

Disclaimer

This note has not been internally reviewed by the DØ Collaboration. Results or plots contained in this note were only intended for internal documentation by the authors of the note and they are not approved as scientific results by either the authors or the DØ Collaboration. All approved scientific results of the DØ Collaboration have been published as internally reviewed Conference Notes or in peer reviewed journals.

Estimation of the QCD background to $W \rightarrow e\nu + jets$

Sailesh Chopra

Panjab University, Chandigarh, India

Rajendran Raja

*Fermi National Accelerator Laboratory,
Batavia, Illinois, 60510*

(April 14, 1994)

© 1994

Abstract

We introduce a new technique to estimate the QCD contribution to the background under $W \rightarrow e + jets$ channel. The method assumes that the E_T of QCD fakes is predominantly due to measurement errors. We demonstrate that this assumption is obeyed remarkably well in DØ. We use this technique to estimate the background due to QCD fakes simultaneously as a function of two variables, namely E_T and the transverse mass. We demonstrate this on a sample of events triggering on the ELE_HIGH trigger. The technique enables us to study the behavior of the background and the signal as a function of the two variables for a variety of requirements on the quality of the electron. Questions of relative trigger efficiency of the background and signal samples do not arise. The method also generates background events that can be used as input to topological searches of the top quark.

I. INTRODUCTION

QCD events can turn up as background to the $W \rightarrow e + \text{Jets}$ channel in one of three ways. Heavy quarks ($b\bar{b}, c\bar{c}$) can produce genuine electrons and a modest amount of \cancel{E}_T . QCD jets can also fake an electron by fragmenting in a highly electro-magnetic mode that can deceive the electron ID algorithm into accepting π^\pm/π^0 overlaps as genuine electrons. Finally, isolated π^0 's can convert early in the central tracker and produce electrons. DØ has in the past used several techniques [1–3] to address the problem of calculating the QCD fake backgrounds to $W + \text{jets}$ data. The problem is a fundamental one in as much as that it affects not only the determination of W and Z cross sections and masses but also the search for the top quark in the electron+jets channel. In the first method [1], an attempt is made to estimate the rate at which QCD jets fake electrons. Electrons are defined to be fakes if they occur back to back with a QCD jet in the transverse direction. The fake rate has been found to depend on event topology. In order to estimate the background to $W+4\text{Jets}$, QCD 5 jet triggers are taken. If one of the jets is identified as an electron, it is eliminated from consideration, since it might well occur in the real $W+\text{Jets}$ channel. Only those events with high EM fraction jets and \cancel{E}_T are considered further and convoluted with a probability to fake an electron. This method suffers from statistical as well as systematical difficulties. It is difficult to estimate the relative efficiency between the signal and the background samples. Also, variation in fake rate with event topology is difficult to parametrize correctly and use effectively. The other two extant methods [2] [3] fit a background curve to the \cancel{E}_T distribution for events with electron triggers. The parametrized background is used to yield a QCD contamination rate as a function of jet multiplicity and \cancel{E}_T . The statistical accuracy of this method becomes limiting when high multiplicities are considered. It is capable of estimating the background in one variable only, in this case \cancel{E}_T . Simultaneous dependence of the contamination rate in \cancel{E}_T and transverse mass cannot be calculated. More important still, none of the existing methods yield QCD background event samples that can be used in topological searches for the top quark.

The method outlined in this D \emptyset note, we believe addresses and to a large extent solves all these problems.

II. METHOD

We select the events from the electron Level 2 filter stream, ELF, where we have required that the electron satisfy the ELE_HIGH trigger. Figure (1) shows the \cancel{E}_T distribution of these events.

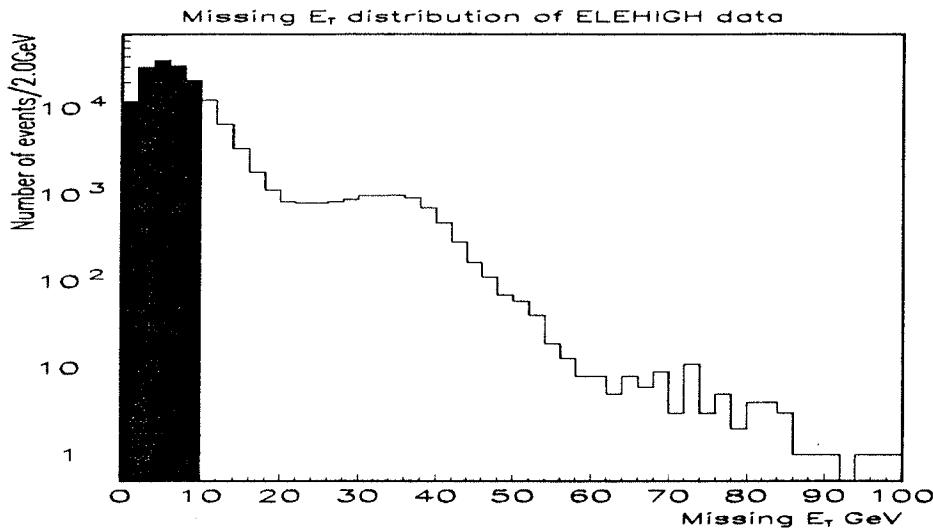


FIG. 1. \cancel{E}_T distribution of events. Shaded region represents events selected for smearing

There are 166379 events in our sample. The only requirement on the electron is that a PELC bank is created. i.e. no H-matrix chisquared cuts. The cluster has to be electromagnetic and there has to be a track in the road. We note that the peak at low \cancel{E}_T is due to the QCD background. The electron due to the QCD background may be real ($b\bar{b}$) or it may be fake, with a jet faking an electron via π^\pm/π^0 overlap. We now hypothesize that the QCD fake events have little genuine \cancel{E}_T and that their \cancel{E}_T is due to the measurement errors introduced by the calorimeter. To test this hypothesis, we select events with \cancel{E}_T below 10 GeV from our sample. This is the shaded region in our plot. There are 130820 events in this low \cancel{E}_T sample. We now zero the \cancel{E}_T of these events by distributing the \cancel{E}_T among

the measured objects (electron, jets and the rest of the event) according to the resolutions of the measurements as per an algorithm described below. We then proceed to treat these zero E_T events as though they have not been measured at all and proceed to smear them by the expected DØ resolutions. Each event is smeared 50 times. We select those events that have a smeared E_T greater than 15 GeV for further analysis. In principle, we are allowed to re-use each event as many times as we desire, as long as the statistical significance of the number of events in our selection does not exceed that of the input sample. Using the high statistics thus generated, we can proceed to examine the QCD contamination rate as a function of the variables in the problem. We note that there are two principal variables in the problem, namely E_T and the transverse mass. Since we independently measure the electron E_T and the P_T of the rest of the event, E_T and the transverse mass are two independent though correlated variables. Using the technique described above, we are for the first time able to examine the QCD fake rate in detail as a function of E_T , transverse mass and jet multiplicity for various electron quality requirements.

A. Detector resolution

We estimate the ability of the detector to measure jet energies and directions by Monte Carlo. We compare the directions of parton jets (PJETS) to RECO jets and thus estimate the energy, η and ϕ resolutions as a function of the energy of the jet. Table (I) gives the resolutions thus estimated for the jets as well as for the electron. In all that follows, we consider .5 cone jets. We assume that the p_T of the “rest of the event”, namely, the energy that is left over after the jets and the electrons are subtracted from the event is measured with the same resolution as the jets. This can be improved with further study.

B. Zero E_T algorithm

We treat the energy resolution of each object in the event (electron, jet or rest of the event as a 2 vector with the same direction as the transverse momentum of the object.

This yields resolution components of the object in question along the x and y axes. We then distribute the x and y \not{E}_T components of the event among the objects in proportion to the resolution component of that object along that axis. This will then result in an event with zero \not{E}_T . Objects such as electrons that are well measured will absorb less of the \not{E}_T components than the objects that are poorly measured such as high E_T jets. The algorithm is stated mathematically as follows.

$$\delta E_T^x = -\frac{\sigma^x}{\sum_{\text{objects}} \sigma^x} \not{E}_T^x \quad (2.1)$$

where δE_T^x is the change in the x component of the energy of an object, \not{E}_T^x is the component of the \not{E}_T along the x axis, and the energy resolution component along the x axis is defined by

$$\sigma^x = \sigma(E) \frac{E^x}{E_T} \quad (2.2)$$

Similar equations hold for the change in the y component of the energy.

C. Proof that the method works

Figure (2) shows a plot of the \not{E}_T of the events thus generated as a result of our smearing, superimposed on the data.

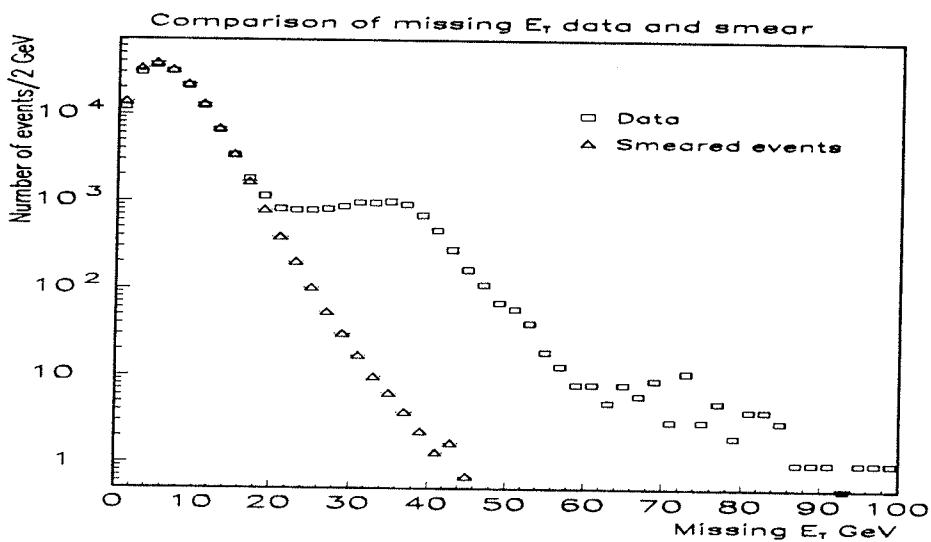


FIG. 2. Comparison of E_T distribution for data and smeared events normalized to data in the range 10 to 20 GeV

It can be seen that in the background region, the agreement is remarkably good. We have adjusted one of the parameters of the resolutions, namely the constant term in the jet resolution from .1848 to .0924 to get slightly better agreement with data. We feel this is justifiable since we are using events that have been already measured and using them (after zeroing the E_T) as though they were unmeasured. This should be eventually done by a global fit of all the resolution parameters to the data to reproduce the E_T distribution of the data. In figure (2), we have normalized the smeared events to data in the E_T range 8.0 GeV to 18.0 GeV. The normalization factor thus determined to apply to the smeared sample is 0.0256702, which is very close to number of data events divided by 50 times the number of smeared events. Figure (3) is the transverse mass of the smeared events superimposed on data. We have normalized the transverse mass distributions in the range $30 \text{ GeV}/c^2$ to $36 \text{ GeV}/c^2$. The normalization factor we determine is .0299846 .

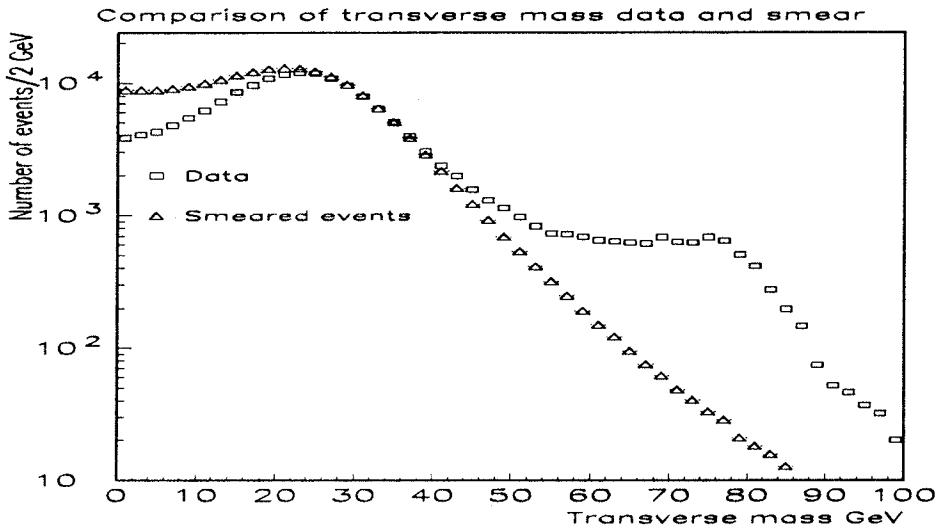


FIG. 3. Comparison of transverse mass distribution for data and smeared events normalized to data in the range 30 to $36 \text{ GeV}/c^2$

The reason why the E_T and the transverse mass normalization factors are not the same is because there is disagreement between the smeared distribution and data in transverse

mass for transverse masses less than $20 \text{ GeV}/c^2$. This is due to the fact that for such low transverse masses, there is a residual correlation between \cancel{E}_T and the electron E_T which will be present in data and not in the smeared events, i.e. there is genuine \cancel{E}_T at low transverse mass that is correlated with the electron E_T direction. However, we can see that the fall off in the transverse mass (m_t) distribution for $m_t > 30.0 \text{ GeV}/c^2$ is well reproduced by the smeared events. The difference between the two normalization constants gives us a measure of the systematic error in the method. For $m_t > 30.0 \text{ GeV}/c^2$, we quote a systematic error in our estimates of 17 %. We would tend to systematically underestimate the contamination due to QCD by a factor which can be on average 17 % if we use the \cancel{E}_T normalization factor for $m_t > 30.0 \text{ GeV}/c^2$. Similarly, we would tend to overestimate the the QCD contamination for $m_t < 30.0 \text{ GeV}/c^2$ systematically. *We should note that since we are normalized in \cancel{E}_T correctly, the systematics only arise when examining differential distributions in transverse mass.*

D. Further cross-checks

We define jet multiplicity as the number of RECO jets of .5 conesize with E_T greater than 15.0 GeV. Figure (4) shows the jet multiplicity for data in the \cancel{E}_T ranges 0.0-7.0 GeV/c and 7.0-15.0 GeV/c. It can be seen that the higher jet multiplicities occur with greater frequency in the higher \cancel{E}_T range. We have taken events with \cancel{E}_T less than 10.0 GeV/c and zeroed the \cancel{E}_T and smeared them. If the hypothesis that the main source of \cancel{E}_T is detector resolution is correct, then the smeared events should also exhibit the above behavior. Figure (5) shows the jet multiplicity for smeared events in the same \cancel{E}_T ranges. It can be seen that the smeared events also exhibit the behavior shown by data , namely, higher multiplicities have preferentially higher \cancel{E}_T . This is just a resolution effect. This cross-check gives us confidence that the QCD contamination can be calculated using this method as a function of jet multiplicities as well.

Another important cross-check is to test whether the E_T distributions of electrons, jets

and “rest of the event” are affected by the smearing procedure or not. Figures (6, 7, 8) show the E_T distributions for electrons jets and rest of the event respectively, for all data, data with E_T less than 10 GeV and smeared events. It is evident that smearing the events that have been measured once already does not alter the distributions significantly. This is to some extent due to the fact that distributions in E_T fall with E_T , however, the resolutions get larger. This tends to compensate to some extent the widening of the distributions. The second lower peak in the electron E_T is due to secondary electrons in the sample. We have taken care to remove Z’s from the sample. Z’s are defined to be those events with at least one “good_electron” which have a pair mass between 85.0 and 95.0 GeV/c^2 .

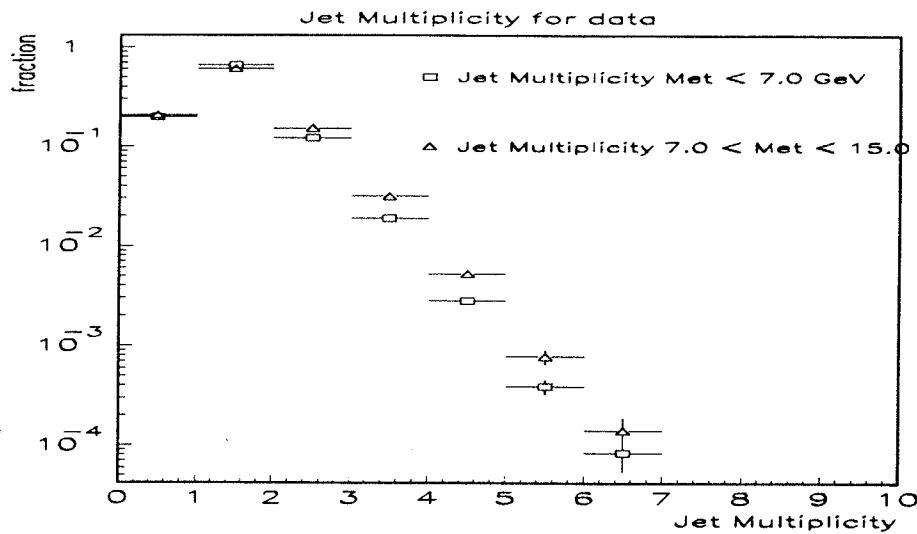


FIG. 4. Comparison of jet multiplicities for different E_T ranges for data

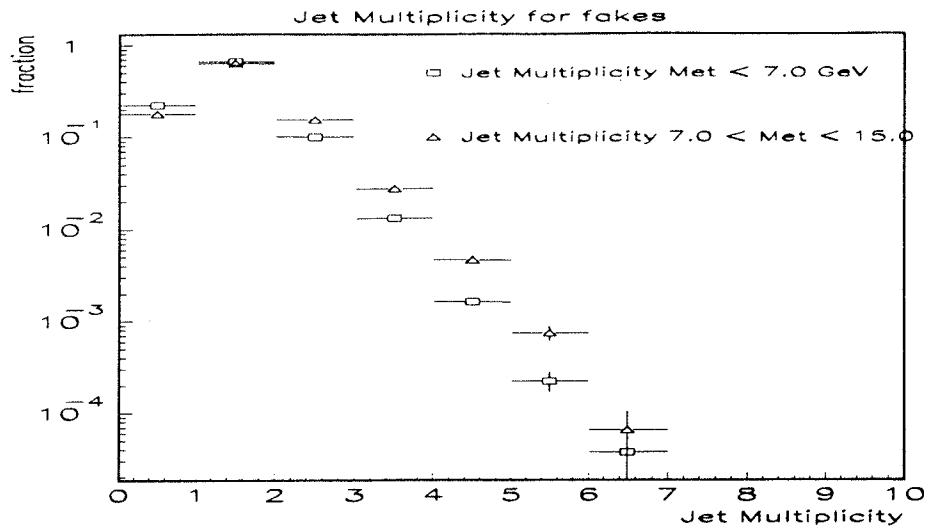


FIG. 5. Comparison of jet multiplicities for different E_T ranges for smeared events

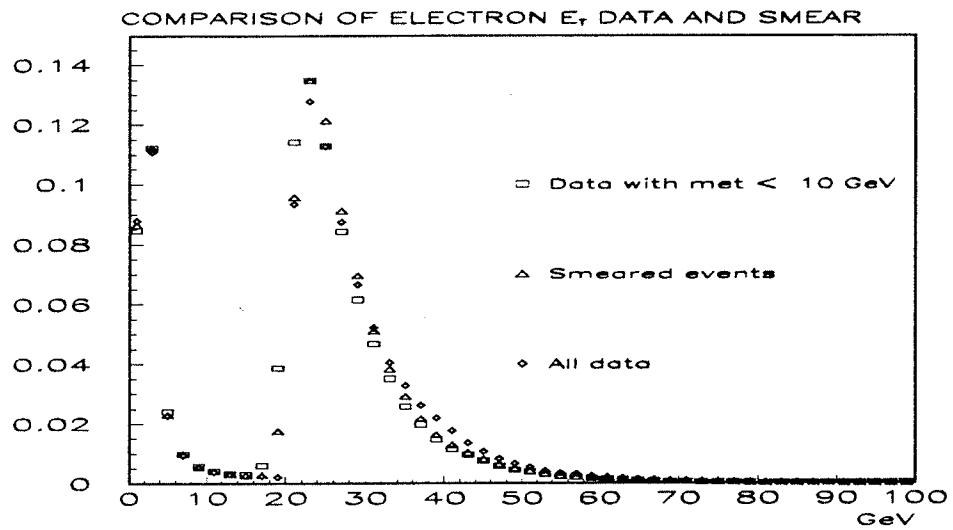


FIG. 6. Comparison of electron E_T distributions for smeared events, all data and data used to smear

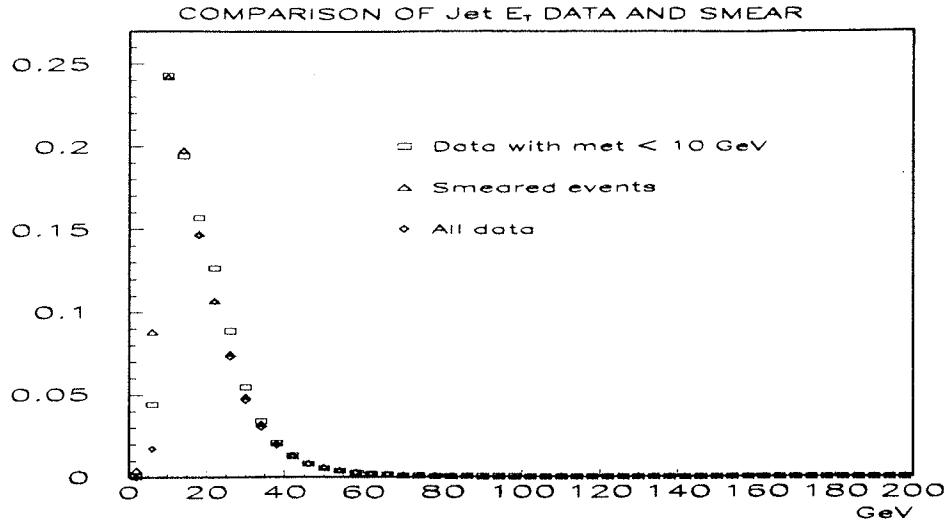


FIG. 7. Comparison of jet E_T distributions for smeared events, all data and data used to smear

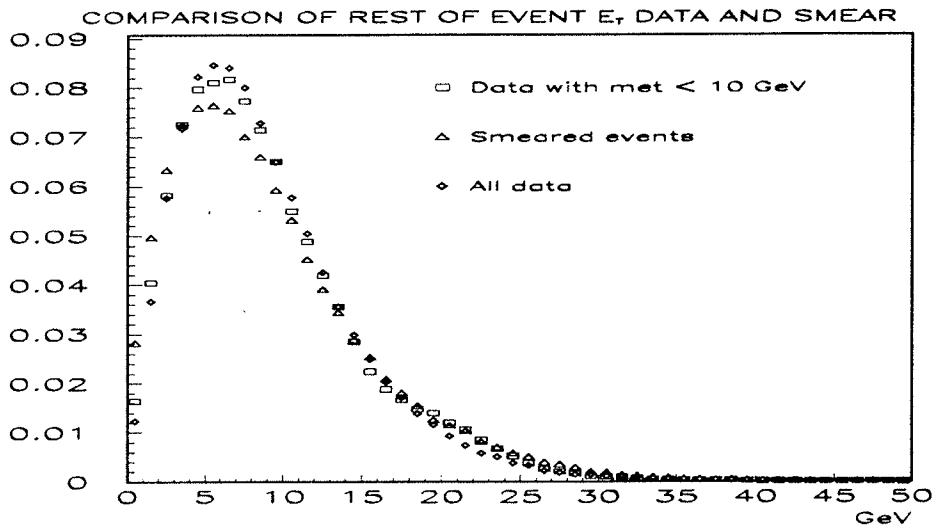


FIG. 8. Comparison of "rest of the event" p_T distributions for smeared events, all data and data used to smear

III. RESULTS

In Figure(9) we attempt to estimate the amount of background from QCD fakes when one plots the transverse mass of the electron and neutrino for E_T greater than 15.0 GeV. We have used only those electrons that pass the D \emptyset tight electron criteria, "Good_electron".

We see that the bump near low transverse masses ($30 \text{ GeV}/c^2$) in the data is actually due to the background which peaks at precisely at this region. We tend to overestimate the background for m_t less than $30 \text{ GeV}/c^2$ as explained above. We intend to fit the background subtracted curve to a W signal Monte Carlo curve that includes τ decays to verify the extent of our background subtraction accuracy.

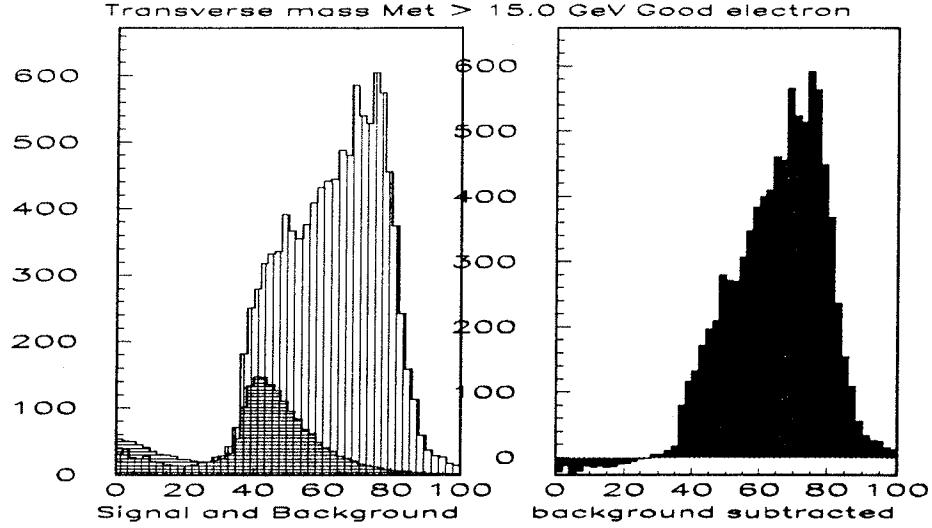


FIG. 9. Background subtraction for transverse mass distribution of events with $\cancel{E}_T > 15.0 \text{ GeV}$ and a good electron

A. Inclusive Transverse mass distributions

We define the QCD contamination rate as the ratio

$$QCD \text{ Contamination} = \frac{\text{estimated number of background events}}{\text{Number of signal and true background events}} \quad (3.1)$$

In regions of low transverse mass, due to the systematics outlined above, it is possible for the estimated number of background events to exceed the total number of true signal and background events. The QCD contamination ratio will exceed 1 in such regions. In figures (10, 11, 12), we explore the dependence of the QCD contamination rate as a function of transverse mass for \cancel{E}_T greater than 15.0, 20.0 and 30.0 GeV respectively. In each figure we show three curves, one where the only electron selection is the existence of a PELC bank,

the second with tight electron cuts (“Good_electron”) and the third with additional single minimum ionizing requirement (MIP1) which demands $\frac{dE}{dx} < 1.5$. It can be seen that the QCD contamination is reduced both by tightening the electron requirement as well as by demanding higher \not{E}_T . We have plotted the error bars of only the first curve in order to make the plot clearer. In all three curves, the contamination decreases as we get near the peak of the W Jacobian. It is worth noting that high p_T W’s sometimes give low transverse masses. For $\not{E}_T > 30.0$ GeV, low transverse masses have a contamination rate of 0.1 and may be usable for the top search.

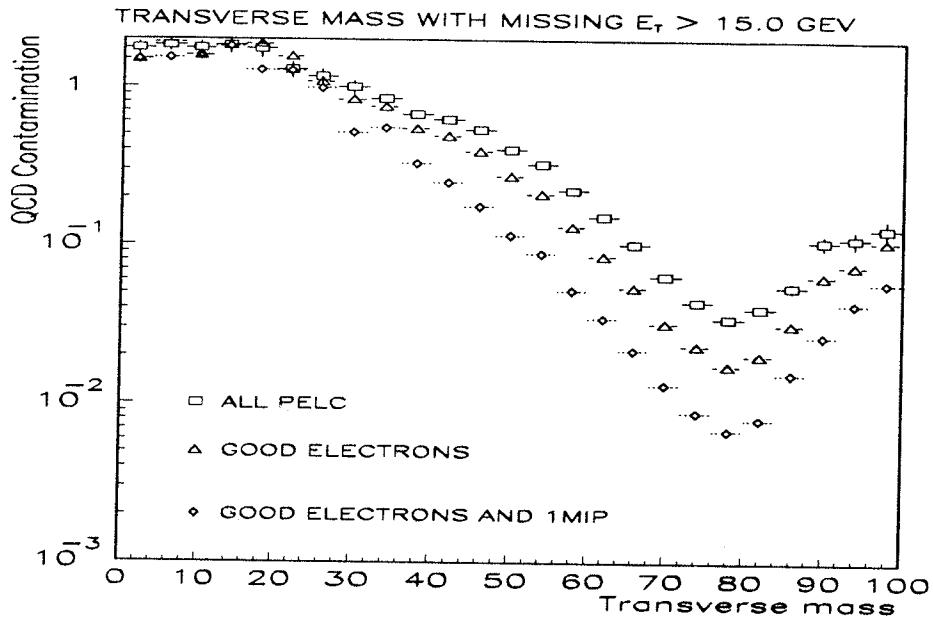


FIG. 10. QCD contamination as a function of transverse mass for $\not{E}_T > 15.0$ GeV for three different electron requirements

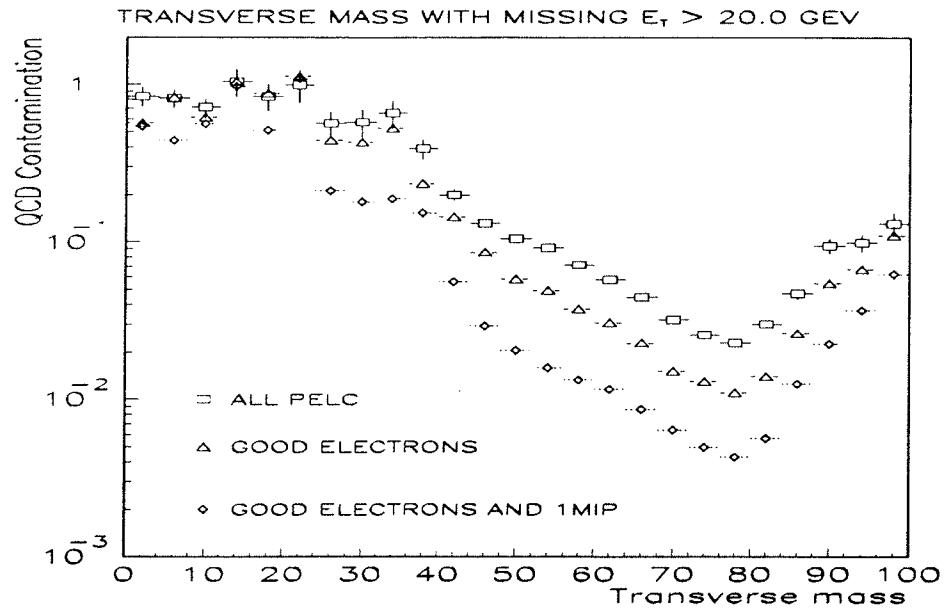


FIG. 11. QCD contamination as a function of transverse mass for $E_T > 20.0$ GeV for three different electron requirements

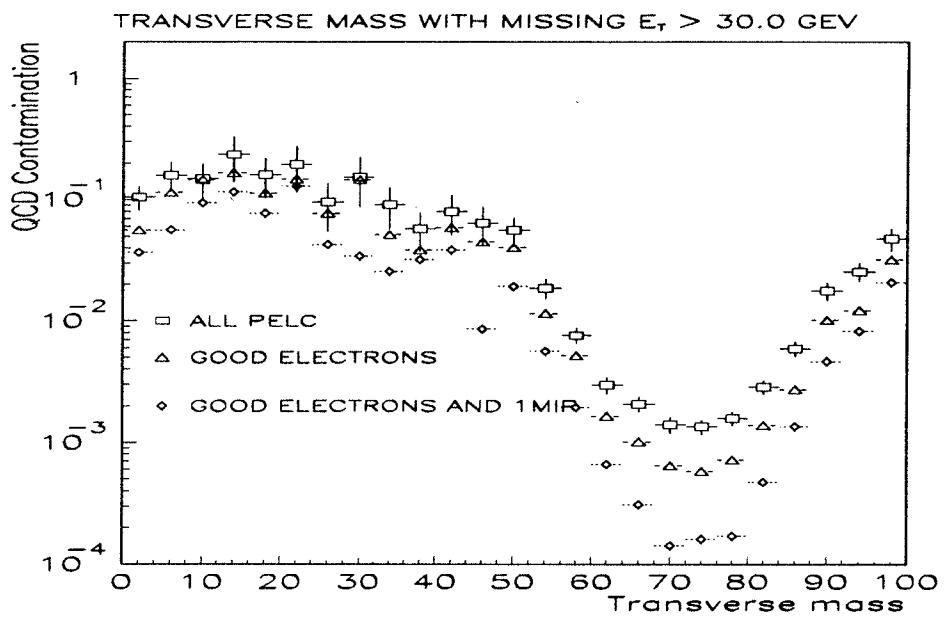


FIG. 12. QCD contamination as a function of transverse mass for $E_T > 30.0$ GeV for three different electron requirements

B. Exclusive Transverse mass distributions

Now we proceed to examine the dependence of the contamination rate as a function of multiplicity and transverse mass. Figures (13,14,15,16,17) show the dependence of the contamination rate as a function of transverse mass for jet multiplicities of 1,2,3,4 and ≥ 5 respectively. On each plot, we show three curves based on the same electron requirements as before. For $W + 4$ jets, there is significant QCD contamination ($\sim 20\%$) for $\cancel{E}_T > 15.0$ GeV. Figures (18,19,20,21,22) show the dependence of the contamination rate as a function of transverse mass for jet multiplicities of 1,2,3,4 and ≥ 5 respectively for $\cancel{E}_T > 20.0$ GeV. Figures (23,24,25,26,27) show the dependence of the contamination rate as a function of transverse mass for jet multiplicities of 1,2,3,4 and ≥ 5 respectively for $\cancel{E}_T > 25.0$ GeV. Figures (28,29,30,31,32) show the dependence of the contamination rate as a function of transverse mass for jet multiplicities of 1,2,3,4 and ≥ 5 respectively for $\cancel{E}_T > 30.0$ GeV. By this \cancel{E}_T cut, the backgrounds are well under control.

C. \cancel{E}_T distributions

In figures (33,34,35,36,37, 38), we show the contamination rate as a function of the \cancel{E}_T distribution for all multiplicities, multiplicity 1,2,3,4 and ≥ 5 jets respectively. All distributions are required to have transverse masses greater than $20 \text{ GeV}/c^2$.

D. Jet multiplicity distributions

In figures (39,40,41,42), we explore the contamination rate as a function of jet multiplicity, for $\cancel{E}_T > 15.0$,20.0, 25.0 and 30.0 GeV respectively. The distributions in Figure (40) can be compared directly with those calculated in [3]). For $\cancel{E}_T > 20$ Gev, we calculate the following QCD contamination rates as a function of exclusive jet multiplicity. Events which possess 1,2,3,4, and 5 jets have contamination rates 0.047 ± 0.001 , 0.158 ± 0.006 , 0.303 ± 0.019 , 0.471 ± 0.054 and 0.451 ± 0.074 respectively. This should be compared to 0.40 ± 0.09 as the

contamination quoted for ≥ 3 jets in [3]. Our statistical errors are much lower, enabling us to explore each multiplicity exclusively.

E. Electron Rapidity distributions

In figures (43,44,45,46), we explore the contamination rate as a function of the electron rapidity, for $\cancel{E}_T > 15.0, 20.0, 25.0$ and 30.0 GeV respectively. It can be seen that the contamination rate increases for high absolute values of electron rapidity, indicating that the QCD background is less central than the W signal.

IV. CONCLUSIONS

To conclude, we introduce a novel method to estimate the QCD contamination to the $W \rightarrow e\nu + \text{jets}$ channel. We have used it to explore the behavior of the background as a function of the two principal kinematic variables of the problem, namely \cancel{E}_T and transverse mass. The events thus generated can be passed to the topological searches (H-Matrix/PDE and Neural networks). The method outlined can be extended to the case where one of the jets has a muon (we need to treat the \cancel{E}_T from this jet somewhat specially). It can also be extended to calculate the QCD background under the Z, by selecting initially events in the range $10 \text{ GeV} < \cancel{E}_T < 20 \text{ GeV}$, where the Z signal should be small, zeroing their \cancel{E}_T and resmearing, similar to the case discussed here. In DØ we have not as yet calibrated the hadronic section of the calorimeter in ϕ and η . Better calibrations might well enable us to cut lower in true \cancel{E}_T .

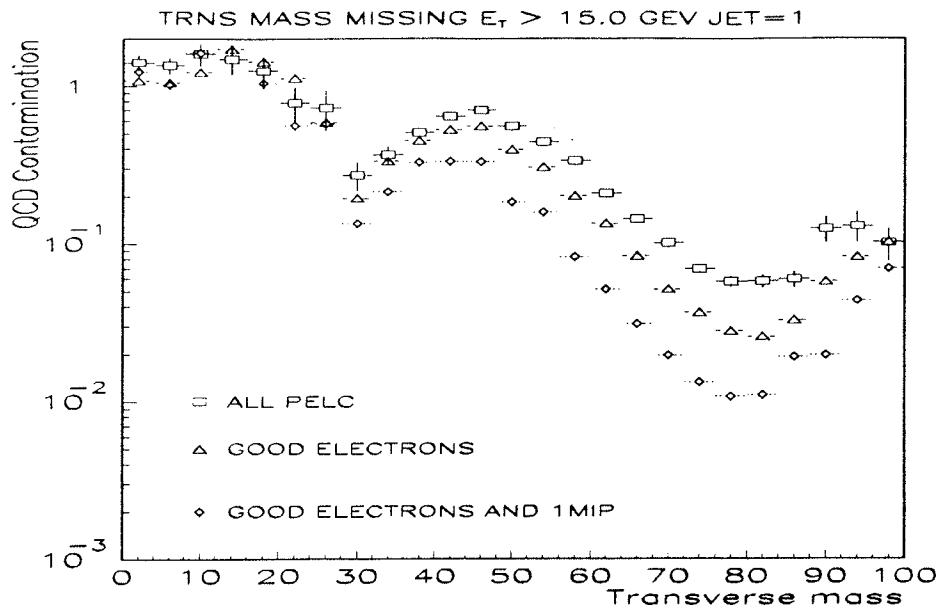


FIG. 13. QCD contamination as a function of transverse mass for $E_T > 15.0$ GeV and number of jets = 1 for three different electron requirements

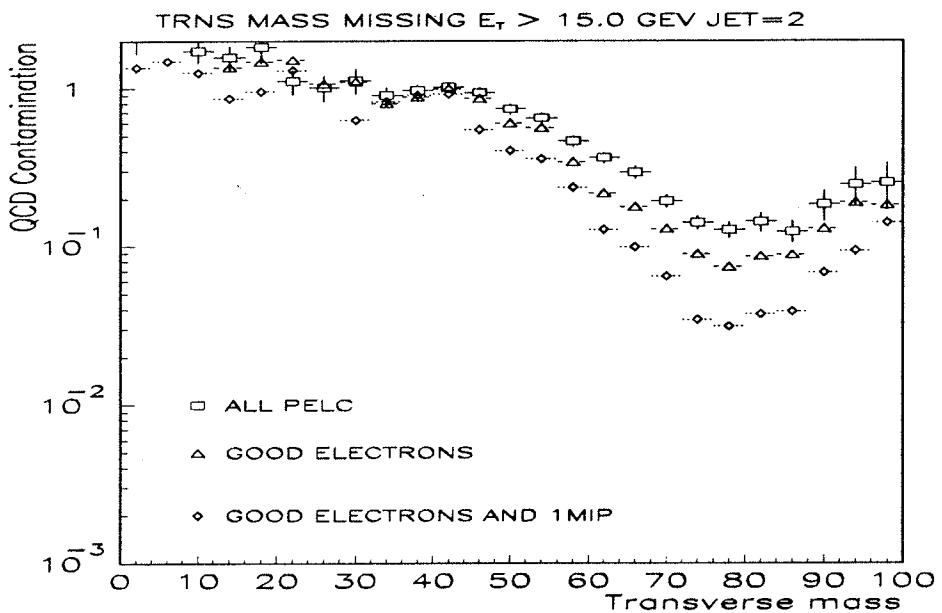


FIG. 14. QCD contamination as a function of transverse mass for $E_T > 15.0$ GeV and number of jets = 2 for three different electron requirements

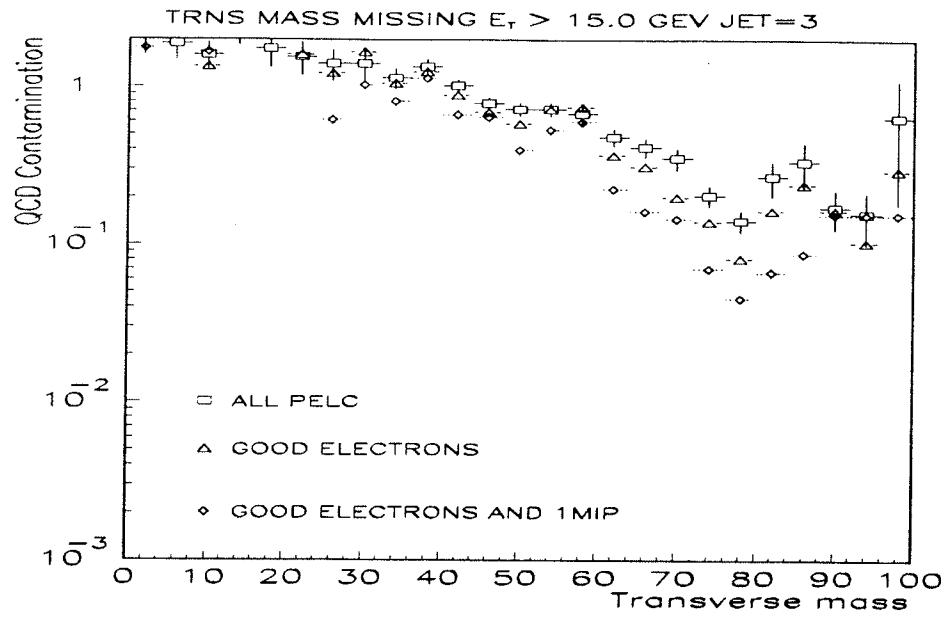


FIG. 15. QCD contamination as a function of transverse mass for $E_T > 15.0$ GeV and number of jets = 3 for three different electron requirements

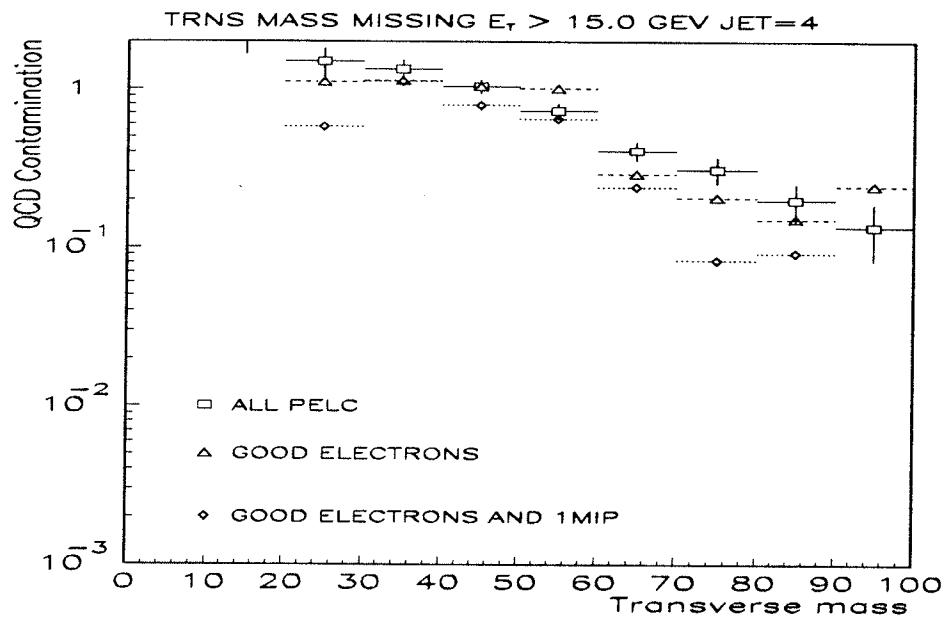


FIG. 16. QCD contamination as a function of transverse mass for $E_T > 15.0$ GeV and number of jets = 4 for three different electron requirements

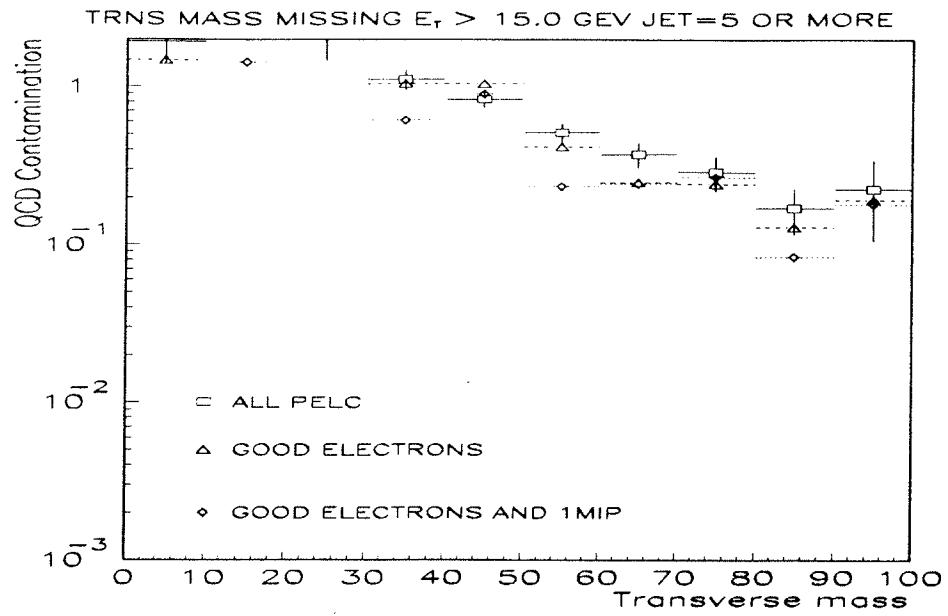


FIG. 17. QCD contamination as a function of transverse mass for $E_T > 15.0$ GeV and number of jets ≥ 5 for three different electron requirements

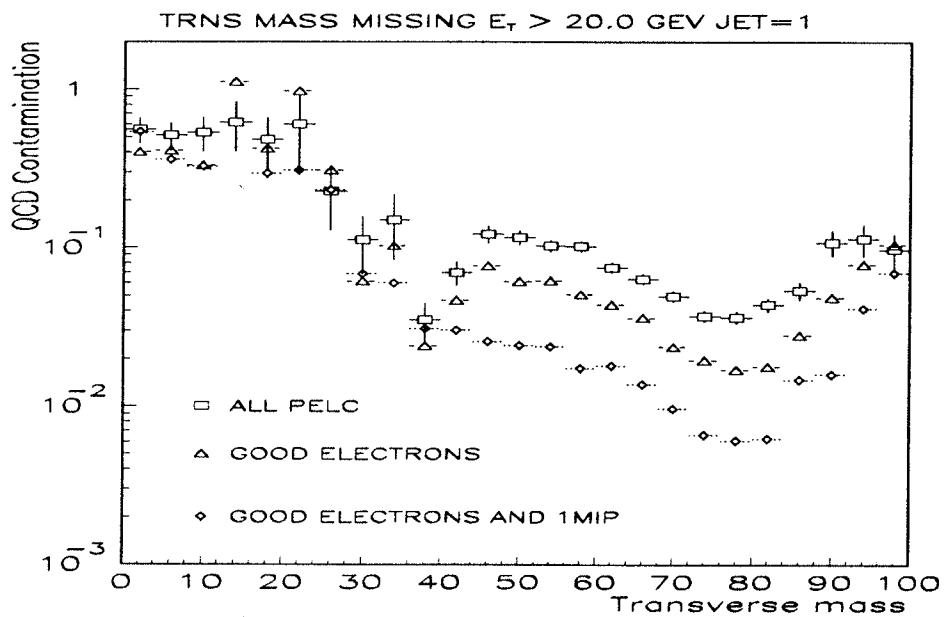


FIG. 18. QCD contamination as a function of transverse mass for $E_T > 20.0$ GeV and number of jets = 1 for three different electron requirements

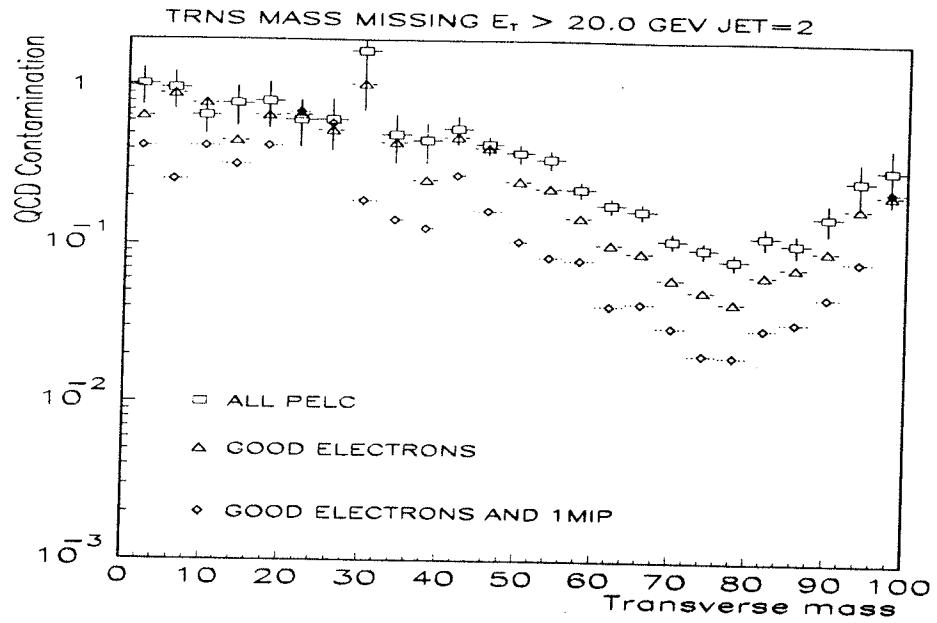


FIG. 19. QCD contamination as a function of transverse mass for $E_T > 20.0$ GeV and number of jets = 2 for three different electron requirements

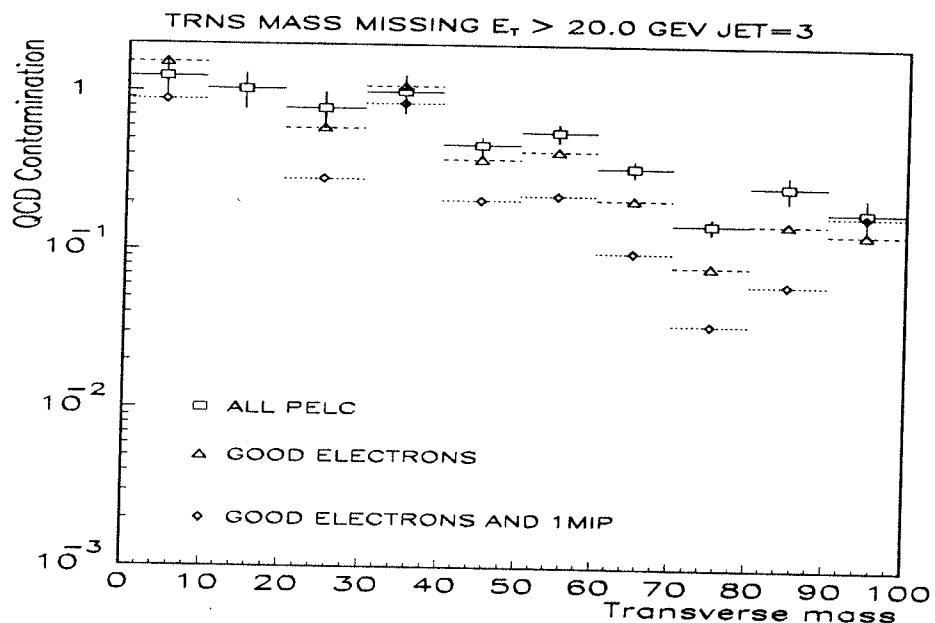


FIG. 20. QCD contamination as a function of transverse mass for $E_T > 20.0$ GeV and number of jets = 3 for three different electron requirements

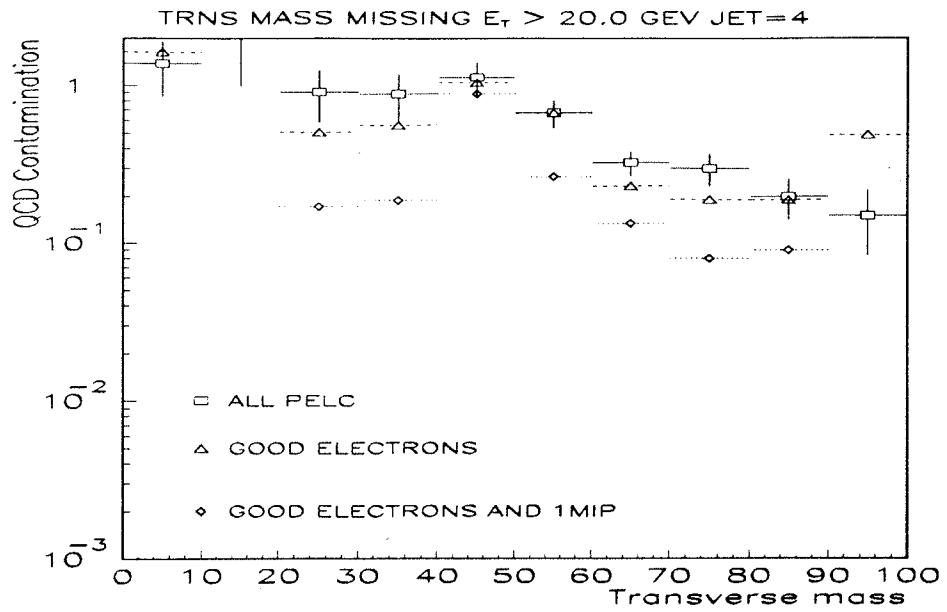


FIG. 21. QCD contamination as a function of transverse mass for $E_T > 20.0$ GeV and number of jets = 4 for three different electron requirements

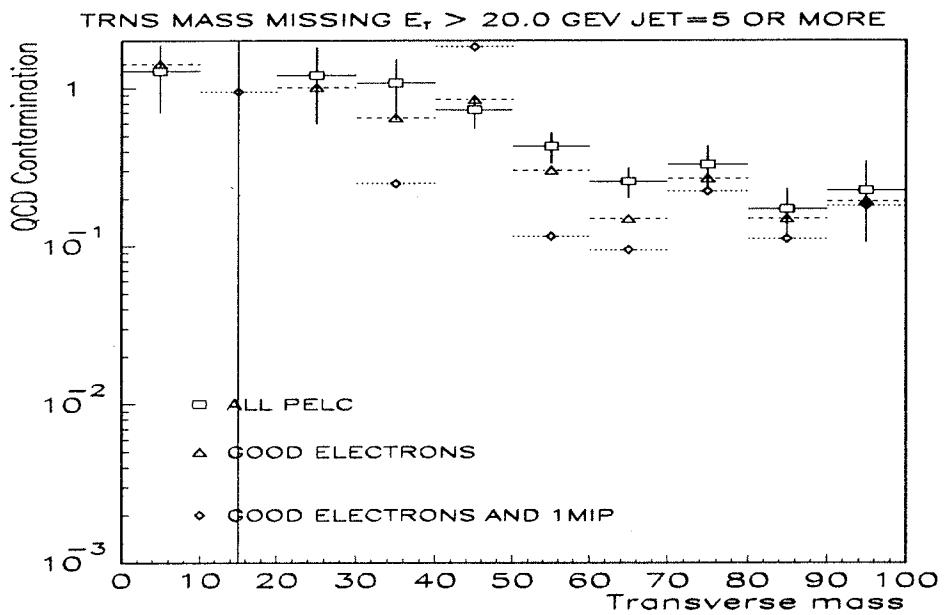


FIG. 22. QCD contamination as a function of transverse mass for $E_T > 20.0$ GeV and number of jets ≥ 5 for three different electron requirements

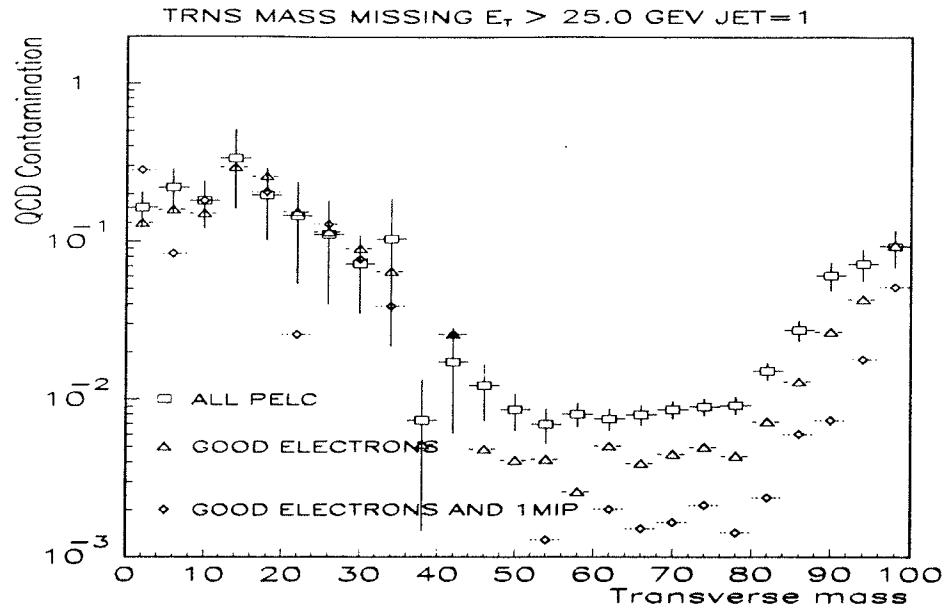


FIG. 23. QCD contamination as a function of transverse mass for $E_T > 25.0$ GeV and number of jets = 1 for three different electron requirements

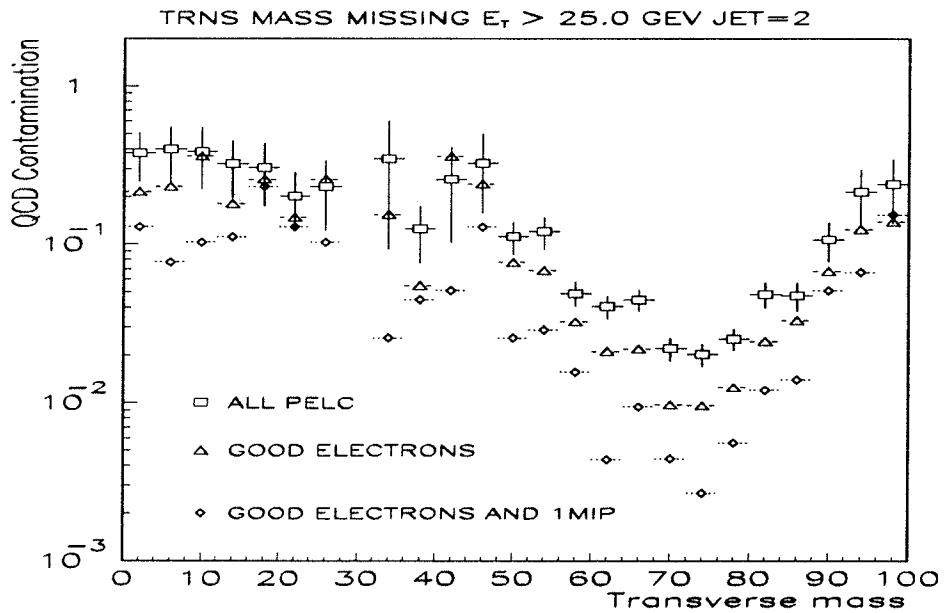


FIG. 24. QCD contamination as a function of transverse mass for $E_T > 25.0$ GeV and number of jets = 2 for three different electron requirements

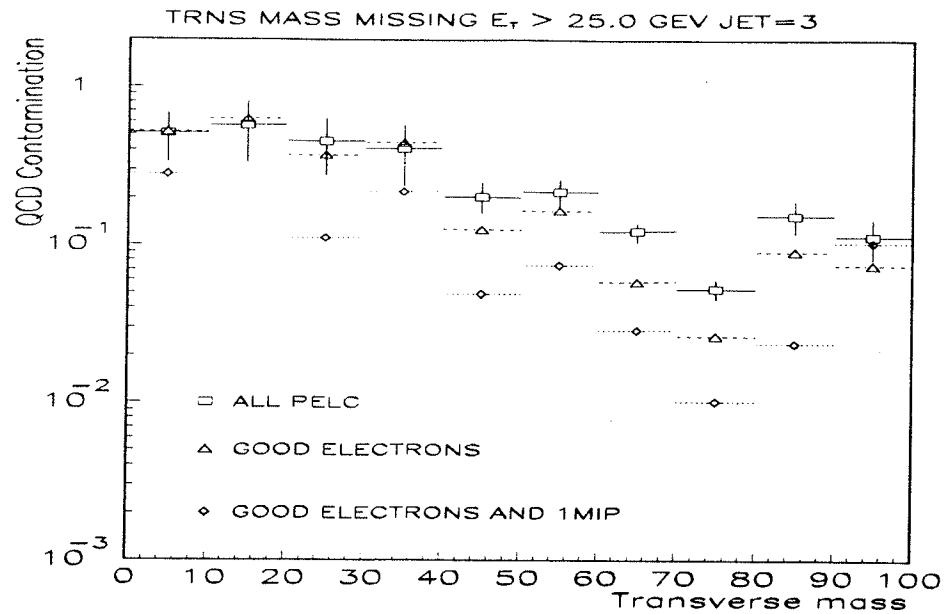


FIG. 25. QCD contamination as a function of transverse mass for $E_T > 25.0$ GeV and number of jets = 3 for three different electron requirements

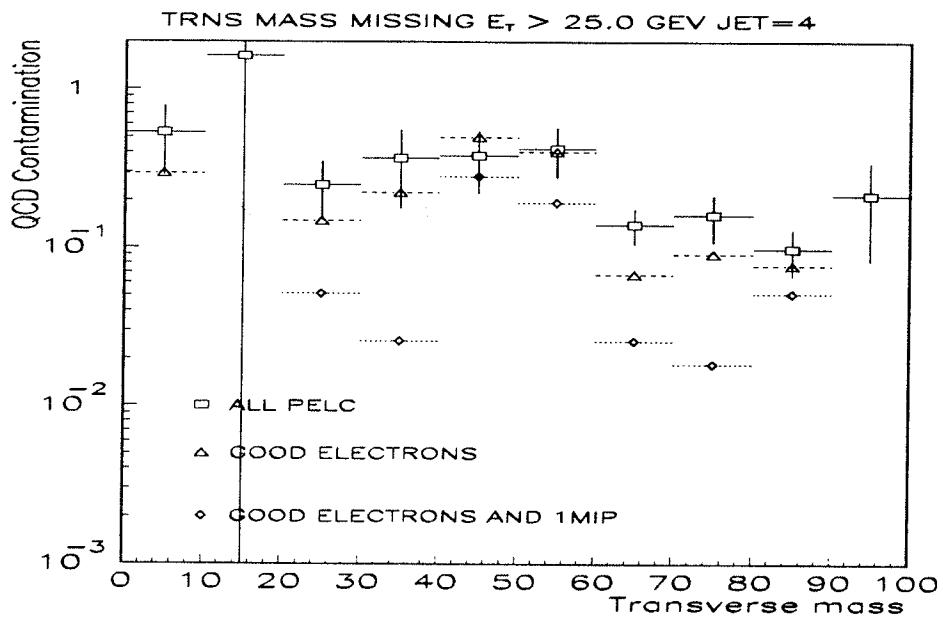


FIG. 26. QCD contamination as a function of transverse mass for $E_T > 25.0$ GeV and number of jets = 4 for three different electron requirements

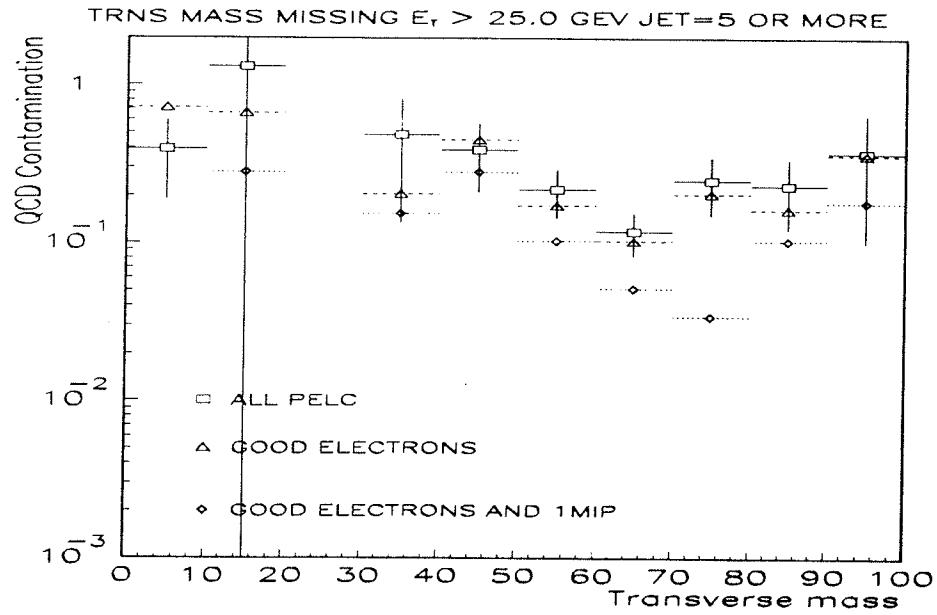


FIG. 27. QCD contamination as a function of transverse mass for $E_T > 25.0$ GeV and number of jets ≥ 5 for three different electron requirements

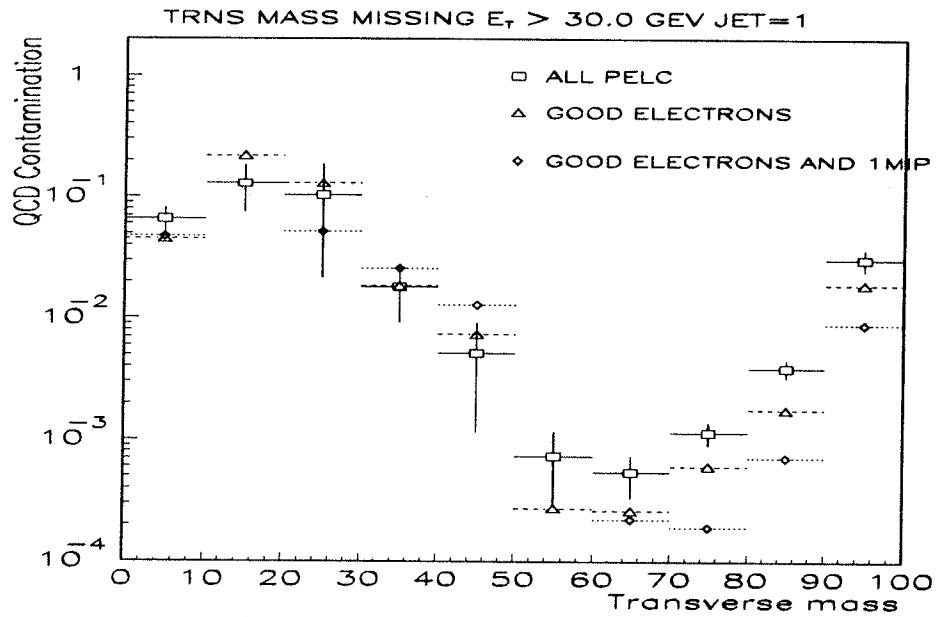


FIG. 28. QCD contamination as a function of transverse mass for $E_T > 30.0$ GeV and number of jets = 1 for three different electron requirements

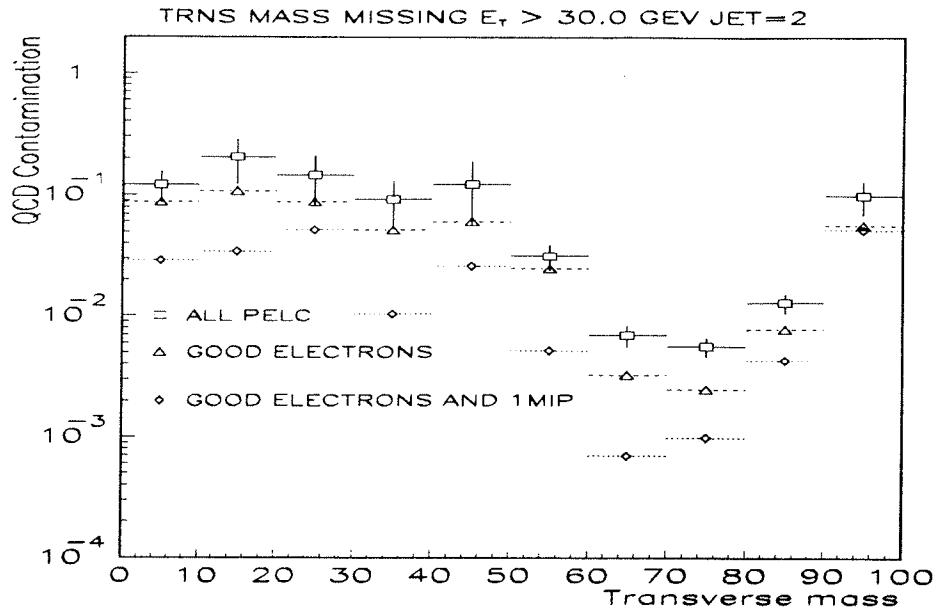


FIG. 29. QCD contamination as a function of transverse mass for $E_T > 30.0$ GeV and number of jets = 2 for three different electron requirements

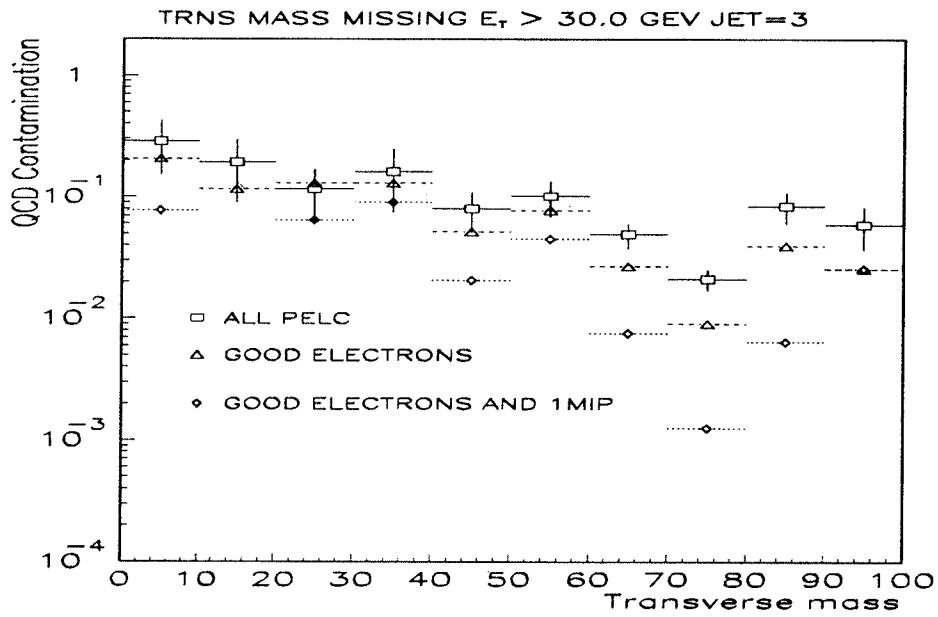


FIG. 30. QCD contamination as a function of transverse mass for $E_T > 30.0$ GeV and number of jets = 3 for three different electron requirements

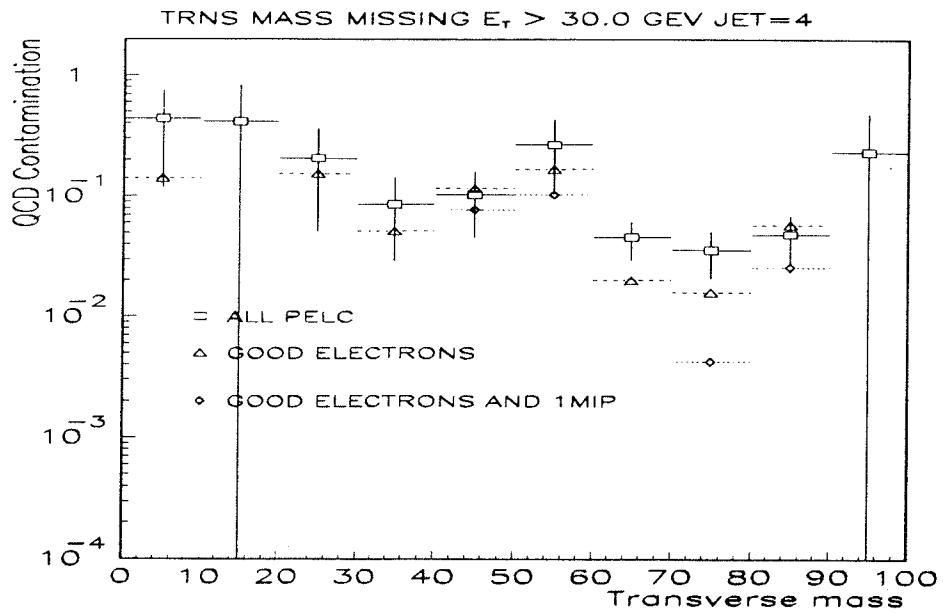


FIG. 31. QCD contamination as a function of transverse mass for $E_T > 30.0$ GeV and number of jets = 4 for three different electron requirements

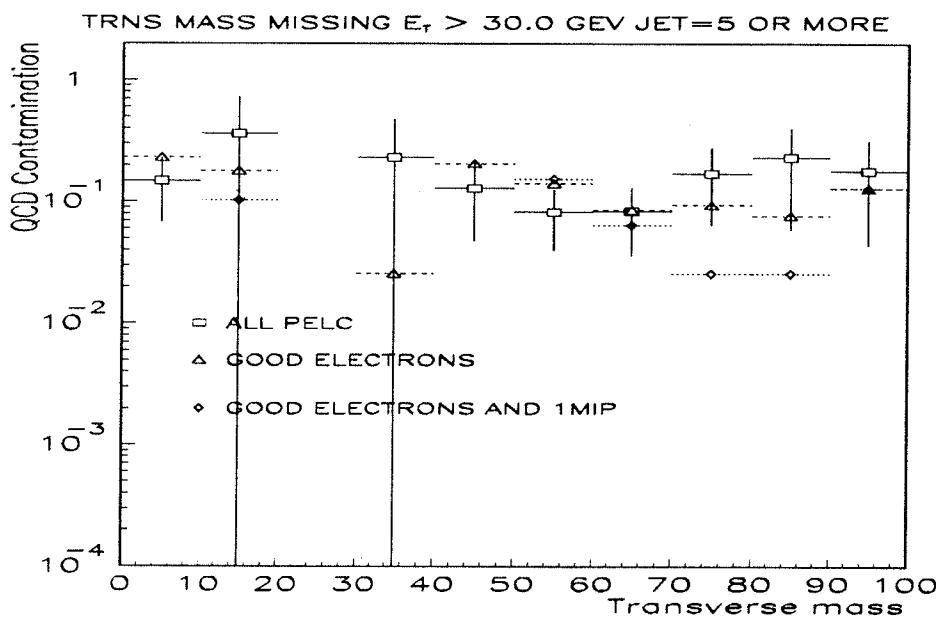


FIG. 32. QCD contamination as a function of transverse mass for $E_T > 30.0$ GeV and number of jets ≥ 5 for three different electron requirements

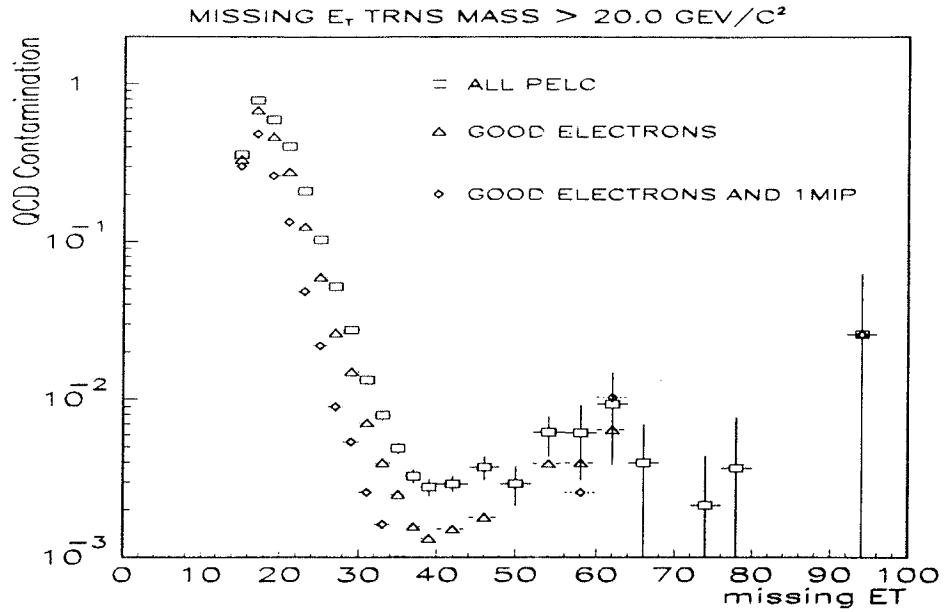


FIG. 33. QCD contamination as a function of E_T for transverse mass $> 20.0 \text{ GeV}/c^2$ for all multiplicities

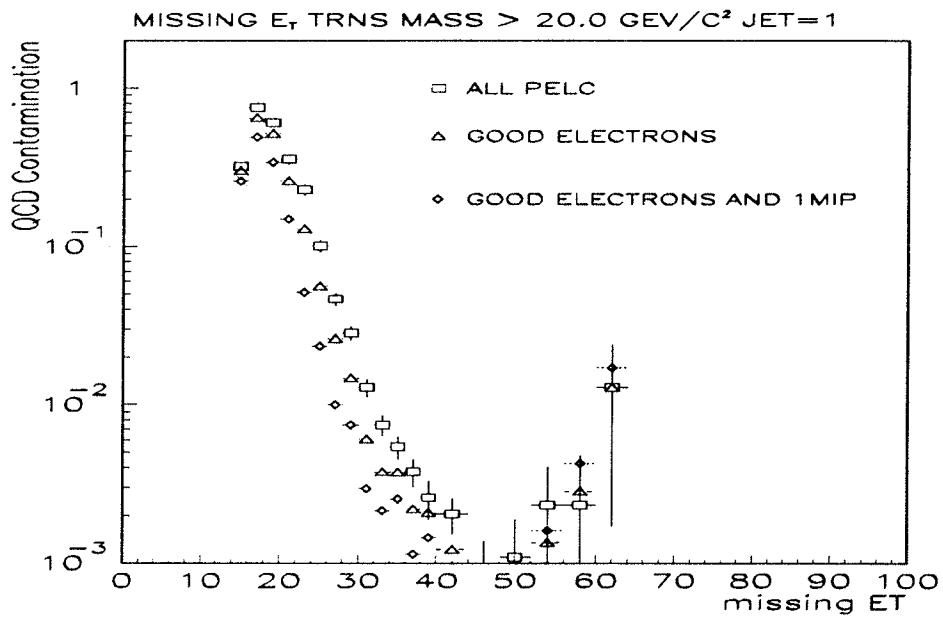


FIG. 34. QCD contamination as a function of E_T for transverse mass $> 20.0 \text{ GeV}/c^2$ for jet=1

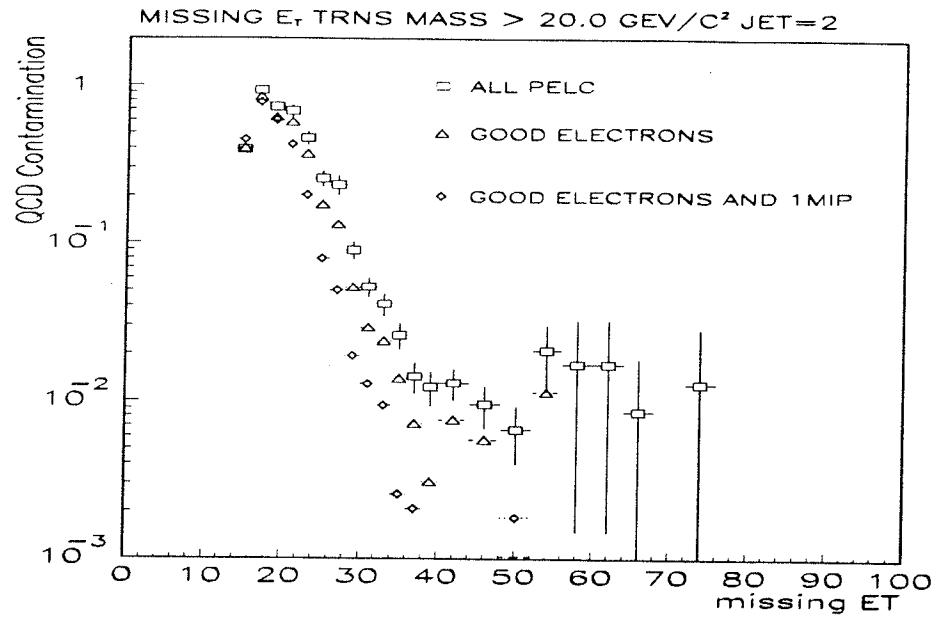


FIG. 35. QCD contamination as a function of E_T for transverse mass > 20.0 GeV/c^2 for jet=2

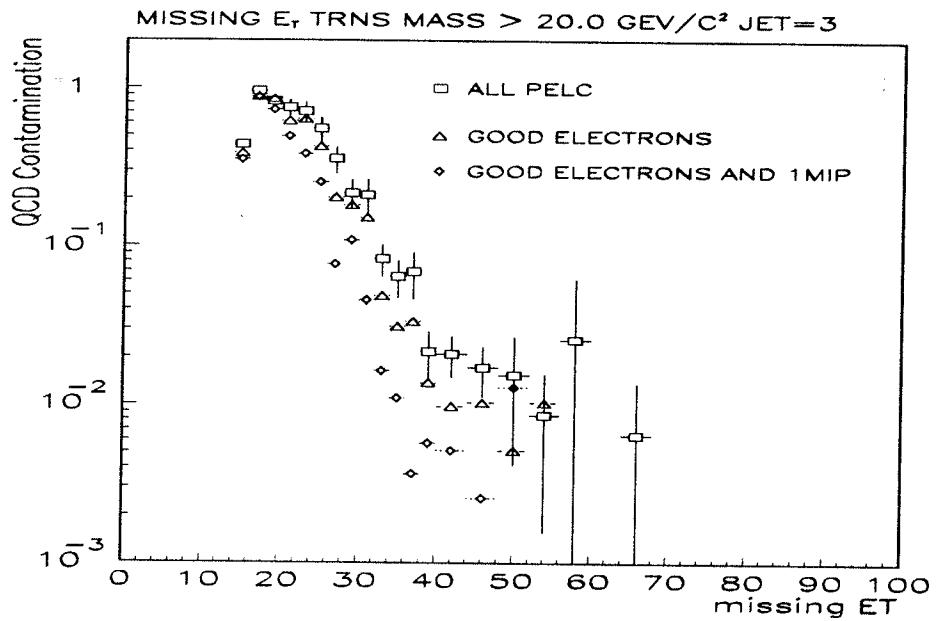


FIG. 36. QCD contamination as a function of E_T for transverse mass > 20.0 GeV/c^2 for jet=3

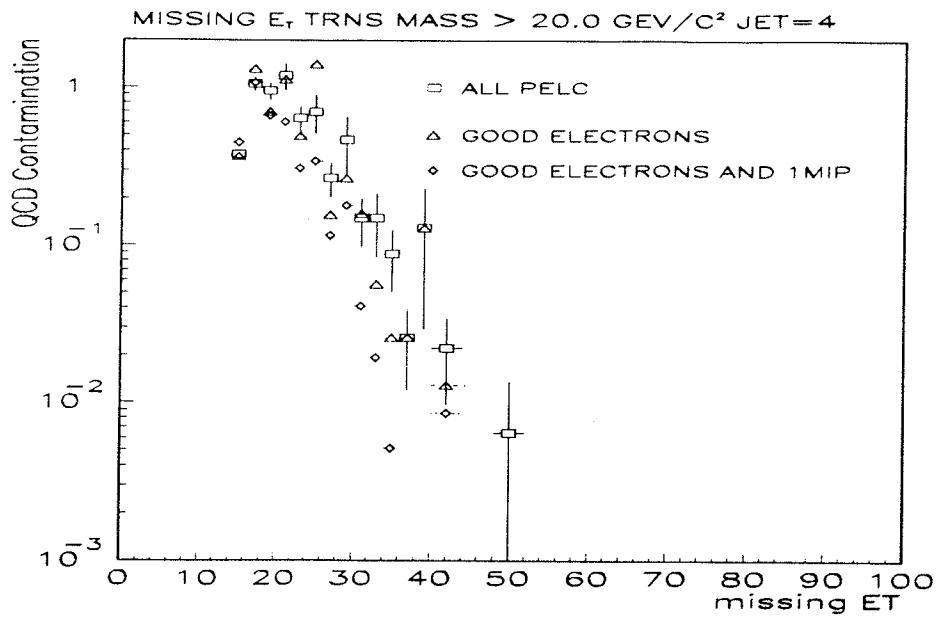


FIG. 37. QCD contamination as a function of E_T for transverse mass > $20.0 \text{ GeV}/c^2$ for jet=4

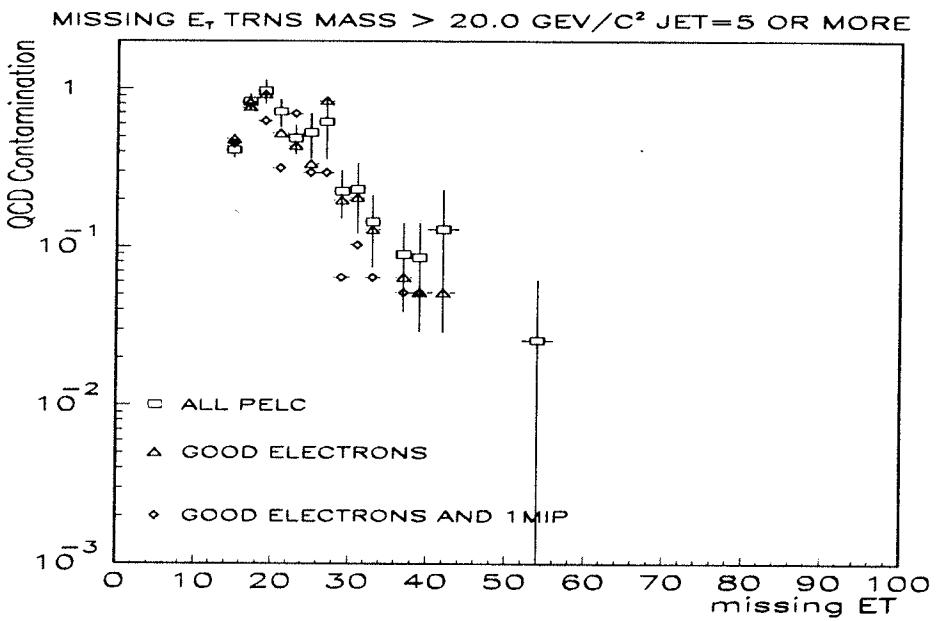


FIG. 38. QCD contamination as a function of E_T for transverse mass > $20.0 \text{ GeV}/c^2$ for jet ≥ 5

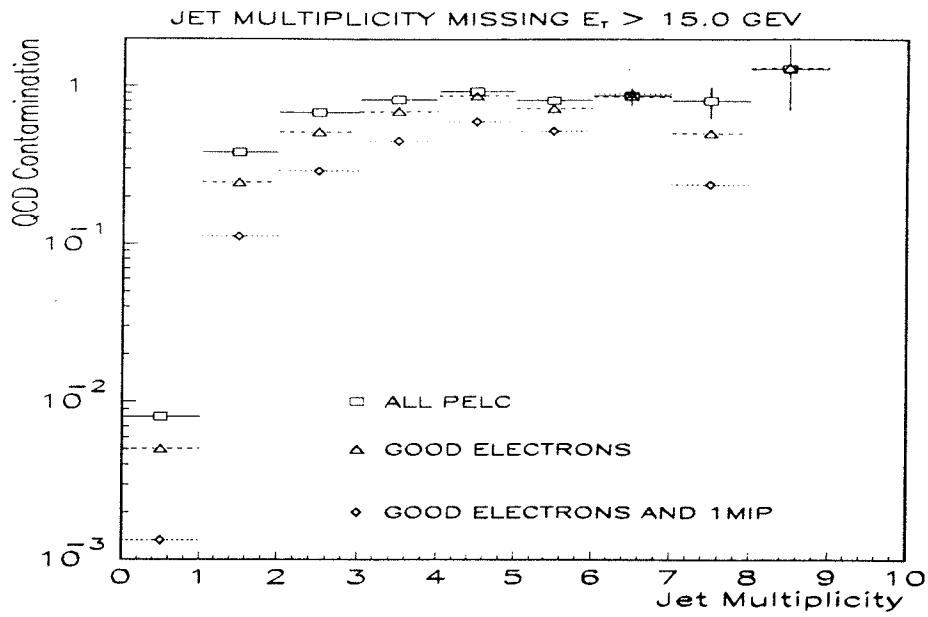


FIG. 39. QCD contamination as a function of multiplicity for $E_T > 15.0\text{GeV}/c^2$

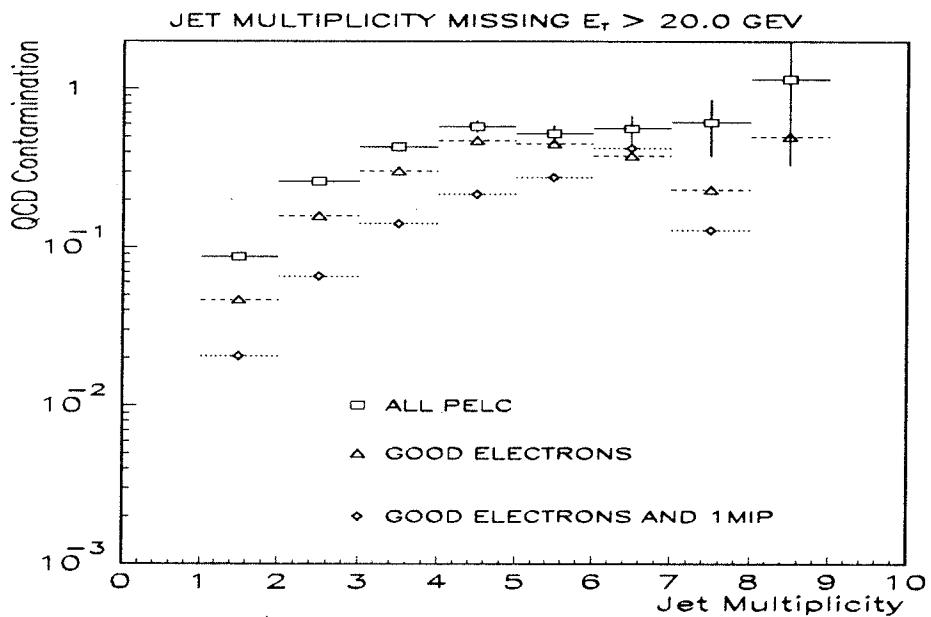


FIG. 40. QCD contamination as a function of multiplicity for $E_T > 20.0\text{GeV}/c^2$

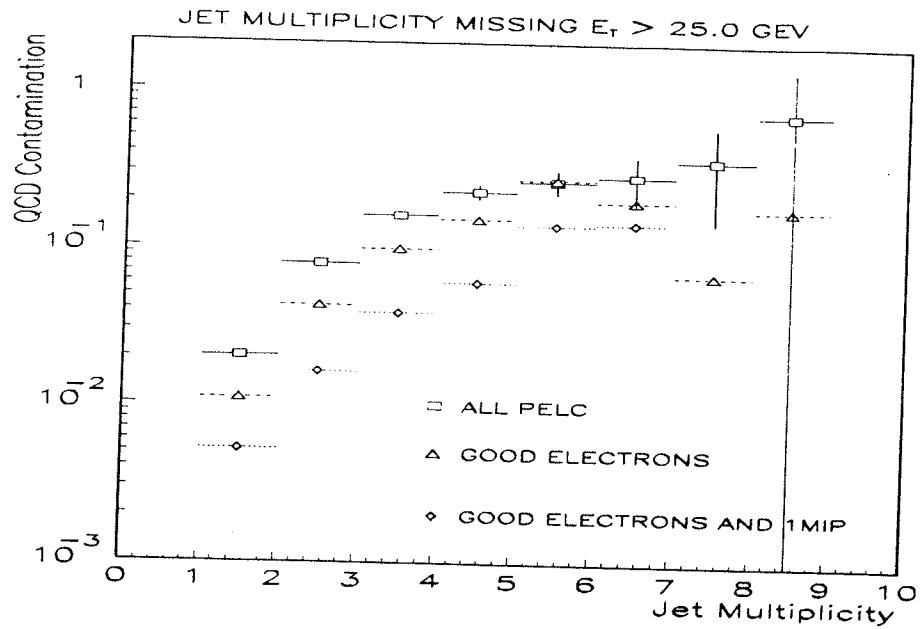


FIG. 41. QCD contamination as a function of multiplicity for $E_T > 25.0 \text{ GeV}/c^2$

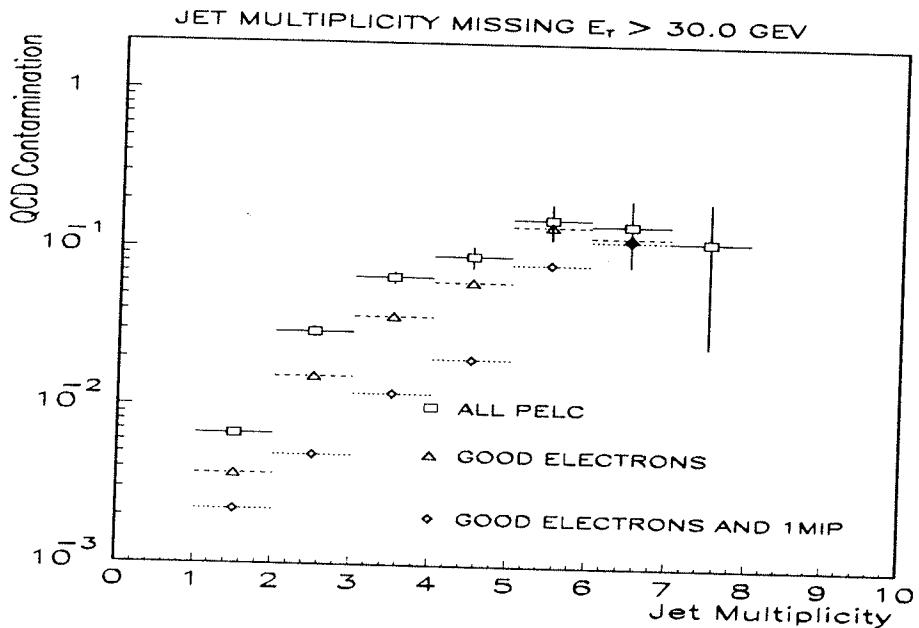


FIG. 42. QCD contamination as a function of multiplicity for $E_T > 30.0 \text{ GeV}/c^2$

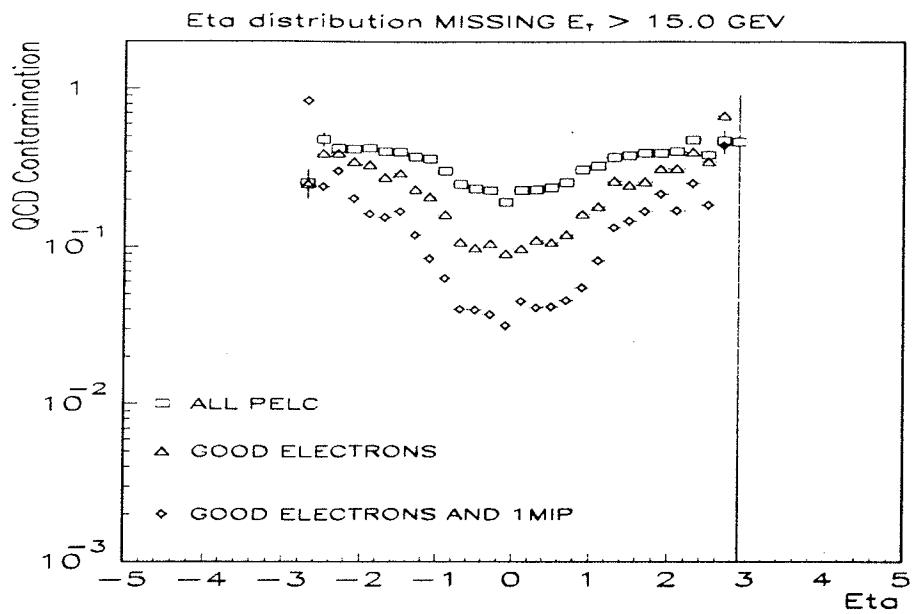


FIG. 43. QCD contamination as a function of electron rapidity for $E_T > 15.0 \text{ GeV}/c^2$

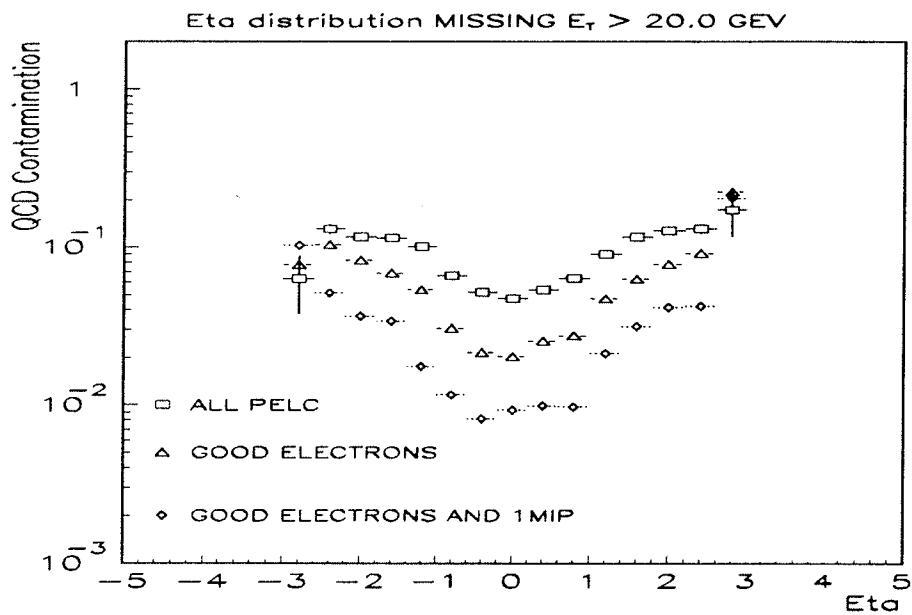


FIG. 44. QCD contamination as a function of electron rapidity for $E_T > 20.0 \text{ GeV}/c^2$

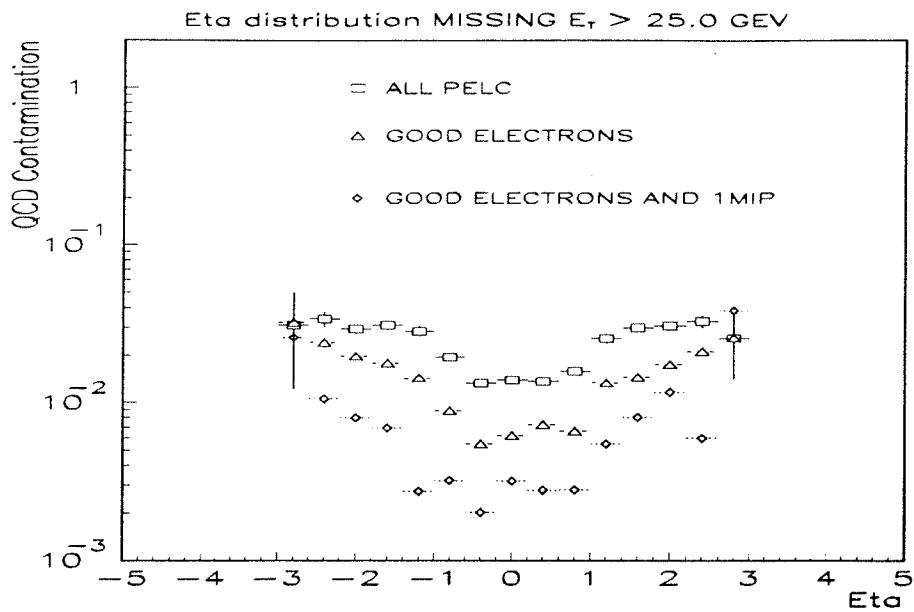


FIG. 45. QCD contamination as a function of electron rapidity for $E_T > 25.0 \text{ GeV}/c^2$

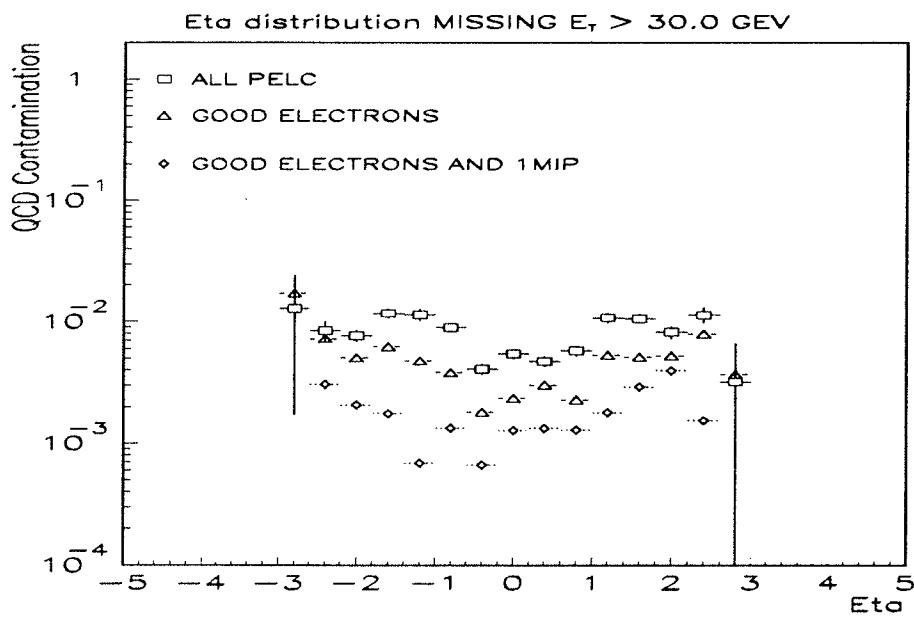


FIG. 46. QCD contamination as a function of electron rapidity for $E_T > 30.0 \text{ GeV}/c^2$

TABLES

Formula	Variable	C	S	N
$\frac{\sigma^2}{E^2} = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2}$	Jet Energy	0.1848	0.8245	0.0
$\sigma^2 = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2}$	Jet Eta	0.03338	0.3352	0.00
$\sigma^2 = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2}$	Jet Phi	0.03291	0.2804	1.106
$\frac{\sigma^2}{E^2} = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2}$	Electron Energy	0.03	0.15	0.0
$\sigma^2 = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2}$	Electron Eta	0.025	0.0	0.0
$\sigma^2 = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2}$	Electron Phi	0.019	0.0	0.0

TABLE I. Jet and electron resolution parameters

REFERENCES

- [1] QCD Backgrounds to Electroweak Signals. A Study of Fake electron and fake Photon Probability. Mirek Fatyga, Dhiman Chakraborty, DØ note 1753
- [2] A Study of Backgrounds to $W \rightarrow e + \nu + \text{jets}$ from Electro- Weak Processes, Jae Yu, DØ note 1685
- [3] QCD background to $W(e\nu) + \text{jets}$, DØ note in preparation, R. Partridge, H. Xu