SNO-STR-94-031

Box Fitter

Tom Andersen, J. Law, J.J. Simpson
University of Guelph June 22, 1994

This document describes the operation and performance of a event fitter for the Sudbury Neutrino Observatory.

Basic Principle of Operation

This fitter calculates a three dimensional probability map for each event: i.e.: A probability density distribution in a region surrounding the event location.

Multiple Box generation

The fitter operates by finding places(boxes) in the detector where photons could have originated from. It does this by starting with a box big enough so that we know that all of the photons in an event came from this box. A box 20m on a side will enclose the PMT array fully, so this is the starting box, known as a the 1st generation box. This box then subdivides it self into 8 subboxes, and each of these boxes checks to see whether some fraction of the photons that made pmt's fire could have originated inside it . If a box survives this test, it then subdivides into 8 subboxes. The process repeats for 6 to 8 generations, giving rise to 1000 to 20000 final generation boxes, each of these boxes being a possible departure point for some fraction of the photons that caused the observed pmt hits.

Note:
Any discussion of time refers to the "departure time"
departure time <==> the time a photon would have had to have left at in order to generate an observed pmt hit.

Box Scoring

The final generation boxes are scored using chi square and the gamma function, giving a probability, or "mass" for each box.
To get chi squared for a box, we first calculate an event time t_o, as seen from that box,(the average of all of the departure times for photons that may have come from the box in question)
and then calculate the chi square using this average time as the event time. Each box then has its own idea of what time the event happened at. Sigma is taken as the tube jitter time (1.6ns) plus the light travel time across a box.

v is the number of degrees of freedom,
X^2 is the chi-square for the box.

$X^2 = \sum ((t - t_0)/\mathrm{Sigma})^2$. Summing over the hits within a 2 Sigma window width(sum over nFittedHits)

v = nFittedHits - 1.

mass = $Q(0.5*v, 0.5*X^2)$, Q is the incomplete gamma function, or the chi-square probability function. (see the discussion in Numerical Recipes in C section 15.1)

Once the masses for each of these thousands of boxes is calculated, We determine the centre of mass, and the principle axes of the resulting mass distribution.
We use the centre of mass as the best fit position, and also find the moment of inertia along each of the three principle axes. These moment of inertia eigenvalues are normalized, so that they can be compared event to event.

Data Interpretation and Results

We ran the fitter on most of the standard data events produced using the Queen's code. It should be noted that there are several parameters that have not been fully explored in order to find the best performance. Speed, as you may have guessed is not phenomenal: dealing with 20k boxes, each of which needs chi - square calculated, among other things makes this fitter slow(seconds per event). This may be sped up some.

This fitter can be used to tell the type of event - see the last few pages of the report for results on separating neutral current from charged current results.

Comments:

Once a probability density distribution is found there are many things that can be done with it that we have not tried. We have not investigated the size of the distributions, other than the observation that they are about the same size as the uncertainty in the fit. Do pmt beta gammas that reconstruct in the central volume of the detector have a distinctive probability distribution? - this may help in removing these events from the central volume. There seems to be lots of things that may be done. The moment distribution for neutral current events looks different from the charged current events.

As an example of the games that can be played with the fitter, we tried a simple trick to get rid of low energy background events from the centre of the tank. : If more than 12% of the mass of an low energy event was at r > 550 cm, we tagged this as probably being from the pmt array, and moved the event outside of the 550 cm mark. (Fancy Fit option in the graphs)

The fitter also lends itself to dealing with multiple events, occurring at the same or different times within the event trigger window.

Questions:

What is required of a fitter in order to do the background elimination that SNO needs?
i.e.: what would be number of pmt beta gamma triggers per day? At what threshold do we have to fit events at in order to do background elimination?

Results: Graphs

The things to note are the differences between the eigenvalue histograms for each type of event, the direction fit, and the performance of the fitter on pmt beta gammas. We have included a set of graphs for five of standard fits from Queens. For each of the five files we show four graphs.

Graph 1: Direction Fit

The direction of an event was fitted by looking at the principle axes of the mass distribution(calculated by diagonalizing the moment of inertia matrix), and picking as the direction of the fit the eigenvector associated with the smallest eigenvalue; i.e. pick the long axis of the mass distribution. This gives us two choices for the direction. The correct one was picked by cheating, but we should be able to pick a direction properly almost all the time using the pmt hit locations, (just taking the side with the most hits). The direction fit then comes about using (almost) no information about the location of the hits (at least directly). Whether or not it proves to be statistically independent of the more standard way of finding direction has not been checked.

Graph 2: Moment of Inertia Eigenvalues

The moments about the three principle axes were normalized, and binned. In all cases these graphs show a non-spherical distribution of probability.
It seems that pmt beta gammas show the most elongated shape.

Graph 3: RFit (two graphs sometimes)

The radius of the fit for each event binned in 10cm bins from 0 at the tank centre to 20m away. The fits that were included on the files from Queens have been binned in the same way for comparison. The graph of RFit for the pmt beta gammas shows a 20 fold improvement over the included fits in the $r < 400$ cm range. It seems that the pmt beta gamma events may have multiple probability density maxima.
The other graph shows the cumulative counts as a function of radius.

Graph 4: Fit Deviation:

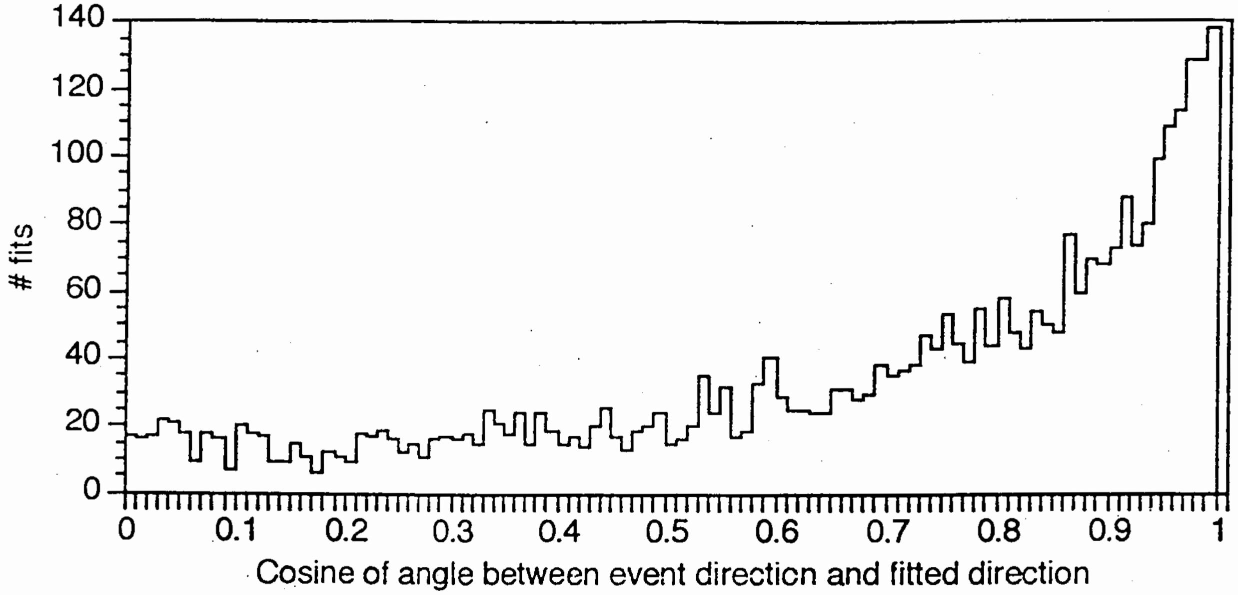
The fit to event distance is shown for the box fitter and the included fits from Queens. Also shown is the distance from the "best box" to the actual event location. The best box was taken as the box with the highest mass out of the thousands of final generation boxes.

Graph Sets:

1)PMT beta gammas with and without fancy fitting option, file: pmt_beta_gamma.std.bin.1
2) Beta Gammas in D2O, with fancy fitting option.
3) Calibration #81 file calib.81.std.bin.1
4) Neutral Current standard. File nc-std.bin.1
5)Charged Current standard. File cc-std.bin.1

Direction Fit

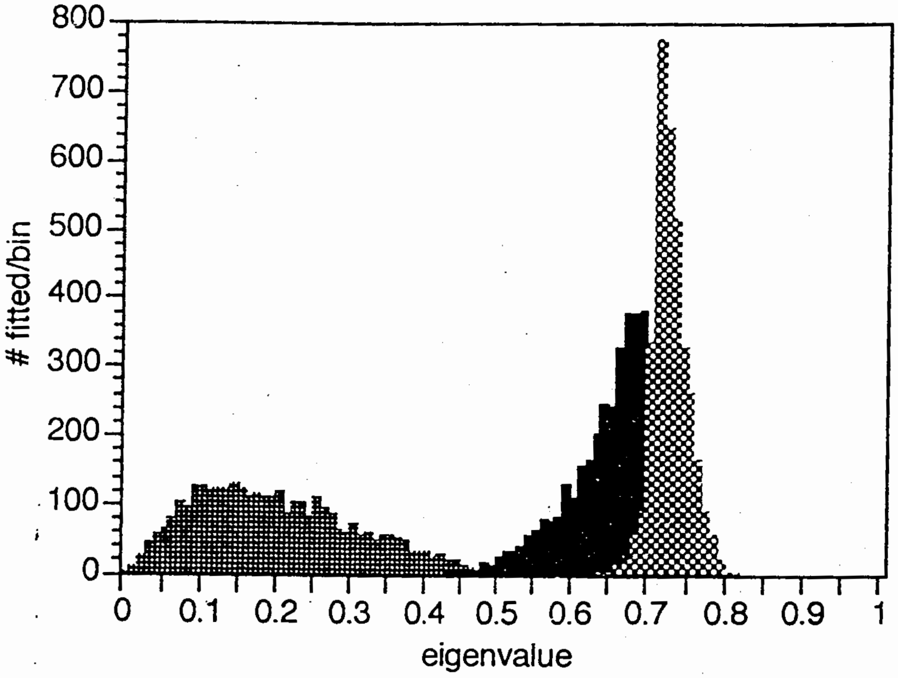
Direction Cosines obtained from time information alone for pmt 208T beta gamma decays
10000 events, 3103 accepted, 100 bins
file pmt_beta_gamma_std.bin.1

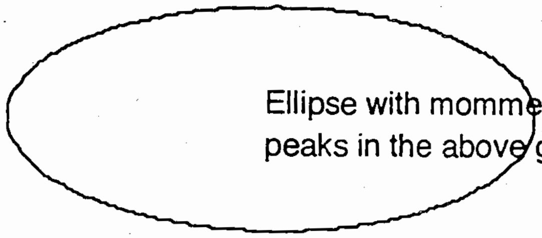


Moment of Inertia PMT beta -gammas

The direction of an event was fitted by finding the principle axes of the probability distribution for each event. The priciple axes emerge as the eigienvectors, the eigenvalues along these axes are the moments of inertia along these axes. To find a fitted direction, the eigenvector associated with the smallest eigenvalue was taken as the direction of the event.

From the graph below, it can be seen that as expected, the probability distibution for an event forms an elongated, - cigar shaped object.

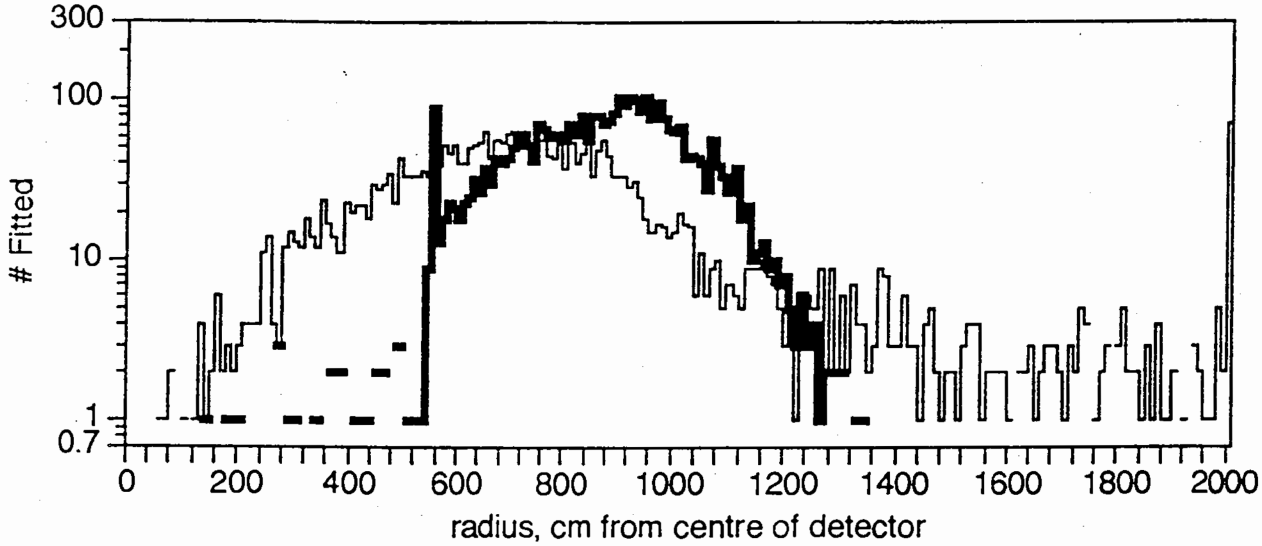


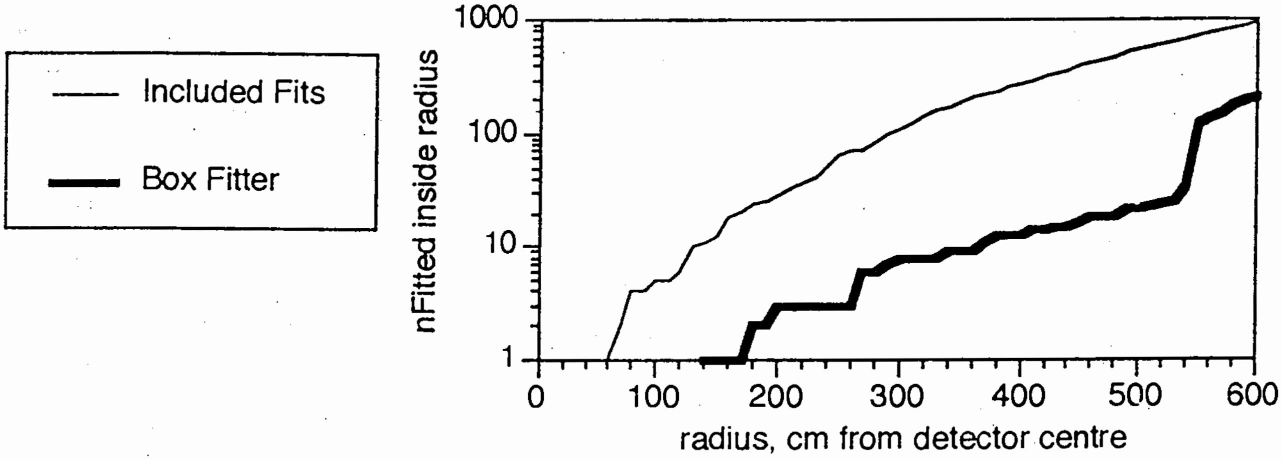


Ellipse with momment of intertia same as the peaks in the above graph.

RFit: PMT Beta - Gammas - Fancy Fitter

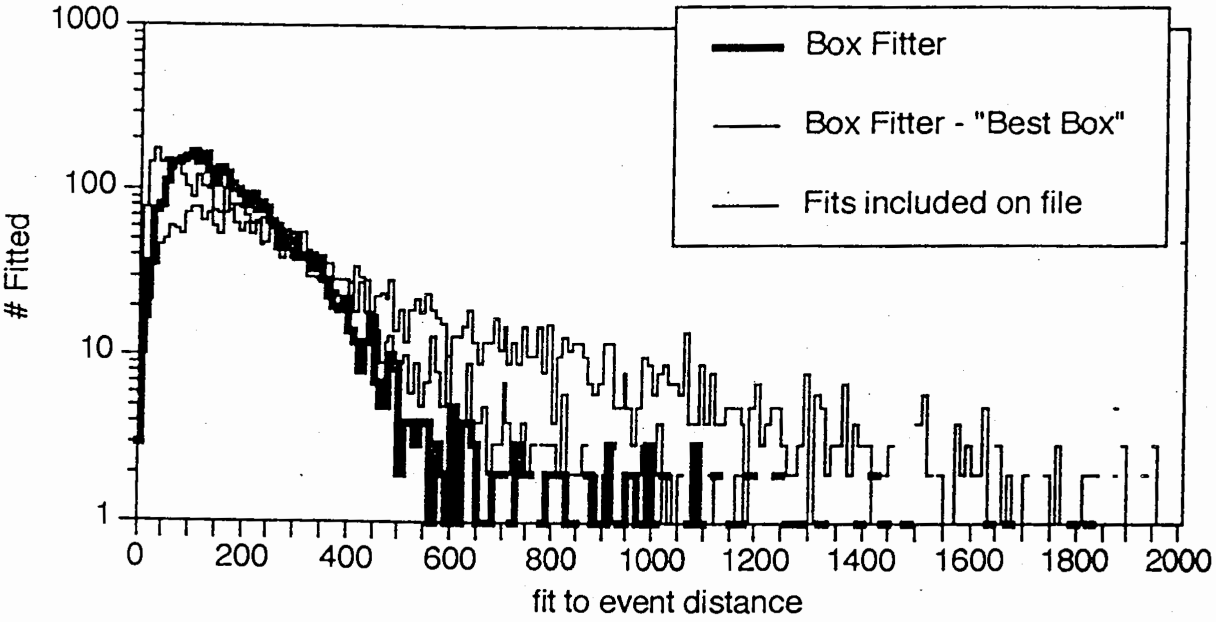
The radius of fit for the file pmt_beta_gamma_std.bin.1
3398 events fitted out of 10000. The 3157 fits that were included in the file are shown as the thin line. The included fits had 278 eventsfitted for R < 400. The box fitter placed 13 events in this region.
The Fancy Fitter option moved any event with nFitted < 13 that had 12% or more of its mass beyond 550 cm to a radius of 550cm, hence the peak at 550cm.





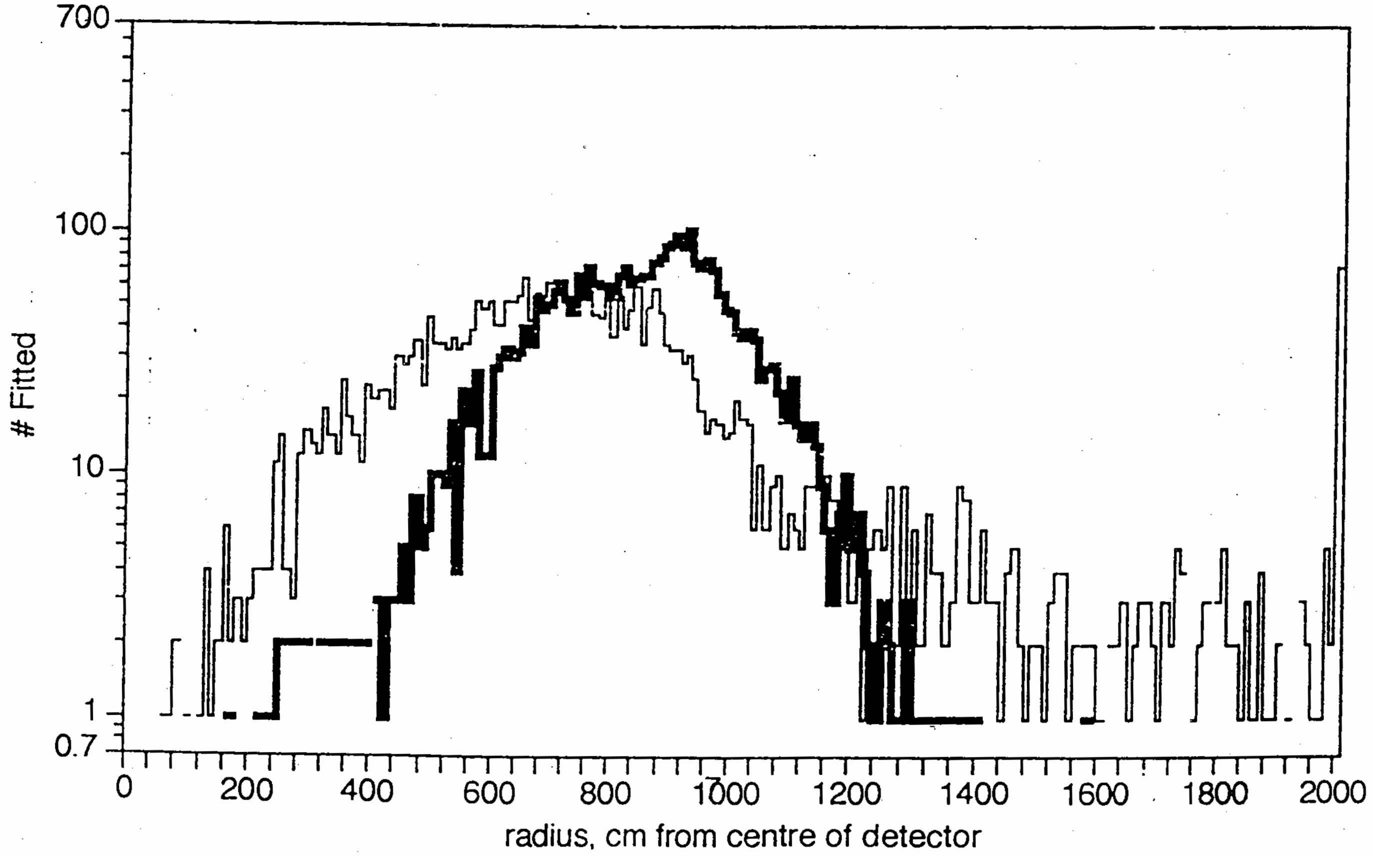
Fit Deviation - Fancy Fitter

Fit Deviation in cm, 10cm/bin.
From the file pmt_beta_gamma.std.bin.1



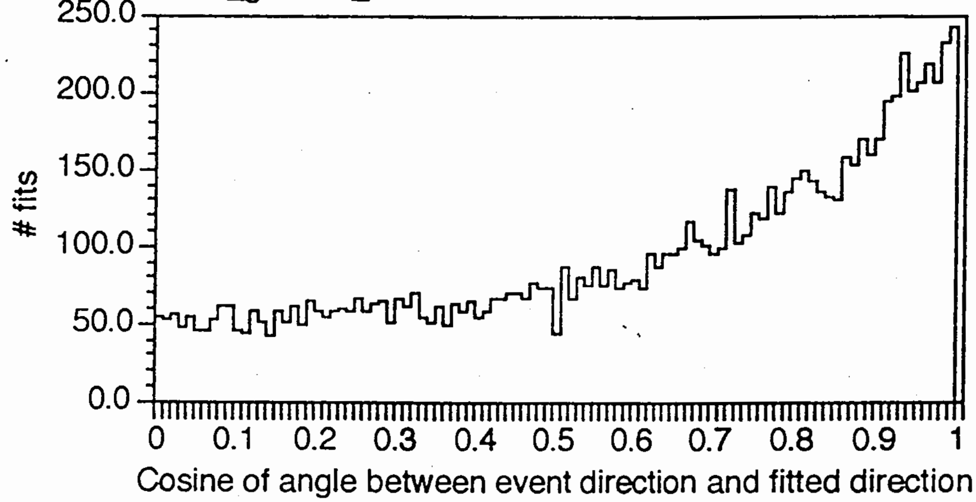
RFit: PMT Beta - Gammas

The radius of fit for the file pmt_beta_gamma_std.bin.1 3103 events fitted out of 10000. Thin line shows the 3157 fits that were included in the file. The included fits had 278 eventsfitted for R < 400. The box fitter placed 23 events in this region.



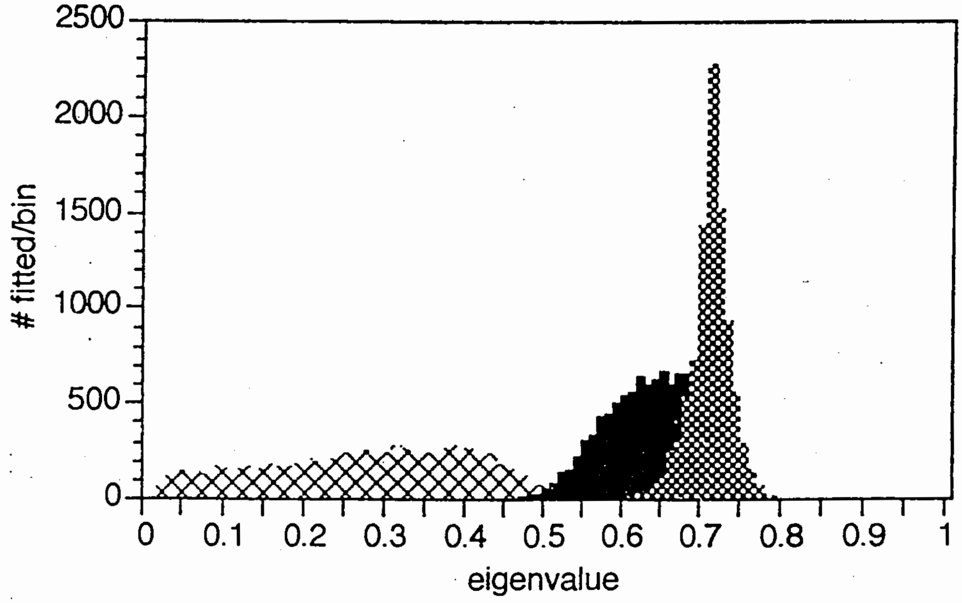
Direction Fit

Direction Cosines obtained from time information alone for
208T beta gamma decays distributed throughout the tank
9900 events, 9392 accepted, 100 bins
file beta_gamma_std.bin.1



Moment of Inertia eigenvalues

208T beta gamma decays distributed throughout the tank
9900 events, 9392 accepted, 100 bins
file beta_gamma_std.bin.1

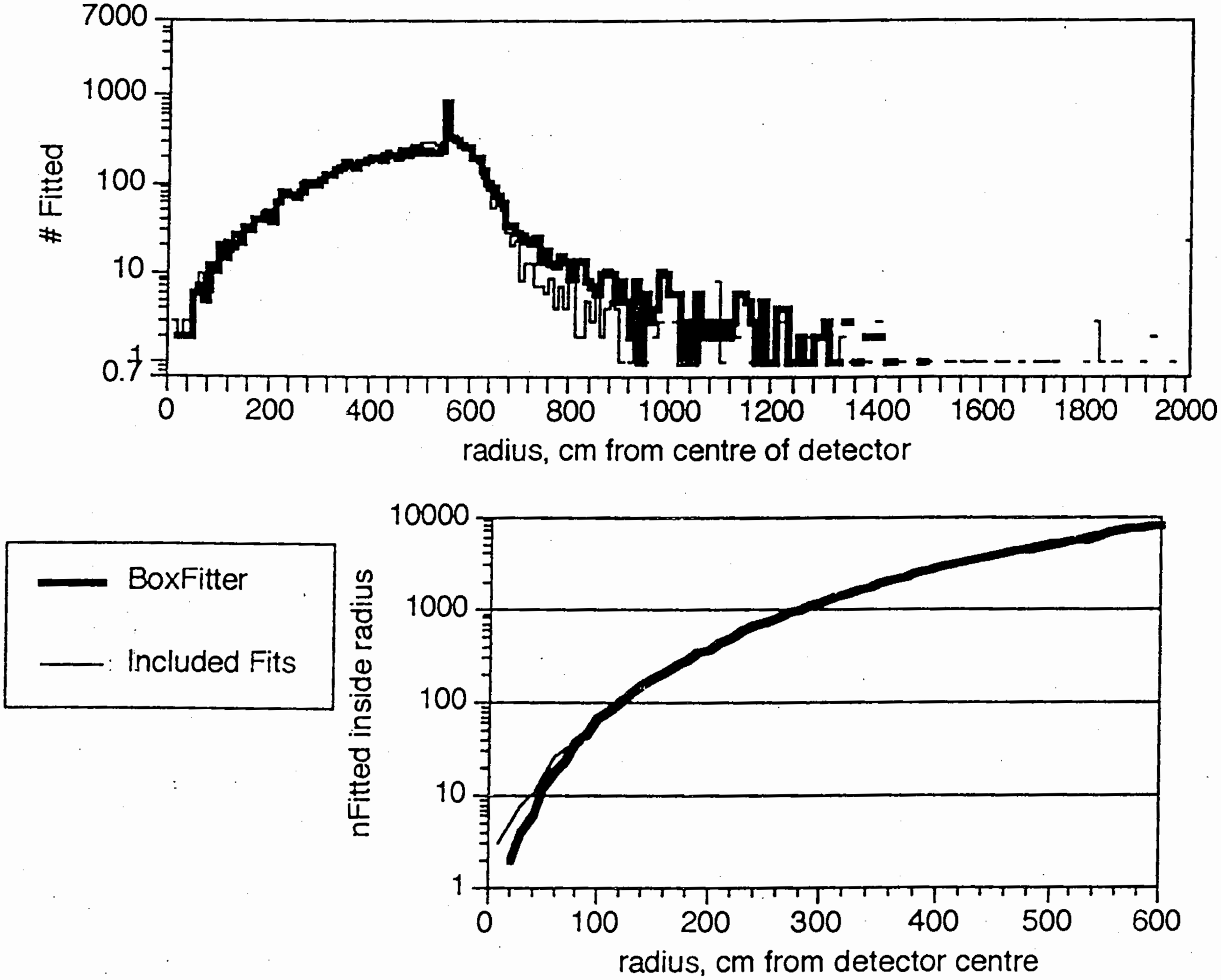


RFit: Beta-Gammas in tank - Fancy Fitter

The radius of fit for the file beta_gamma_std.bin.1

9392 events fitted out of 9900 The box fitter placed 2831 events in the $r \leq 400$ region.

The Fancy Fitter option moved any event with nFitted < 13 that had 12% or more of its mass beyond 550 cm to a radius of 550cm, hence the peak at 550cm.



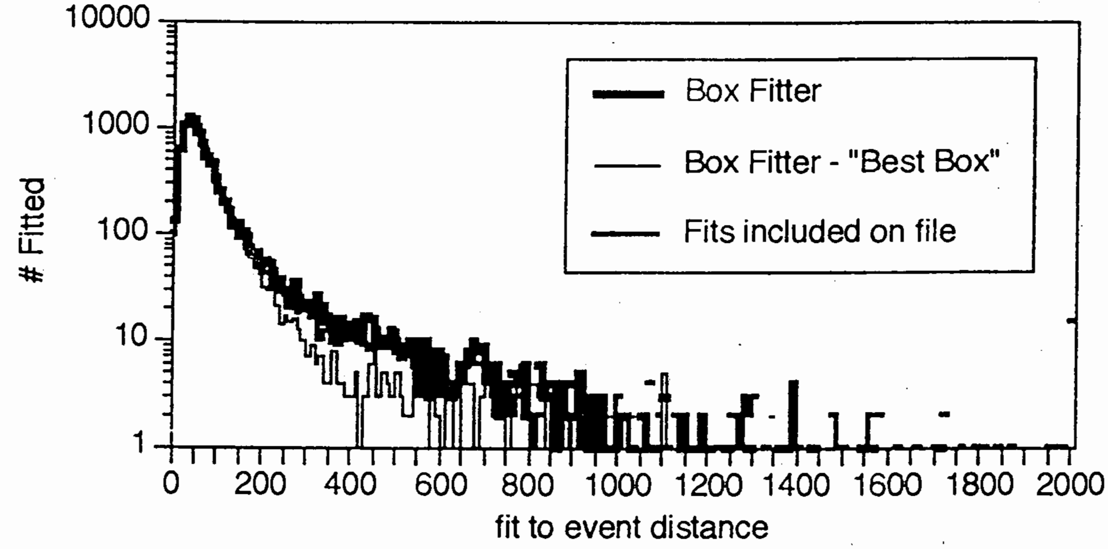
Fit Deviation - Fancy Fitter beta-gammas in D_2O

Fit Deviation in cm, 10cm/bin.

The highest peak is the box fitter fitting using the centre of mass of the probabilty distribution.

The fits included in the file are shown

For the file beta_gamma.std.bin.1

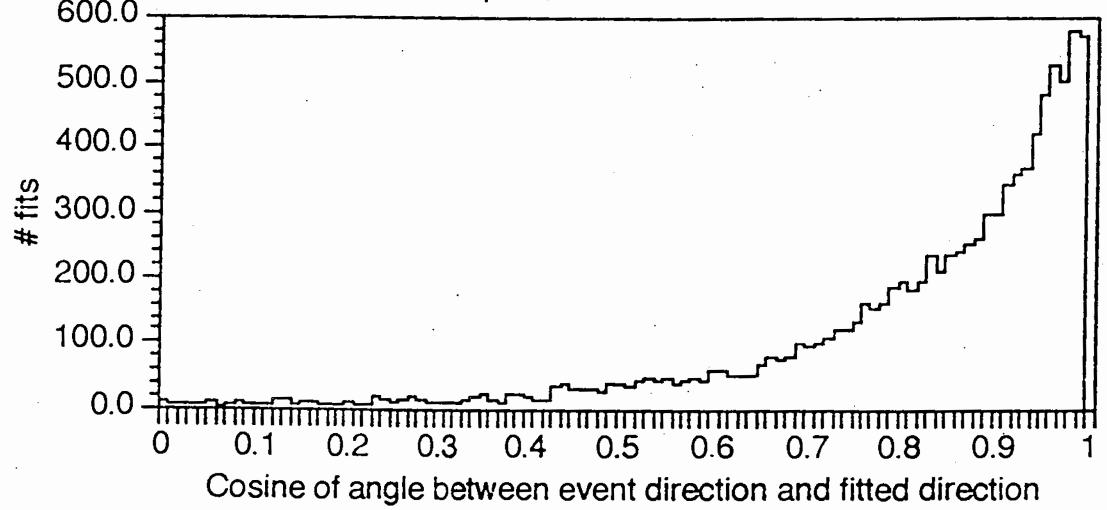


Direction Fit calib 81

7.5Mev electrons at 0, 400, 0

Direction Cosines obtained from time information alone for 7.5MeV electrons at 0, 400, 0 in the tank

9999 events, 9999 accepted, 100 bins

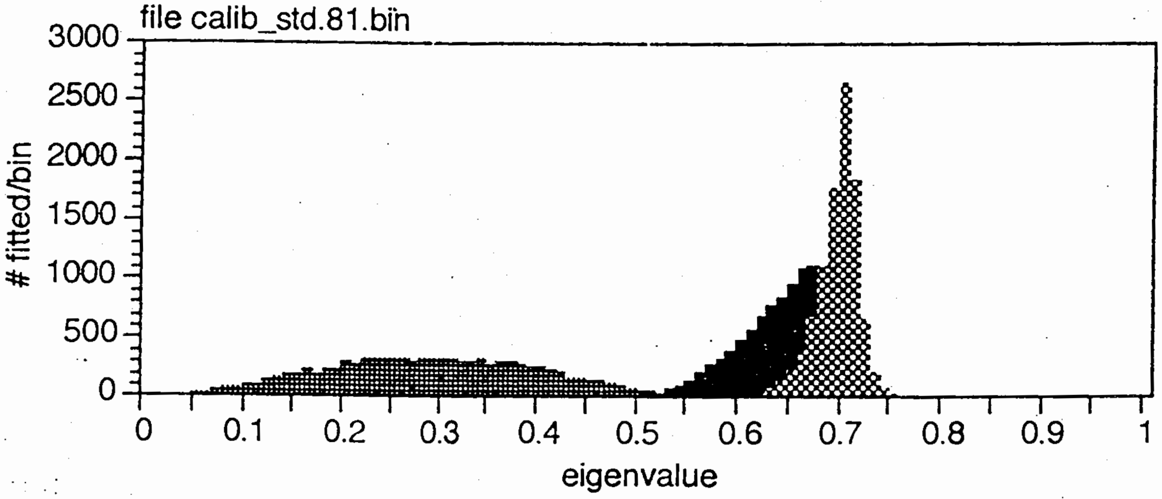


Moment of Inertia eigenvalues

moment eigienvalues for 7.5MeV electrons at 0, 400, 0 in the tank

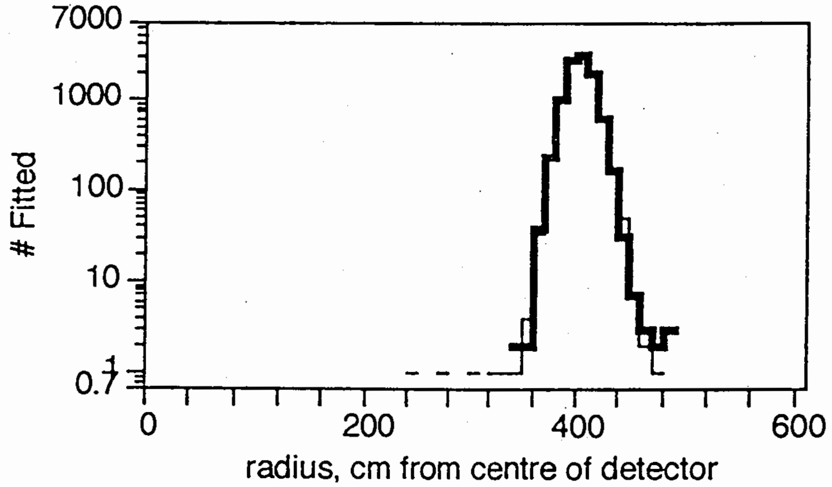
9999 events, 9999 accepted, 100 bins

file calib_std.81.bin



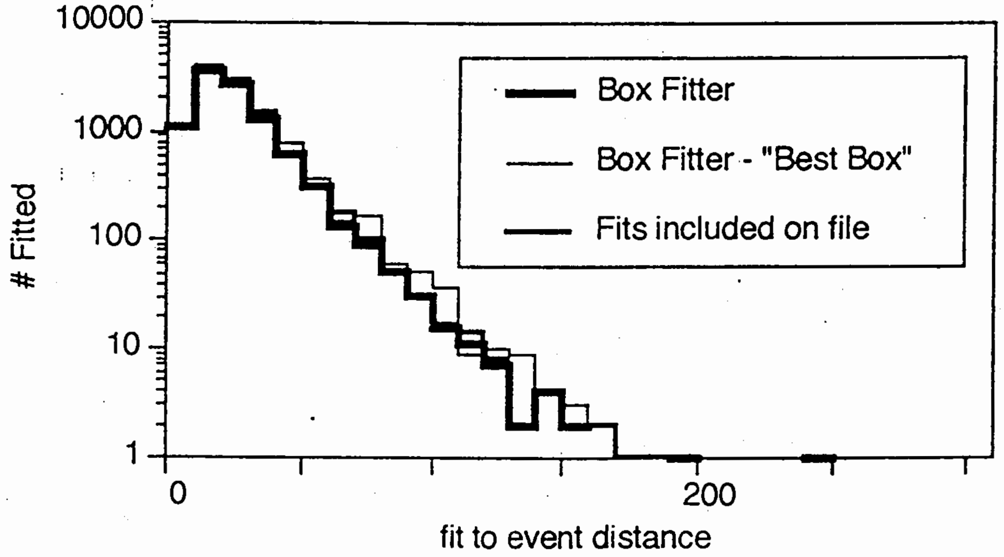
RFit: Electrons in tank

rFit for 7.5MeV electrons at 0, 400, 0 in the tank
9999 events, 9999 accepted, 100 bins
file calib_std.81.bin



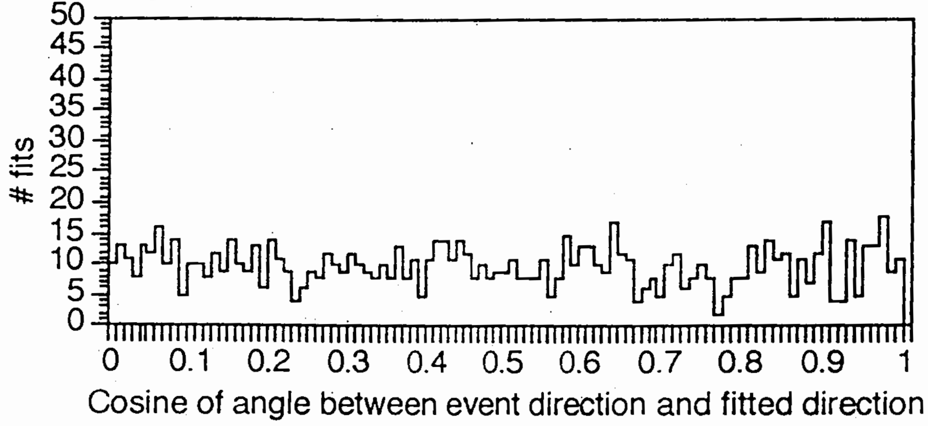
Fit Deviation 7.5MeV electrons in D_2O

FitDeviation for 7.5MeV electrons at 0, 400, 0 in the tank
9999 events, 9999 accepted, 100 bins
file calib_std.81.bin



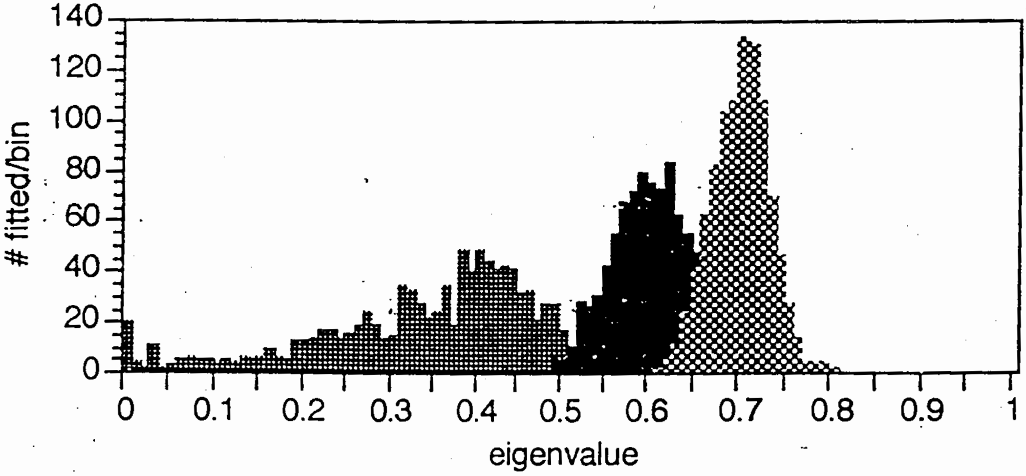
Direction Fit neutral current

Direction Cosines obtained from time information alone for
1000 neutral current events, (no direction, as expected)
file nc_std.bin.1



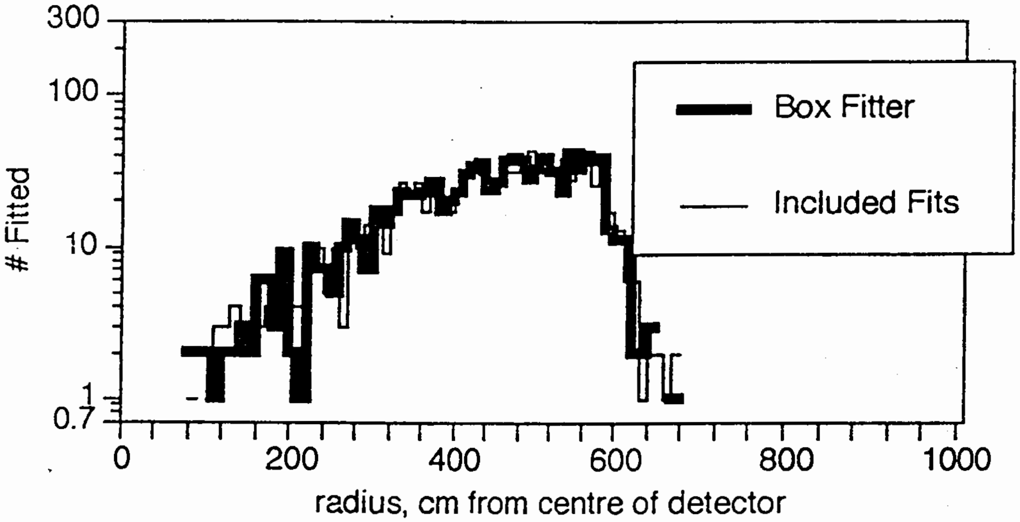
Moment of Inertia

neutral current events show little directionality



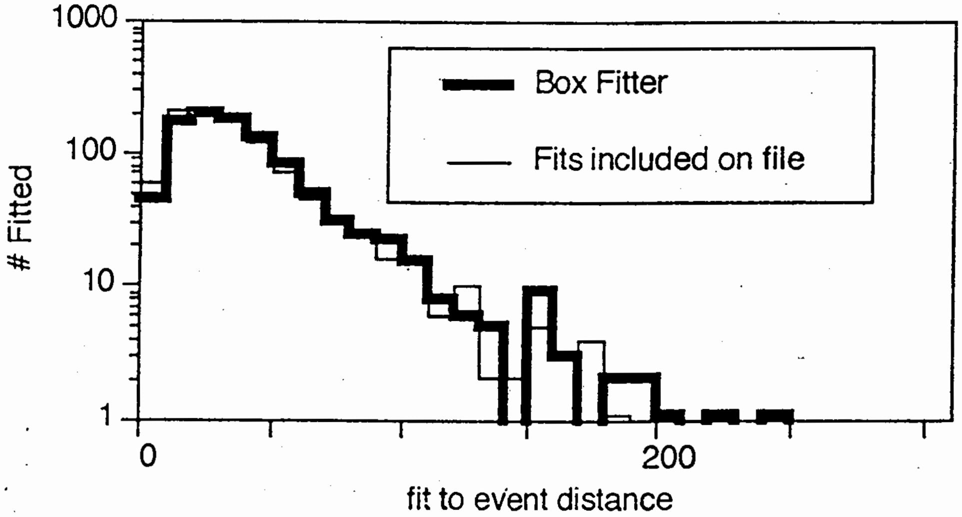
RFit: Neutral Current Events

The radius of fit for the file nc_std.bin.1
999 events fitted out of 1000



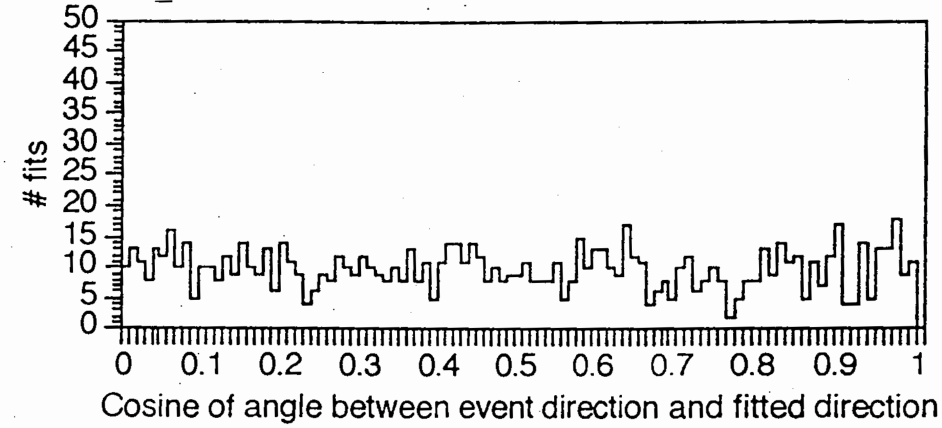
Fit Deviation Neutral Current Events

Fit Deviation in cm, 10cm/bin.
Very similar to included fits



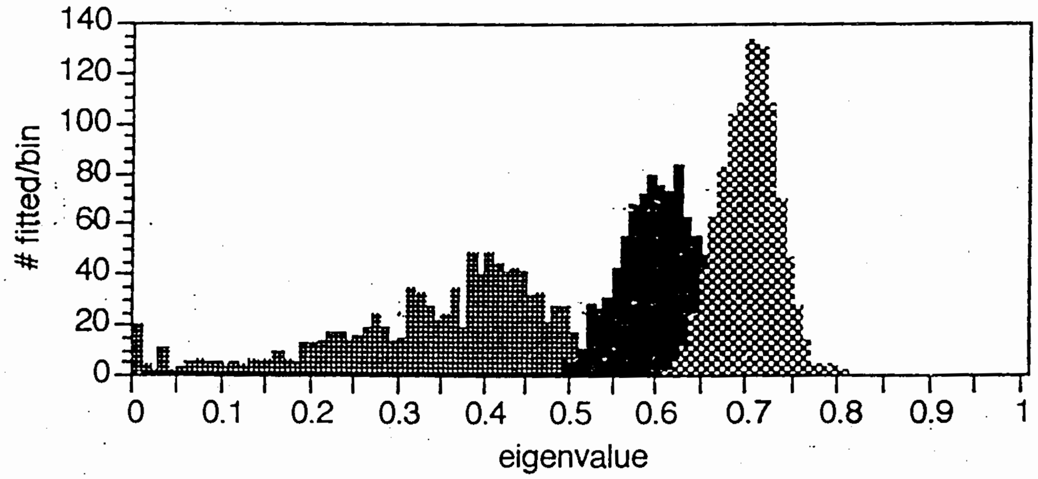
Direction Fit neutral current

Direction Cosines obtained from time information alone for
1000 neutral current events, (no direction, as expected)
file nc_std.bin.1



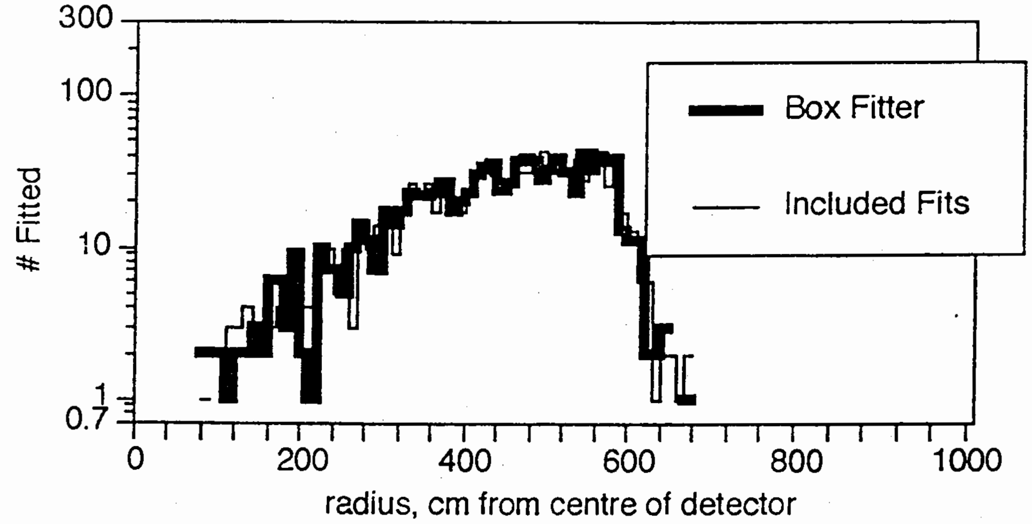
Moment of Inertia

neutral current events show little directionality



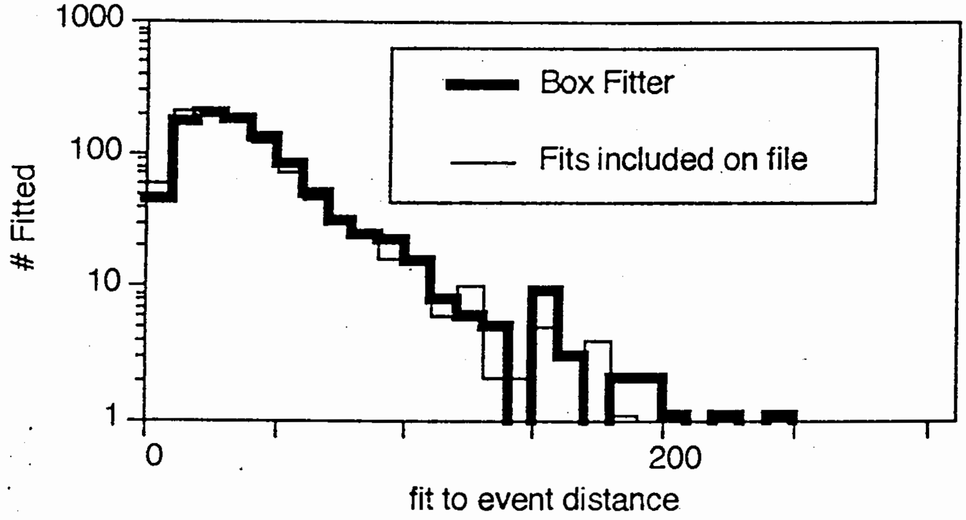
RFit: Neutral Current Events

The radius of fit for the file nc_std.bin.1
999 events fitted out of 1000



Fit Deviation Neutral Current Events

Fit Deviation in cm, 10cm/bin.
Very similar to included fits



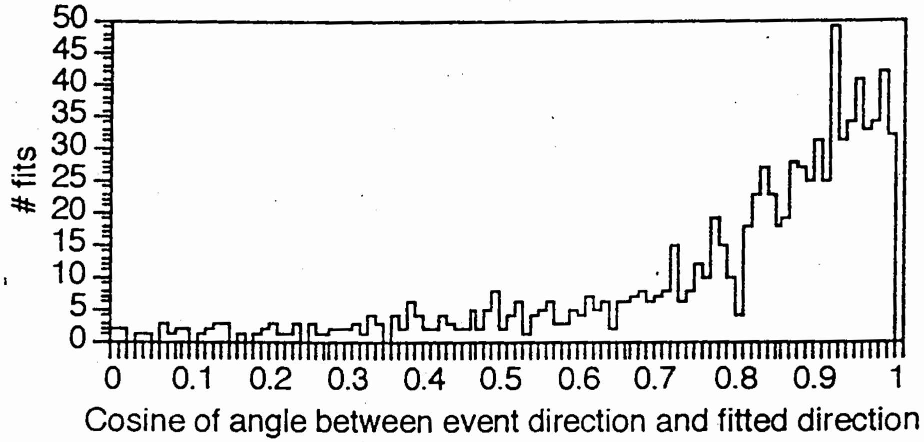
Direction Fit

charged current events in the D_2O

Direction Cosines obtained from time information alone for charged current events in the tank

900 events, 900 accepted, 100 bins

file cc_std.bin

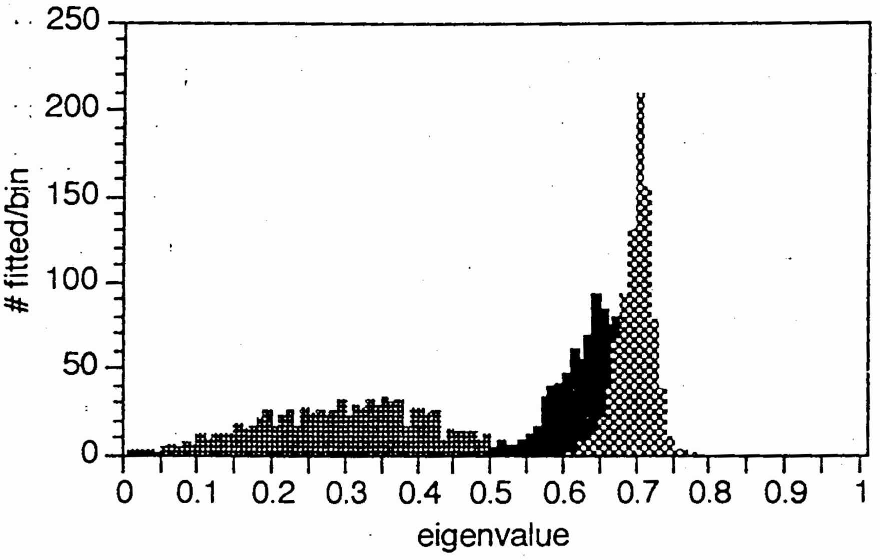


Moment of Inertia eigenvalues

moment eigienvalues for charged current events in the tank

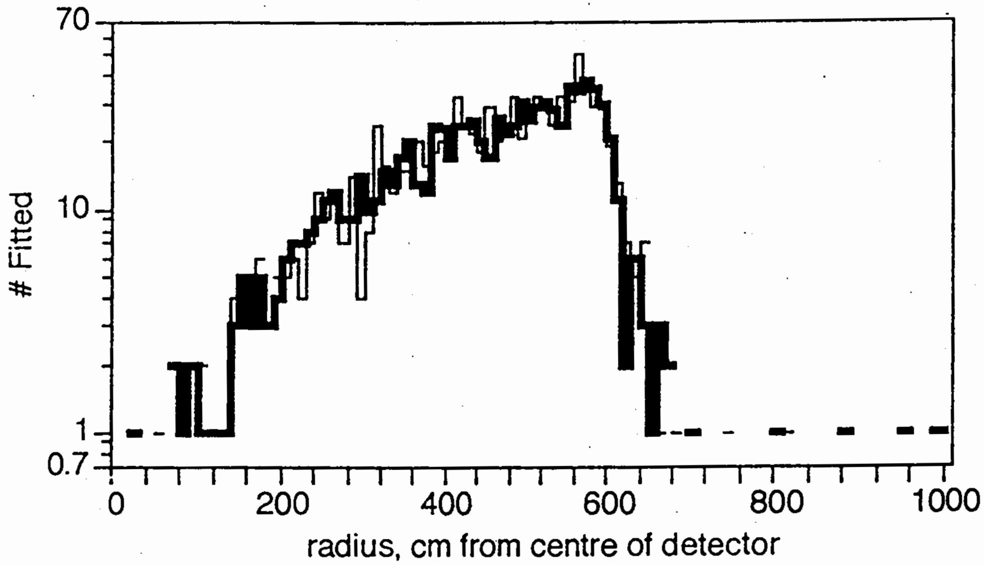
900 events, 900 accepted, 100 bins

file cc_std.bin



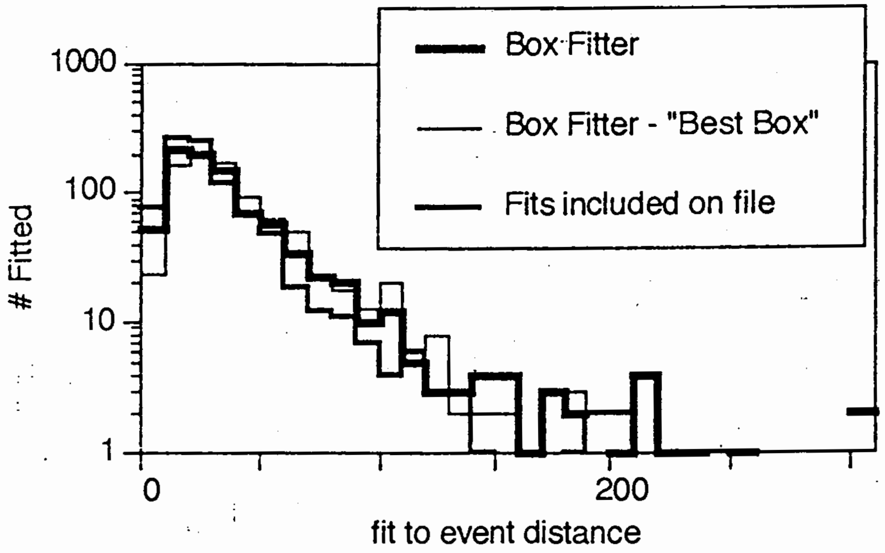
RFit: Charged Current

rFit for charged current events in the tank
900 events, 900 accepted, 100 bins



Fit Deviation: Charged Current

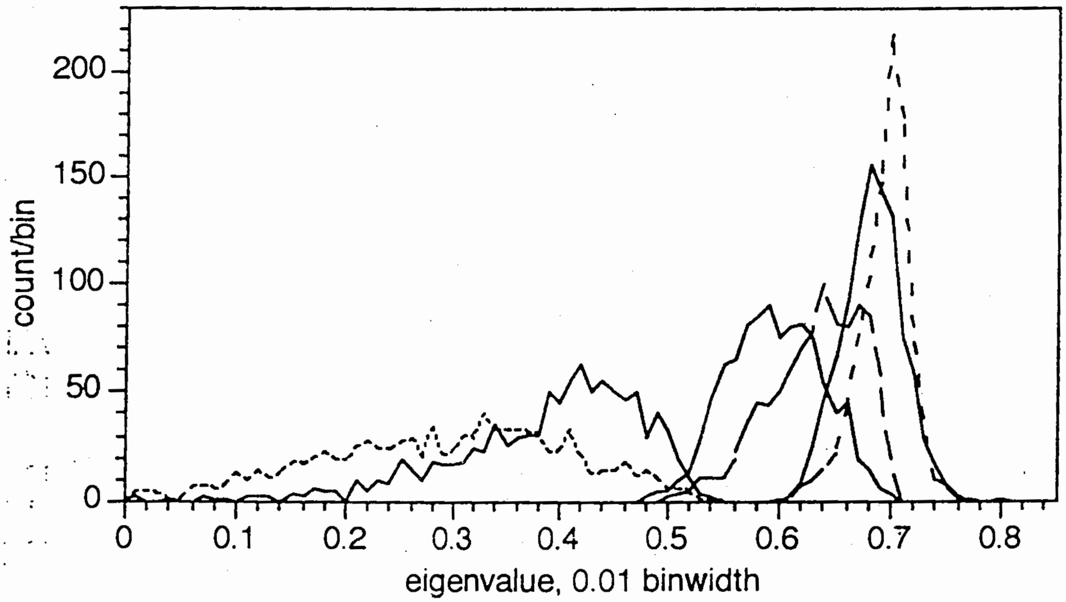
FitDeviation for charged current events in the tank
900 events, 900 accepted, 100 bins
file cc_std.bin



Event Discrimination using the eigenvalue distribution: Charged Current - Neutral Current Discrimination

The plots of eigenvalues on the previous pages show that different types of events have different distributions of eigenvalues. This difference can be used to distinguish events on an event by event basis, or on a statistical basis.
The difference in the eigenvalue histograms between neutral current events and charged current events is likely due to the two electron events that make up a large fraction of the neutron capture reaction on Cl. A two pronged event naturally defines a plane, and a plane will have three distinct eigenvalues. The graph below shows the differences between the eigenvalue distribution for the two kinds of events. It is just the histograms presented earlier in this document put on top of one another to show the differences between the two types of events. The three solid lines show the distribution obtained for the neutral current events, and the three dotted ones show the charged current results.

NC - CC Eigenvalue distribution:



In order to distinguish neutral current events from charged current events, we modeled the six peaks in each of the above graph using double guassians. As each event was fitted, and the eigenvalues found, the event was considered to be a neutral current event if the event multiplicities summed over each eigeinvalue favoured the hypothesis that the event was a neutral current one.

Example (perhaps best describes it)
Imagine a hundred events with an eigenvalue set of (0.43, 0.61, 0.71) now of these hundred events, how many would be neutral current ones?
At 0.43 75 nc, 25 cc
At 0.61 50nc, 50 cc
At 0.71 35nc, 65 cc
So there are more neutral current events with eigenvalue sets like this than cc events, so pick the event to be a nc event.

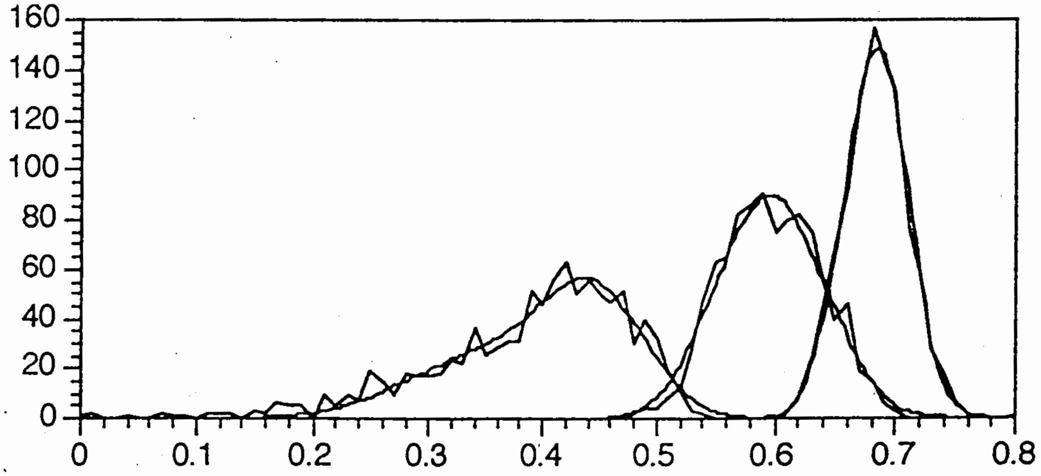
The success we have had using this method on the two Queen's cc and nc events files is that 70% of the events in each file were correctly identified. (1391 correct of 1994 events fitted)

The other way to determine the event type is to not look on an event by event basis, but rather to deconvolute an eigienvalue distribution into two components, using standard techniques. We have not yet tried this but it should work.

Eigenvalue Distribution Model:

The eigenvalue distribution was modeled in this study using a double guassian for each peak, but it seems the best way to use the method would be to use the monte carlo data from a high statistics run, and then do the fitting on a different set of high statistics events. Since we only had one set of files to work with, we figured that by using a smooth analytical approximation to the eigenvalue set was the best way to approximate the above process.

Neutral Current Functions



Charged Current Functions