N.Armenise and G.Iaselli : A METHOD TO CLASSIFY NEUTRINO
EVENTS ACCORDING TO THEIR COMPLETENESS

A METHOD TO CLASSIFY NEUTRINO EVENTS ACCORDING TO THEIR COMPLETENESS

N. Armenise and G.Iaselli
Dipartimento di Fisica dell'Università di Bari, and INFN - Sezione di Bari

ABSTRACT

Complete neutrino events are separated from the total sample with a discriminant analysis in a many-fold space described in the text. Two new powerfull variables are found which discriminate with high efficiency and tag the event type as far as the completeness is concerned.

Introduction

Energy correction for neutrino(antineutrino)experiments has always been a major problem.Many methods have been employed to take care of loss of neutral energy in these interactions.

All of them act mainly on the longitudinal and trasverse momentum balance between the hadronic shower and the muon produced in the interaction.

Their efficiency has never been fully tested,and although the average statistical result may be promising,they introduce systematic biases by over-estimating the correction for certain classes of events and under-estimating it for others.

Moreover,the non selective application of these methods to the total sample results in the 'correction' of complete events,which may be present in a non neglibible ammount,and which may be distorted by this operation.

Alternatively, we propose a method to discriminate quasi-complete events from the total sample and classify the non completeness of the remaining events.

This method relys on the selection power of two new variables, and on employing a statistical discriminant technique to separate classes of neutrino interactions in a many-fold space defined by kinematical and dynamical variables.

We apply this method to a sample of Monte Carlo generated antineutrino- neon interactions. The Monte Carlo was designed for the Cern Bubble Chamber neutrino and antineutrino experiment WA59 [1].

Preliminary Classification

Monte Carlo events are used to construct two separate data banks corresponding to true quantities,which are the ones that would be measured in an ideal perfect detector,and visible

quantities,which are obtained after the event has been filtered through the performance of the simulated bubble chamber.

We define

$$\text{Etrue} = P_{\parallel\mu} + P_{\parallel h}^{T}$$

$$\text{Evis} = P_{\parallel\mu} + P_{\parallel h}^{V}$$

where 'T' stands for 'true','V' for 'visible' ,'h' for 'hadronic',and '$\parallel$' represents the projection along the neutrino direction.

Smearing effects are not included in the computation of Evis (they will be taken into account later),and therefore the observed muon longitudinal visible momentun is assumed to be equal to the corresponding true quantity.

If Z is the fraction of visible energy (Z=Evis/Etrue) then

$$P_{\parallel\mu} + P_{\parallel h}^{V} = Z(P_{\parallel\mu} + P_{\parallel h}^{T})$$

which gives

$$P_{\parallel h}^{T} = P_{\parallel h}^{V}/Z + P_{\parallel\mu}(1-Z)/Z$$

and therefore

$$P_{\parallel h}^{T} > P_{\parallel\mu}(1-Z)/Z \qquad (1)$$

holds for each event with $P_{\parallel h}^{V} > 0$.

If we now apply a correction C to the longitudinal visible hadronic momentum $(P_{\parallel h}^{C} = CP_{\parallel h}^{V})$,in the assumption of a good correction $(P_{\parallel h}^{C} \sim P_{\parallel h}^{T})$ equation (1) may be trasformed into

$$C > (P_{\parallel\mu}/P_{\parallel h}^{V})(1-Z)/Z \qquad (2)$$

The quantity $(P_{\parallel\mu}/P_{\parallel h}^{V})(1-Z)/Z = Clim$ can be made calculable in the real case of events for which Etrue is not known;in fact,if the assumption $P_{\parallel h}^{C} = P_{\parallel h}^{T}$ is valid,we expect Ecorr = Etrue and therefore,defining Zc = Evis/Ecorr,condition (2) can be rewritten

as

$$C > (P_{\mu} /P_{\mu}^{v})(1-Zc)/Zc = XClim$$

In the following R and Rc will denote respectively $P_{\mu k}^{v} /P_{\mu k}^{T}$ and $P_{\mu k}^{c} /P_{\mu k}^{T}$,where $P_{\mu k}^{c}$ is computed according to the Heilmann method [2] which has been commonly used in neutrino(antineutrino)experiment.

Figure 1 shows R versus XClim.A very strong correlation can be seen between the completeness of an event and its XClim value.
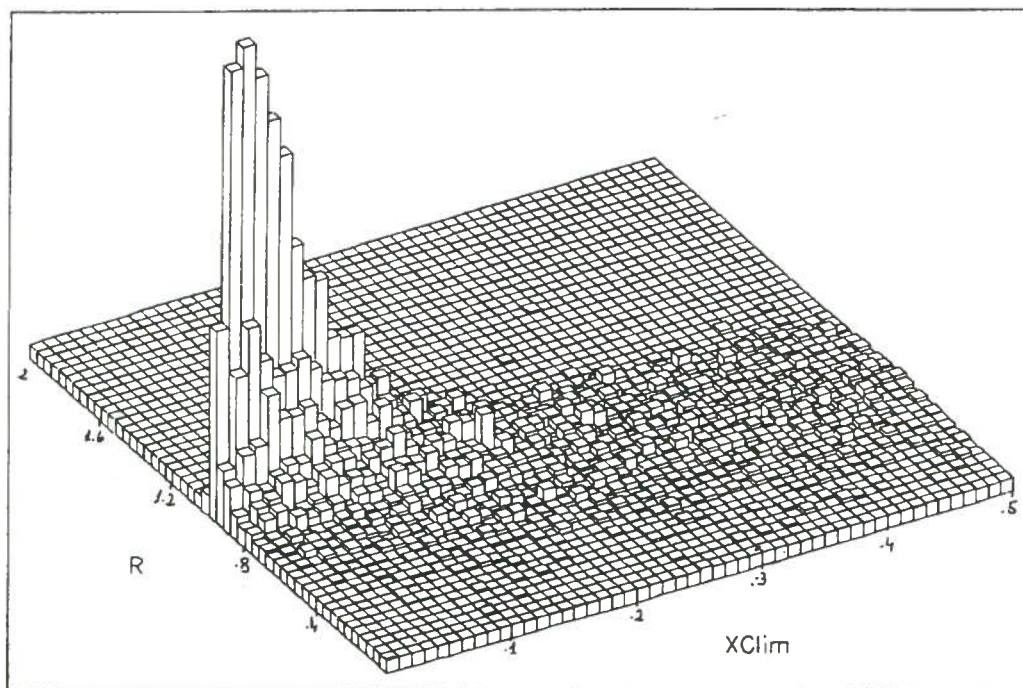

Fig. 1 - Lego plot of R versus XClim.

So this new variable selects a category of quasi-complete events (XClim very low) and classifies the remaining events in bands of increasing non-completeness .

Although XClim is built up from the correction factor C it is ,
nevertheless,more sensible then the correction itself for the

purpose of classification.

Moreover figure 2 shows the ratio R plotted versus $P_{\perp_\mu}/P_{/\!\!/_h}^{\vee}$ (=Rt).
It is evident a large eccess of well balanced events in the
region $P_{\perp_\mu} < P_{/\!\!/_h}^{\vee}$,i.e. Rt < 1.

We can therefore attempt to isolate a sample of quasi-complete
events by asking that XClim < .1 and Rt < .8 . More then 35% of
event falls in this category. Figures 3a and 3b show their R and
Rc distributions.   Thus we see that is more accurate to retain
these events as they are rather then to correct them with the
Heilmann method.

In fact,although the mean value of the Rc distribution is very
close to 1, its r.m.s. is quite large and a large fraction of
the events fall in the region Rc > 1.

Discriminant Analysis

From the previous analysis we learnt that it is possible to
classify neutrino events in the two dimensional space of XClim
and Rt.

The encouraging results so far obtained suggest that it might be
worthwhile   attempting   a   discriminant   analysis in an n-
dimensional space defined by all the measurable kinematical and
dynamical variables of the interaction.

Transverse and   longitudinal   muon   momentum,transverse   and
longitudinal   hadronic   momentum,charged   and   neutral
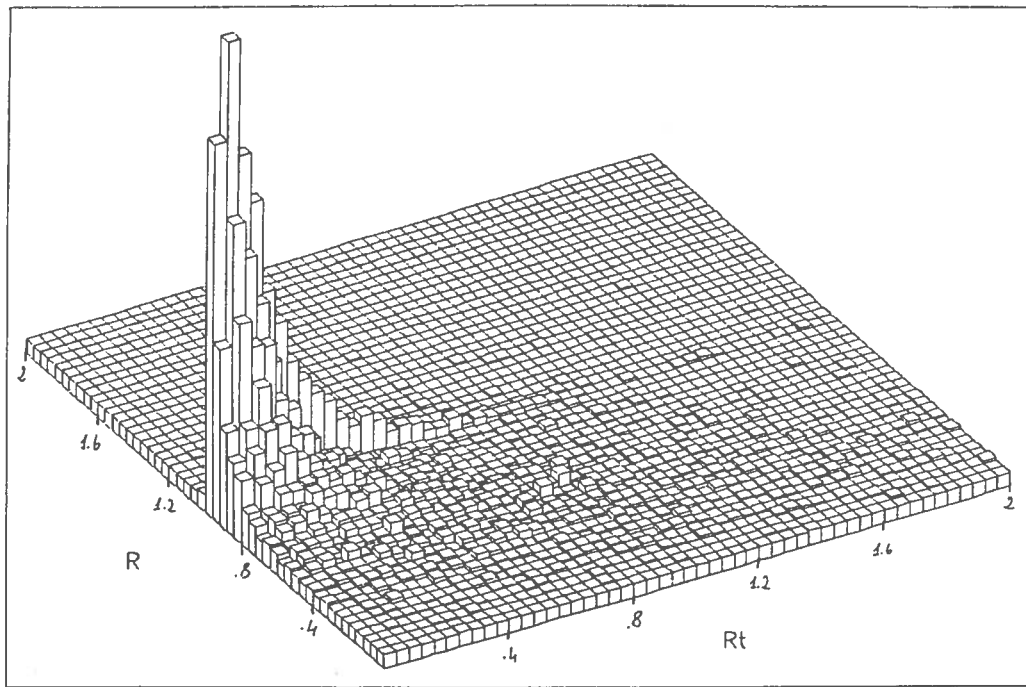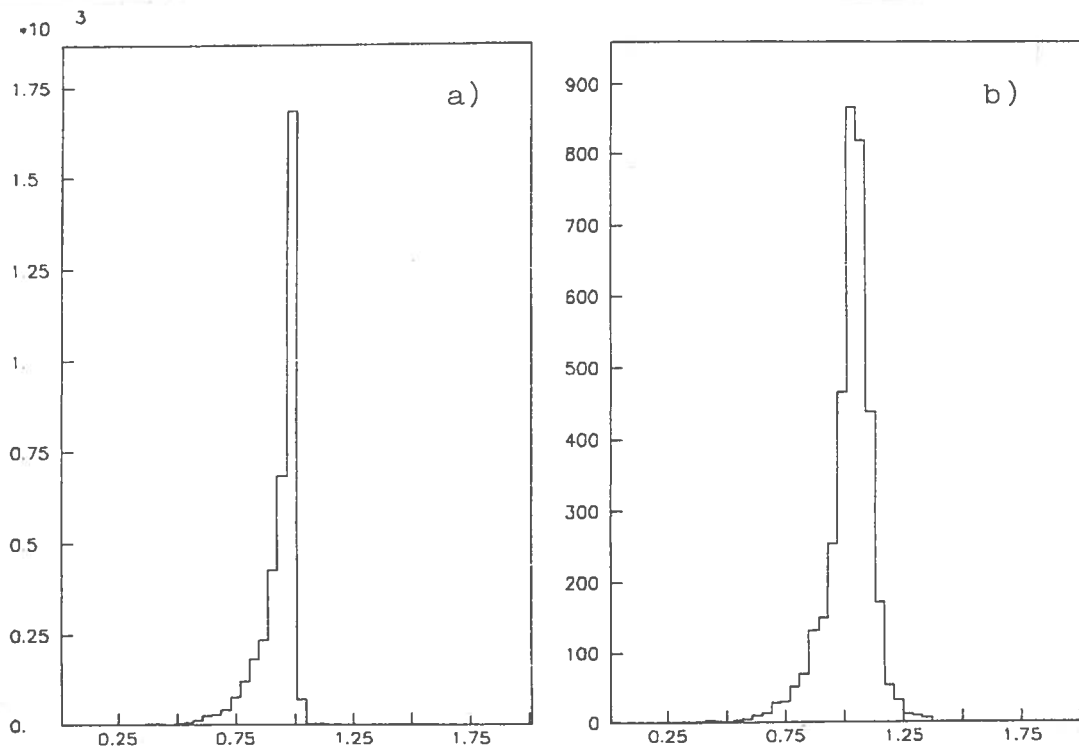
Fig. 2 - Lego plot of R versus Rt.

Fig. 3 - R and Rc distributions (Fig. a and b) for events with
$\overline{\text{XClim}}$ < .1 and Rt < .8.

energy,corrected hadronic momentum ,number of charged and neutral prongs,XClim and Rt themselves,may all enter in the definition of this many-fold space.

To perform such an analysis we use the SPSS statistical package[3]. An event is a-priori defined as complete if R > .97 .Events with R > 1 are present in a negligible fraction (smearing effect are not taken into account yet),and therefore they do not baise the results of the analysis.

SPSS is then asked to classify complete events in the n-dimensional space defined above,and to give the efficiency of the discrimination.

The best linear combination of the n variables is chosen ,such that in the space defined by this new variable,events are classified according to their completeness.

For each event we can compute a score Fd,and the discrimination of quasi-complete events is obtained by asking that this score has a value below a suitable lower limit,that we fix at −.2 . This operation leaves us with 26% of the initial events.

Figure 4 shows R versus Fd for the total sample of events. Although the efficiency of the selection is very high,not all the complete events are discriminated.Some of them migrate towards the high score region and escape our selection.A second SPSS iteration on the remaining events enable us to discriminate a further 2% of quasi-complete events.We therefore end up with 28% of the initial sample classified as quasi-complete.Figure 5a-5b show R and Rc distributions for this quasi complete sample.

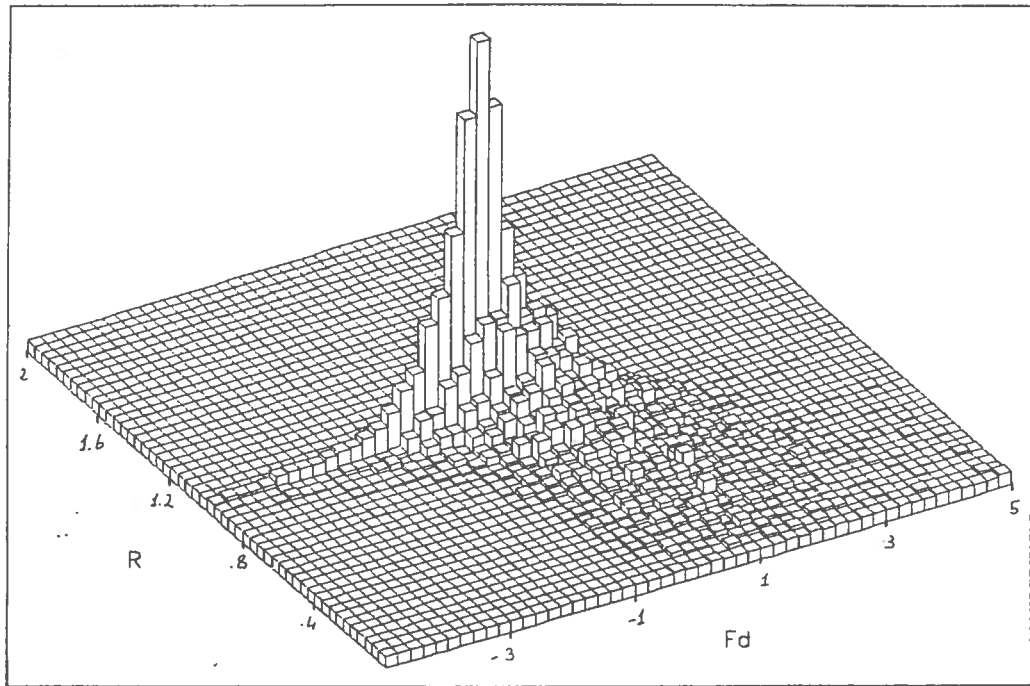Again it is evident how powerfull this method is when comparing
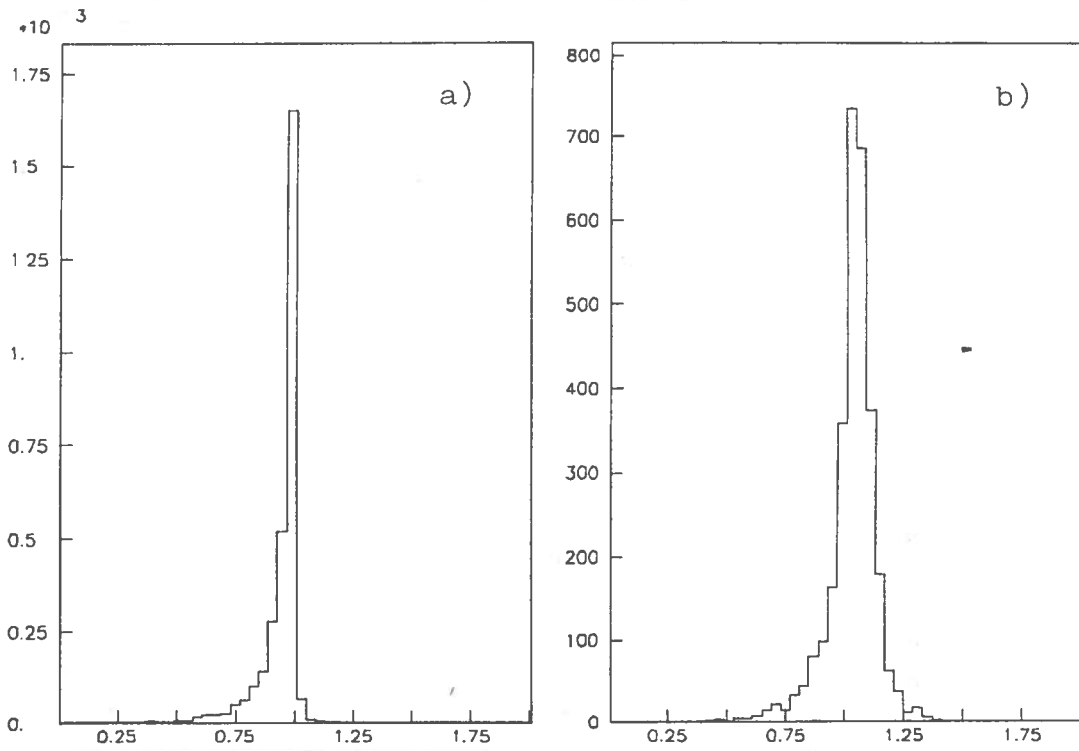
Fig. 4 - Lego plot of R versus Fd.

Fig. 5 - R and Rc distributions (Fig. a and b) for events with Fd < -.2.

it with the Heilmann correction,which over-estimates the energy
of a large fraction of the events.

The discriminant analysis confirms that the classification of
neutrino interaction is possible. Moreover the distribution of
XClim for the events not falling in the quasi-complete sample,has
its most probable value just beyond .1 confirming that our
earlier selection in XClim (XClim < .1) was opportune.

Treatment of non complete events

We are now left with a sample of non complete events.This
category is not homogeneous and contains events with different
degrees of completeness.

In the following only this sample will be considered.A
preliminary classification into subsamples may be achieved in
terms of the variable Rt.Events very much incomplete are mainly
at high values of Rt,as shown in figures 6a-6b,where the
distributions of R for events with respectively Rt < .8 and Rt >
.8 are given. The relative contributions to these subsamples is
4 to 1.

While XClim proves to be very powerfull in discriminating
complete events,Rt is as good in separating very non complete
events.

A further classification may be obtained for the events with Rt
<.8, whose Fd score may also help to discriminate bands of non
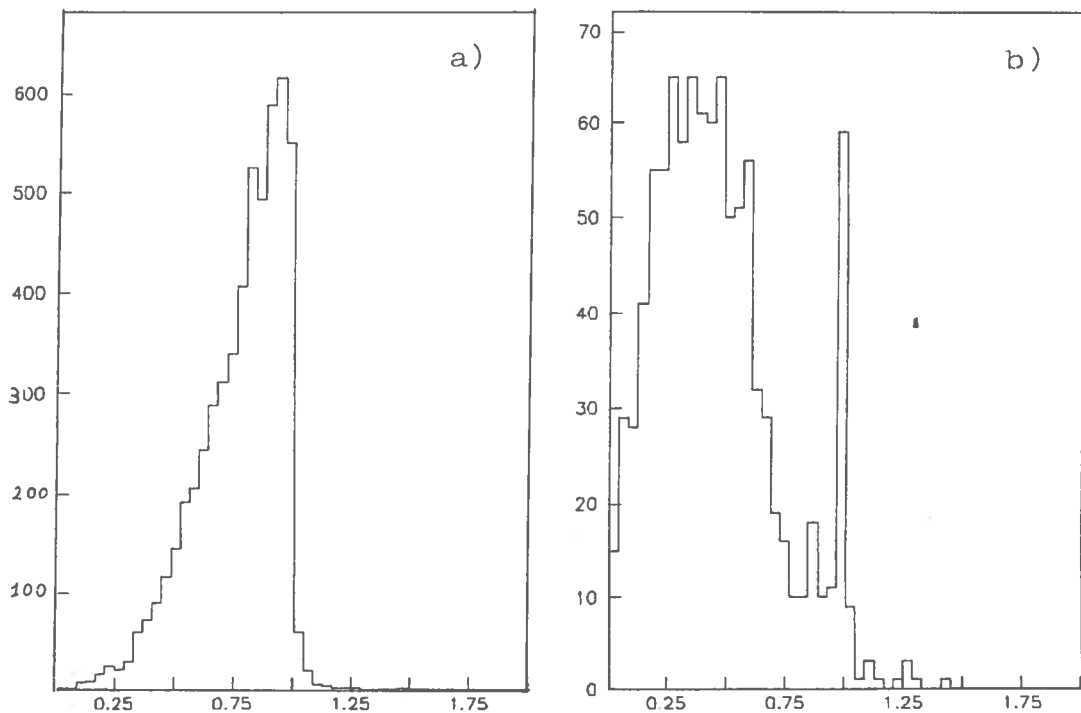completeness.We consider events in the three different windows

Fig. 6 - R distributions for non-complete events (as defined in the text) with Rt < .8 (Fig. a) and Rt > .8 (Fig. b).

−.2 < Fd < .4 , .4 < Fd < 1.2 , Fd > 1.2 , and for each of them we plot R (figures 7a,7b,7c).The maximum of the distribution approaches 1 as Fd get smaller .

This suggests that a correction depending on Fd could be constructed for this class of events.

One can also study the efficiency of the Heilmann method in correcting these incomplete events.Again the discriminant analysis may be employed.Events are divided into three bins of the variable Rc : Rc < .95 , Rc > 1.04 and .95 <Rc< 1.04 . The discrimination of this last subsample (good Heilmann corrected events) is attempted.

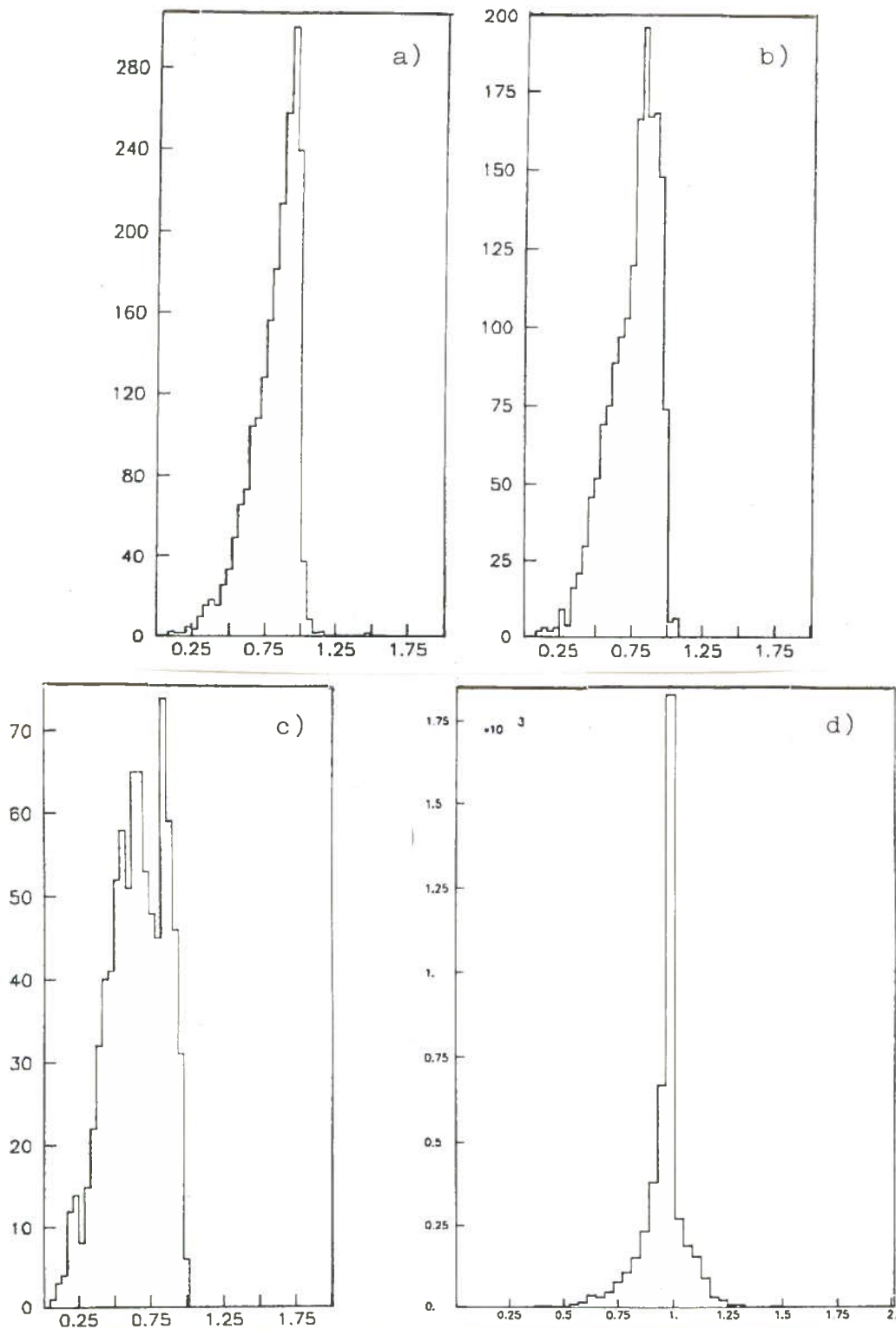SPSS finds two new score functions Fdx,Fdy,in term of which this

Fig. 7 - R distributions of non-complete events (Fig. a,b,c) in different bands of Fd (as defined in the text), and Rc distribution (Fig. d) for events selected or reconstructed as complete (Rc is set equal to R for quasi-complete events).

discrimination may be achieved,selecting an appropriate window in the two-dimensional space defined by Fdx and Fdy.

About 13% of the total sample (20% of the non-complete sample) falls into this new category of non-complete events for which the Heilmann correction works successfully.This sample may now be added to the quasi-complete one.We therefore end up with 41% of the events selected,or reconstructed,as complete.Their Rc distribution (Rc is set equal to R for quasi-complete events) is shown in figure 7d.

Effect of introducing the smearing

So far no smearing has been assumed in our sample of Monte Carlo events. We now introduce smearing effects to study their influence on the potentiality of the method.

Measurament errors on the longitudinal momentum for the hadrons and for the muons are introduced with gaussian low.

Figures 8a,8b show the distributions of R and Rc obtained using smeared events,after the selection in Fd has been done. The same cut defined previously has been used for Fd. About 25% of the initial sample falls now in this category (to be compared to 28% obtained previously),and still the improvement relative to the Heilmann correction is evident.

We conclude therefore that the influence of the smearing on the classification of incomplete events in bands of Fd is not
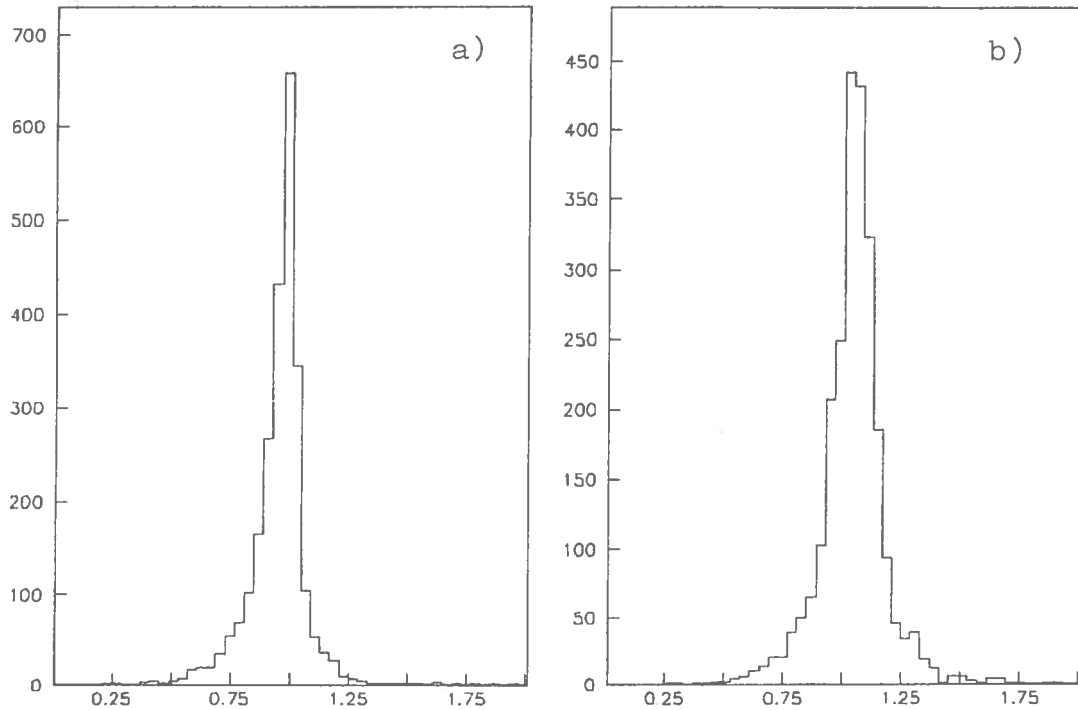
critical.

Fig. 8 - R and Rc distributions (Fig. a and b) for events with Fd < -.2, after introducing smearing effects.

Conclusion

The discriminant analysis proves to be a powerfull method to discriminate quasi-complete events and to classify the non-completeness of the remaining ones.

It has been applied to a Monte Carlo generated sample of antineutrino interactions,where, event by event, the degree of completeness was known a-priori.

In the real case of experimental neutrino (antineutrino) interactions the same information is not directly avaible,but a sample of quasi- complete events can be extracted using a selection on the variables XClim and Rt,as described in the paper.

A criteria to define the sensible value of XClim could be to assume that all the events below 2Max(XClim) are quasi-complete. This sample,therefore,could the imput to be given to the discriminant procedure.

Acknowledgments

References

[1] P.Kasper,WA59 Collabotation,internal note

[2] H.G.Heilmann ,Bonn Int.Note, WA-21-Intl (1978)

[3] SPSS,Statistical Package for the Social Sciences,
        McGraw-Hill Book Company, 1975