

ISTITUTO NAZIONALE DI FISICA NUCLEARE

IT8500279

Sezione di Bari

INFN/AE-84/10
2 Ottobre 1984

N.Armenise and G.Jaselli : 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 powerful 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 negligible amount, 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 relies 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

$$\begin{aligned} E_{true} &= P_{\mu} + P_{h}^T \\ E_{vis} &= P_{\mu}^V + P_{h}^V \end{aligned}$$

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

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

If Z is the fraction of visible energy ($Z = E_{vis}/E_{true}$) then

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

which gives

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

and therefore

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

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

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

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

The quantity $(P_{\mu} / P_{h}^V)(1-Z)/Z = C_{lim}$ can be made calculable in the real case of events for which E_{true} is not known; in fact, if the assumption $P_{h}^C = P_{h}^T$ is valid, we expect $E_{corr} = E_{true}$ and therefore, defining $Z_c = E_{vis}/E_{corr}$, condition (2) can be rewritten

as

$$C > (P_{\mu}^{\nu} / P_{\mu}^{\tau}) (1 - Z_c) / Z_c = X_{Clim}$$

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

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

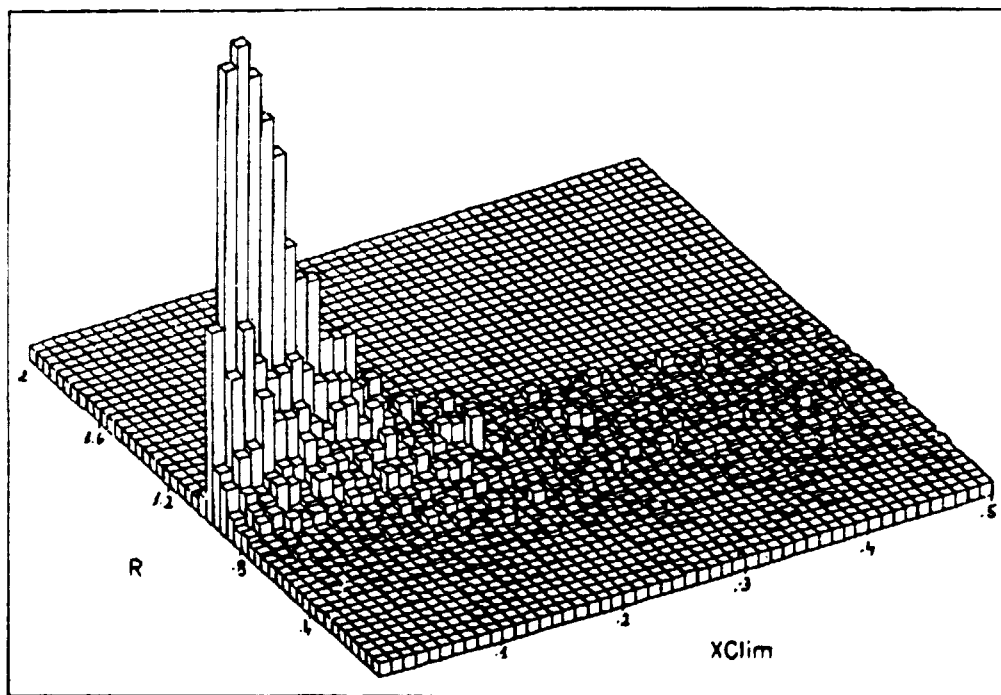


Fig. 1 - Lego plot of R versus X_{Clim}.

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

Although X_{Clim} is built up from the correction factor C it is , nevertheless, more sensible than the correction itself for the

purpose of classification.

Moreover figure 2 shows the ratio R plotted versus $P_{\perp}^{\nu} / P_{\perp}^h$ ($=R_t$).

It is evident a large excess of well balanced events in the region $P_{\perp}^{\nu} < P_{\perp}^h$, i.e. $R_t < 1$.

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

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

Discriminant Analysis

From the previous analysis we learnt that it is possible to classify neutrino events in the two dimensional space of X_{Clim} and R_t .

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

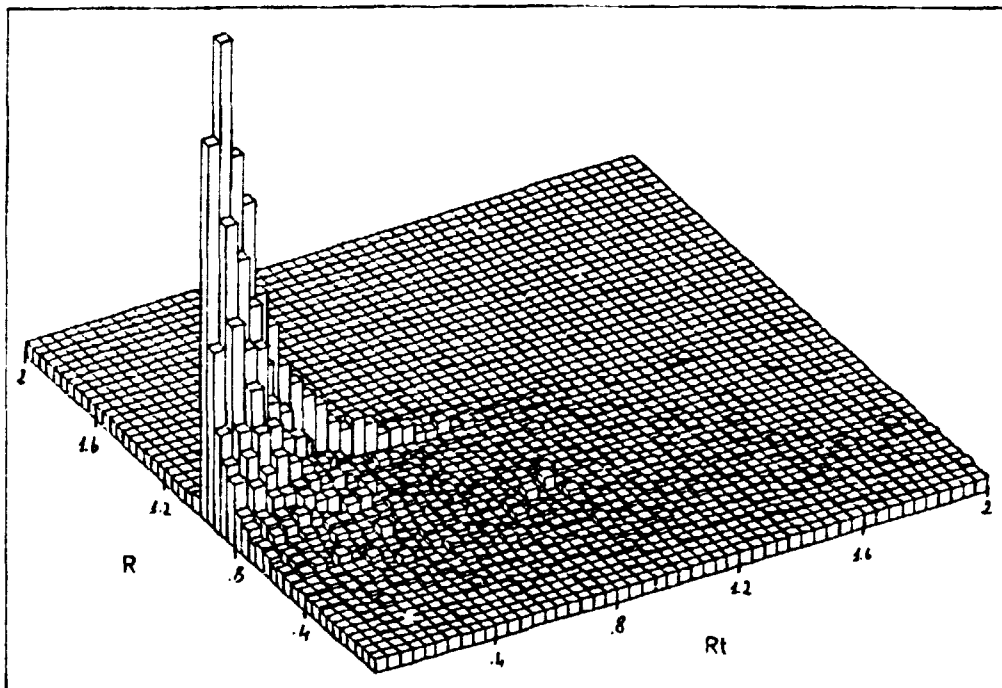


Fig. 2 - Lego plot of R versus Rt.

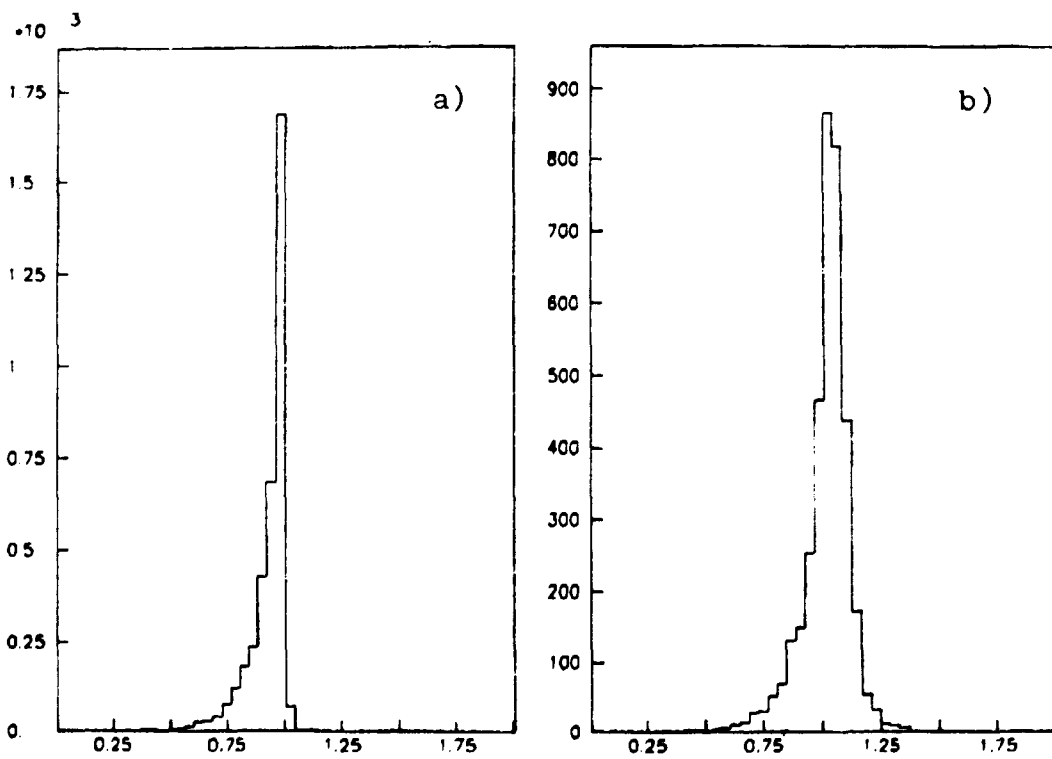


Fig. 3 - R and Rc distributions (Fig. a and b) for events with $\bar{X}_{Clim} < .1$ and $R_t < .8$.

energy, corrected hadronic momentum, number of charged and neutral prongs, X_{Clm} and R_t 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 bias 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 F_d , 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 F_d 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 R_c distributions for this quasi complete sample.

Again it is evident how powerful this method is when comparing

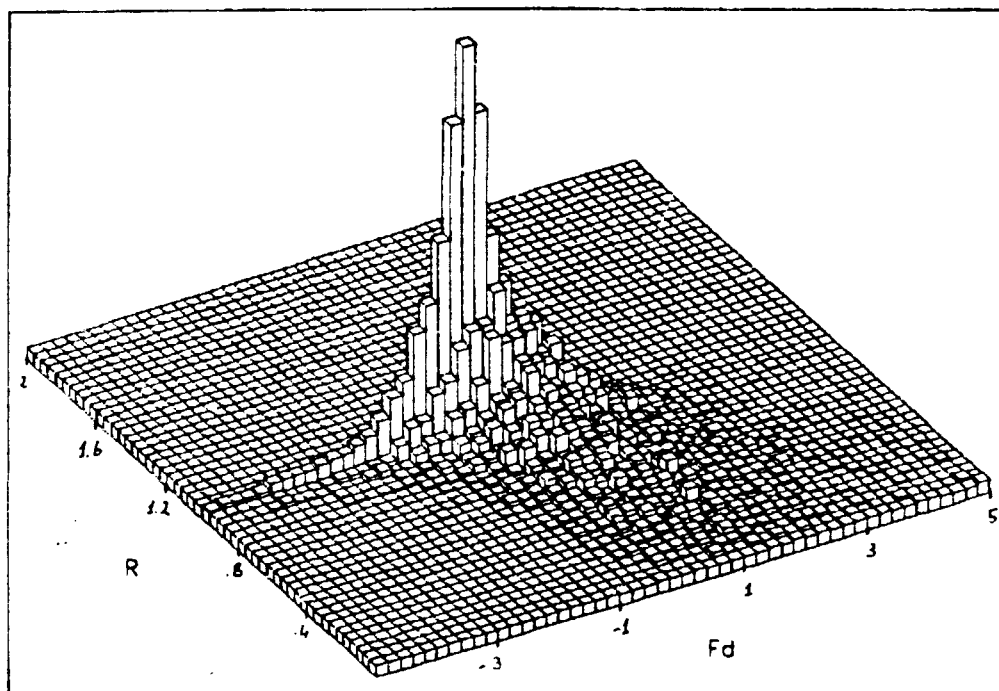


Fig. 4 - Lego plot of R versus Fd.

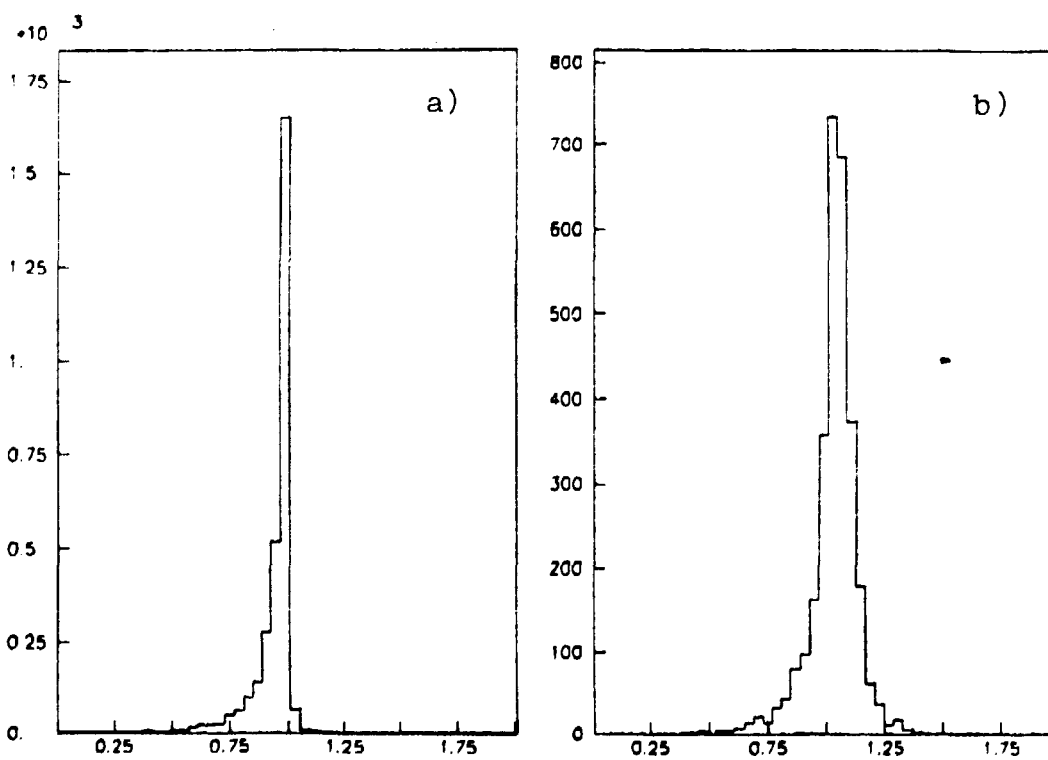


Fig. 5 - R and Rc distributions (Fig. a and b) for events with $Fd < -0.2$.

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 X_{Clim} for the events not falling in the quasi-complete sample, has its most probable value just beyond .1 confirming that our earlier selection in X_{Clim} ($X_{Clim} < .1$) 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 R_t . Events very much incomplete are mainly at high values of R_t , as shown in figures 6a-6b, where the distributions of R for events with respectively $R_t < .8$ and $R_t > .8$ are given. The relative contributions to these subsamples is 4 to 1.

While X_{Clim} proves to be very powerful in discriminating complete events, R_t is as good in separating very non complete events.

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

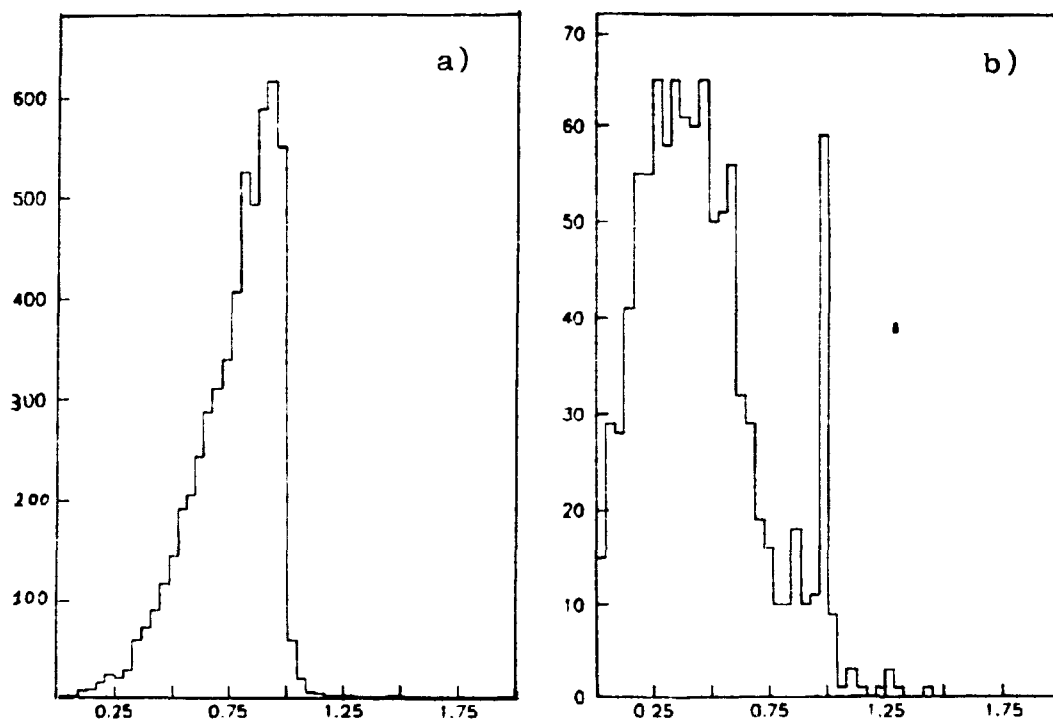


Fig. 6 - R distributions for non-complete events (as defined in the text) with $R_t < .8$ (Fig. a) and $R_t > .8$ (Fig. b).

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

This suggests that a correction depending on F_d 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 R_c : $R_c < .95$, $R_c > 1.04$ and $.95 < R_c < 1.04$. The discrimination of this last subsample (good Heilmann corrected events) is attempted.

SPSS finds two new score functions F_{dx} , F_{dy} ,in term of which this

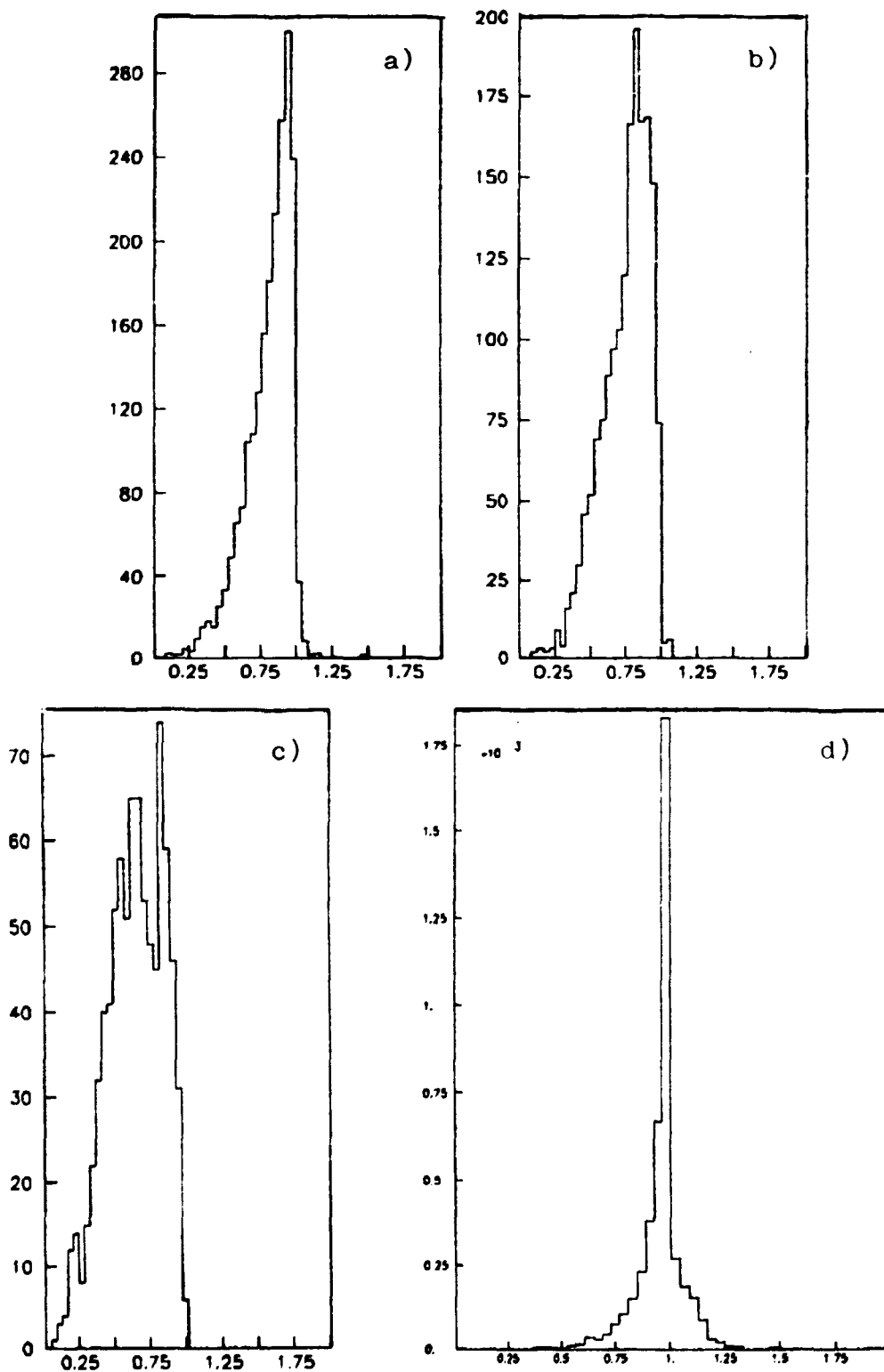


Fig. 7 - R distributions of non-complete events (Fig. a,b,c) in different bands of F_d (as defined in the text), and R_c distribution (Fig. d) for events selected or reconstructed as complete (R_c is set equal to R for quasi-complete events).

discrimination may be achieved, selecting an appropriate window in the two-dimensional space defined by F_{dx} and F_{dy} .

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 R_c distribution (R_c 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.

Measurement errors on the longitudinal momentum for the hadrons and for the muons are introduced with gaussian law.

Figures 8a, 8b show the distributions of R and R_c obtained using smeared events, after the selection in F_d has been done. The same cut defined previously has been used for F_d . 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 F_d is not

critical.

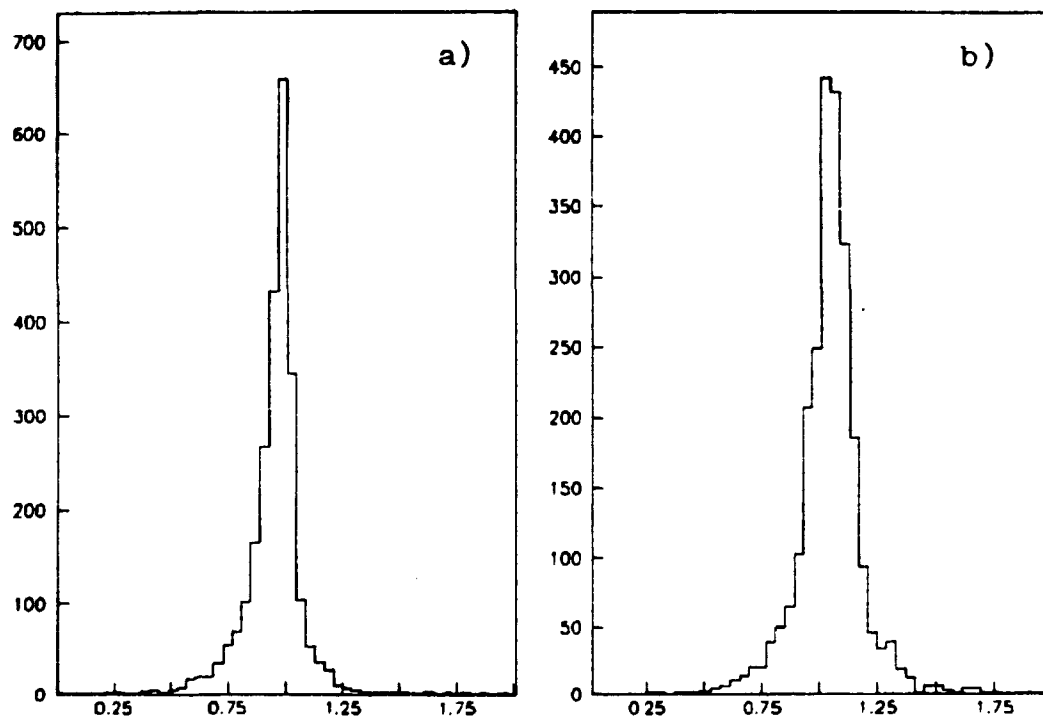
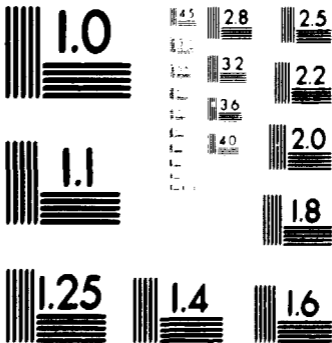


Fig. 8 - R and Rc distributions (Fig. a and b) for events with $F_d < -0.2$, after introducing smearing effects.

Conclusion

The discriminant analysis proves to be a powerful 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.



MICROCOPY RESOLUTION TEST CHART
 NATIONAL BUREAU OF STANDARDS
 STANDARD REFERENCE MATERIAL 1010a
 (ANSI and ISO TEST CHART No. 2)

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

A criteria to define the sensible value of X_{Clim} could be to assume that all the events below $2Max(X_{Clim})$ are quasi-complete. This sample, therefore, could be the input to be given to the discriminant procedure.

Acknowledgments

We wish to thank P.Kasper for making available his Monte Carlo and for many useful discussions about this manuscript.

References

- [1] P.Kasper, WA59 Collaboration, internal note
- [2] H.G.Heilmann, Bonn Int.Note, WA-21-Int1 (1978)
- [3] SPSS, Statistical Package for the Social Sciences,
McGraw-Hill Book Company, 1975