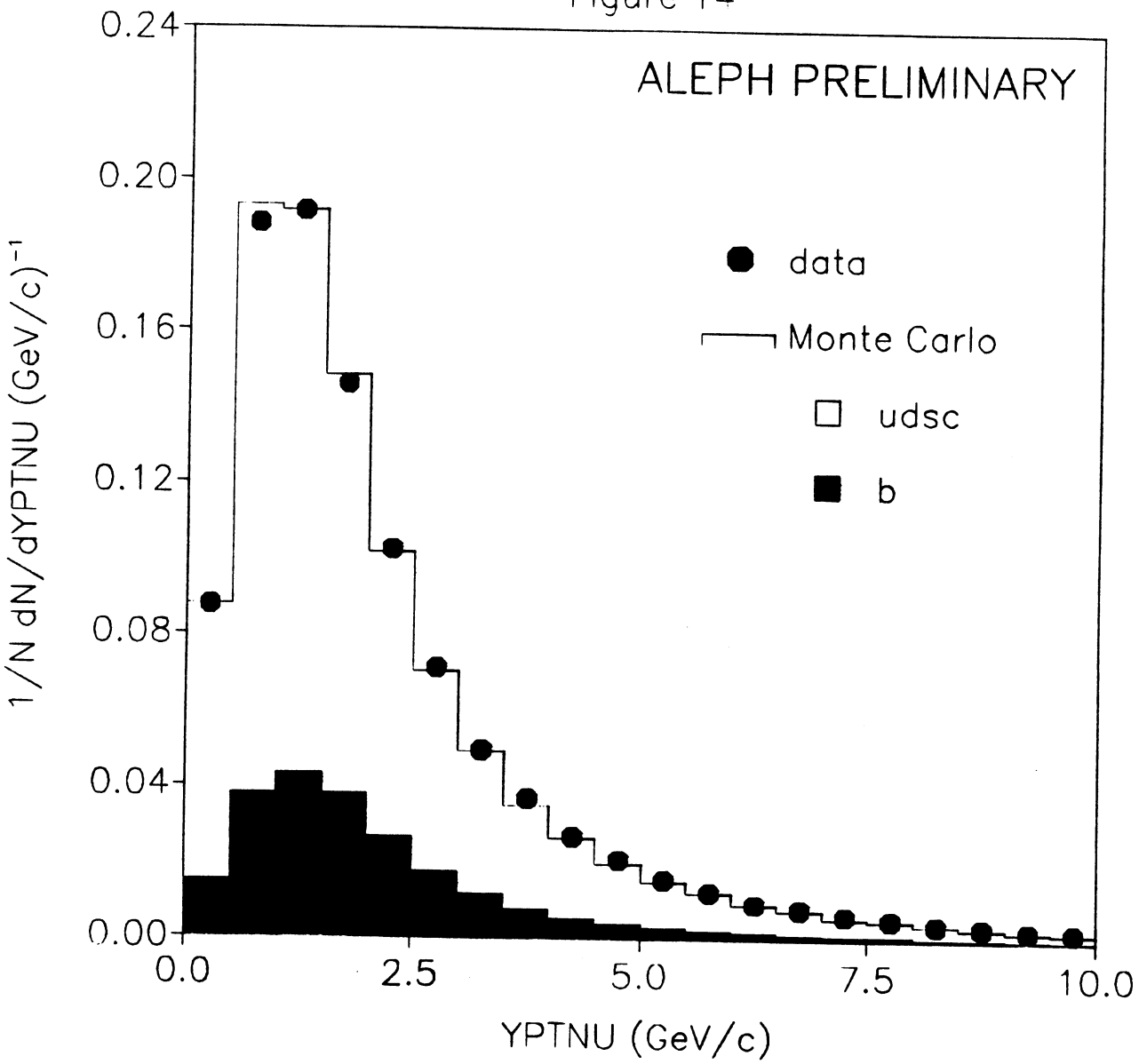

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.24
0.20
0.16
0.12
0.08
0.04
0.00
0.0
2.5
5.0
7.5
10.0
1/N dN/dYPTNU (GeV/c)$^{-1}$
YPTNU (GeV/c)

Figure 15

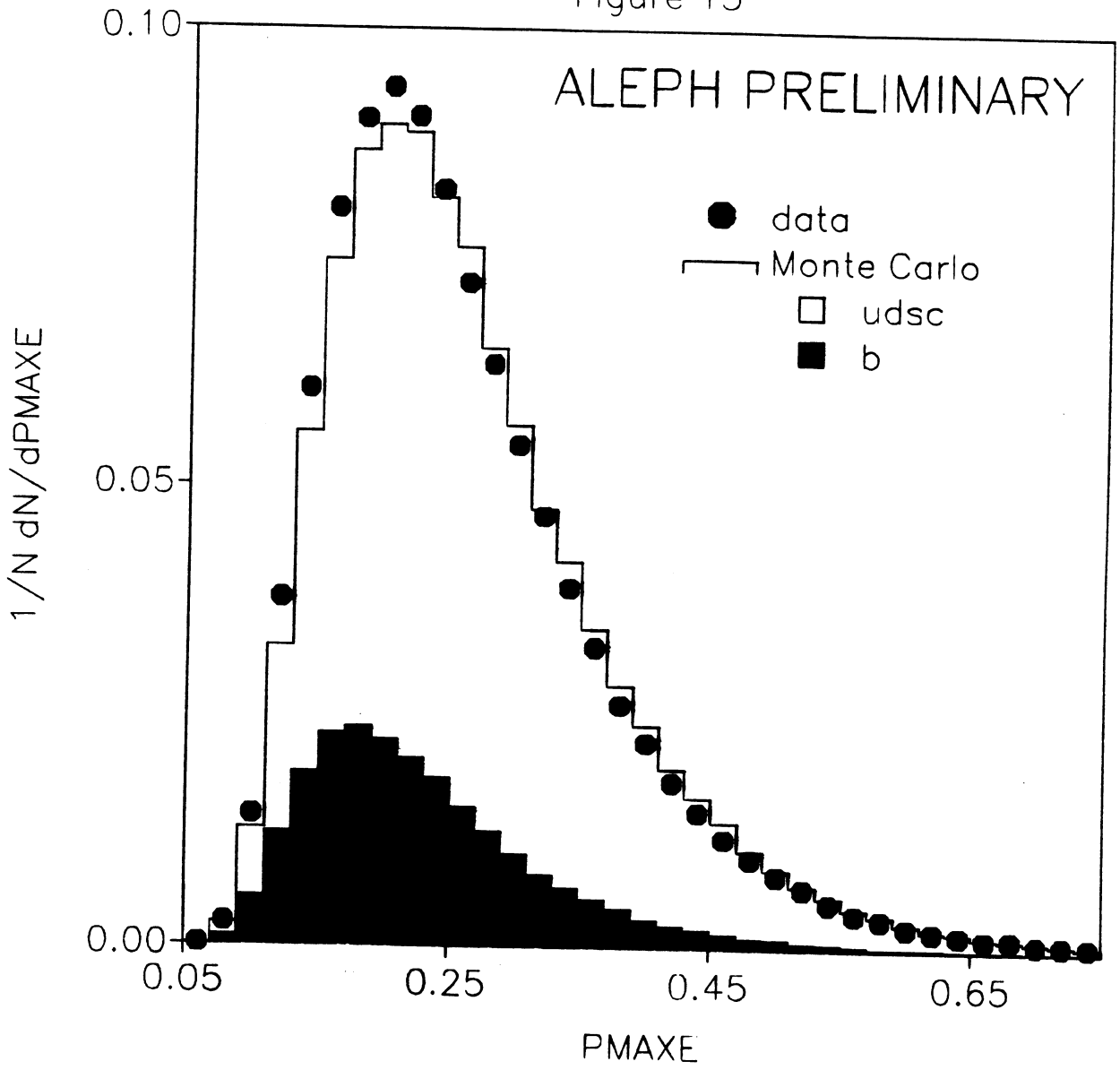

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.10
0.05
0.00
0.05
0.25
0.45
0.65
1/N dN/dPMAXE
PMAXE

Figure 16

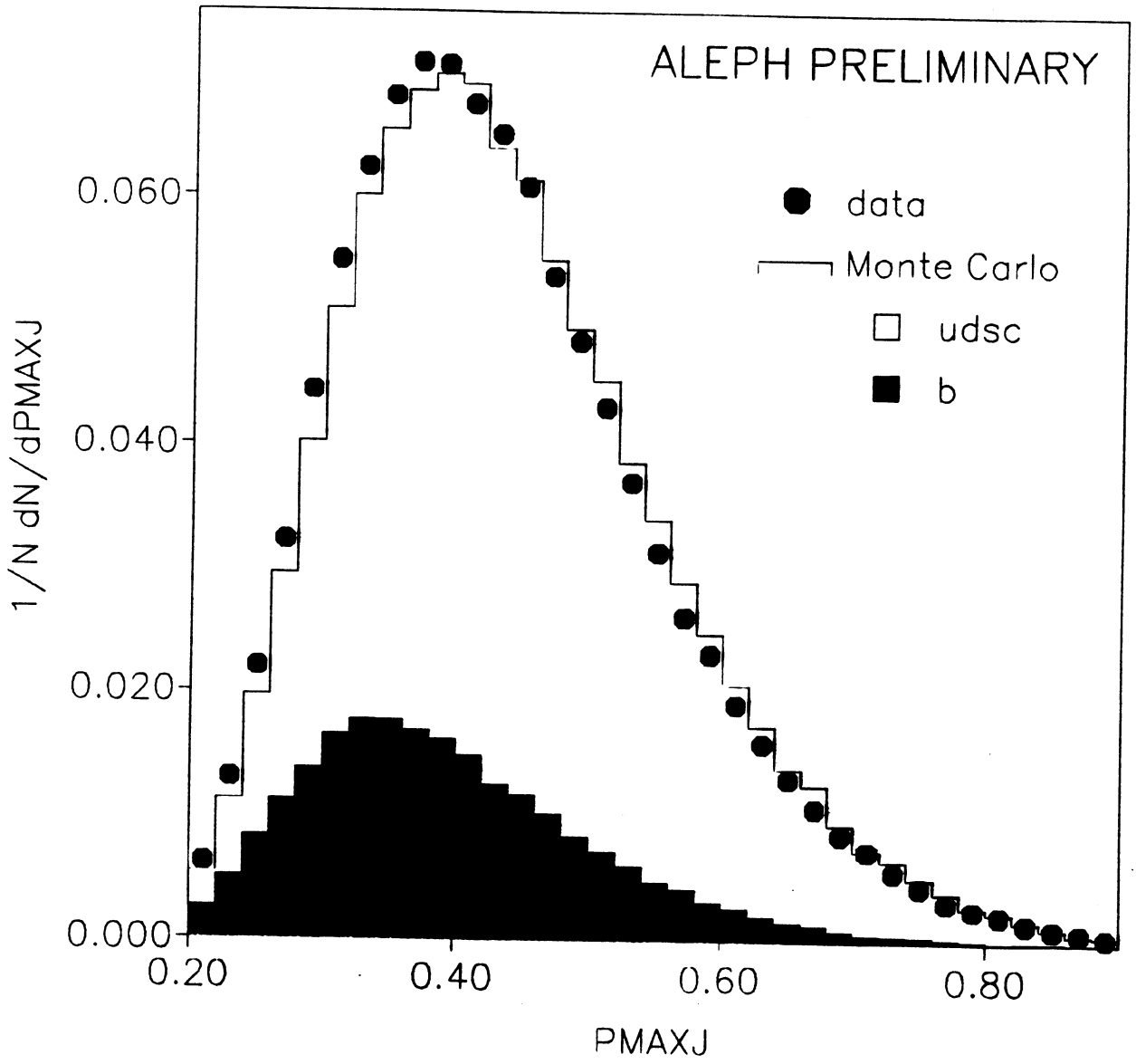

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.060
0.040
0.020
0.000
1/N dN/dPMAXJ
0.20
0.40
0.60
0.80
PMAXJ

Figure 17

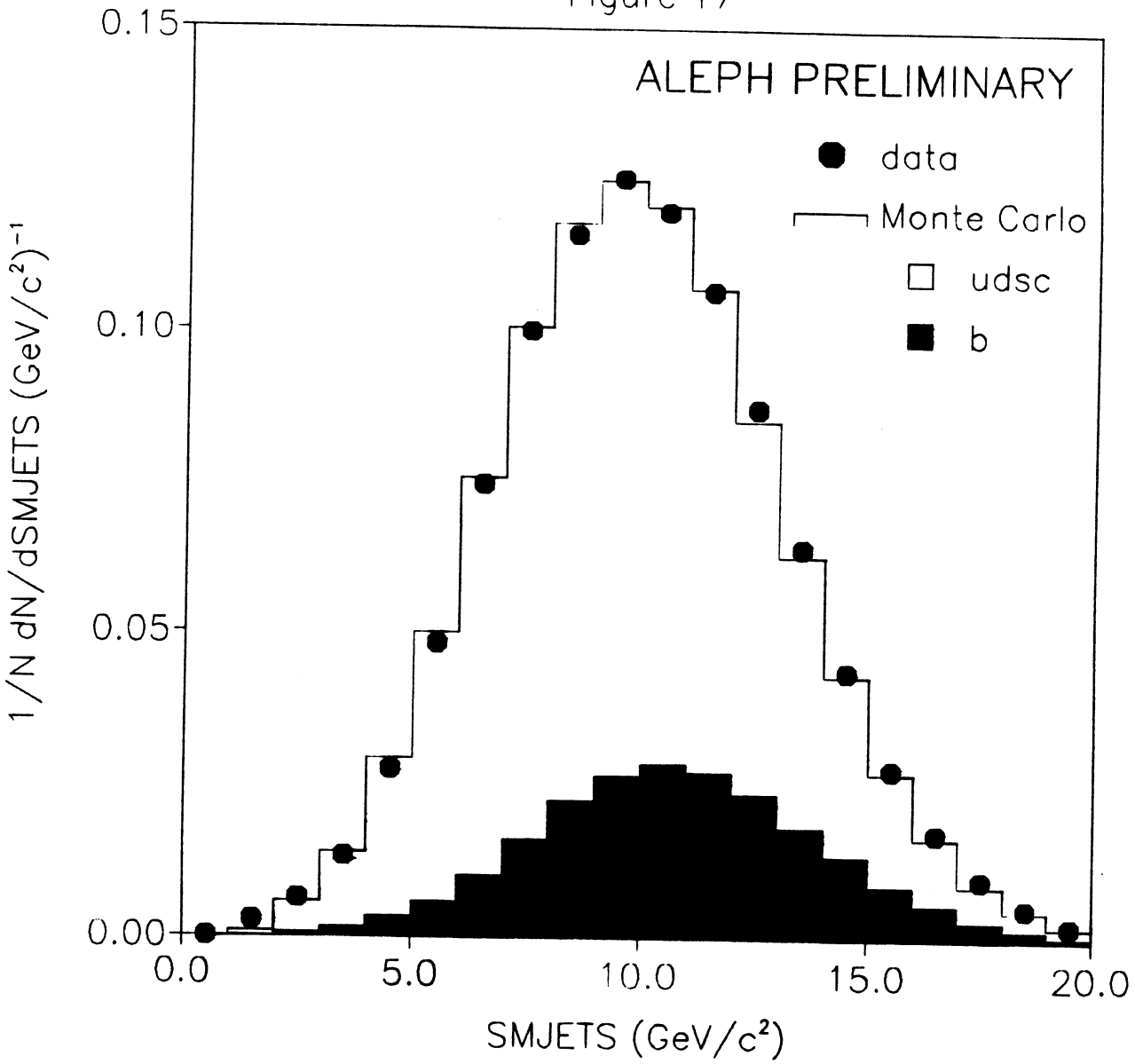

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.15
0.10
0.05
0.00
0.0
5.0
10.0
15.0
20.0
1/N dN/dSMJETS (GeV/c²)⁻¹
SMJETS (GeV/c²)

Figure 18

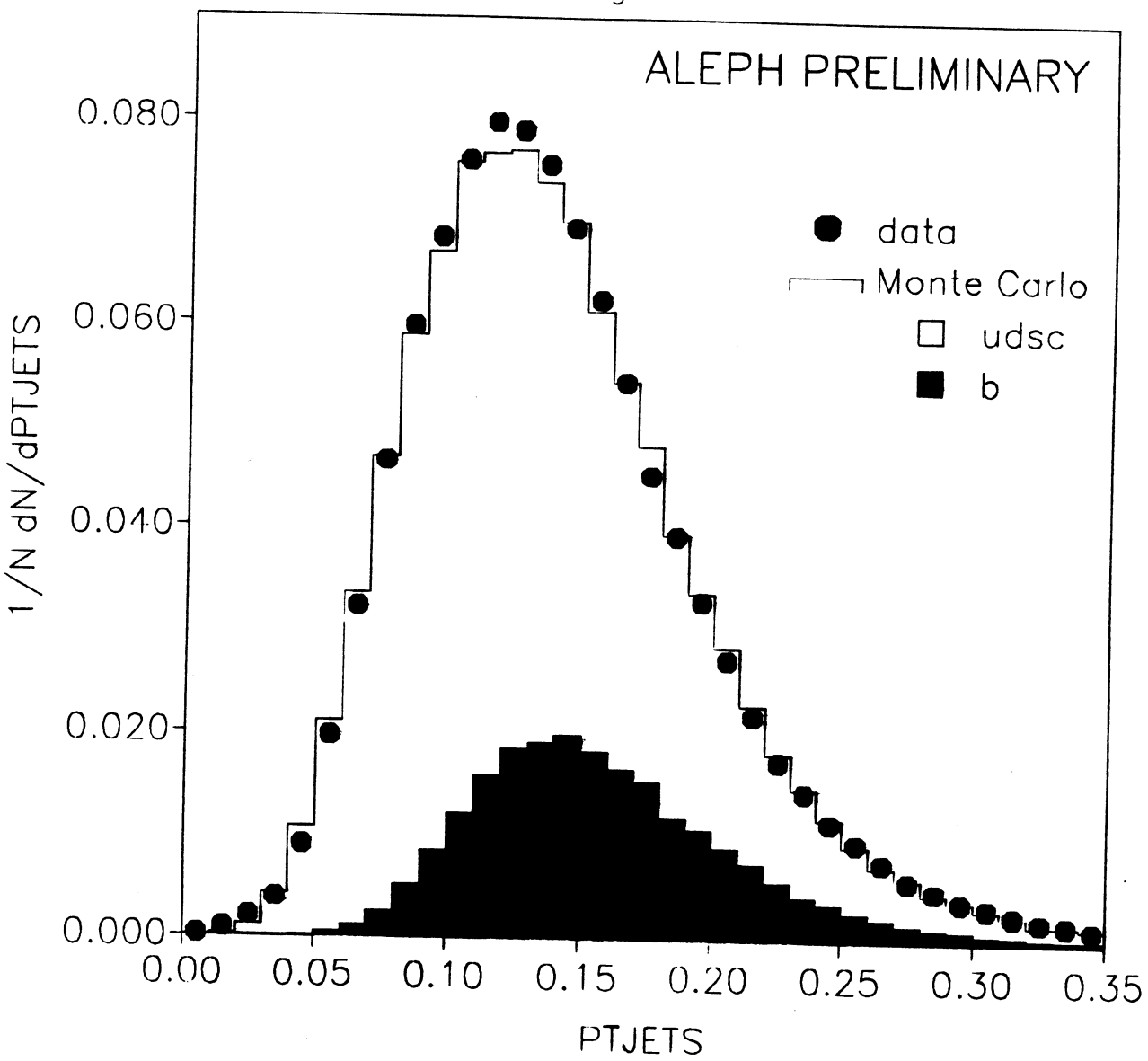

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
1/N dN/dPTJETS
0.080
0.060
0.040
0.020
0.000
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
PTJETS

Figure 19

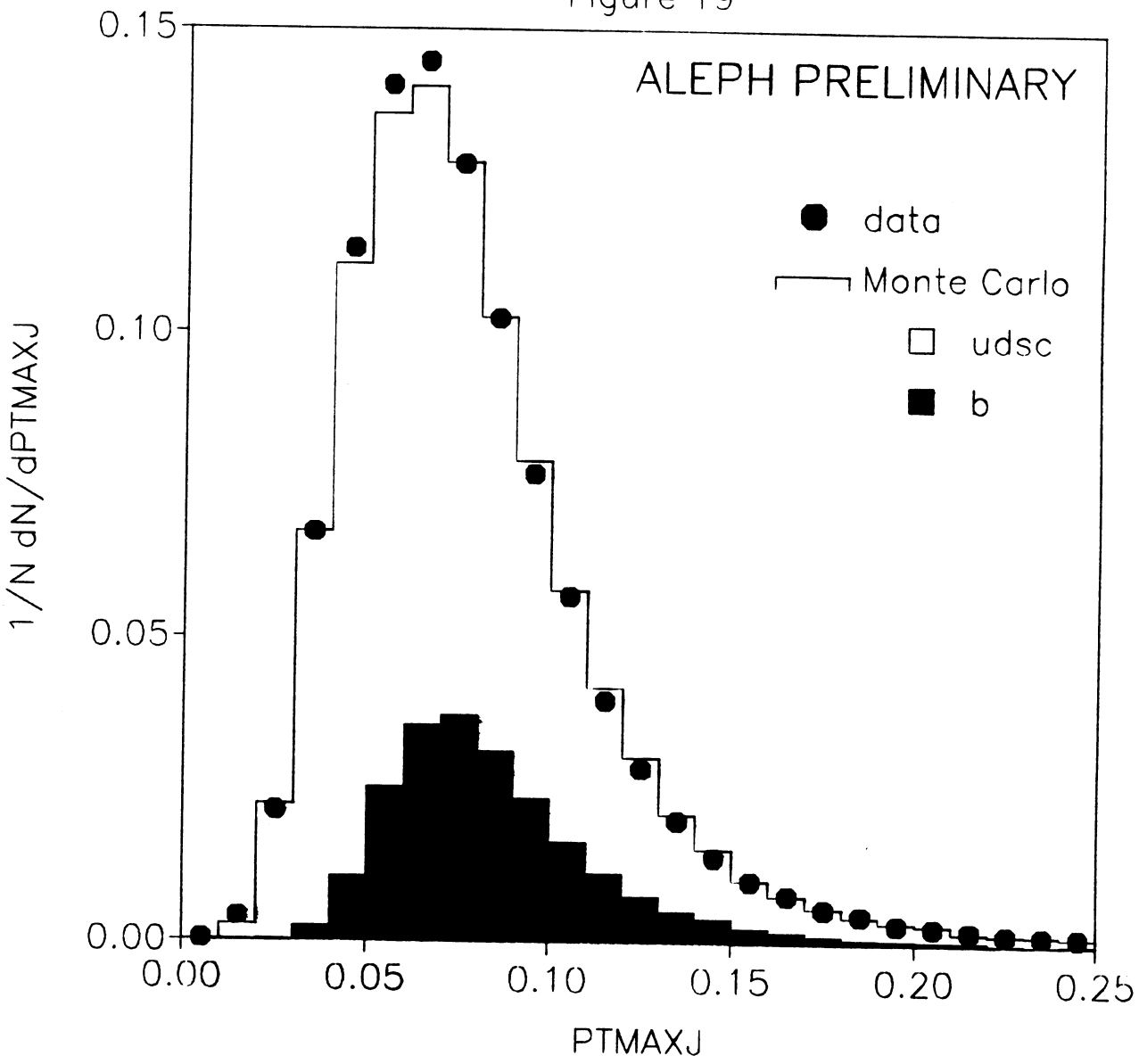

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
1/N dN/dPTMAXJ
PTMAXJ
0.00
0.05
0.10
0.15
0.20
0.25
0.00
0.05
0.10
0.15

Figure 20

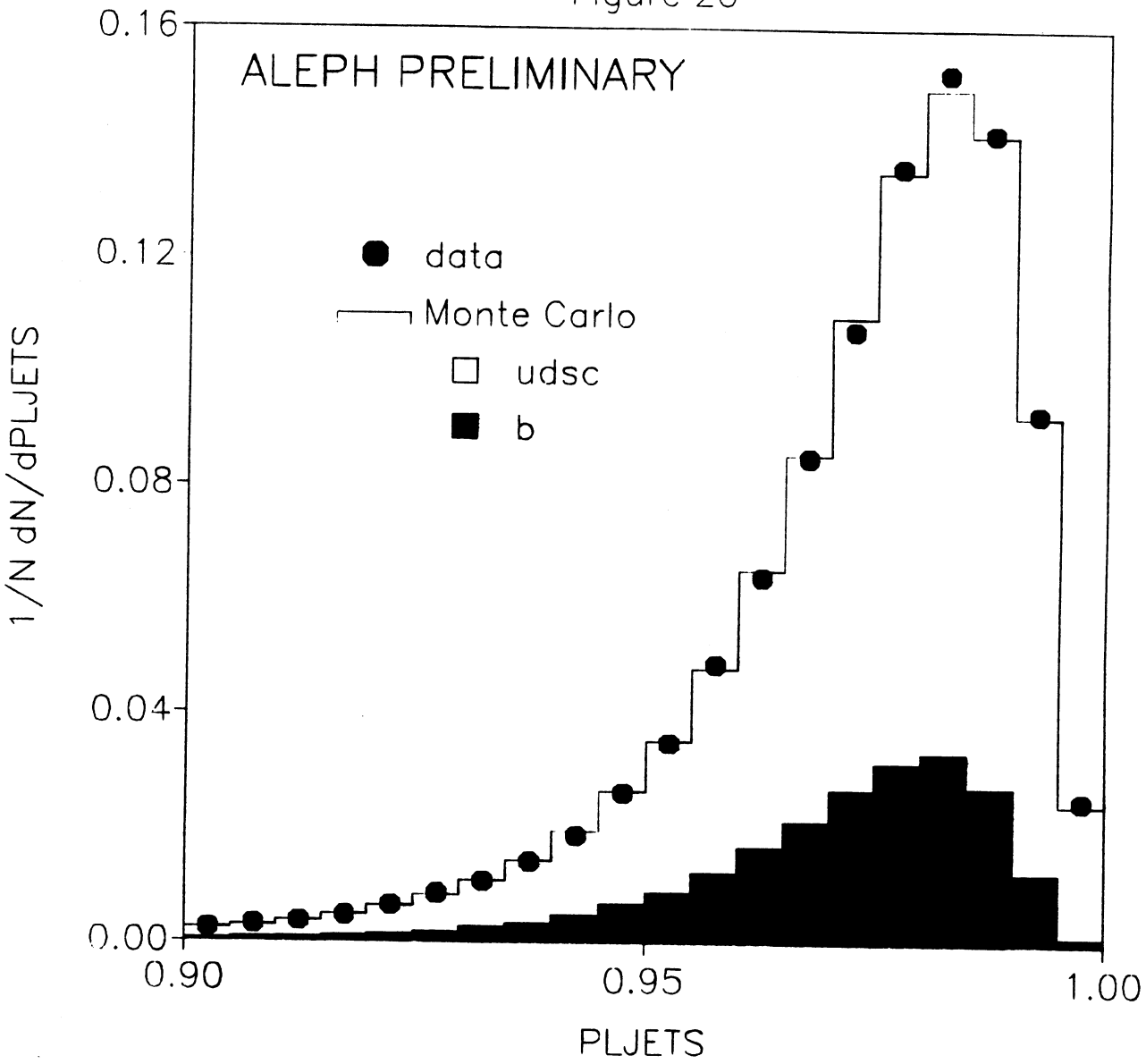

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.16
0.12
0.08
0.04
0.00
1/N dN/dPLJETS
0.90
0.95
1.00
PLJETS

Figure 21

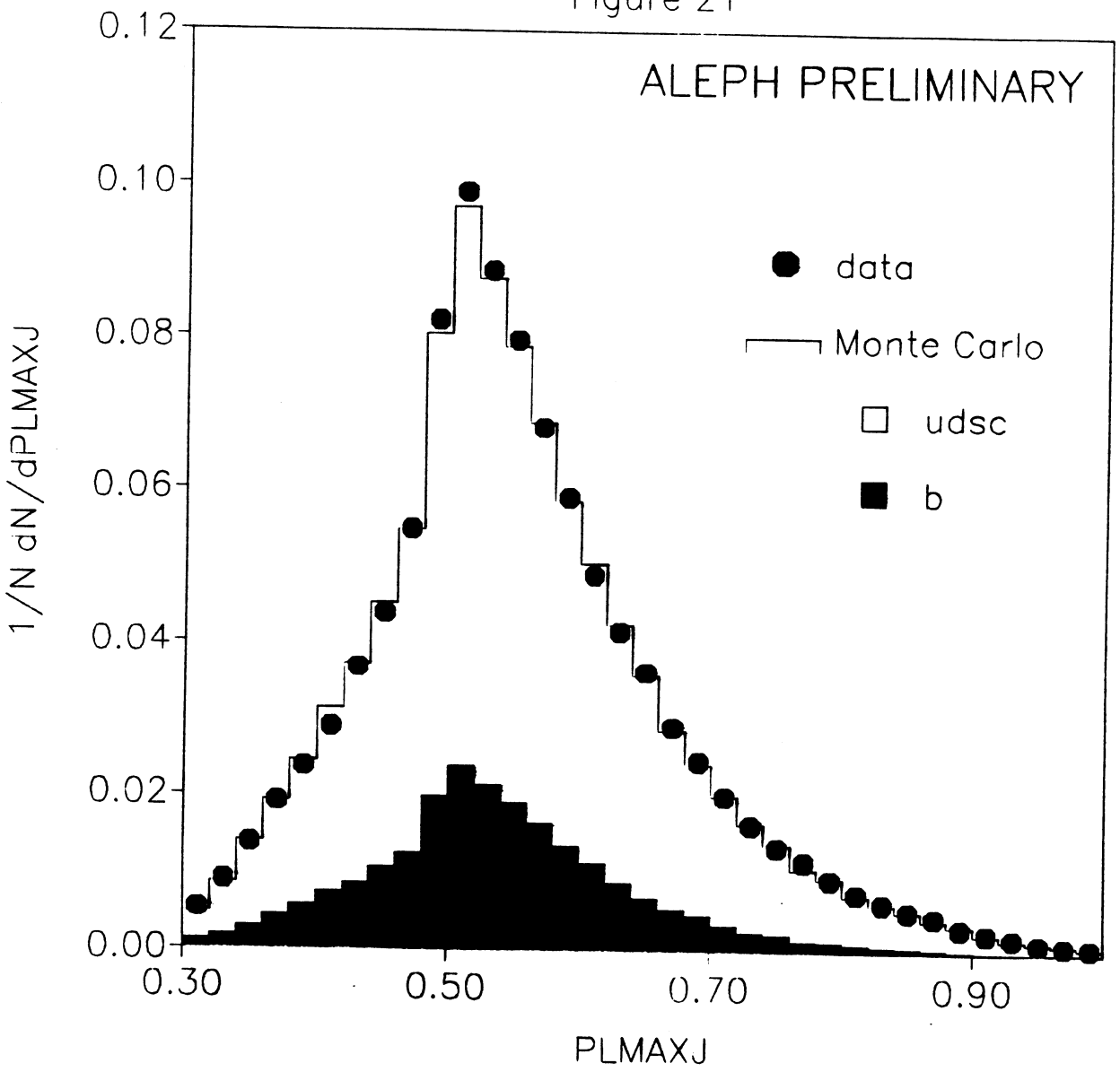

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
1/N dN/dPLMAXJ
PLMAXJ
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0.30
0.50
0.70
0.90

Figure 22

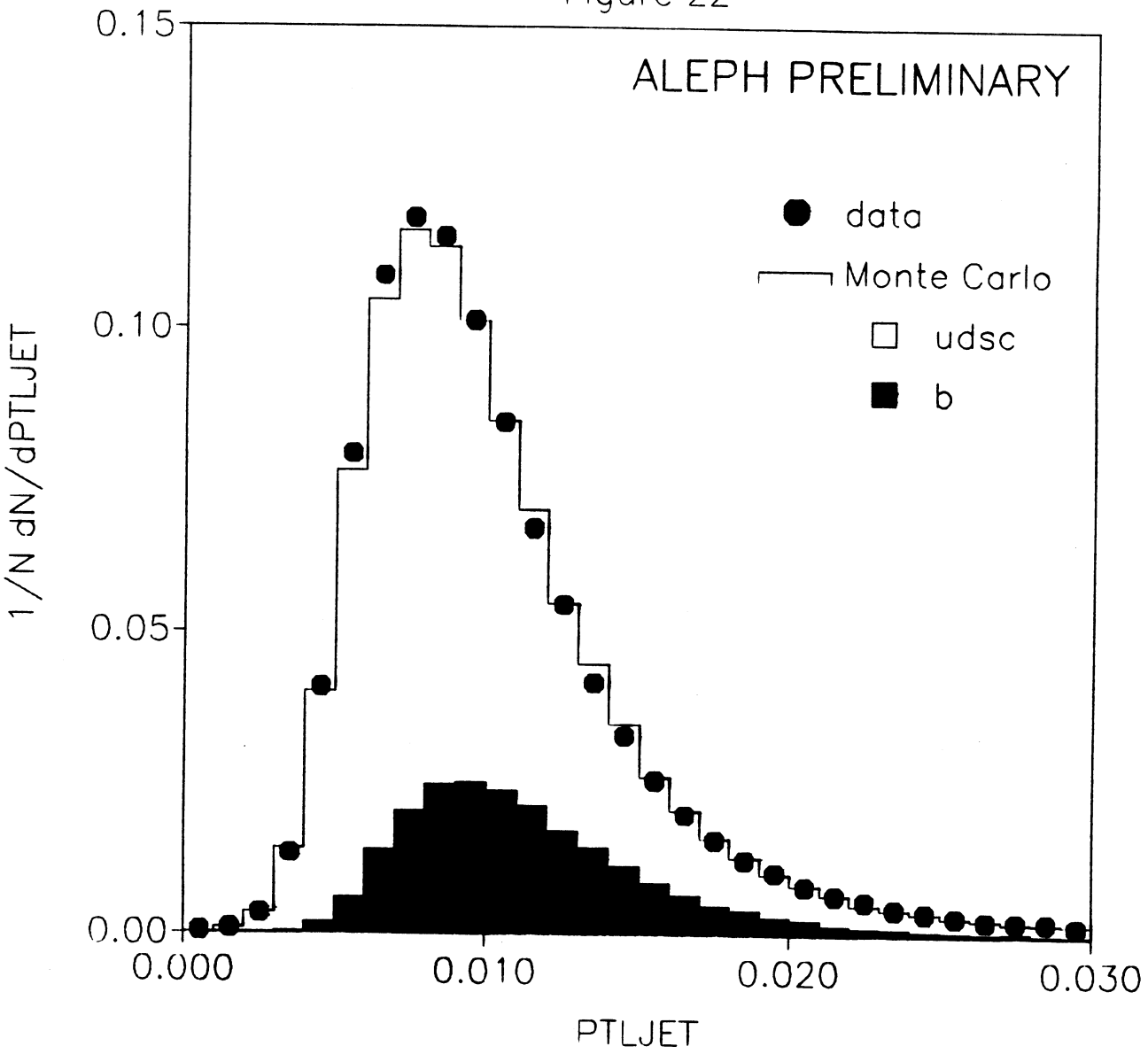

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.15
0.10
0.05
0.00
0.000
0.010
0.020
0.030
1/N dN/dPTLJET
PTLJET

Figure 23

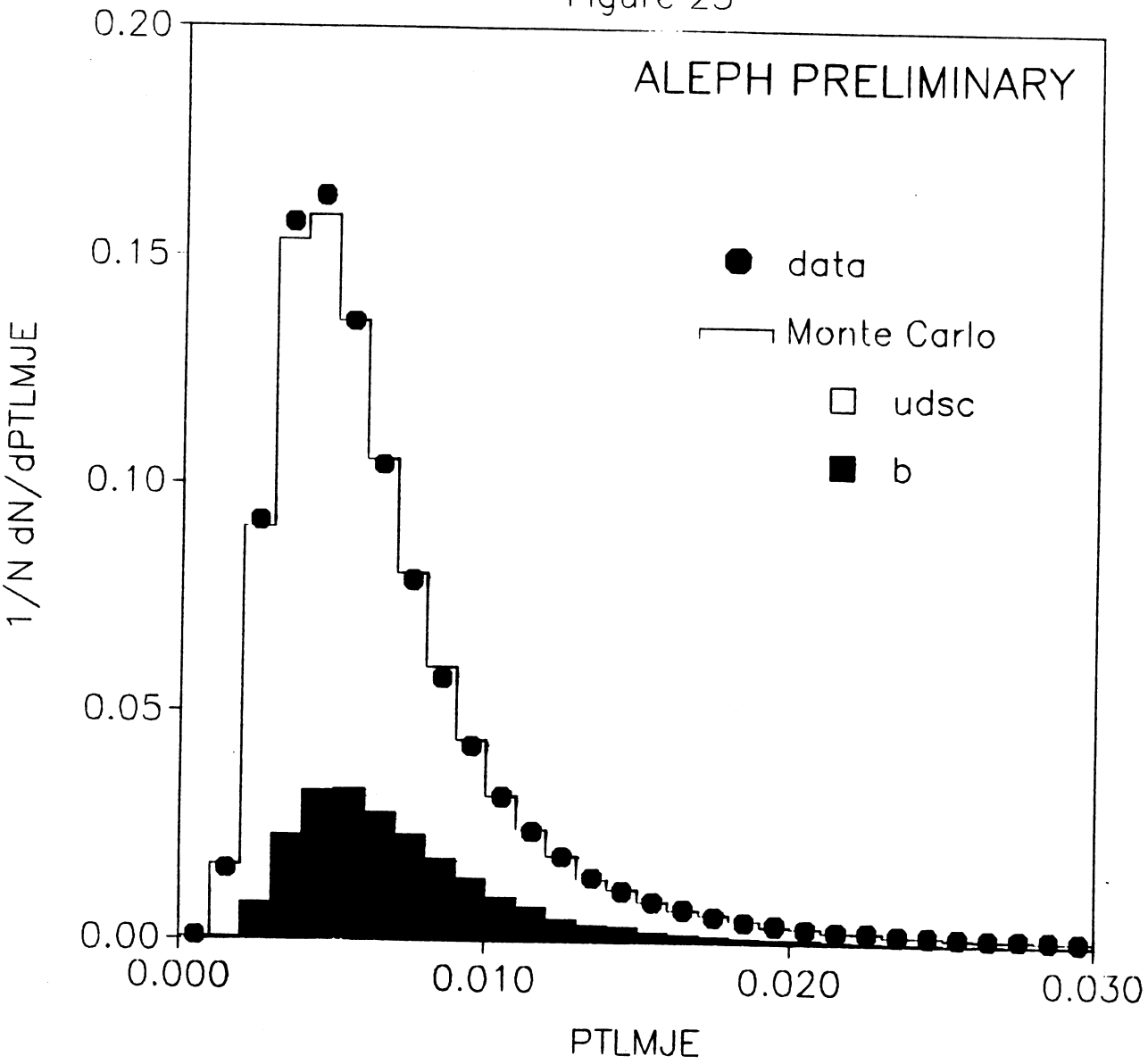

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.20
0.15
0.10
0.05
0.00
0.000
0.010
0.020
0.030
1/N dN/dPTLMJE
PTLMJE

Figure 24

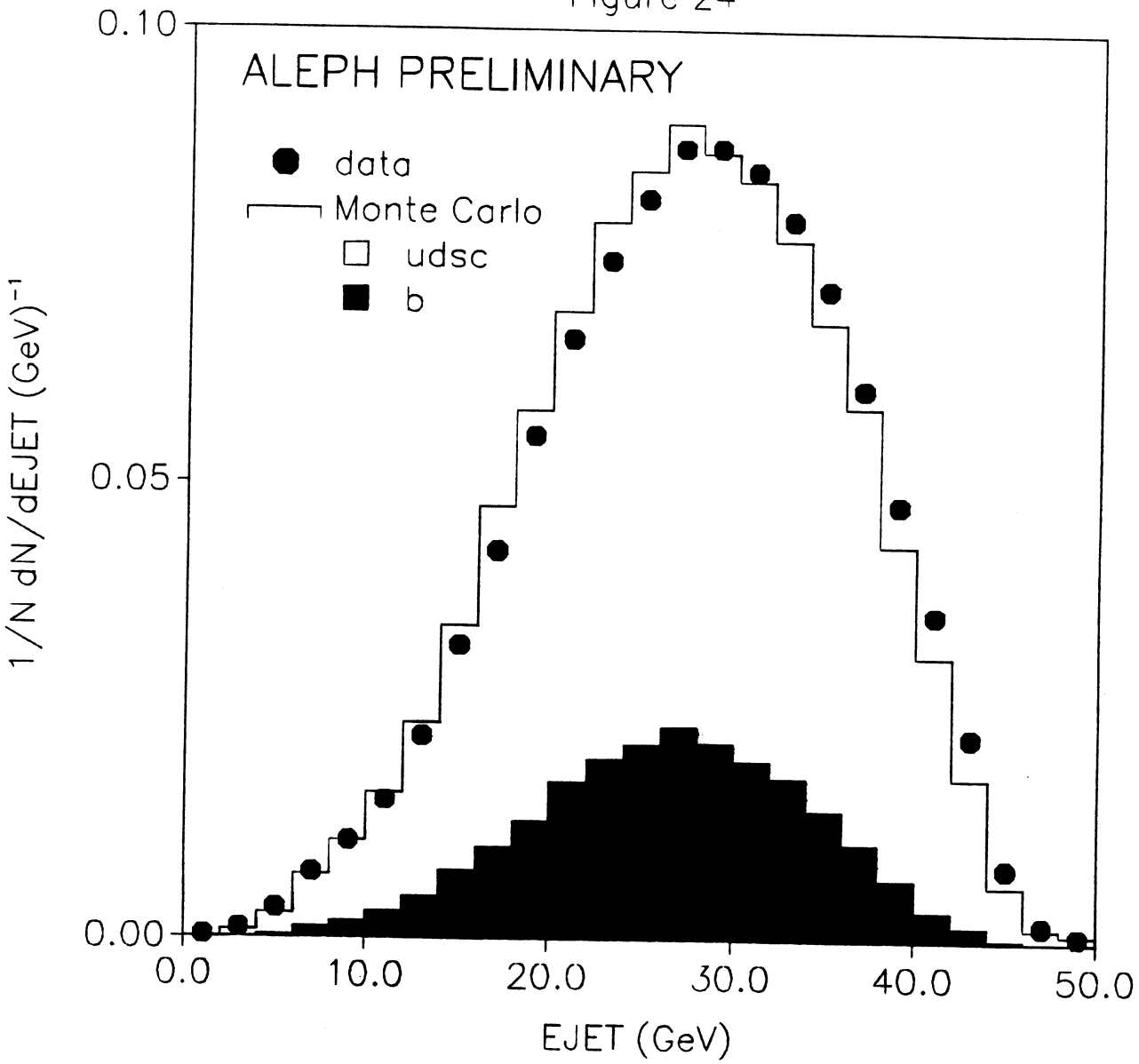

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.10
0.05
0.00
1/N dN/dEJET (GeV)$^{-1}$
0.0
10.0
20.0
30.0
40.0
50.0
EJET (GeV)

Figure 25

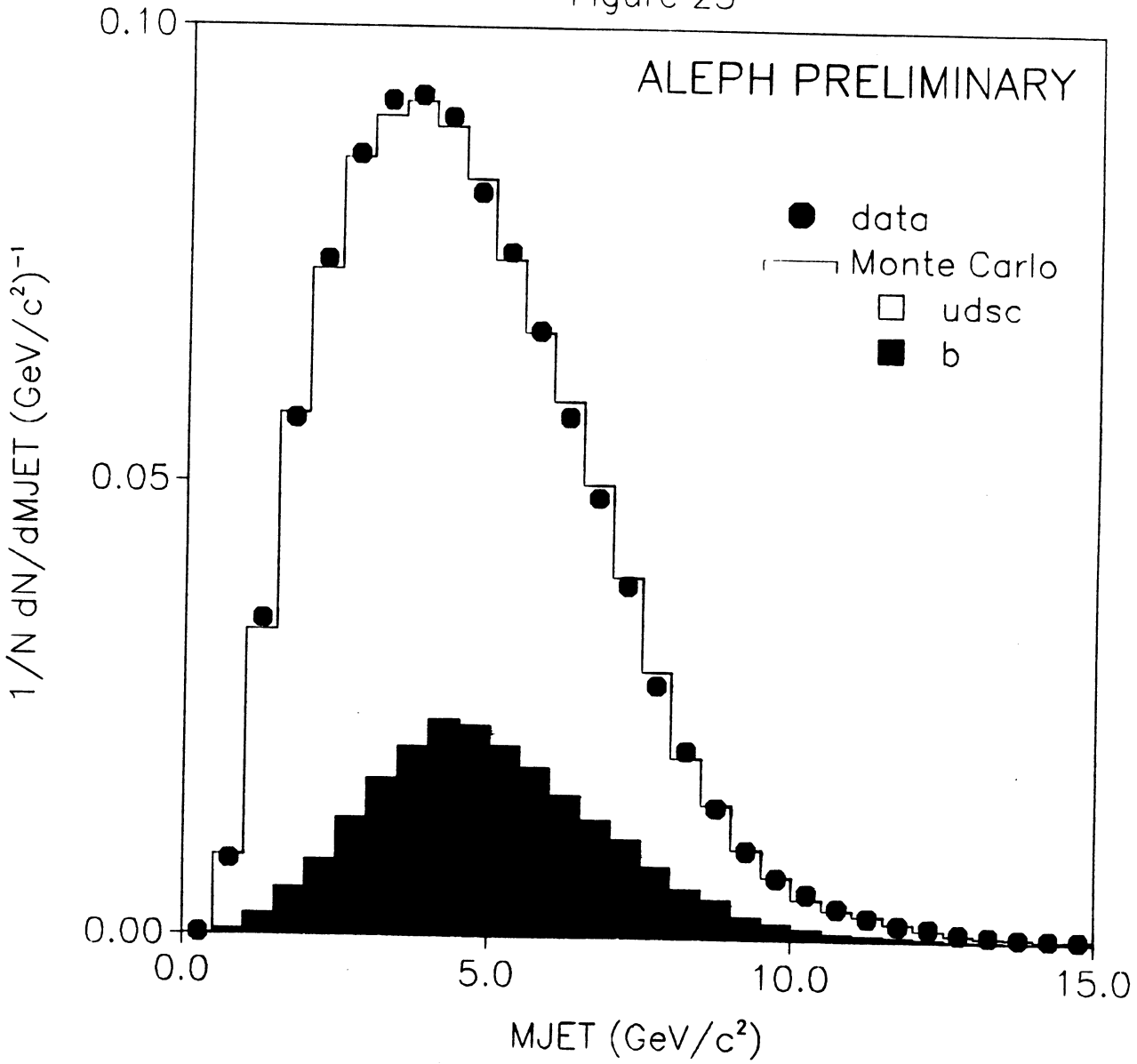

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.10
0.05
0.00
0.0
5.0
10.0
15.0
1/N dN/dMJET (GeV/c²)⁻¹
MJET (GeV/c²)

Figure 26

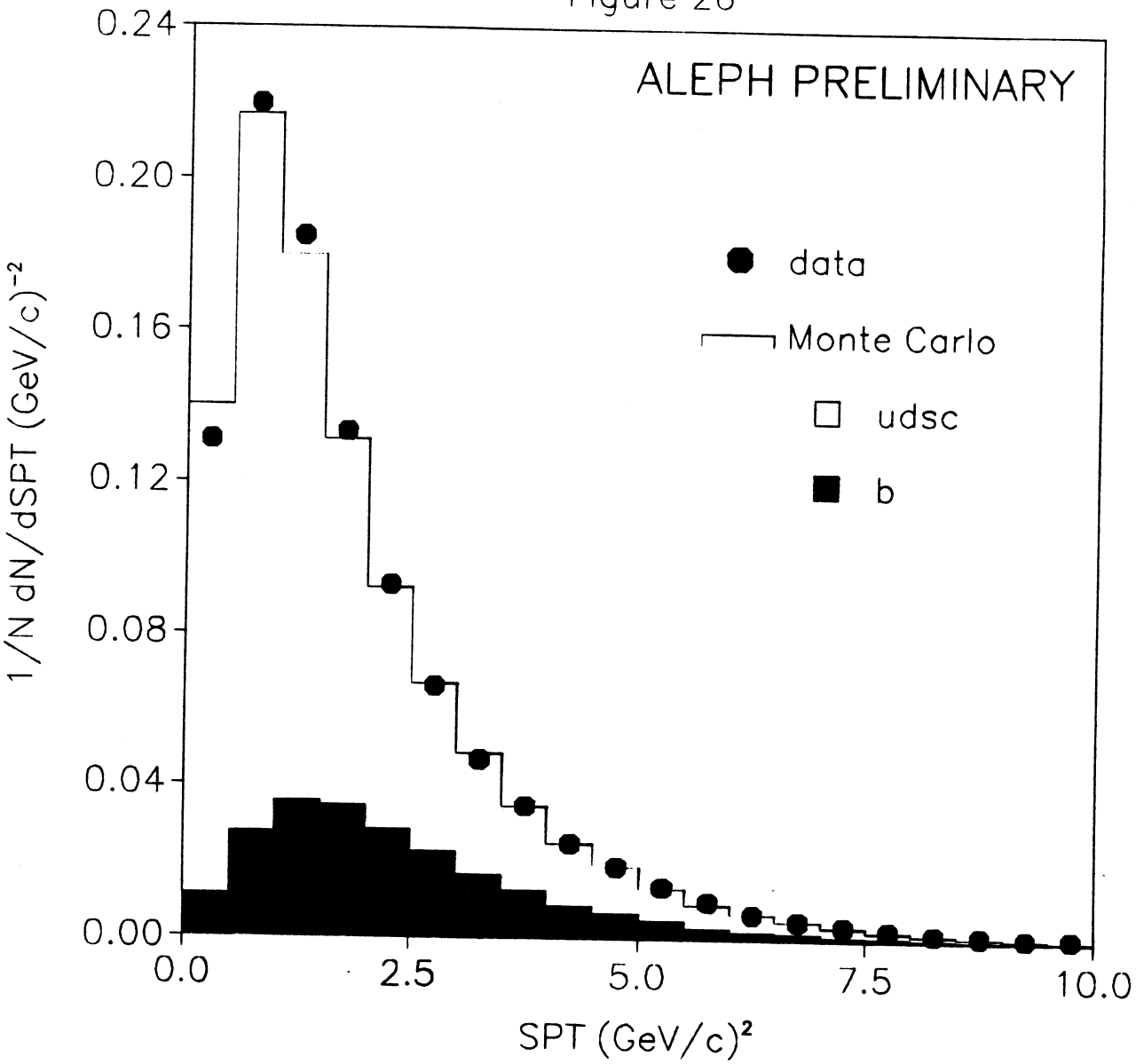
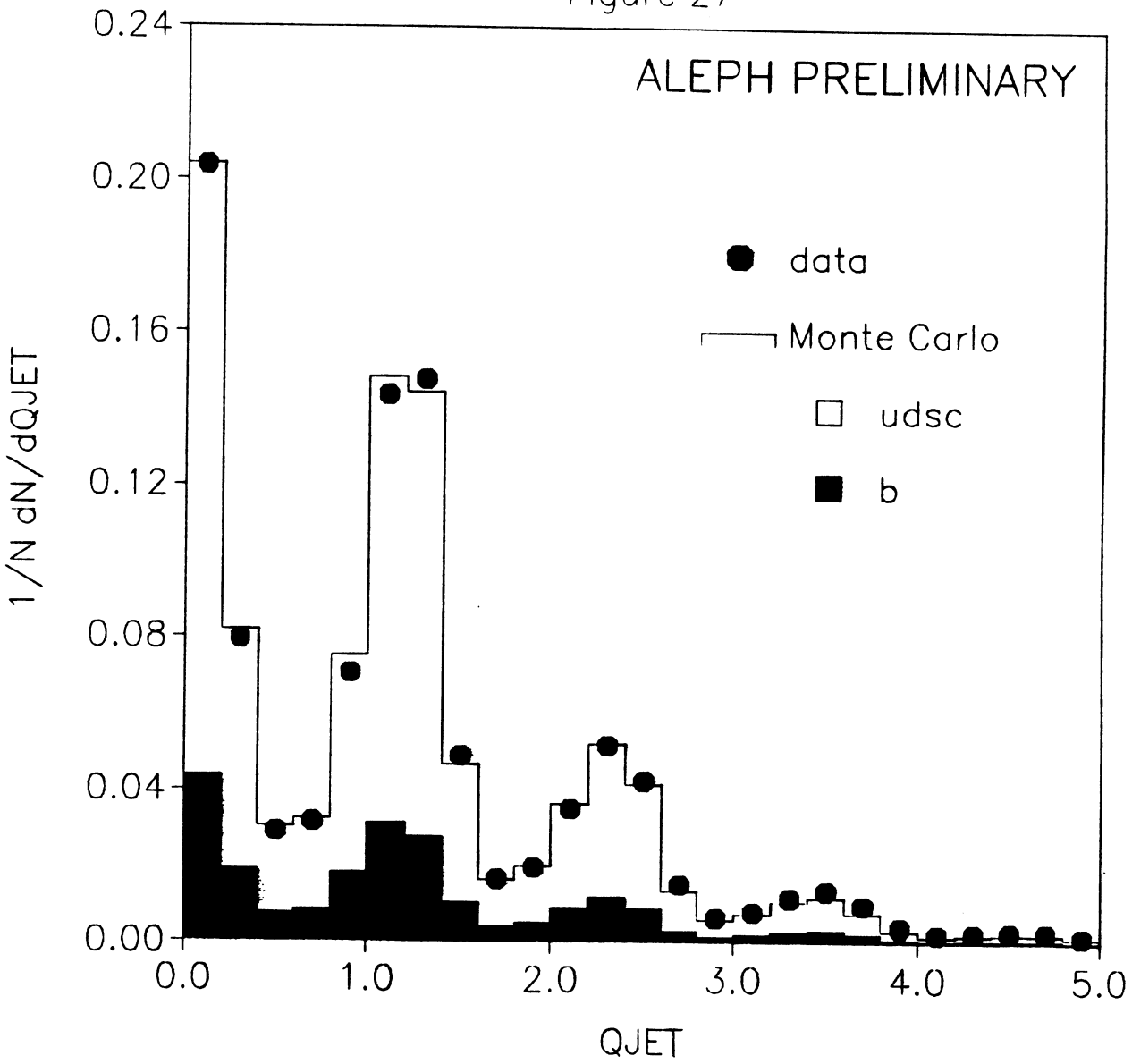

ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
0.24
0.20
0.16
0.12
0.08
0.04
0.00
0.0
2.5
5.0
7.5
10.0
1/N dN/dSPT (GeV/c)⁻²
SPT (GeV/c)²

Figure 27


ALEPH PRELIMINARY
data
Monte Carlo
udsc
b
1/N dN/dQJET
QJET
0.24
0.20
0.16
0.12
0.08
0.04
0.00
0.0
1.0
2.0
3.0
4.0
5.0

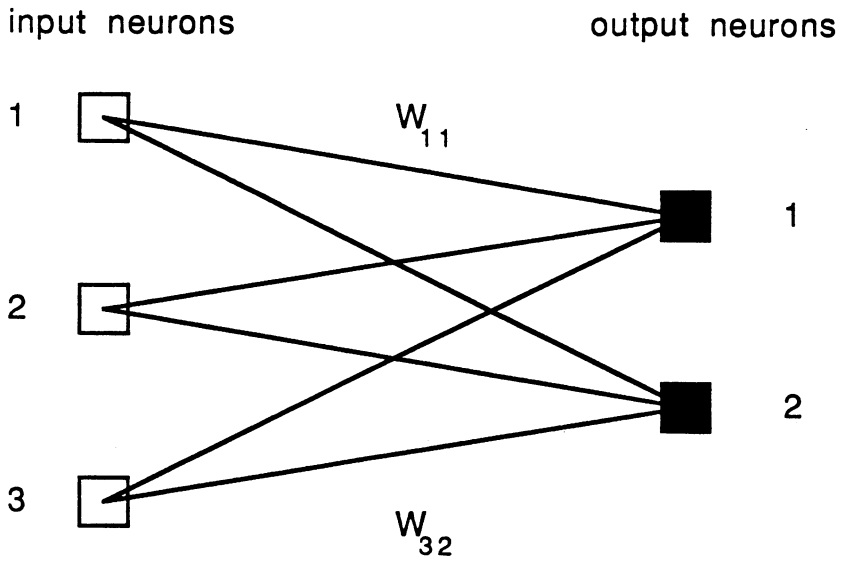
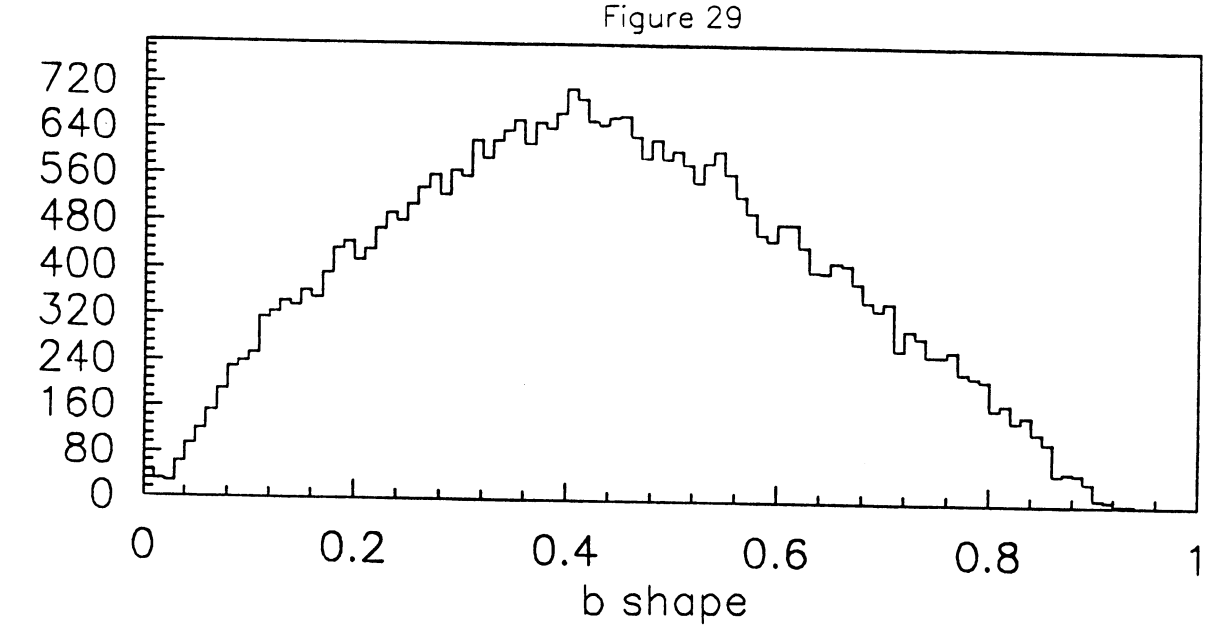


Fig. 28

Figure 29


720
640
560
480
400
320
240
160
80
0
0
0.2
0.4
0.6
0.8
1
b shape

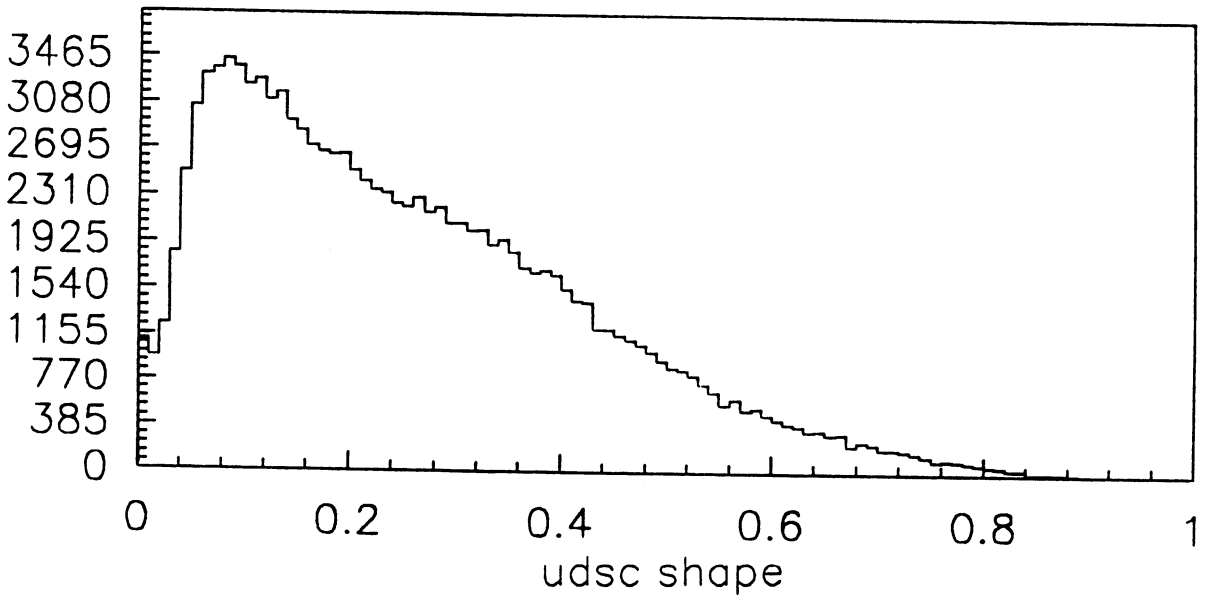

3465
3080
2695
2310
1925
1540
1155
770
385
0
0
0.2
0.4
0.6
0.8
1
udsc shape

Figure 30

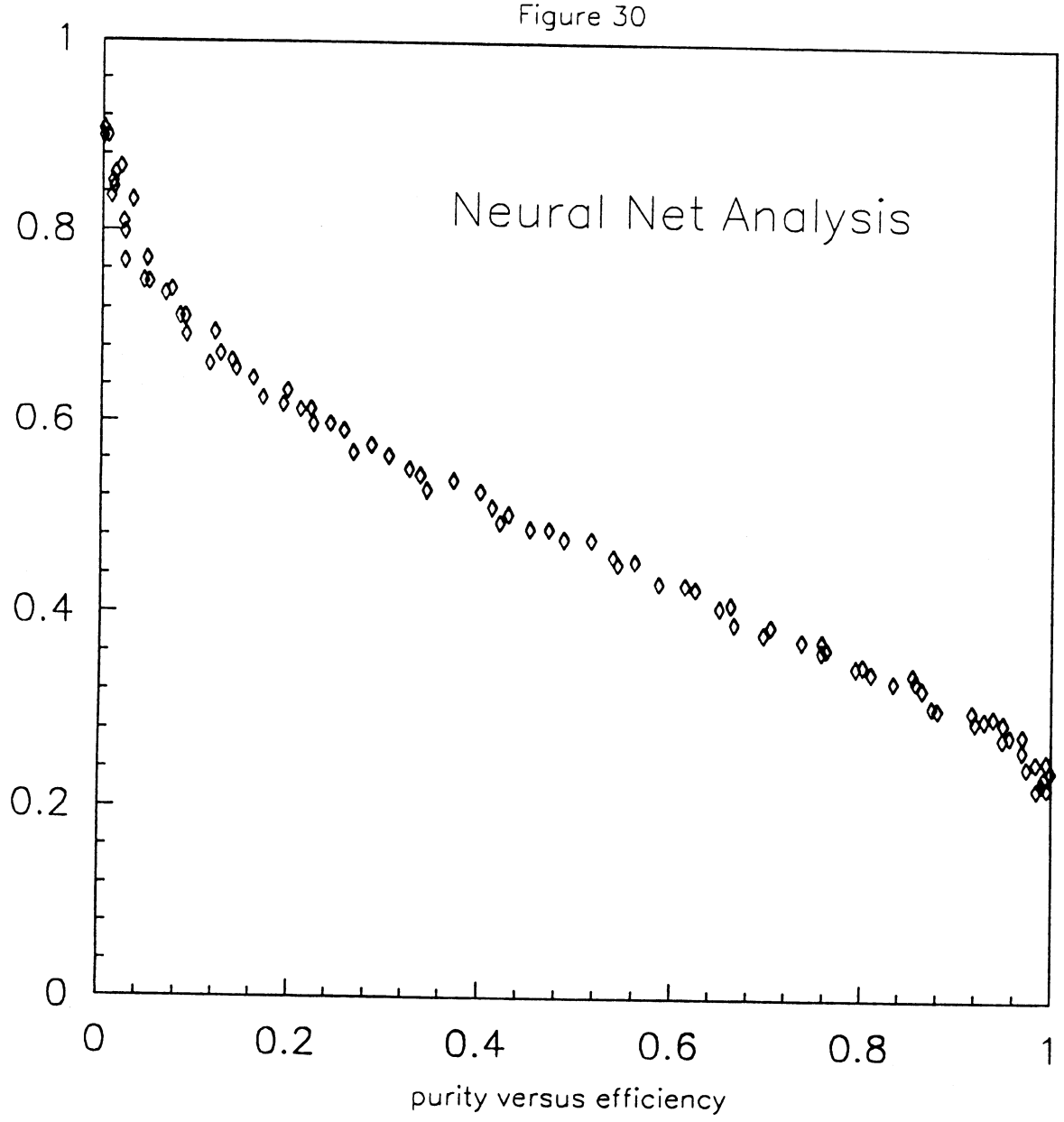

Neural Net Analysis
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
purity versus efficiency

Figure 31 a

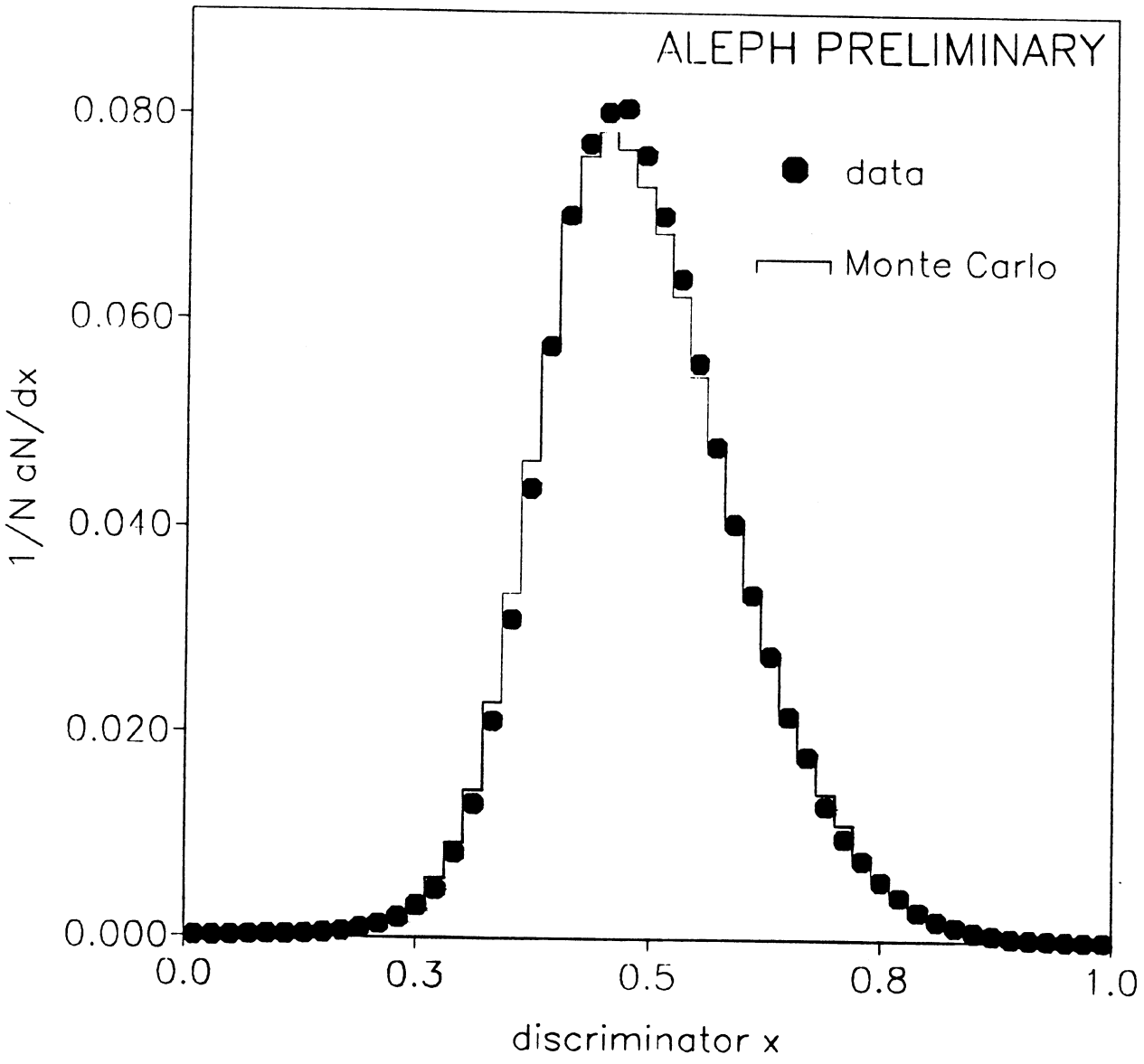

ALEPH PRELIMINARY
data
Monte Carlo
0.080
0.060
0.040
0.020
0.000
1/N dN/dx
0.0
0.3
0.5
0.8
1.0
discriminator x

Figure 31 b

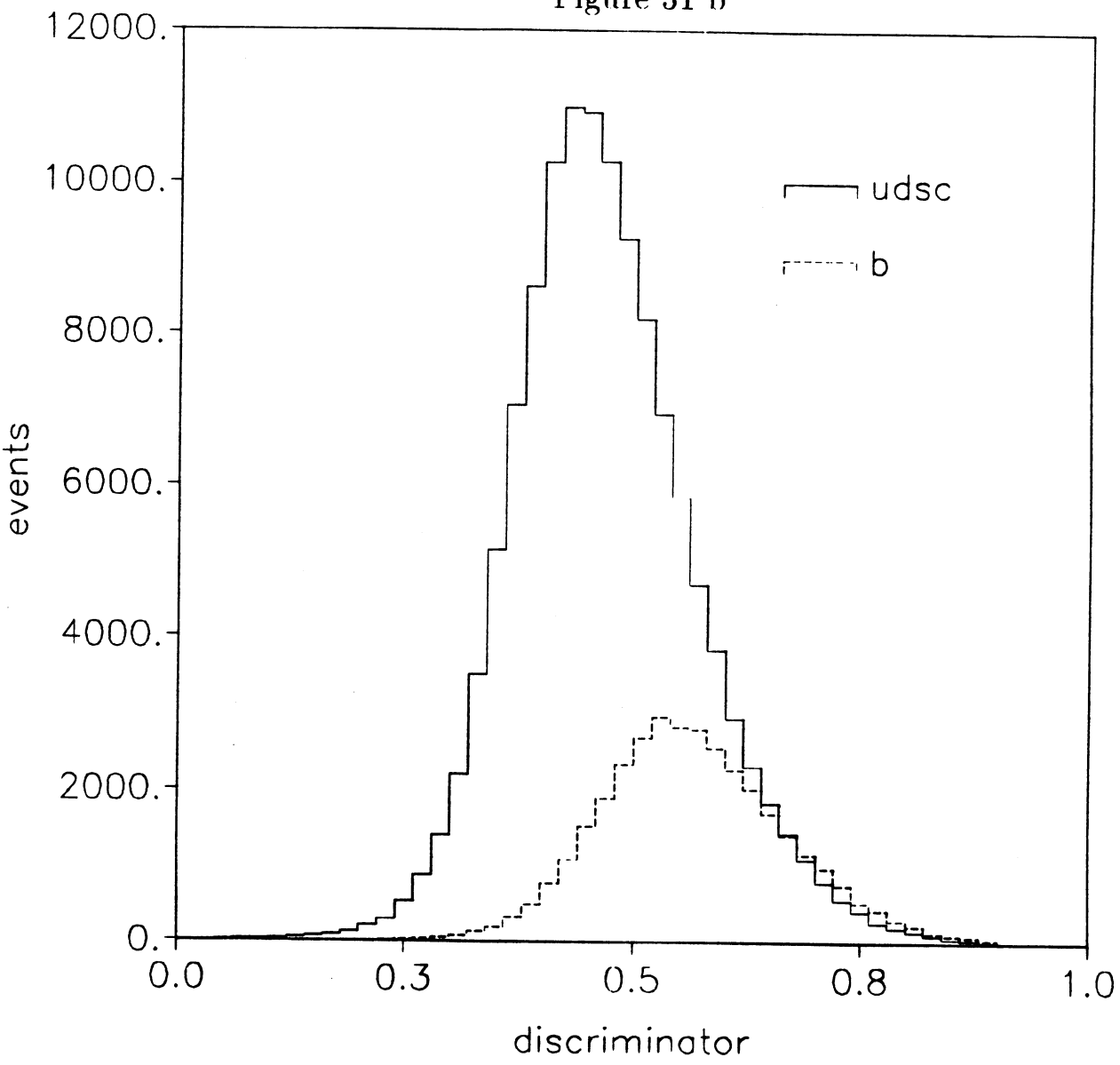

12000.
10000.
8000.
6000.
4000.
2000.
0.
events
udsc
b
0.0
0.3
0.5
0.8
1.0
discriminator

Figure 32

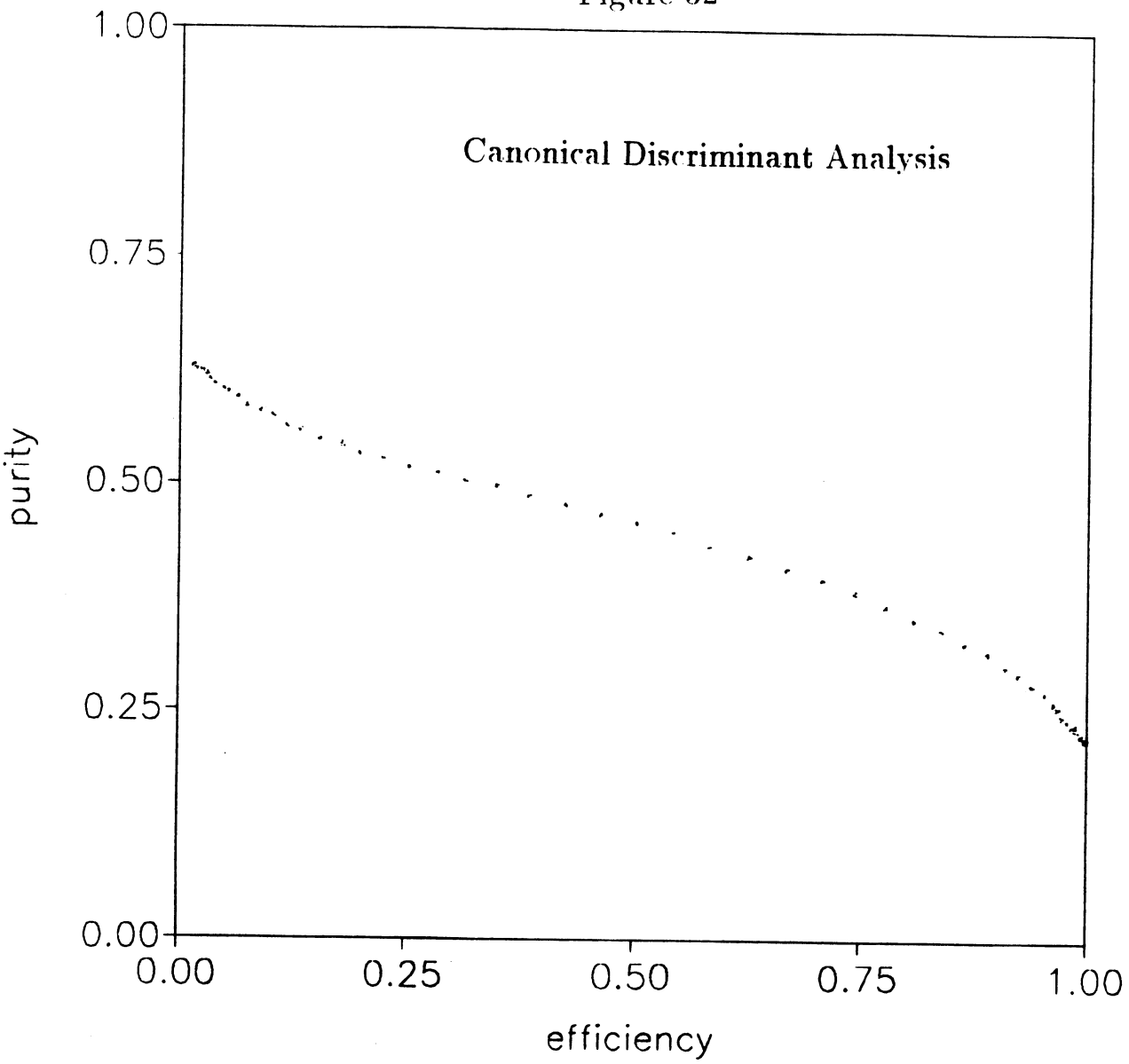
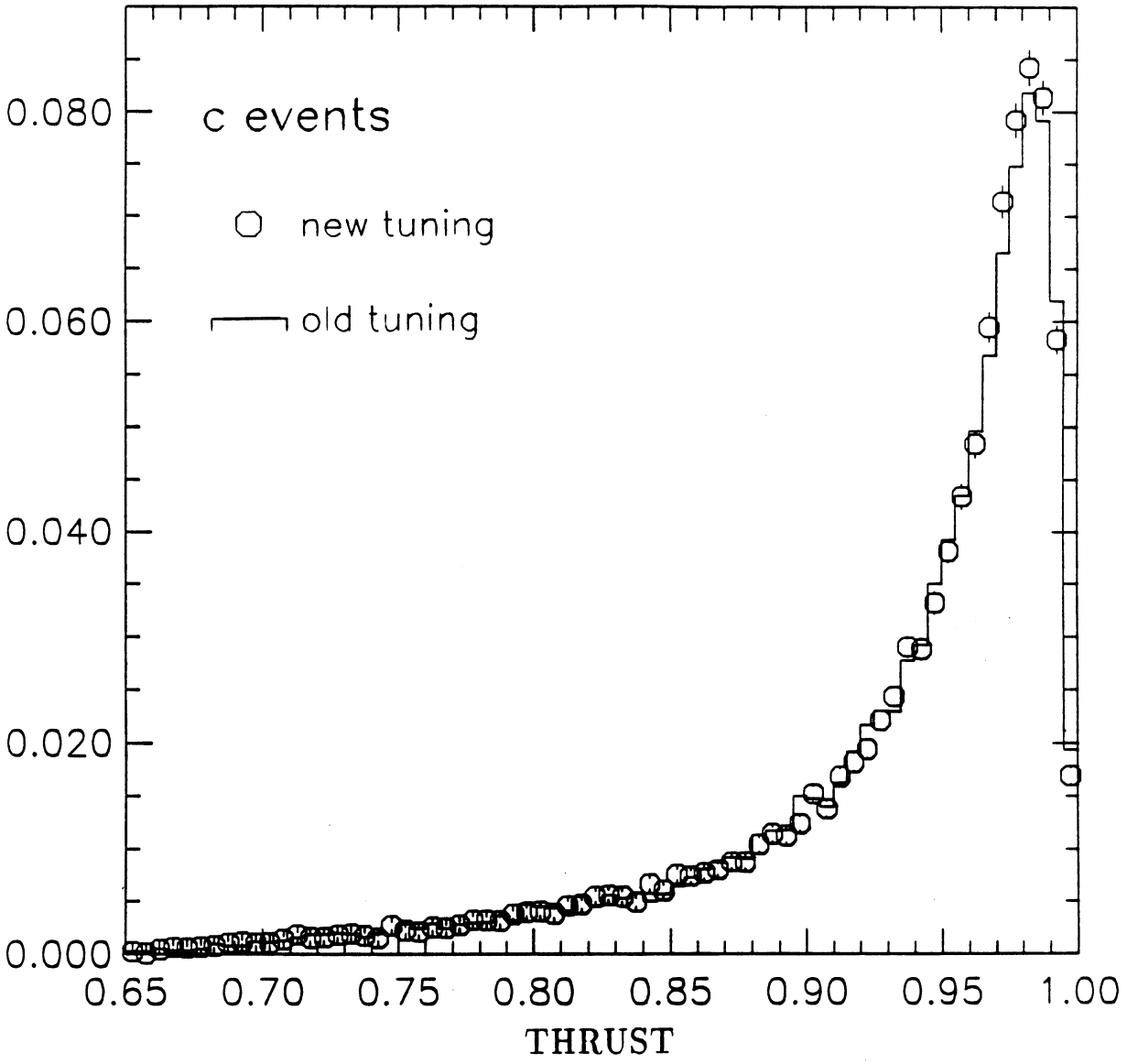

Canonical Discriminant Analysis
1.00
0.75
0.50
0.25
0.00
purity
0.00
0.25
0.50
0.75
1.00
efficiency

Figure 33


c events
new tuning
old tuning
0.080
0.060
0.040
0.020
0.000
0.65
0.70
0.75
0.80
0.85
0.90
0.95
1.00
THRUST