Olsson

JADE-Computer Note 51
11.5.1981
M.C. Goddard

A General Routine for the Fast Reconstruction of Jet Events

In the analysis of jet events it is important to be able to reconstruct the jet axes as accurately and as fast as possible. The high reconstruction accuracy is clearly a prerequisite for discriminating tests of the underlying parton dynamics. Just as important is computational speed. Since many thousands of jet events, both data and Monte-Carlo, have to be analysed. An algorithm is not practical if its CPU time is more than a fraction of a second per event.

I describe in this note a general routine for the reconstruction of an arbitrary number of jets. This algorithm reconstructs the jet axes as accurately as any existing method and is two to one hundred times faster[1]. This dramatic increase in the speed comes about because the standard procedure of partitioning the event is not used. Instead an iterative method is employed to reconstruct each jet in the event. This gives an extremely fast algorithm with a CPU time only weakly dependent on event multiplicity or number of jets.

In the first part of this note the algorithm will be described and compared to existing jet reconstruction methods. I shall describe in general terms these standard algorithms for two, three and four jet reconstruction. This discussion will emphasize their limitations and the necessity for a new approach will become clear. Next, results obtained with this algorithm will be given. The accuracy in reconstructing the parton axes in space and in determining the jet energies will be presented. Then, the CPU time required will be compared to the time taken by existing methods. Finally, detailed instructions will be given on how to use this routine.

Jet Reconstruction

Assume we have a set C_K of N_K particles which are hypothesized to belong to jet number K. Let $\vec{T}_k$ be that function of the momenta which is chosen as the estimator of the original parton axis. This algorithm takes for $\vec{T}_k$ the thrust axis of jet K:

$$\vec{T}_K = \sum_{i \in C_K} \vec{P}_i$$

This is the direction $\hat{N}$ which maximizes the quantity $\sum_{i \in C_K} (\vec{p}_i \cdot \hat{N})$. It has the advantage that it is linear in the momenta and therefore not affected by particle decays. A measure of the correctness of the assignment of C_K to jet K is then :

$$|\vec{T}_K| = \frac{\vec{T}_K \cdot \vec{T}_K}{|\vec{T}_K|} = \sum_{i \in C_K} \vec{P}_i \cdot \hat{T}_K$$

The correctness of the reconstruction of the whole event is measured by:

$$T \equiv \sum_K |\vec{T}_K| = \sum_K \sum_{i \in C_K} (\vec{P}_i \cdot \hat{T}_K)$$

which is, apart from a normalization factor, an extension of the definition of triplicity[2] .

Standard algorithms for the reconstruction of M jets[2,3] use the method of partitioning the particles of the event into M (non-empty) subsets. This generates all possible sets C_K and the partition which maximizes the chosen function T determines the jet axes. The number of different partitions of N objects into M (non-empty) subsets is:

$$S_n^{(m)} = \frac{1}{m!} \sum_{k=1}^{m} (-1)^{m-k} \binom{m}{k} k^n \qquad (1)$$

$$\sim m^{n-1}/(m-1)! \qquad \text{for } n \gtrsim 10$$

The CPU time needed for one partition is typically 50 microsec. For two jet events this means a fraction of a sec per event for multiplicity 15 but hundreds of seconds for multiplicity 25. Each three jet event would require hours of CPU time. However because three jet events are planar, they can be projected onto the event plane without distorting the underlying jet structure. Then the standard procedure is to use contiguous partitions of the projected momenta[3]. The number of contiguous partitions is:

$$\frac{1}{m!}\binom{n}{m}$$

$$\sim n^3/(3!)^2 \qquad \text{for} \quad n \gtrsim 10$$

leading again to a fraction of a second per event. For a four jet event of multiplicity 35 the number of partitions is, according to equation 1, 4.9×10^{19}. Thus a CPU time of 10^{13} years would be necessary to reconstruct the jet axes. This would result in an unacceptably long delay between data taking and publication. This difficulty cannot be circumvented, as was possible for three jet events, because four jet events are not planar and therefore contiguous partitions cannot be used. The only other existing four-jet algorithm uses three-jet reconstruction in conjunction with a Lorentz Transformation and takes several seconds per event[4]. The conclusion is that the standard methods are not practical for high multiplicity events or for reconstructing more than three jets. The basic problem is that the number of partitions is such a rapidly increasing function of the multiplicity and the number of jets.

This algorithm does not partition the event but uses an iterative method which I will now describe.

Assume we have an initial (zeroth) approximation for the jet axes: $\vec{T}_k^{(0)}$

How these $\vec{T}_k^{(0)}$ are obtained will be described later. Each particle in the event is then assigned to the closest jet axis. These form the set of particles associated with the jet axis for the first approximation: $C_k^{(1)}$.

The corresponding jet axis for the first approximation is then:

$$\vec{T}_k^{(1)} = \sum_{i \in C_k^{(1)}} \vec{p}_i$$

This procedure is repeated giving a $C_k^{(\ell)}$ and $\vec{T}_k^{(\ell)}$ for the lth iteration:

$$\vec{T}_k^{(\ell)} = \sum_{i \in C_k^{(\ell)}} \vec{p}_i$$

$$C_k^{(\ell)} = \left\{ \vec{p}_i \;\middle|\; \text{such that } \vec{p}_i \cdot \hat{T}_k^{(\ell-1)} \text{ is a max}, \; k=1,\dots m \right\}$$

The iteration is terminated when

$$\cos^{-1}\left[\hat{T}_k^{(\ell)} \cdot \hat{T}_k^{(\ell-1)} \right] < \delta \quad , \text{for each } k.$$

for each jet. For the results shown here δ was one degree and this typically required four or five iterations.

How well the reconstructed axes correspond to the real jet axes depends on how good a choice for the zeroth approximation we can find. In general the iteration will converge on the real jet axis K if $C_K^{(1)}$ contains most of the fast particles from the fragmentation of parton K and none from any other. This means that the $\vec{T}_K^{(o)}$ should be chosen such that if $\vec{T}_{K'}^{(o)}$ is the closest of all the $\vec{T}_K^{(o)}$ to parton jet K then the closest parton jet to $\vec{T}_{K'}^{(o)}$ is jet K. In other words there should be an unambiguous one to one correspondence between $\vec{T}_K^{(o)}$ and real jet axes. The set of $\vec{T}_K^{(o)}$ for reconstructing M jets is called a 'configuration'. The configurations are chosen to meet the condition above based on Monte Carlo studies of jet events. For two jet events there is only one configuration and this is where $\vec{T}_1^{(o)}$ is the fastest particle in the event and $\vec{T}_2^{(o)} = -\vec{T}_1^{(o)}$. Two configurations are used for three jet events and 10 for four jet events[5].

After convergence of the iteration we have a set of $\vec{T}_K$ from each configuration. That set with the maximum value of T is chosen as the final reconstructed jet axes.

Although this algorithm can reconstruct an arbitrary number of jets. The present version is coded for only up to four jets. This is sufficient for presently available energies. To extend it to more than four jets one needs to simply include the necessary configurations.

Results

The effectiveness of the algorithm in reconstructing the jet axes can be determined using Monte Carlo events[6]. Since the directions and energies of the partons are known they can be directly compared with the reconstructed quantities. In this manner the resolution in reconstructing the energy and the error in the reconstructed jet directions can be obtained.

Spatial Reconstruction

The statistical nature of the Field Feynman fragmentation[7] used in the Monte Carlo produces an inherent limitation to the accuracy with which the original parton axes can be determined. Figure 1 shows the angle between the generated parton axis and $\vec{T}_K$ calculated using all of the fragmentation products from the jet. In practical situations where neutrinos and K^0_L are not detected the distributions have RMS's of about 2,4 and 6 degrees for 2,3 and 4 jet events. There is also a long tail for 3 and 4 jet events. These distributions are important because they show the best agreement one can hope for between reconstructed and generated jet axes (within the Field Feynman fragmentation model).

1) Thrust Axis Determination

The determination of the thrust axis of an event is simply the reconstruction of the jet axes assuming an underlying two jet structure with the constraint that the reconstructed jets be back to back.

Figure 2 compares the error in reconstructing the thrust axis for 2 jet events using this algorithm with that from the standard thrust algorithm[8]. The error in reconstructing the thrust axis is slightly better using the present method. This is because to keep the CPU time reasonable the standard algorithm uses only a subset of the particles in an event. Figure 3 shows that the two reconstructed axes are seldom more than three degrees apart. Figure 4 is a scatter plot of the error in reconstructing the thrust axis using this algorithm versus the error using the standard method. The thrust values using the two methods agree to within .01. Similar results are obtained for the reconstructed thrust axis in 3 jet events.

2) Two Jet Reconstruction

A unique advantage of this algorithm is that for two jet events each jet is reconstructed independently. This has many potentially interesting applications, one of which I shall now describe.

Because of radiation in the initial state, the parton axes in two jet events are in general not 180 degrees apart. Figure 5 shows the angle between the generated parton axes in two jet events versus the energy of the radiated photon in the initial state[9]. Figure 6 compares the angle between the visible jet directions from the Monte Carlo[10] and the angle between the reconstructed jets . The reconstructed angle is then used to determine the energy of the initial state photon. Figure 7 compares this reconstructed energy with the photon energy predicted by the Monte Carlo. The figure shows only the result for medium energy photons although the distributions are normalized to the same number of events in the complete spectrum. For photon energies close to zero the reconstructed energy is extremely sensitive to the angle between the jets and therefore the reconstruction is not accurate in this region. Since the algorithm is primarily intended to reconstruct two jet events where the jets are nearly back to back it does not efficiently reconstruct the hard part of the photon spectrum where the jets are very close together. This could be easily rectified by including an additional configuration for these cases. However, for medium energy photons figure 7 shows that the reconstructed spectrum agrees quite well with the Monte Carlo prediction.

3) Three and Four Jet Reconstruction

The errors in reconstructing the jet directions for three jet and four jet events are determined by matching the Monte Carlo axes to the reconstructed axes. This is achieved by permuting the ordering of the Monte Carlo axes. That permutation which minimizes the differences between the reconstructed and Monte Carlo axes summed over all jets in the event determines the correspondence. The errors in the spatial reconstruction of the jets determined in this way are shown in figure 8. Using the triplicity method[13] gives results indistinguishable from figure 8. Figure 9 is shown so that the reader may compare it to the similar distribution in ref. 4. One concludes from this comparison that the reconstruction errors of the two algorithms are essentially the same.

The long tails in the error distributions occur when the fragmentation products of jets overlap to a significant extent. This happens either when the partons are close together or when the visible energy of a jet is quite small.

Energy Reconstruction

1) Three Jet Events

The raw reconstructed jet energy for jet K, defined as $E_K^R = \sum_{i \in C_K} |\vec{p}_i|$ is not a good estimator of the original parton energy. This is because the visible jet energy after fragmentation is systematically less than the parton energy, due to neutrinos and K_L^0 escaping detection. In addition, if the jets are not well separated the particles on the boundaries of the jets will be wrongly assigned. On the other hand, the jet directions are not systematically wrong. Given the angle θ_K opposite jet K, the jet energies are uniquely determined[11,12):

$$E_K = \sqrt{s}\left[\sin\theta_K/(\sin\theta_1+\sin\theta_2+\sin\theta_3)\right]$$

Figure 10 shows the resolution obtained in the reconstruction of the jet energies. The result is considerably better using the above equation rather than the raw reconstructed energies. The energy reconstruction has a sigma of about 1 GeV.

Figure 11 is the parton energy from three jet events compared with the reconstructed jet energy. The reconstruction is not particularly good below a parton energy of about 2 GeV and above 13 GeV. At low energies the jet is very broad and the jet direction is badly determined. When one jet has almost the beam energy the other two jets are very close together and the jet directions are again badly reconstructed. Figure 12 shows that the maximum reconstructed jet energy gives a much better estimate of the underlying parton thrust than does the ordinary thrust value. This means one can get a much better separation of three jet events by reconstructing the jet axes than by using the (non-perturbative) thrust value.

2) Four Jet Events

For a four jet event there is no constraint analogous to that for the energies of three-jet events. The jet energies are obtained by simply scaling the raw reconstructed energies by the quantity alpha:

$$\alpha = \sqrt{s} / \sum_K E_K^R$$

where E_K^R is the raw reconstructed energy of jet K. Figures 13 and 14 show the resolution obtained. The sigma for reconstructing the jet energies is about 2 GeV.

Timing

Table 1 compares the CPU time required by this algorithm to that needed by others. For two-jet (or thrust) reconstruction I compare with the standard thrust routine used in JADE[8]. The time needed for three-jet reconstruction is compared with the triplicity routine used in JADE[13] and by TASSO[3]. The four-jet reconstruction is compared with TASSO's algroithm[4].

The table shows that the present algroithm is approximately fifty times faster for thrust axis determination from two to one hundred times faster for three-jet reconstruction and ten to one hundred times faster for four-jet reconstruction. There is only a very weak (approximately linear) dependence on event multiplicity.

How to Use the Routine

The program exists on the libraries 'F11LHO.JADEGS/L'.

1) Input

The user must fill the following common block before each call to the subroutine:

```
    COMMON/SENSE/PP(4,100),INUM
```

The array PP contains INUM particles to be used in the reconstruction. The momenta in the X,Y,Z directions for track K should be stored in PP(1,K), PP(2,K), PP(3,K). In location four the user may store a track index. It will be copied into the output common block of the corresponding track.

The program is invoked by 'CALL MCGJET(NJET,Y)' where NJET and Y are both input arguments. NJET is the number of jets to be reconstructed and Y is a normalized vector in the direction of the event plane normal. For all the results here I have taken Y to be the eigenvector corresponding to the smallest eigenvalue of the sphericity tensor. Y need not be filled for NJET = 2.

?

call spray

2) Output

Normalized vectors corresponding to the reconstructed jet axes are stored in the:
COMMON/MARKET/PAR(3,4)
Where PAR(1,K),PAR(2,K),PAR(3,K) are the X, Y, Z direction cosines of reconstructed jet axis K.

The track assignment for each jet axis is in the:
COMMON/COLD/IJ(4), PTH(3,100,4),IPJ(100,4)

IJ(J) is the number of tracks associated with jet J
PTH(1-3,K,J) contains the X,Y,Z momentum components for the KTH track associated with jet J. K runs from 1 to IJ(J).

IPJ(K,J) is the (input) track index for track K of jet J.

The thrust axis and thrust value are stored in the:
COMMON/CRUST/AXIS(3), THR.
The thrust axis is determined (using a balanced set of vectors).
For any call to MCGJET with NJET > 2. When NJET = 2 only the particles in the COMMON/SENSE/ are used.

AX 1 Sphericity
AX 2
AX 2 1
2
AX 3
3 1 - 3
3 -1

References and Footnotes

1) I do not consider cluster-finding algorithms which search for clusters of particles in events rather than a pre-determined number of jet axes.

2) S. Brandt, H. Dahmen, Z. Physik C1, 61 (1979)
3) S. Wu, G. Zobernig, Z. Physik C2, 107 (1979)
4) S. Wu, DESY 80/127 (1980)

5) The two configurations for three-jet events are:
$$\vec{T}_1^{(o)} = \vec{x}\ ,\ \vec{T}_{2,3}^{(o)} = -\vec{x} \pm \vec{z}$$
plus one configuration with the above axes reversed.
$\vec{X}$ is the thrust axis (previously determined from TWO-Jet reconstruction) and $\vec{Z}$ is in the event plane perpendicular to $\vec{X}$.

6) A.Ali, E. Pietarinen, G. Kramer, J.Willrodt, Phys. Lett. 93B, 155 (1980)
7) R. Field, R. Feynman, Nucl. Phys. B136,1 (1978)

8) Coded by E. Elsen

9) F. Berends, R. Kleiss, DESY 80/73 (1980)

10) The visible jet direction is the vector sum of the visible fragmentation products.

11) J. Ellis, M. Gaillard, G. Gross, Nucl. Phys. B111, 253 (1976)

12) This equation is valid only if all angles are less than 180 degrees.
13) Coded by T. Kobayashi

Figure Captions

1) The angle between the generated parton axis and the axis determined using all of the fragmentation products of the jet, for two-jet (A), three-jet (B) and four-jet (C) events. The open circles correspond to neutrinos and K_L^0 not being included in the axis determination.

2) The difference between the generated two-jet axis and the reconstructed thrust axis for the standard thrust algorithm (closed circles) and this algorithm (open circles).

3) The difference between the thrust axis determined using the two algorithms.

4) The error in reconstructing the thrust axis using this algorithm vs the error using the standard algorithm.

5) The angle between the parton axes vs the energy of the initial state photon for two-jet events.

6) The angle between the visible jet directions. The circles are the results from reconstruction and the histogram is the Monte Carlo prediction.

7) Energy of the initial state photon. The circles are the reconstructed energy and the histogram is the Monte Carlo prediction. The two distributions are normalized to the same number of events in the complete spectrum.

8) The error in the spatial reconstruction of the jet axes for three jet (a) and four-jet (b) events determined by matching the reconstructed axes to the Monte Carlo parton axes.

9) The angle between the reconstructed jet axis and the closest visible jet direction, for four-jet events.

10) The difference between the reconstructed jet energy and the energy of the parton given by the Monte Carlo for three-jet events. The open circles correspond to using raw reconstructed energy and the closed circles are using the energy constraining equation.

11) Parton energy spectrum for three-jet events. The closed circles show the reconstructed jet energy and the histogram is the parton energy from the Monte Carlo.

12) The difference between the reconstructed thrust and the parton thrust from the Monte Carlo for three-jet events. The open circles show the result using the (non-perturbative) thrust and the closed circles using the maximum reconstructed jet energy.

13) The difference between the reconstructed jet energy and the energy of the parton given by the Monte Carlo for four-jet events. The open circles correspond to using raw reconstructed energy and the closed circles are using the energy scaling equation.

14) Parton energy spectrum for four-jet events. The closed circles show the reconstructed jet energy and the histogram is the parton energy from the Monte Carlo.

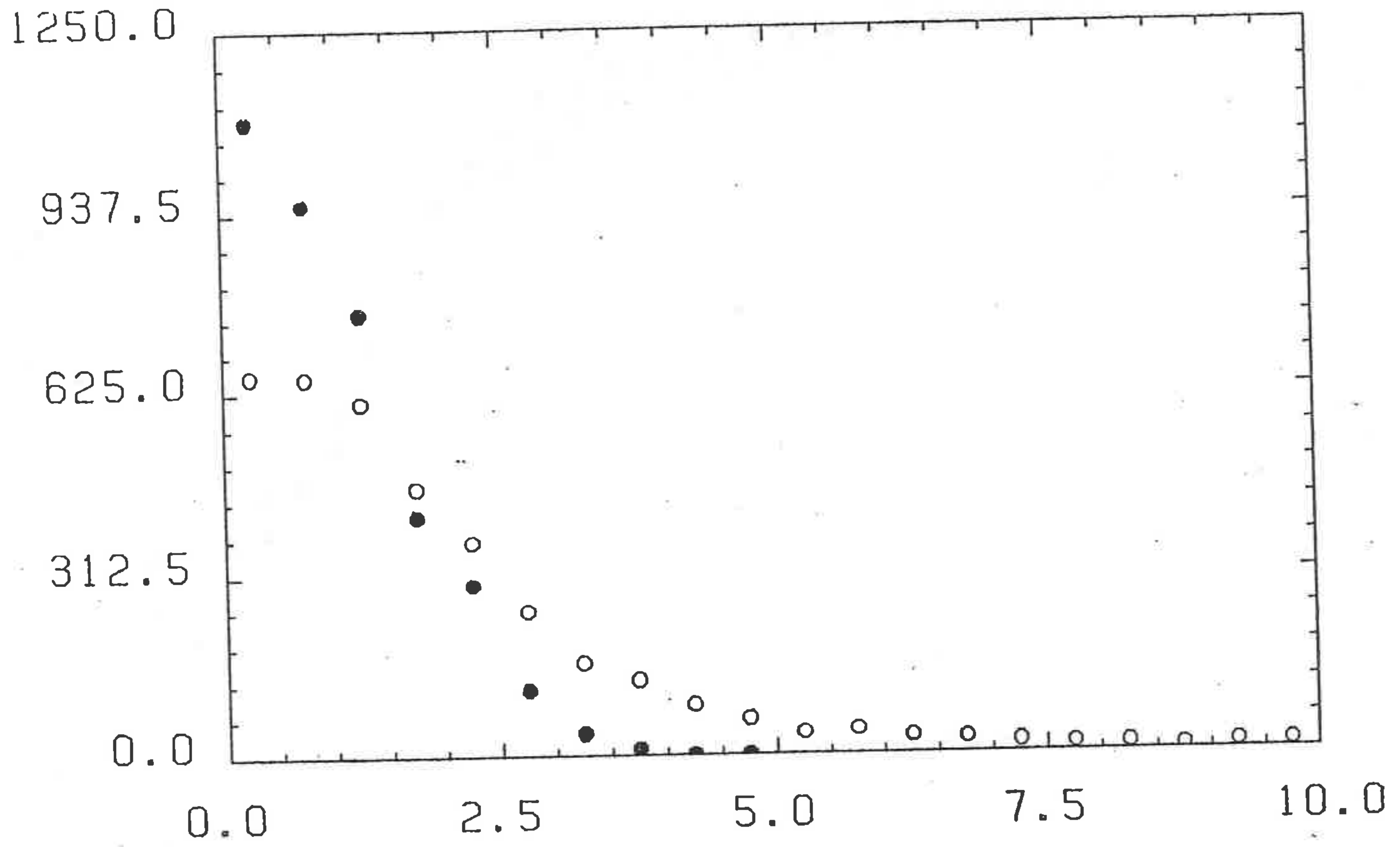
Table 1

Number of Events Reconstructed per Sec

	Multiplicity				
	15	19	22	28	35
2-jet Thrust	460 9	390 7	350 7	300 7	
3-jet Triplicity Gen Spher	70 30 2	55 14	45 9 0.7	35 4	
4-jet TASSO 4-J	42	32	28 0.5	22	 0.1

1(a)

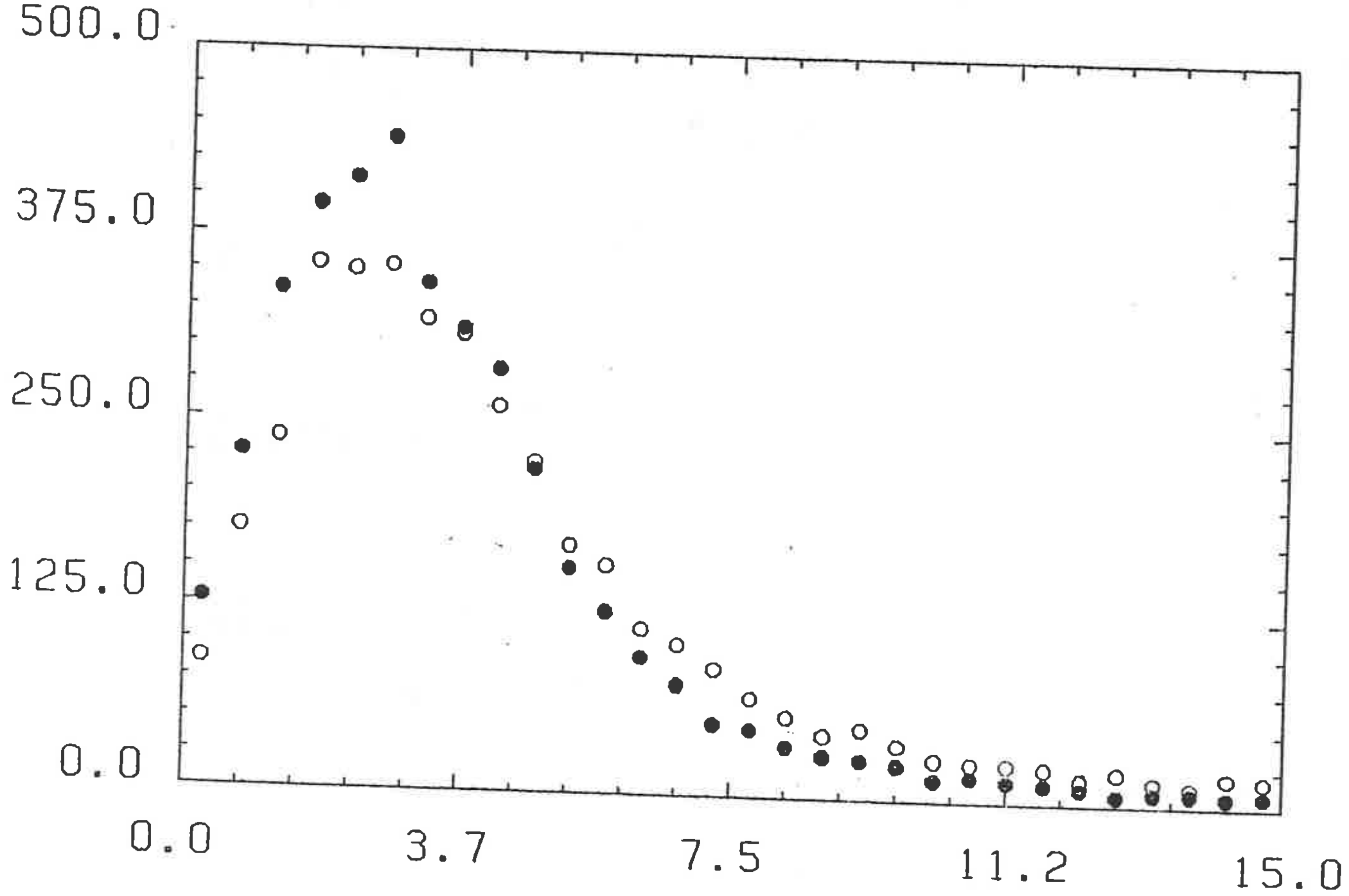
DSN=F11G0D.GEP.ERR1
05/03/81 KA 1 1
02.21.35 KB 1 5
KC 0 0
NSYM -2 -12



TWØ-JET
ERRØR IN RECØNSTRUCTING JET AXIS FRØM FRAG PRØDUCTS (G=0)

1(b)

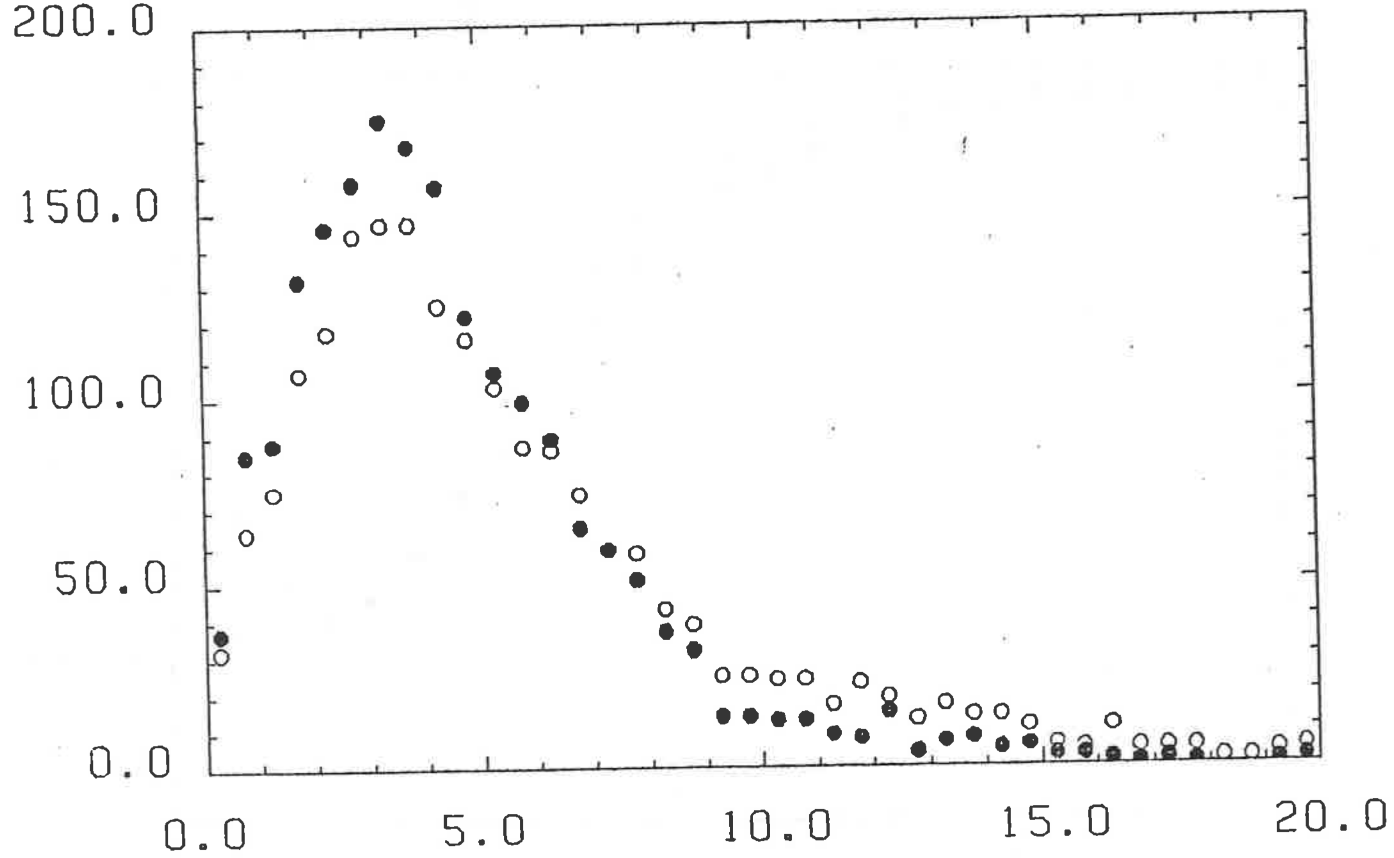
DSN=F11GØD.GEP.ERR3
05/03/81 KA 1 1
02.30.16 KB 1 5
KC 0 0
NSYM -2 -12



THREE-JET
ERRØR IN RECØNSTRUCTING JET AXIS FRØM FRAG PRØDUCTS (G=0)

1(b)

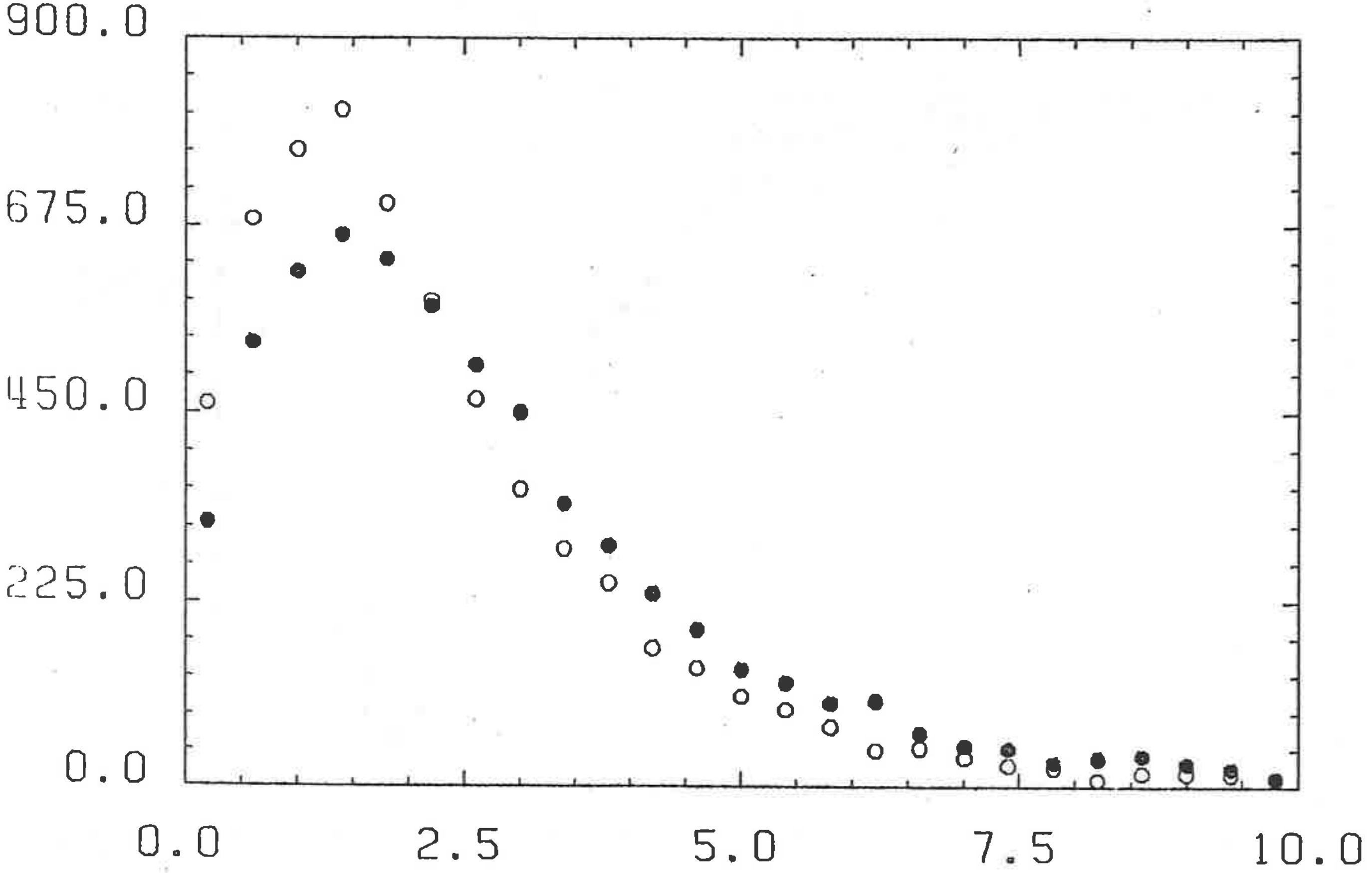
DSN=F11G0D.GEP.ERR4
05/03/81 KA 1 1
02.37.04 KB 1 5
KC 0 0
NSYM -2 -12



FOUR-JET
ERROR IN RECONSTRUCTING JET AXIS FROM FRAG PRODUCTS (G=0)

1 (c)

DSN=F11G0D.GEP.T2J4V
05/03/81 KA 1 1
05.18.50 KB 506 507
KC 0 0
NSYM -12 -2

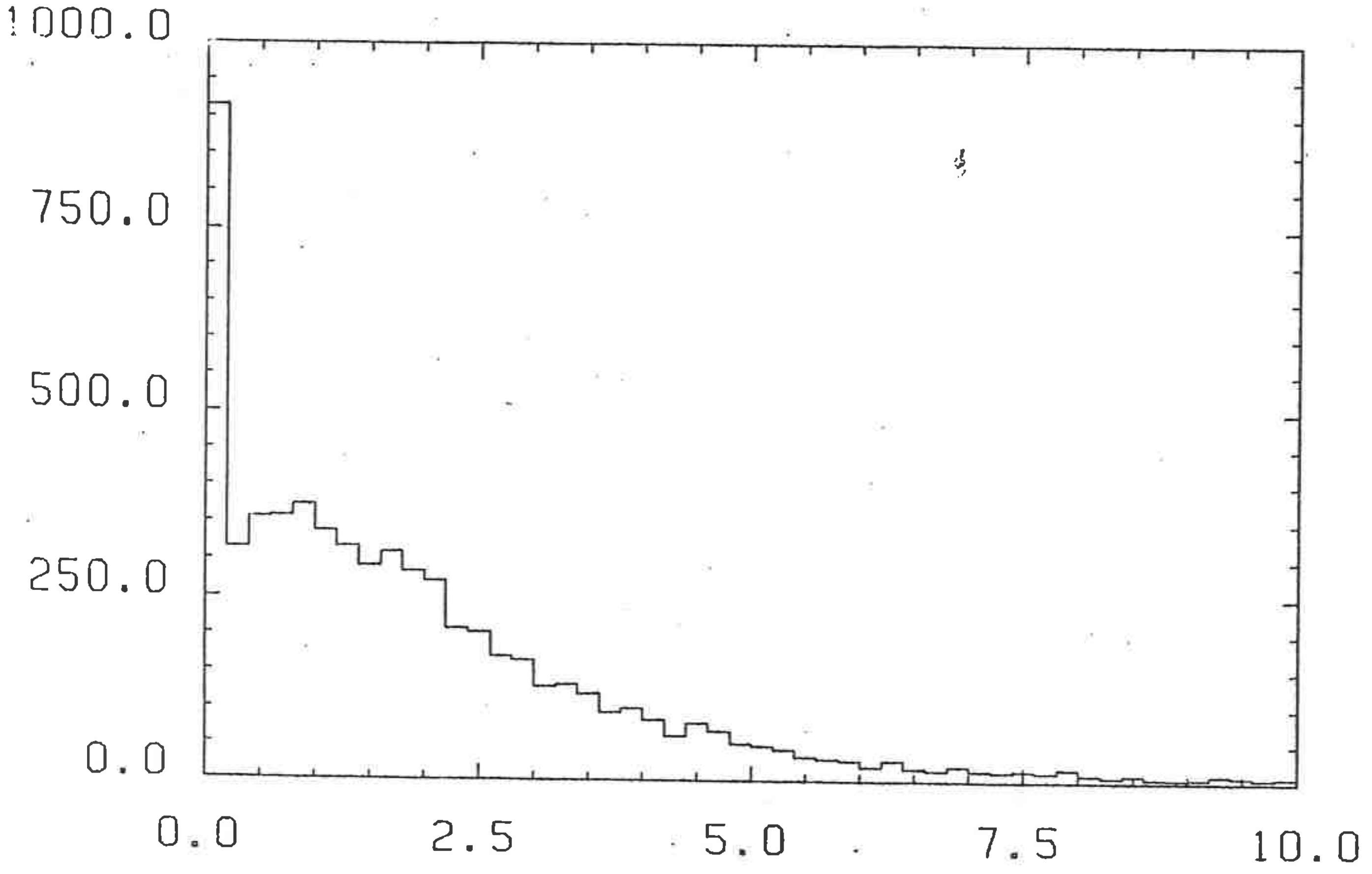


TWO-JET 4 VECT
ERR THRUST REC

2

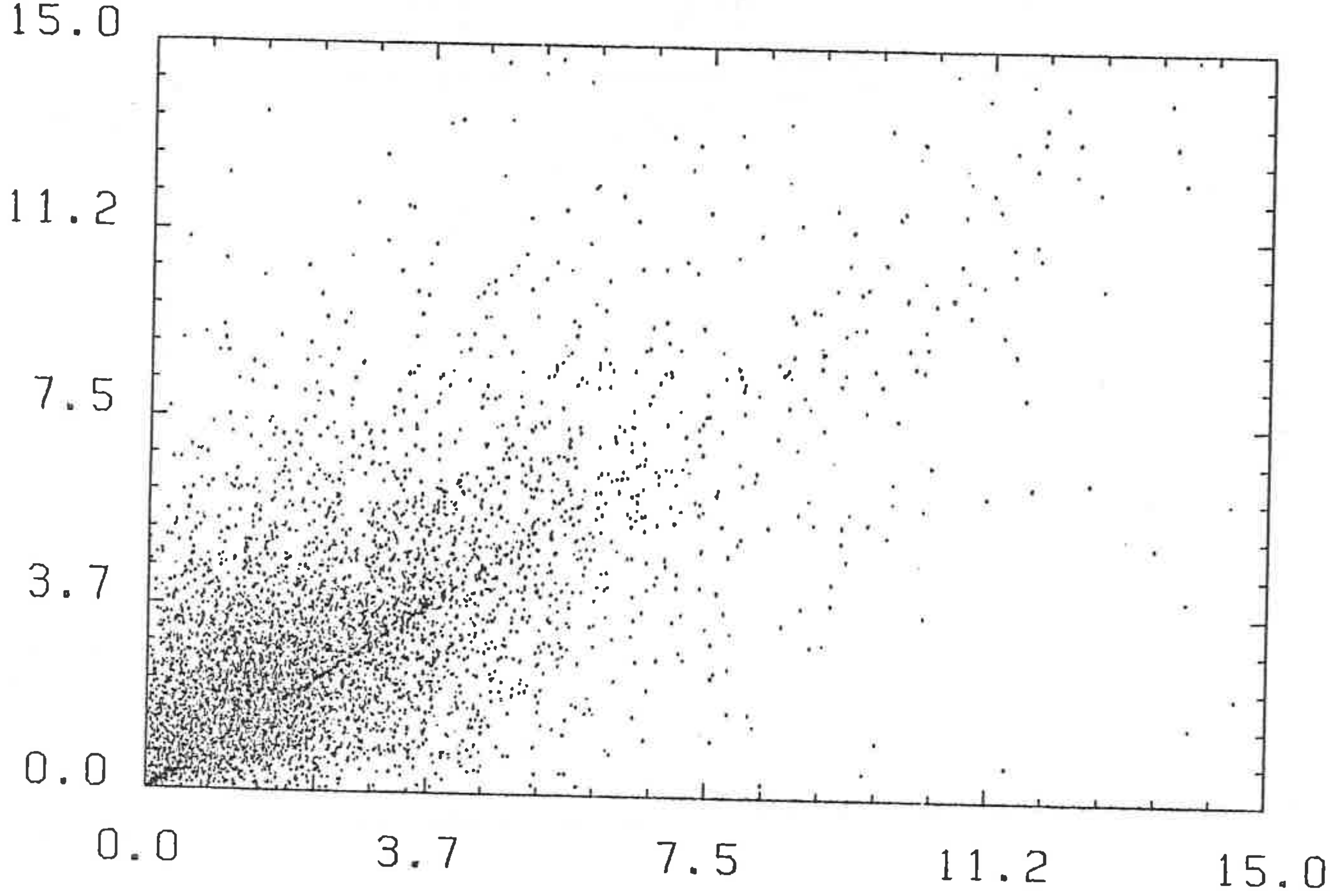
3

DSN=F11G0D.GEP.T2J4V
05/03/81 KA 1
05.04.43 KB 502
KC 0
NSYM 10



TWO-JET 4 VECT
DIFF AXES

DSN=F11G0D.GEP.T2J4V
05/03/81 KA 2
05.21.37 KB 508
KC 0
NSYM -3

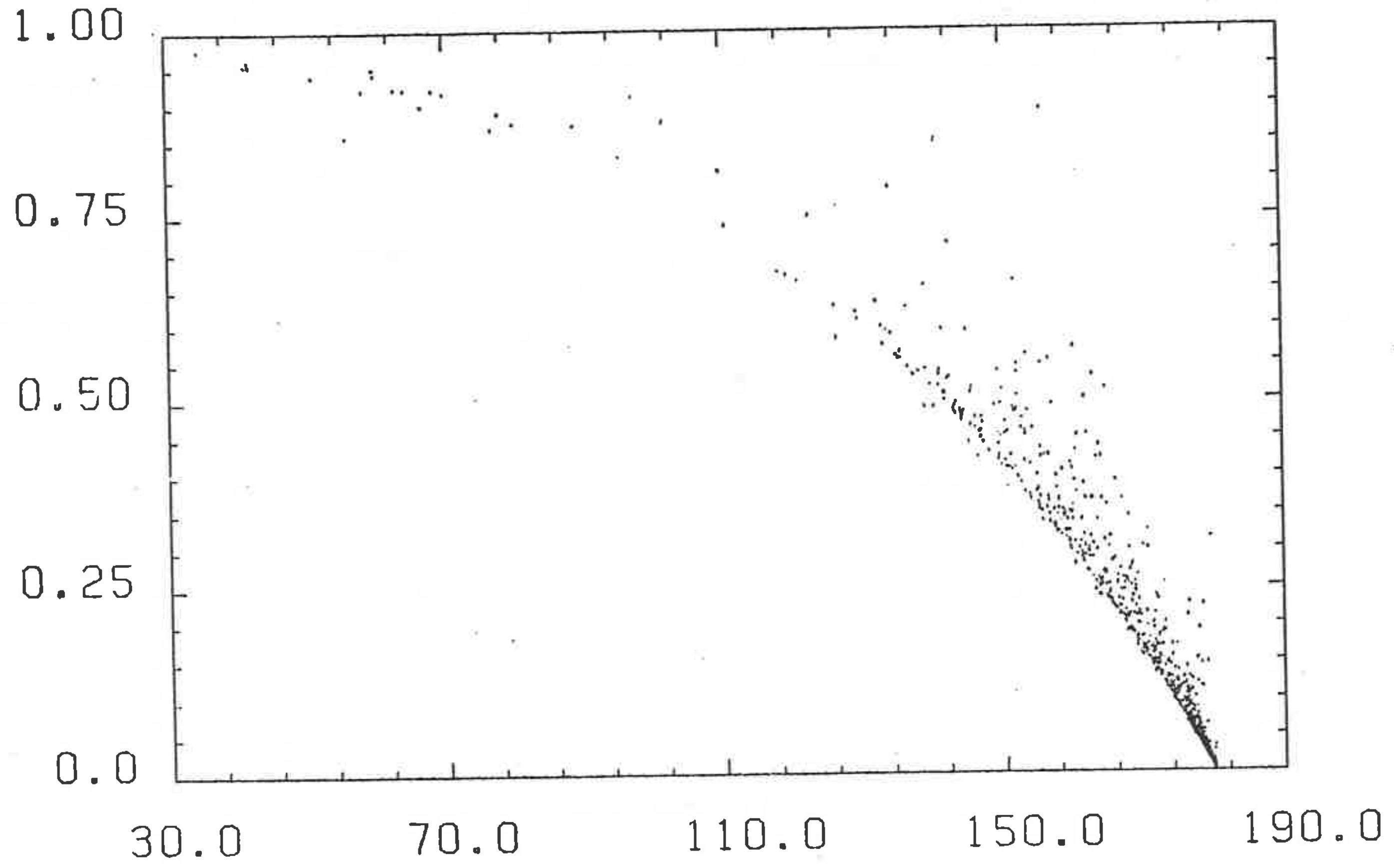


TWØ-JET 4 VECT
ERR THR REC VS ERR THR

4

5

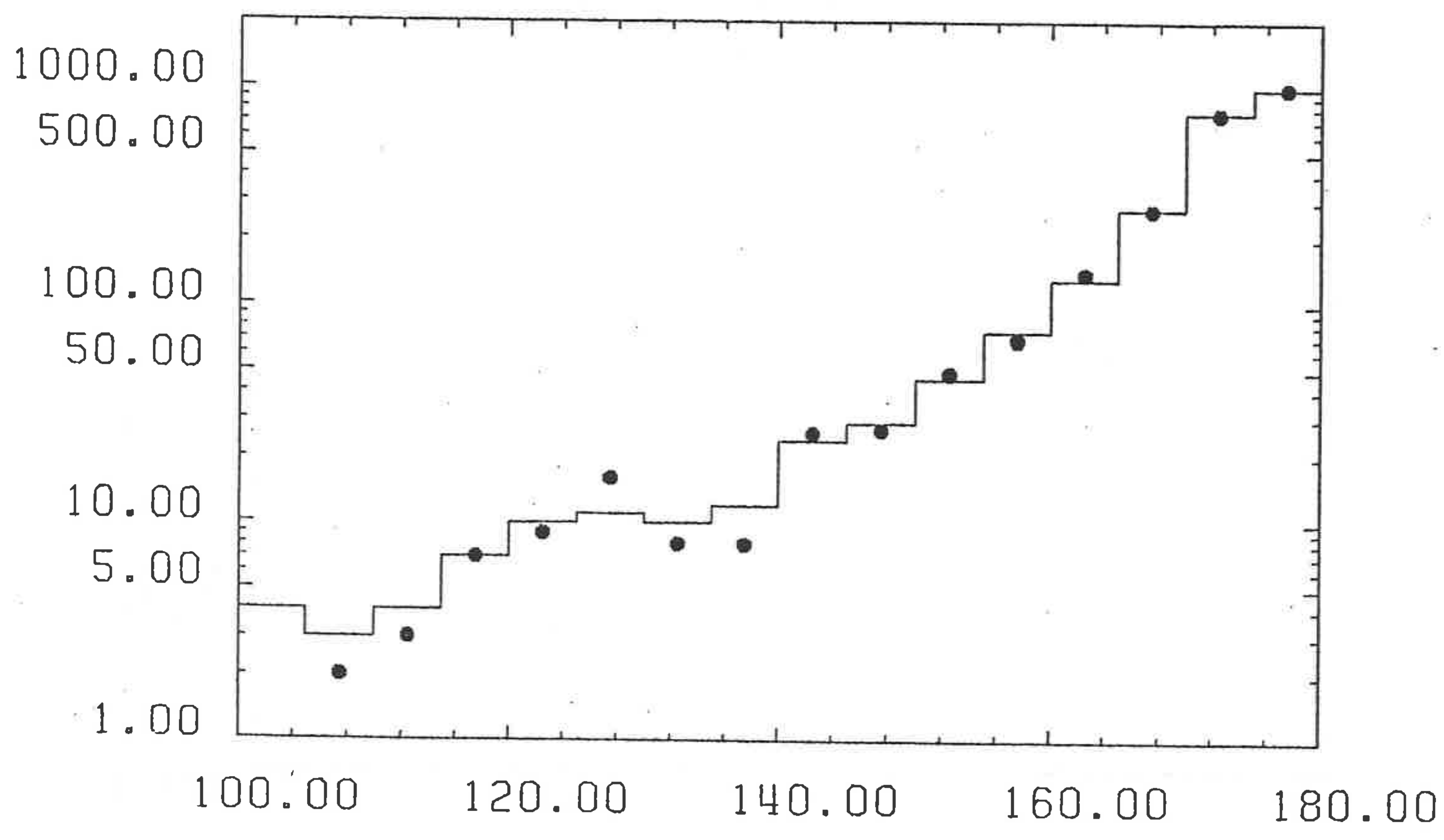
DSN=F11G0D.GEP.T2JTRK
05/03/81 KA 2
06.35.23 KB 511
KC 0
NSYM -3



TWØ-JET TRACKED
ANGLE MC VS ERAD

6

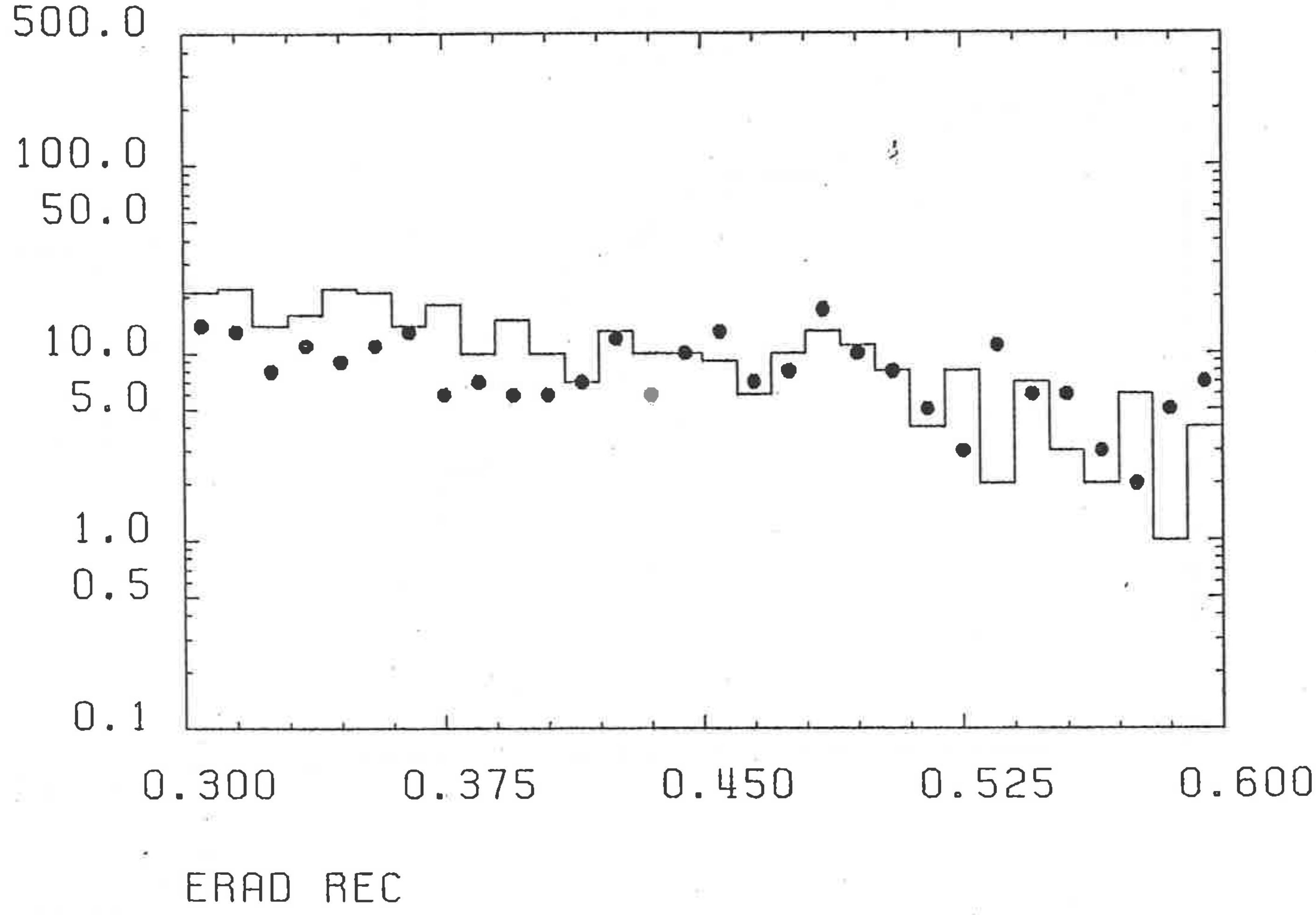
DSN=F11G0D.GEP.T2JTRK4V
11/03/81 KA 1 1
23.47.51 KB 498 1499
KC 0 0
NSYM -2 10



TWO-JET 4V
ANGLE VIS MC

7

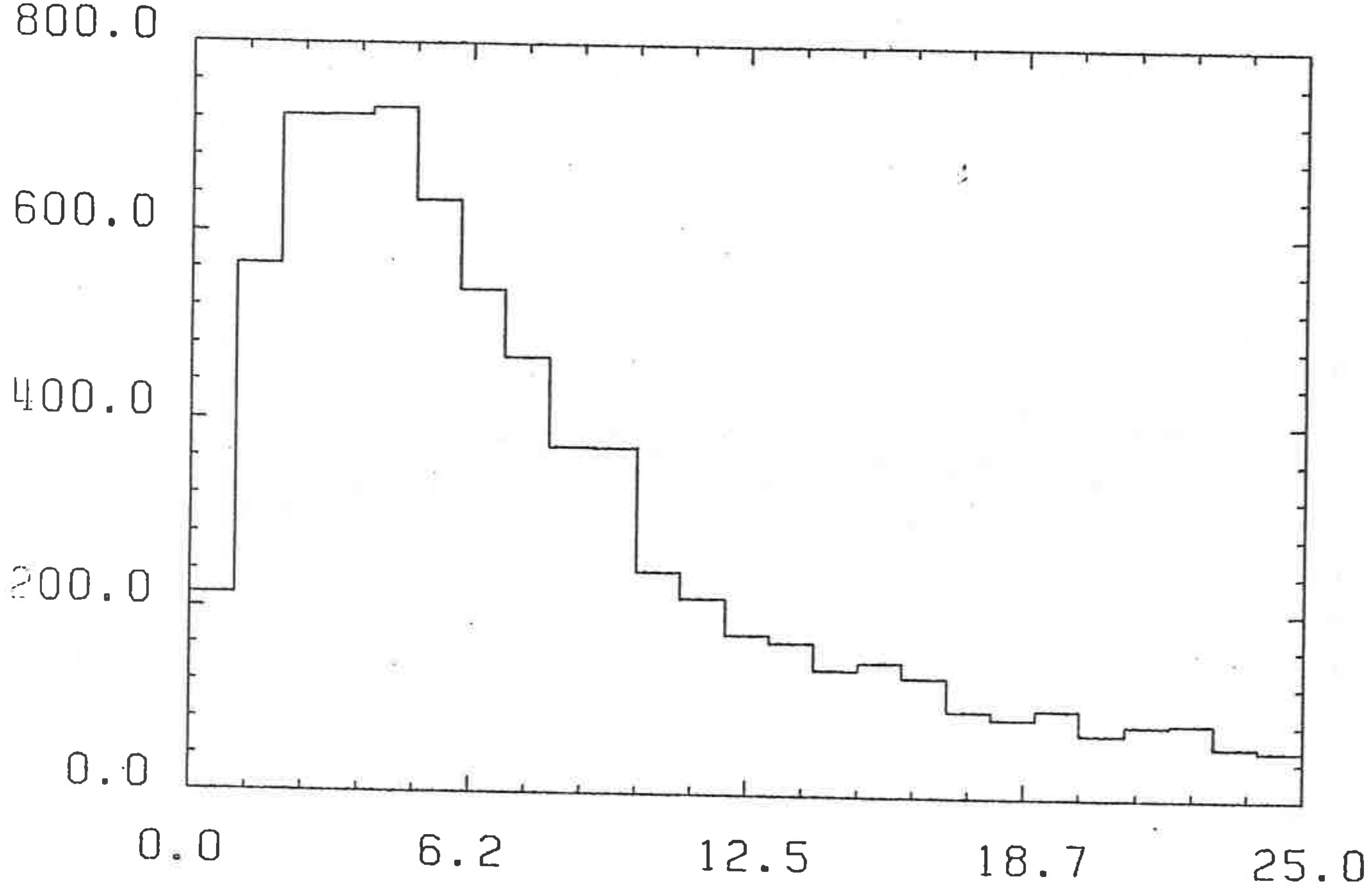
DSN=F11GOD.GEP.T2JTRK4V
12/03/81 KA 1 1
15.42.14 KB 176 689
KC 0 0
NSYM 10 -2



7

8(a)

SN=F11G0D.GEP.T3J4V
5/03/81 KA 1
6.59.26 KB 20
KC 0
NSYM 10



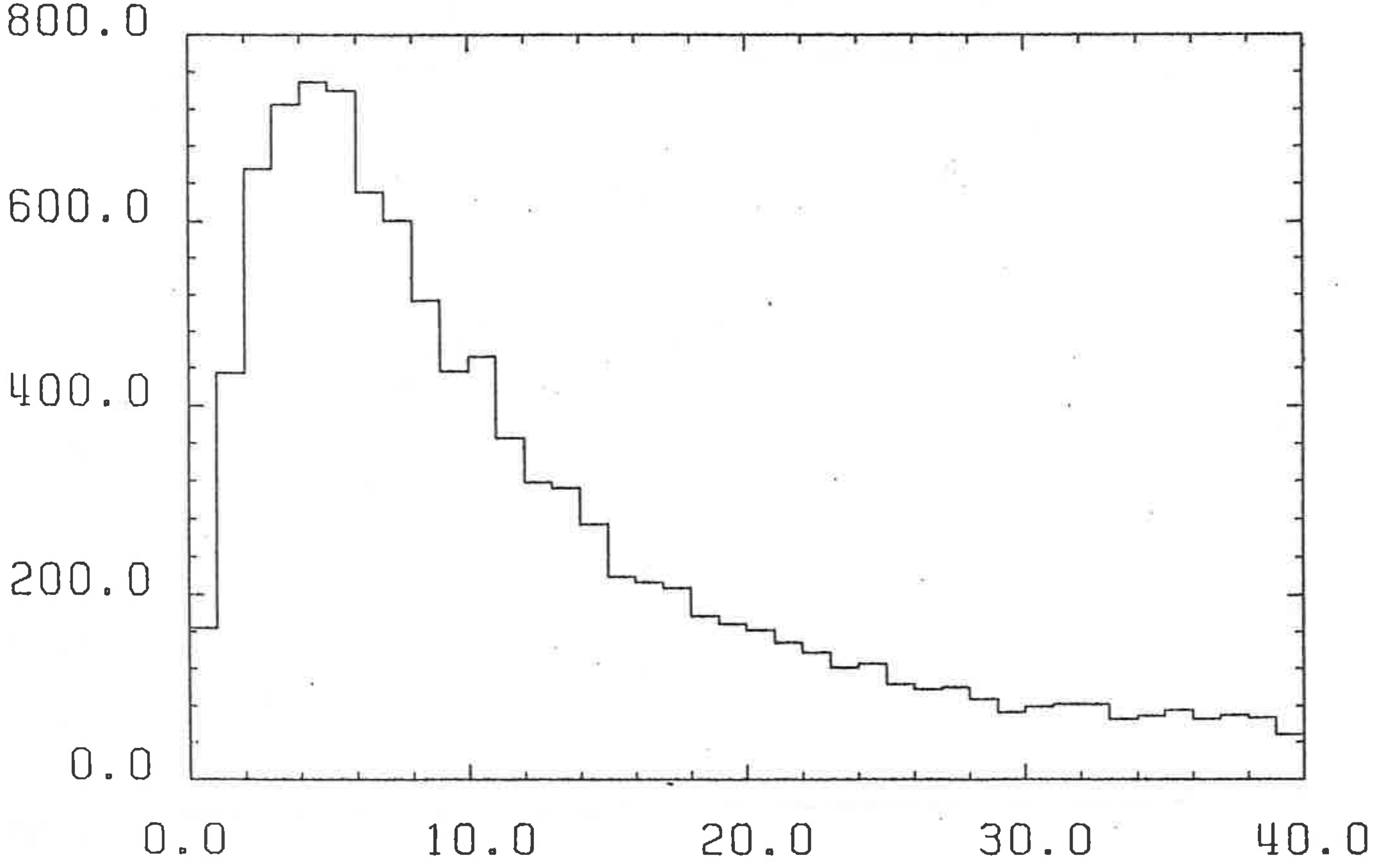
THREE-JET 4V
ERRØR (WITH MATCHING)

8 (a)

8(b)

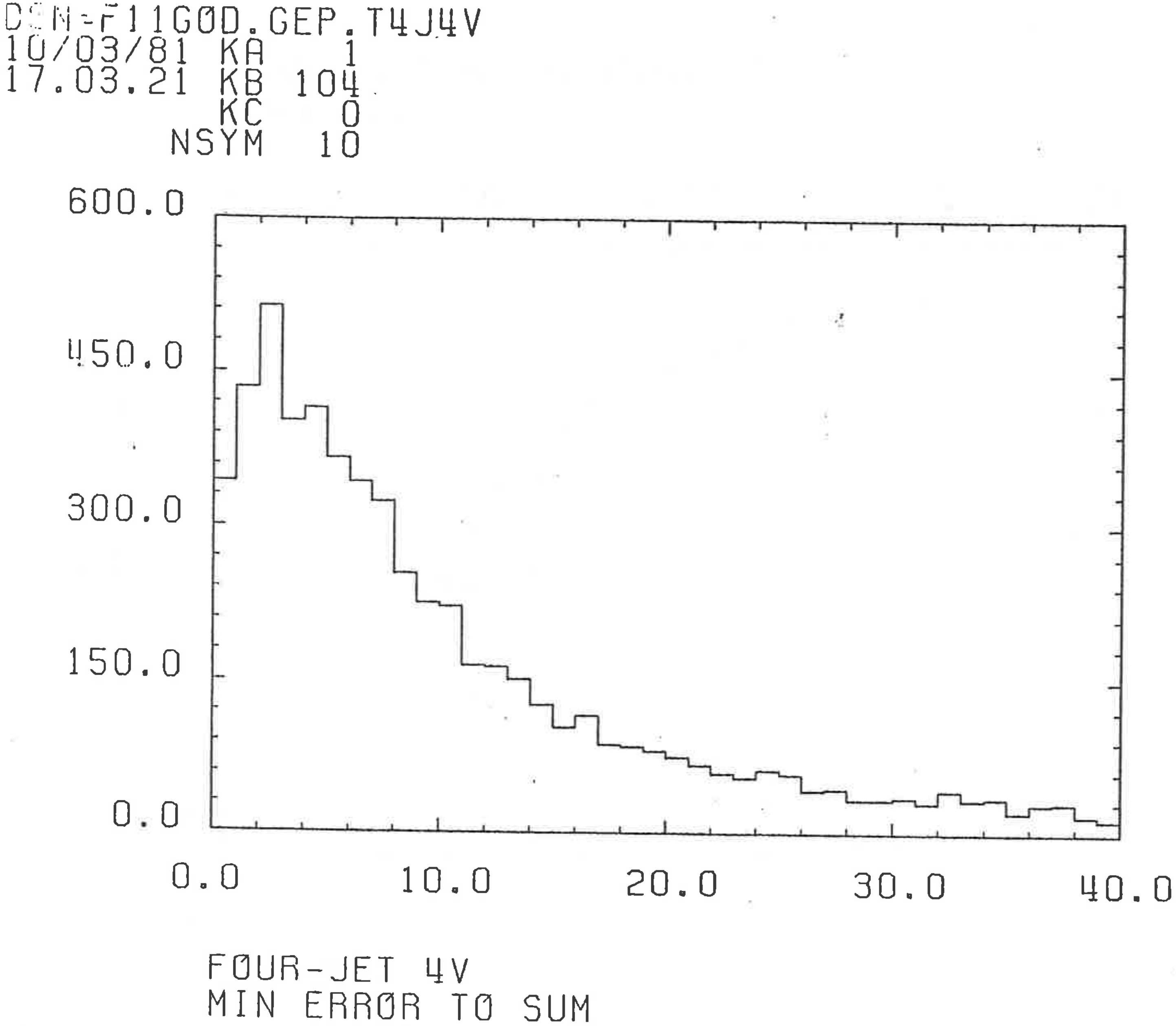
8(b)

DSN=F11G0D.GEP.T4J4V
05/03/81 KA 1
14.14.00 KB 20
KC 0
NSYM 10

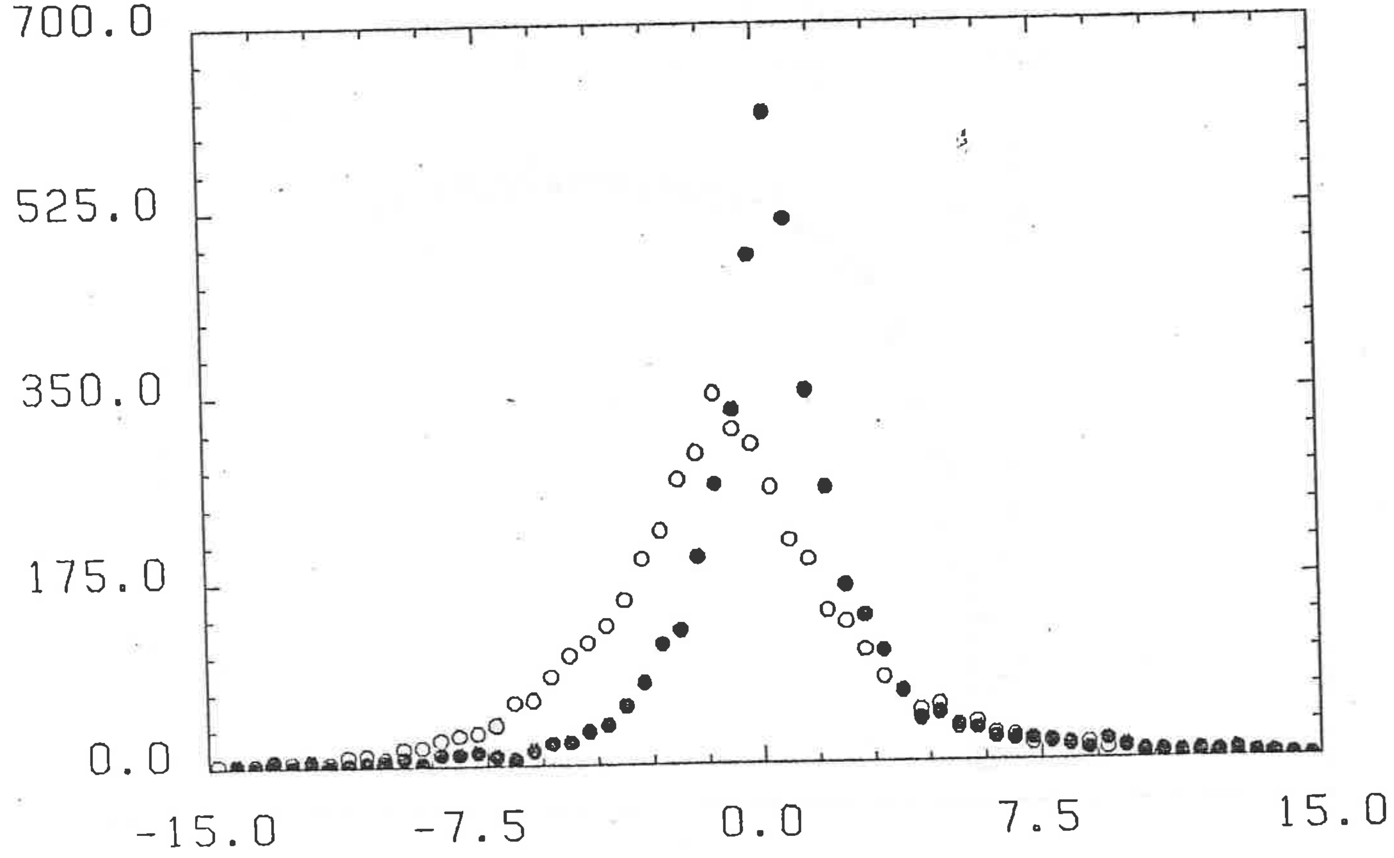


FØUR-JET 4V
ERRØR (WITH MATCHING)

9


DSN=F11G0D.GEP.T4J4V
10/03/81 KA 1
17.03.21 KB 104
KC 0
NSYM 10
600.0
450.0
300.0
150.0
0.0
0.0
10.0
20.0
30.0
40.0
FØUR-JET 4V
MIN ERRØR TØ SUM

N=F11G0D.GEP.T3JTRK
/05/81 KA 1 1
.13.05 KB 59 158
KC 0 0
NSYM -12 -2



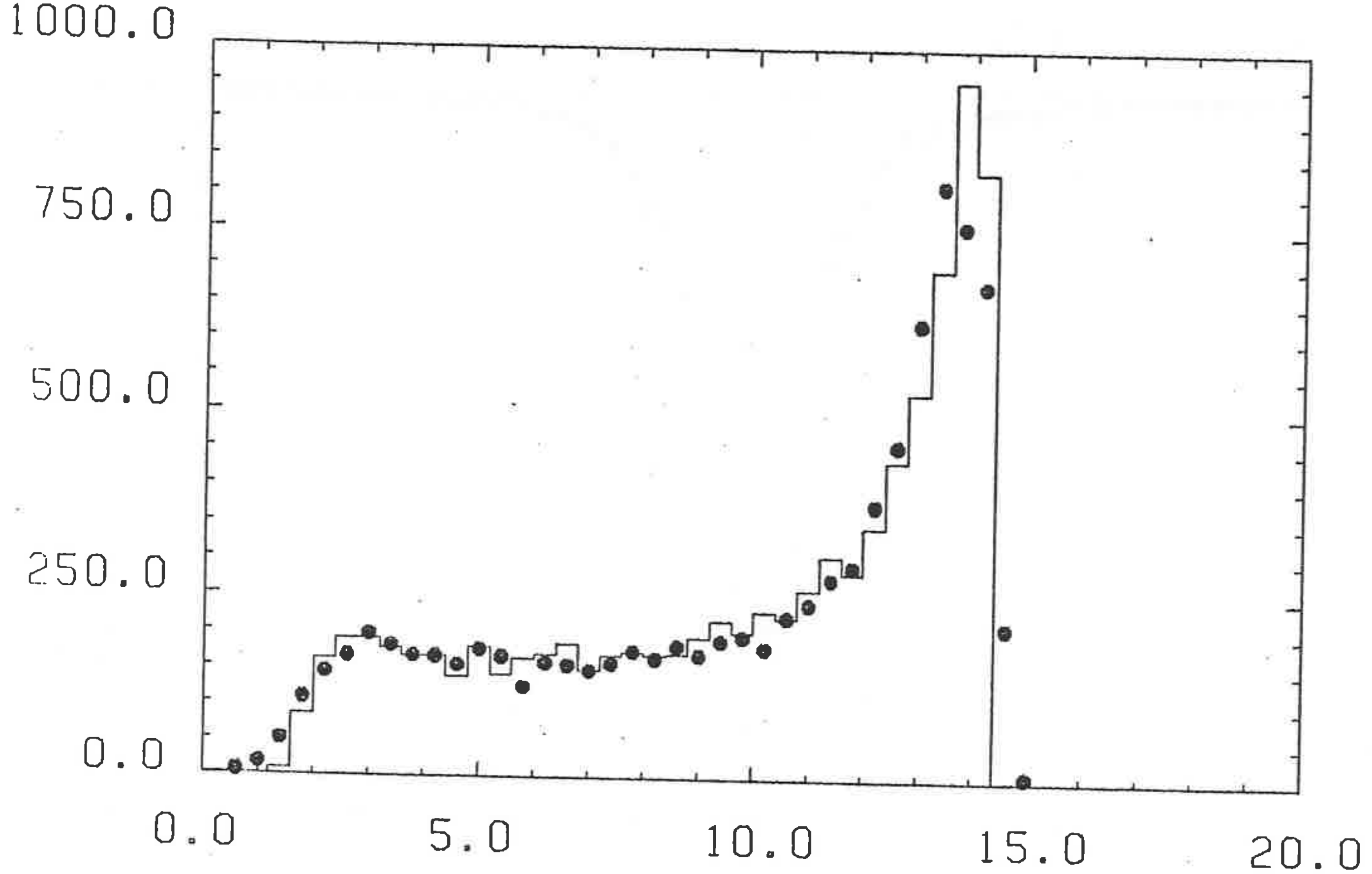
THREE-JET TRACKED
~~VIS JET E~~ EREC-PARE

10

START OF PAGE 1

11

DSN=F11GOD.GEP.T3J4V
06/04/81 KA 1 1
14.36.00 KB 9 12
KC 0 0
NSYM 10 -2



THREE-JET 4V
PARTON E

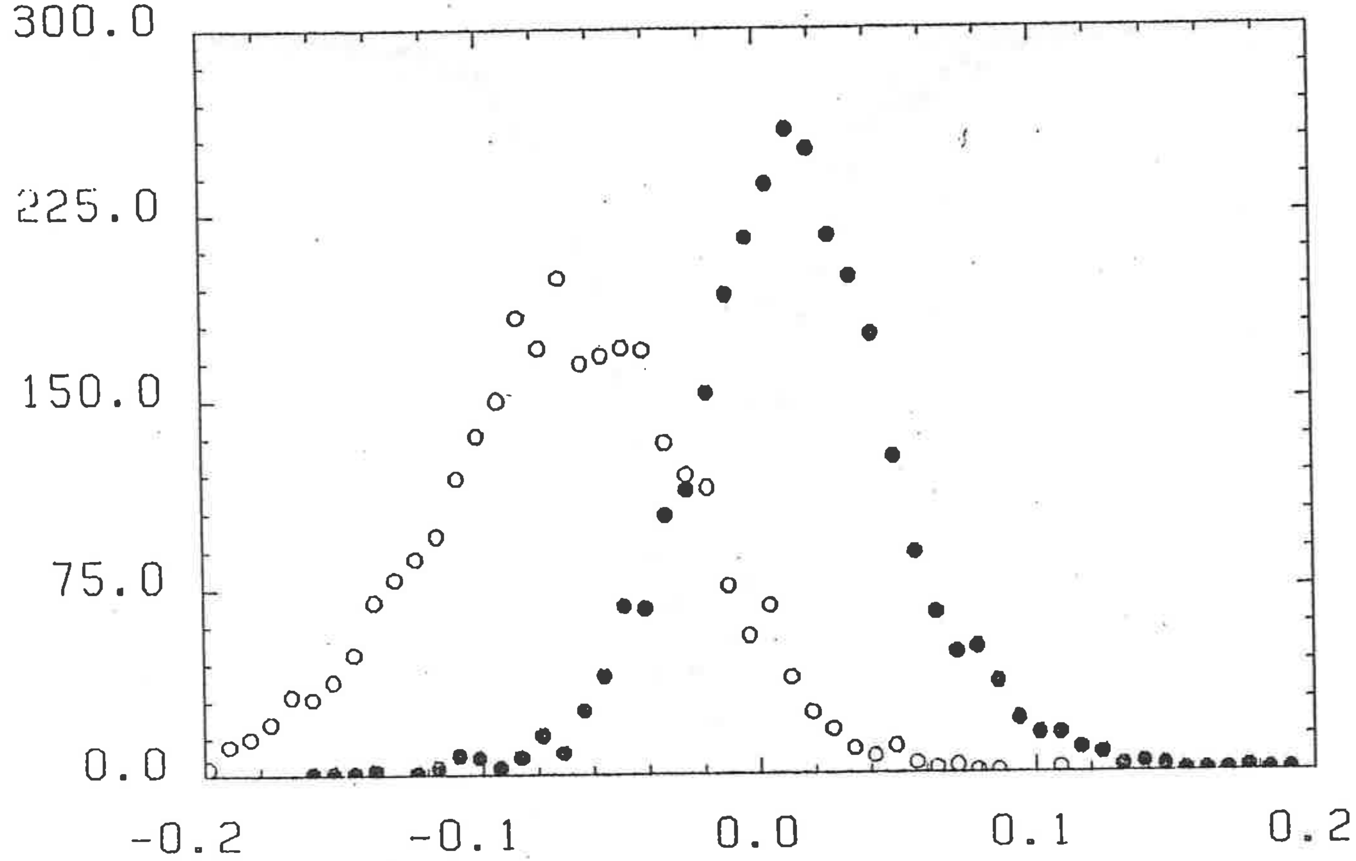
11

END OF PAGE 1

START OF PAGE 1

12

DSN=F11G0D.GEP.T3J4V
05/03/81 KA 1 1
07.25.13 KB 287 288
KC 0 0
NSYM -12 -2



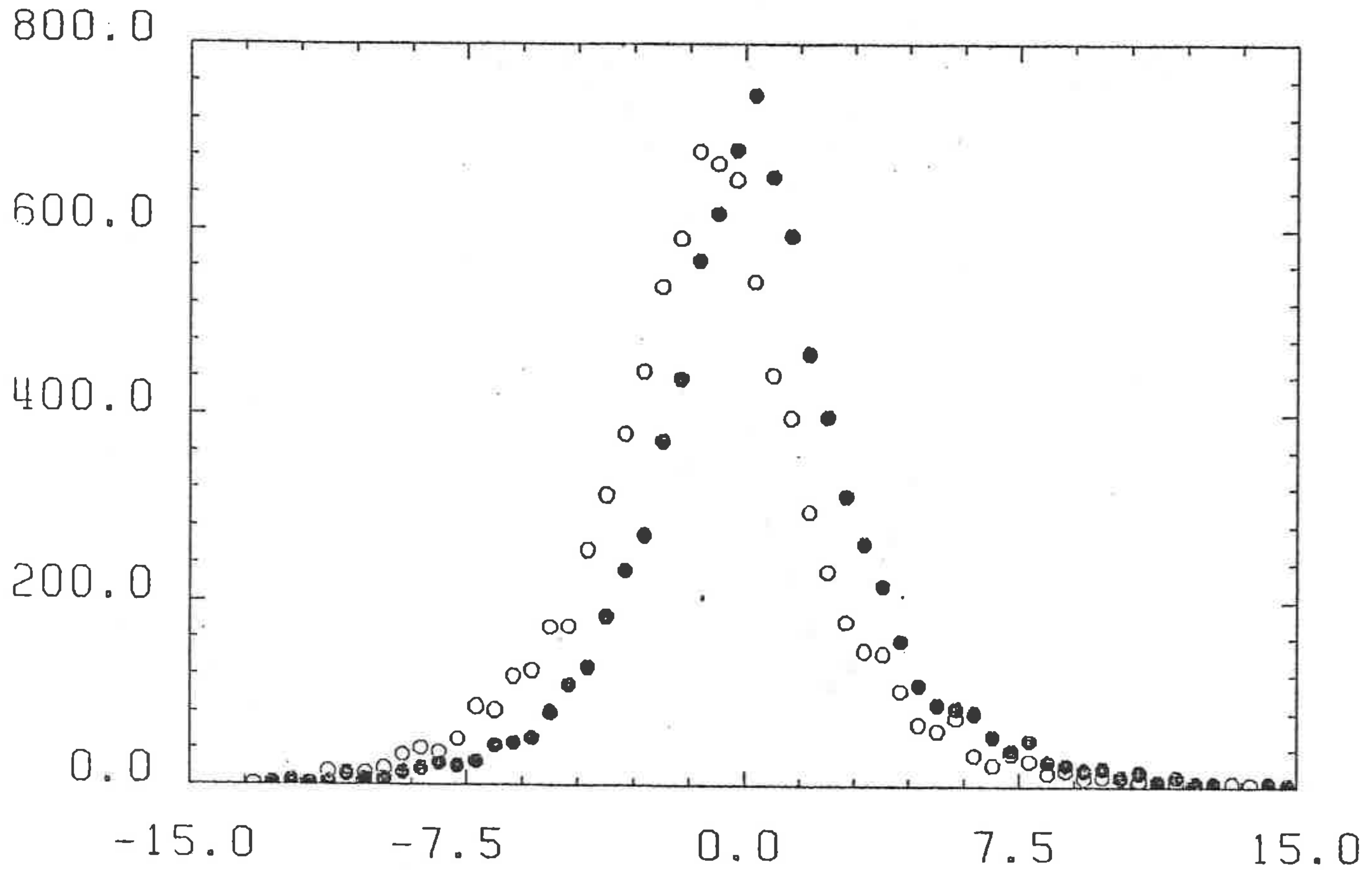
THREE-JET 4V
THRUST - THPAR

12

13

13

```
SN=F11GØD.GEP.T4JTRK
5/05/81  KA    1    1
1.15.10  KB   59  158
         KC    0    0
       NSYM  -12   -2
```

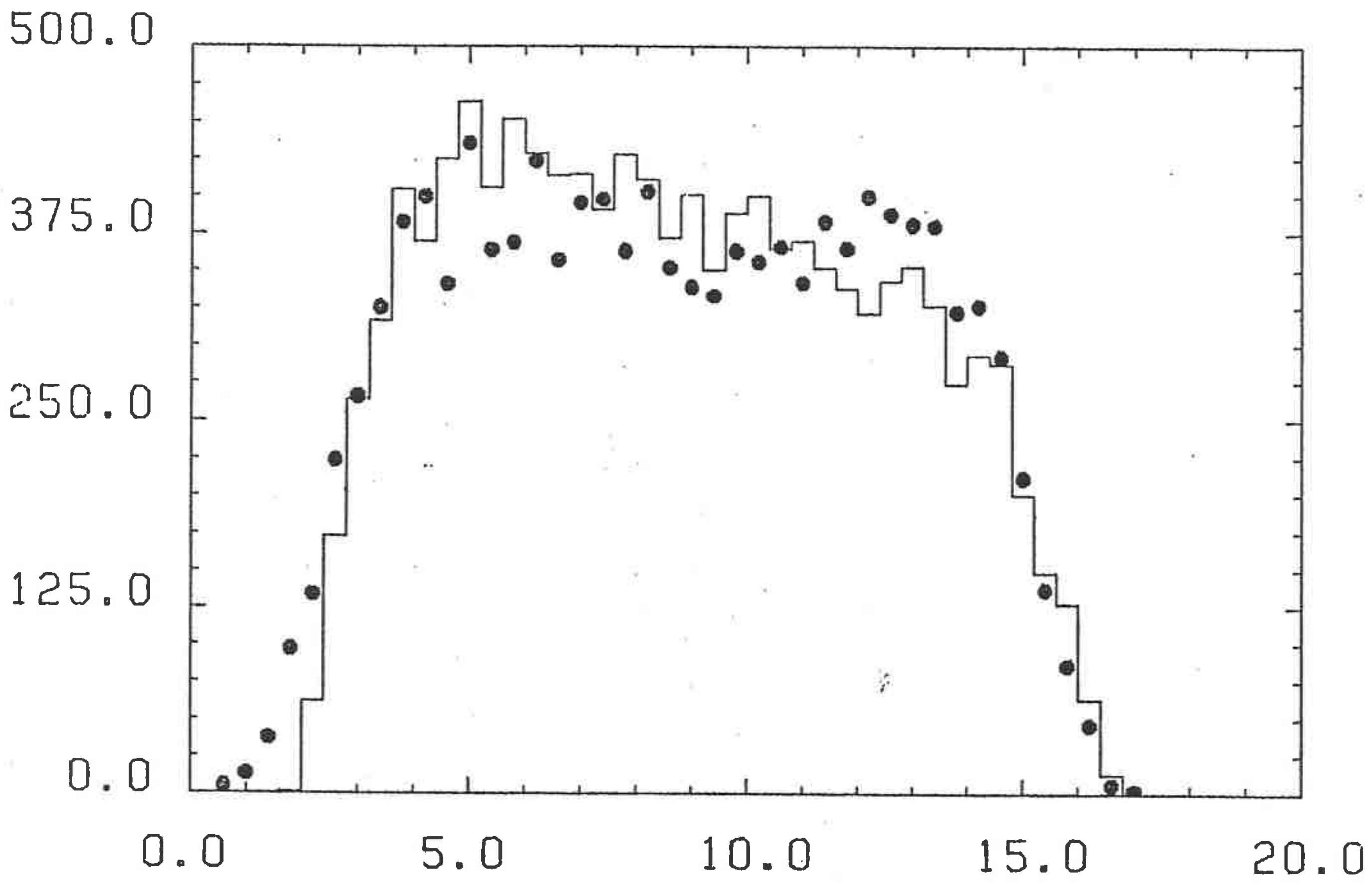


FØUR-JET TRACKED

~~VIS-JET-E~~ EREC-PAR E

14

DSN=F11GOD.GEP.T4J4V
10/03/81 KA 1 1
17.06.22 KB 9 12
KC 0 0
NSYM 10 -2



FØUR-JET 4V
PARTØN E