# Use of Angular Correlation Functions to Separate CC and NC Events in Salt
SNO-STR-2002-004

William Heintzelman
University of Pennsylvania

April 30, 2002

## Introduction

In two previous papers I discussed the use of angular correlation functions to identify background events [1], and to distinguish neutral current (NC) events from single-electron-initiated events in salt [2]. This paper presents the results of subsequent work on the problem of separation of CC events from NC events, and supersedes [2]. (The results in [2] were subsequently found to be over-optimistic because the events I had used to calculate various standard functions were not independent from the sample used to study separation.) Much work has been done by others on event separation: see [3], [5], and [6].

This paper concentrates on methods for using angular correlation functions to separate CC and NC events, both on an event-by-event basis and by estimating the fraction of each type of event in a mixed sample ("statistical separation"). However, other statistics characterizing the distribution of Čerenkov light were studied, and the separation effectiveness of the various statistics was compared.

The definition and calculation of the angular correlation function is discussed in [1]. The angular correlation function as used in this work is in essence the distribution of the cosines of the angles between all possible pairs of the hit PMTs in an event. The event location estimated by one of the standard fitters is used as the vertex in the angle calculations. In the work reported here, the path fitter was used to determine the vertex. For an isotropic distribution of PMT hits, the angular correlation function is a uniform distribution in the cosine of the pair angle. Plots of correlation functions in this paper are normalized so the function for an isotropic distribution of PMT hits would everywhere be equal to one. However, the distributions are calculated in bins, and in most calculations are normalized appropriately for a discrete probability distribution with the sum of the bin values equal to one.

## The Residual Time Cut

In earlier work, I studied event separation using two subsets of the hit PMTs of an event: the first subset included all the hit PMTs; the second only those passing a residual time cut (RTC). However, as expected, separation using PMTs passing the RTC was found to be uniformly superior to that using all PMTs, and all the results in this paper were calculated with the RTC.

The residual time cut differs from the time exclusion window applied in the work reported in Reference [1], and is now more similar to the PMT hit cut applied in the SNOMAN Classifier. A time residual is calculated for each hit PMT, given by

$$t_{res} = t_{hit} - t_{fit} - d_{fit}/v, \tag{1}$$

where

$$\begin{aligned} t_{hit} &= \text{PMT hit time;} \\ t_{fit} &= \text{fitted event time;} \\ d_{fit} &= \text{distance between fitted event position and PMT;} \\ v &= \text{speed of light (21.75 cm/ns).} \end{aligned}$$

An 8 ns sliding window is imposed on the residual times, positioned to include the greatest number of PMT hits. PMT hits that fall outside the window are excluded from the calculations.

For comparison, in the cut used in the SNOMAN Classifier, PMT hits are excluded for which

$$|t_{hit} - t_{fit} - d_{fit}/v| < 200cm/v \tag{2}$$

$$(200 \text{ cm/v} = 9.2 \text{ ns})$$

## Monte Carlo Generation of Events

Four independent sets of events were generated to study CC/NC separation. All events were generated uniformly inside the AV in salt. The analysis was restricted to events that had total $N_{hit}$ less than 86, $N_{hit}$ after RTC between 41 and 70, and that reconstructed inside the AV.

The events were simulated retaining all of the Čerenkov photons (cerfac = 1 in SNOMAN) and with the full 3D PMT simulation. The PMT noise rate was 560 Hz. The fitted event positions, directions, and times, and the lists of hit PMTs and PMT hit times were written to intermediate files for input to subsequent computations, in which the path fitter outputs were used.

Two large sets of 163525 CC and 136277 NC events (after the above cuts) were used to derive standard correlation functions and standard distributions of the other statistics used for event separation (referred to subsequently as the standard sets). Two smaller sets of 29639 CC and 24740 NC events comprised those for which separation was studied (referred to subsequently as the study sets).

Figures 1 to 4 are the total $N_{hit}$ distributions for all standard events, the total $N_{hit}$ distributions after the event cuts, the distribution of the fraction of PMTs cut by the RTC, and the radial distributions of the standard events.

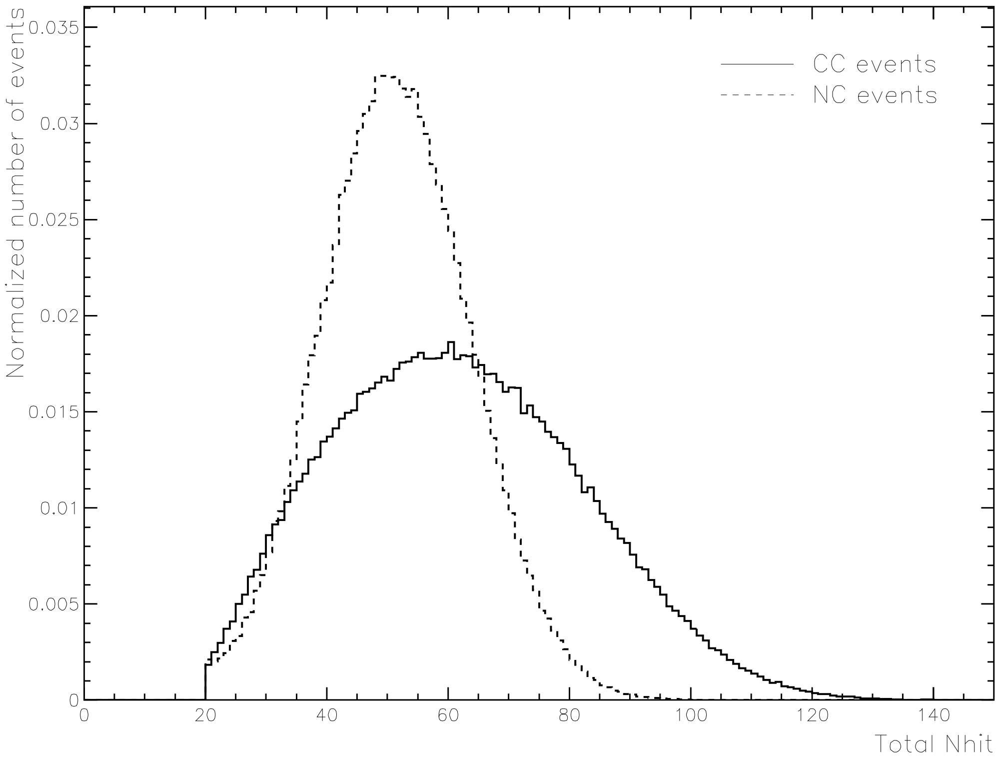


Figure 1: Distribution of Total $N_{hit}$ for All Standard Events

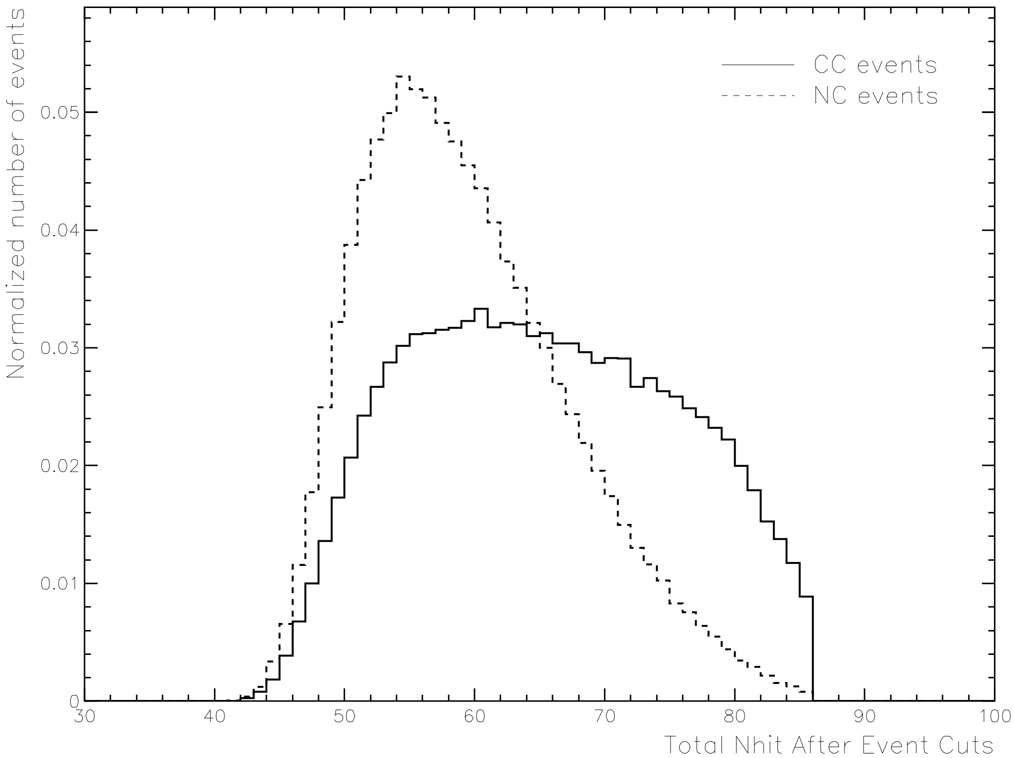


Figure 2: Distribution of Total $N_{hit}$ for Retained Events

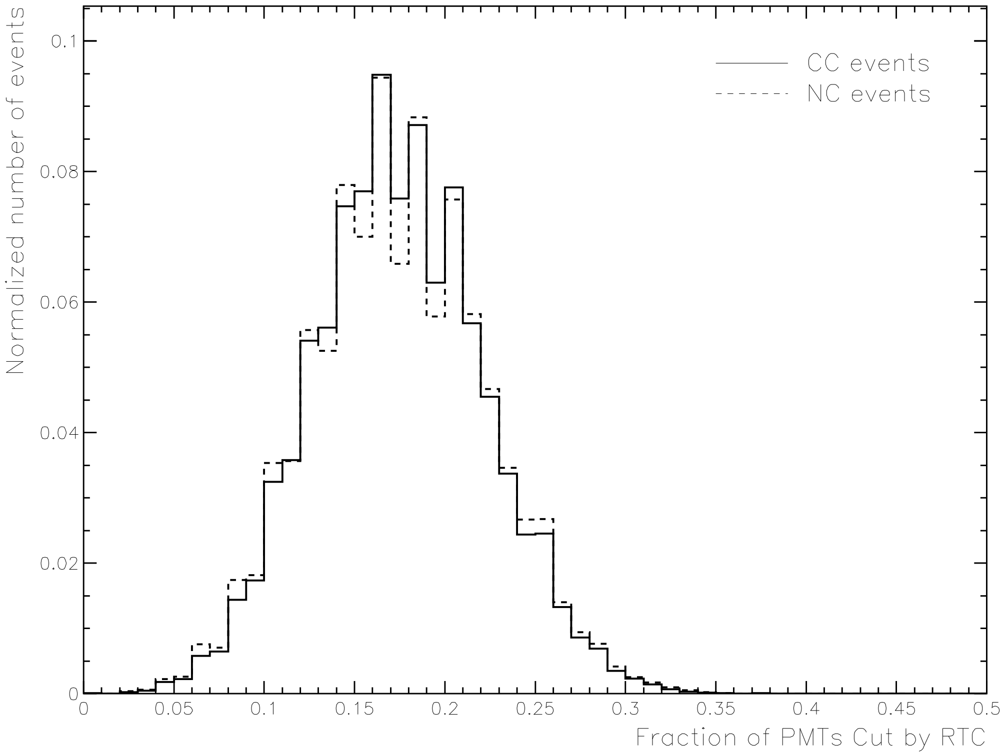


Figure 3: Fraction of PMTs Cut for Retained Events

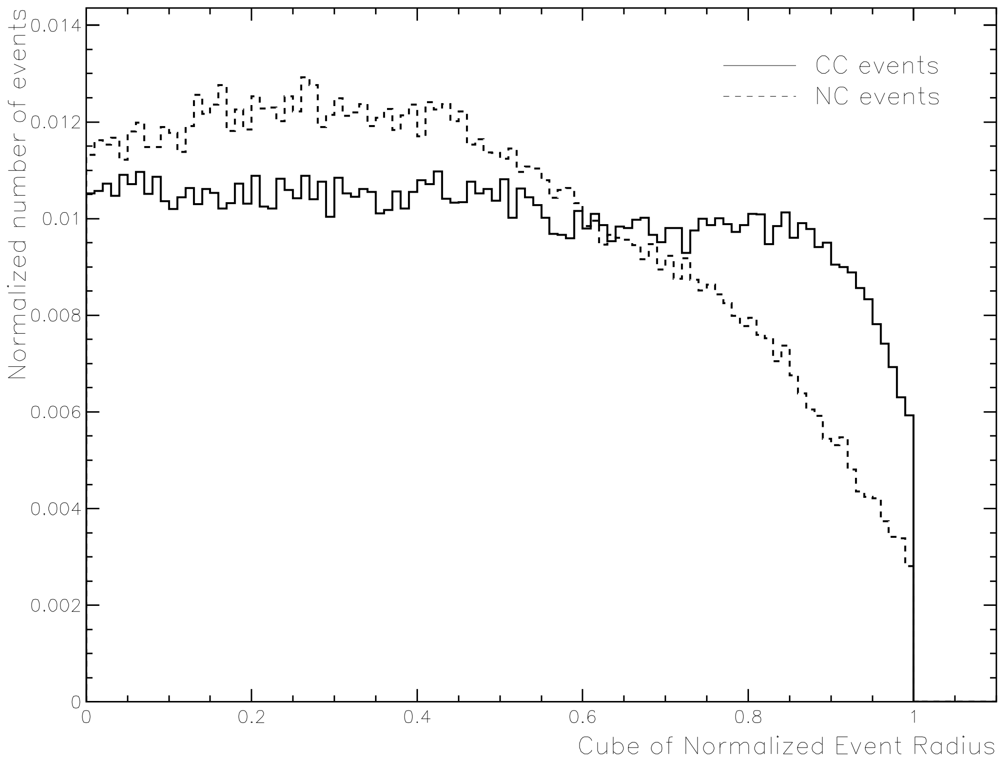


Figure 4: Radial Distributions of Retained Events (units of r/600)

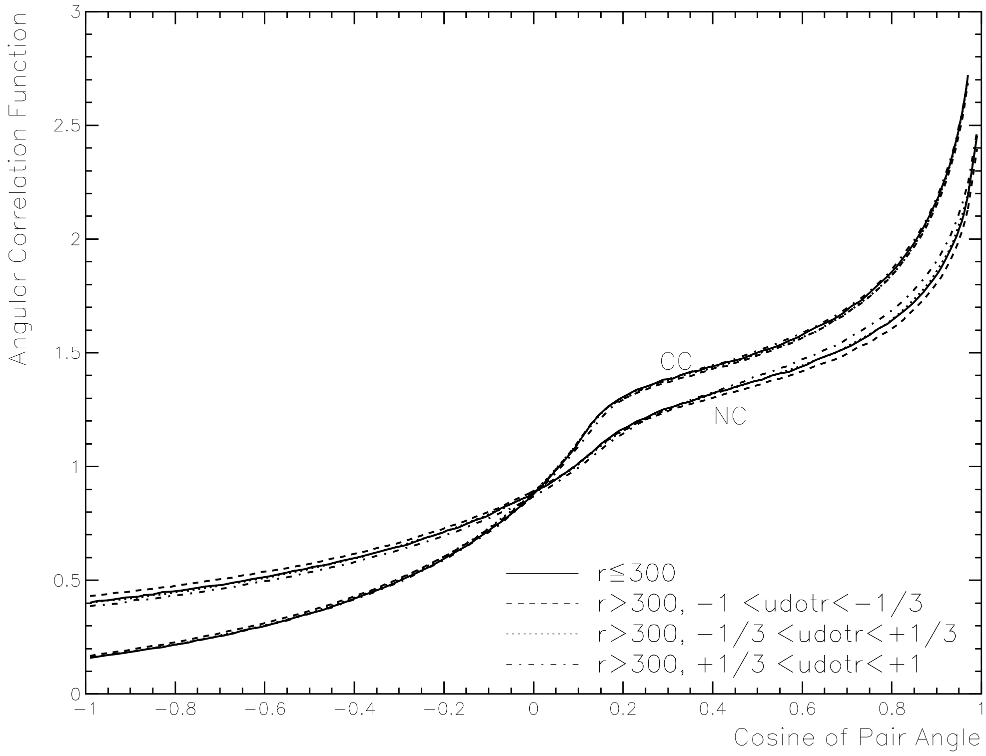


Figure 5: Average Correlation Functions with RTC

## Calculation of Standard Functions

Standard angular correlation functions and associated covariance matrices were calculated from the standard events as a function of $N_{hit}$ in bins of width three, and within each $N_{hit}$ bin in sub-bins with the following four sets of conditions:

$$r \leq 300 \tag{3}$$

$$300 < r \leq 600 \ ; \ -1 \leq \hat{u} \cdot \hat{r} \leq -1/3 \tag{4}$$

$$300 < r \leq 600 \ ; \ -1/3 < \hat{u} \cdot \hat{r} \leq 1/3 \tag{5}$$

$$300 < r \leq 600 \ ; \ 1/3 < \hat{u} \cdot \hat{r} \leq 1 \tag{6}$$

(See Reference [1] for discussion of the methods used.)

Figure 5 shows the average angular correlation functions for the standard CC and NC events (with no $N_{hit}$ cut applied). As expected, the NC angular correlation functions are flatter than those of the CC events. We see also that there is a small dependence on $\hat{u} \cdot \hat{r}$, which is more pronounced for the NC events. Note that because many of the other statistics that are used to separate event types vary with the angular correlation function, they also depend on $\hat{u} \cdot \hat{r}$.

## Event-by-Event Separation

Event-by-event separation was studied first, although it was thought doubtful that the light distributions of CC and NC events are sufficiently different to do this effectively.

The following paragraphs describe various approaches to event-by-event separation.

### Correlation Function Likelihood Ratio

The likelihood ratio test is a standard method for hypothesis testing. If one can calculate the probabilities of occurrence of an observed event's angular correlation function under the two alternative hypotheses (CC or NC), one can use the logarithm of the likelihood ratio to categorize the event. That is, let $P_{CC}$ be the probability of occurrence of the observed angular correlation function under the assumption that it is an CC event, and $P_{NC}$ the probability of occurrence under the assumption that it is a NC event. We calculate the logarithm of the likelihood ratio, $\mathcal{L} = \log(P_{CC}/P_{NC})$, and categorize events with $\mathcal{L}$ above some value as CC events and events with $\mathcal{L}$ below that same value as NC events.

As discussed above, average correlation functions and associated covariance matrices were calculated from the standard events. In doing this, it is assumed that the bin values of the correlation function have a multivariate normal distribution. If $x$ is the vector of bin values of an independent event's correlation function, with the assumption that the event is of type i (either CC or NC), $x$ then has the distribution

$$p(x|i) = \frac{1}{(2\pi)^{\nu/2}\sqrt{\det V_i}} \exp\left[-\frac{(x-\mu_i)^T V_i^{-1}(x-\mu_i)}{2}\right] \tag{7}$$

where $\nu$ is the number of bins in the correlation function minus one, $\mu_i$ is the vector of expected bin values and $V_i$ is the covariance matrix of the bin values for event type i, calculated from the standard set of events. The logarithm of the likelihood ratio is

$$\mathcal{L} = \log\frac{p(x|CC)}{p(x|NC)} = \frac{1}{2}\left[\chi^2_{NC} - \chi^2_{CC} + \log(\det V_{NC}) - \log(\det V_{CC})\right] \tag{8}$$

where $\chi^2_i = (x-\mu_i)^T V_i^{-1}(x-\mu_i)$ has been referred to previously (in [1]) as the correlation chi-square. In practise the two terms involving the determinants of the covariance matrices have been found to be almost equal, so this likelihood ratio test is equivalent to a test based on the difference of correlation chi-squares relative to standard functions for the two types of events. Figure 6 shows the distribution of $-2\mathcal{L}$ for the events in the study sets.

### Signed Kolmogorov-Smirnov Statistic

A method was developed based on a modification of the Kolmogorov-Smirnov statistic commonly used in a non-parametric test of the difference between distributions. For two sets of data binned identically the K-S statistic is given by

$$\max_{n=1}^{N}\left|\sum_{i=1}^{n} g_i - \sum_{i=1}^{n} f_i\right| \tag{9}$$

where $f_i$ and $g_i$ are the numbers of events in bin i for the two distributions, and N is the number of bins. In the K-S test, one is only interested in obtaining a measure of the

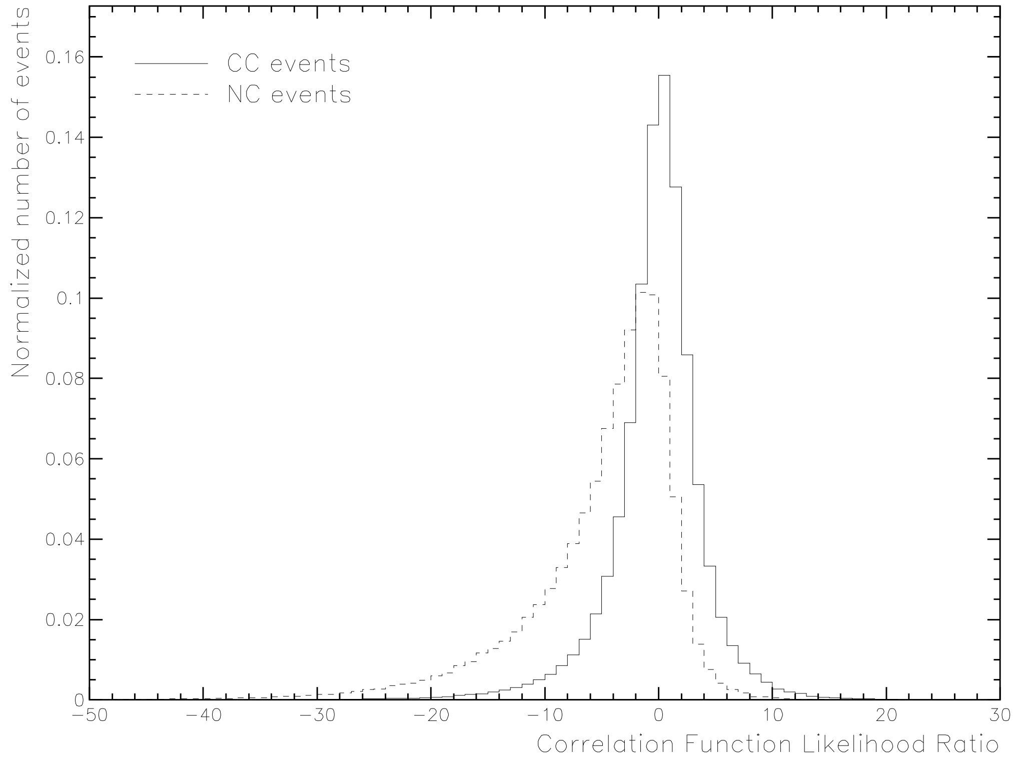


Figure 6: Correlation Function Likelihood Ratio for Study Events

magnitude of the difference between two distributions, which is why the absolute value appears in the above definition. But in our comparisons, we are interested in whether one correlation function is “steeper”or “flatter” than another. If we calculate the K-S statistic relative to a standard function, and then give it a sign depending on whether its greatest absolute value was attained when the difference in partial sums was positive or negative, we see that if an event correlation function is generally flatter than the standard function, the signed K-S statistic will be positive, and if the event function is steeper, the signed K-S statistic will be negative. Signed K-S statistics were calculated for an event relative to both the CC and the NC standard functions, and summed. An event whose correlation function is steeper than both standard functions (probably a CC event) will tend to have a large negative value; an event flatter than both standard functions (probably an NC event) will tend to have a large positive value; and an event with correlation function generally between the two standard functions will tend to have a small (either positive or negative) value. The distribution of the signed K-S sum is shown in Figure 7 for events in the study set.

### Correlation Function Inner Product with Ideal Ring

From the perspective of the distinctiveness of the light pattern produced, an ideal event would have only a single electron, which undergoes no noticeable scattering and proceeds in virtually a straight line until it slows to below the threshold for Čerenkov light production. The photons from such an event intersect in a narrow ring with a sphere

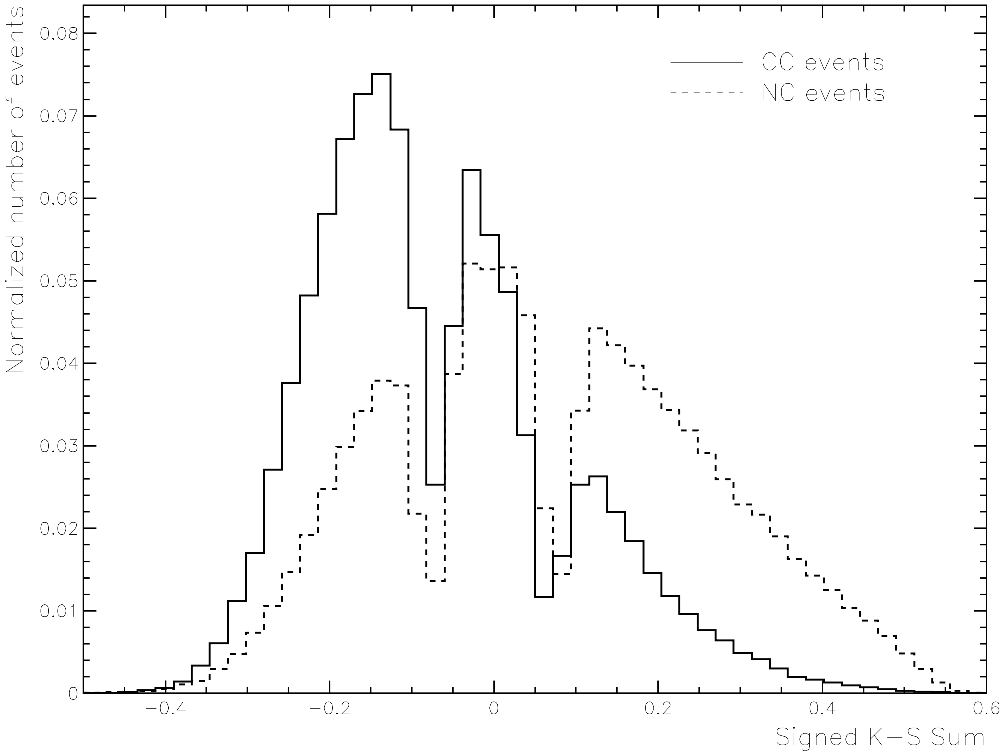


Figure 7: Distribution of the Signed K-S Sum

centered at the event's location. Figure 8 shows angular correlation functions computed by numerical integration for ideal patterns consisting of rings with a Čerenkov cone half-angle of 42 degrees and with various widths. (Note how the bump in these functions appears in a muted form in the correlation functions of Figure 5.) The light patterns produced by real ES, CC, and NC events are all in some sense degraded versions of the ideal. Because NC events on average deviate more from the ideal than the other two types, we should be able to use a measure of deviation from the ideal to separate NC events from CC events.

Several ways of doing this were tried. For purposes of these computations, both the event correlation functions and the ideal functions were calculated in 100 bins over the interval (-1,+1). One method was an unweighted sum of the squares of the bin differences between the event and the ideal functions. The second was a weighted sum, where the squared bin difference was divided by the number of hit-PMT pairs in the bin for the event under consideration. The third was an inner product computation, in which the values in the different bins are treated as components of a vector and normalized so the sum of their squares is one. If $\varepsilon_i$ and $\rho_i$ are the components of the event and the ideal correlation functions respectively, the inner product is $\sum_{i=1}^{100} \varepsilon_i \rho_i$. The unweighted sum of squares method and the inner product method were found to give nearly identical separation effectiveness, but the weighted sum was less effective. The inner product method was retained because of its computational simplicity. Ring widths of 4, 8, 16, 24, and 32 degrees for the ideal function were tried. Although

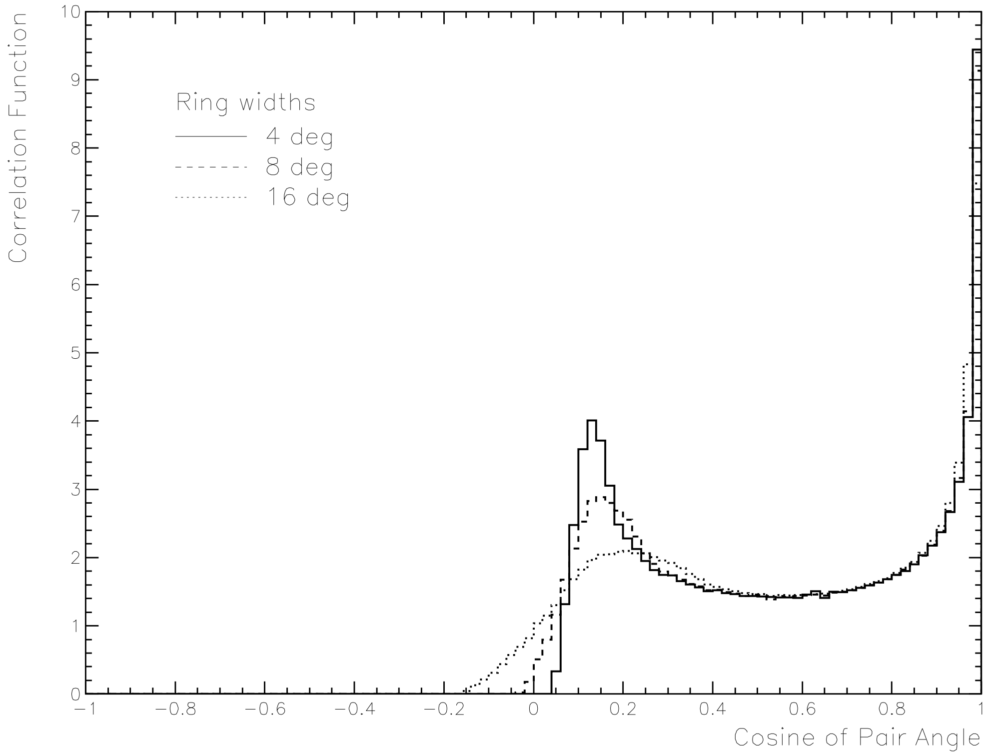


Figure 8: Correlation Functions for Ideal Ring Patterns

the values of the inner product changed of course, there was little difference in the effectiveness of separation of CC and NC events. The results shown in the remainder of this report were calculated with the 8-degree ring. Figure 9 shows the distribution of the correlation function inner product with the 8-degree ideal ring for events in the study set.

**Correlation Function Inner Product with CC/NC Difference**

If $\vec{x}$, $\vec{a}$, and $\vec{b}$ are unit vectors in a 2- or 3-dimensional space, one can use the quantity $\vec{x} \cdot \vec{a} - \vec{x} \cdot \vec{b} = \vec{x} \cdot (\vec{a} - \vec{b})$ as an indicator of the closeness of the vector $\vec{x}$ to $\vec{a}$ compared to its closeness to $\vec{b}$. This idea was applied in an n-dimensional space to compare correlation functions. The difference between the average correlation functions of the standard sets of CC and NC events was calculated in 100 bins, with the result shown in Figure 10. Just as with the correlation function for an ideal ring discussed in the preceding section, this function was treated as a vector in a 100-dimensional space and normalized so the sum of its squares is one. The inner product of the resulting vector with the analogous vector for individual events was calculated. Its distributions for CC and NC events are shown in Figure 11.

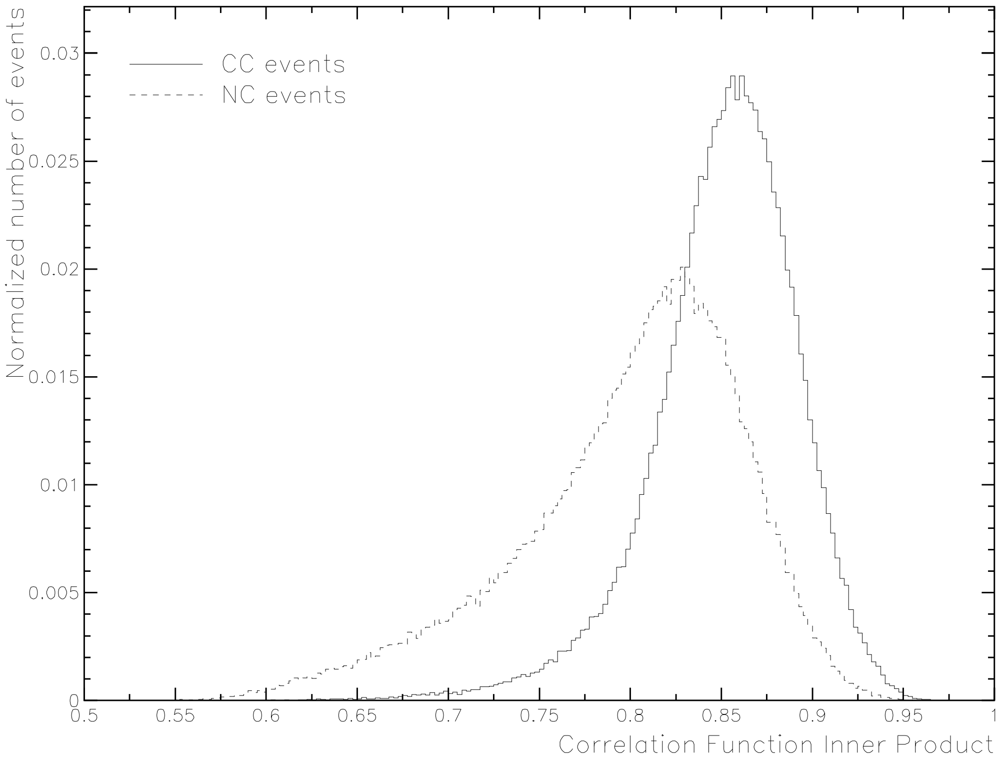


Figure 9: Inner Product with Ideal Ring Correlation Function

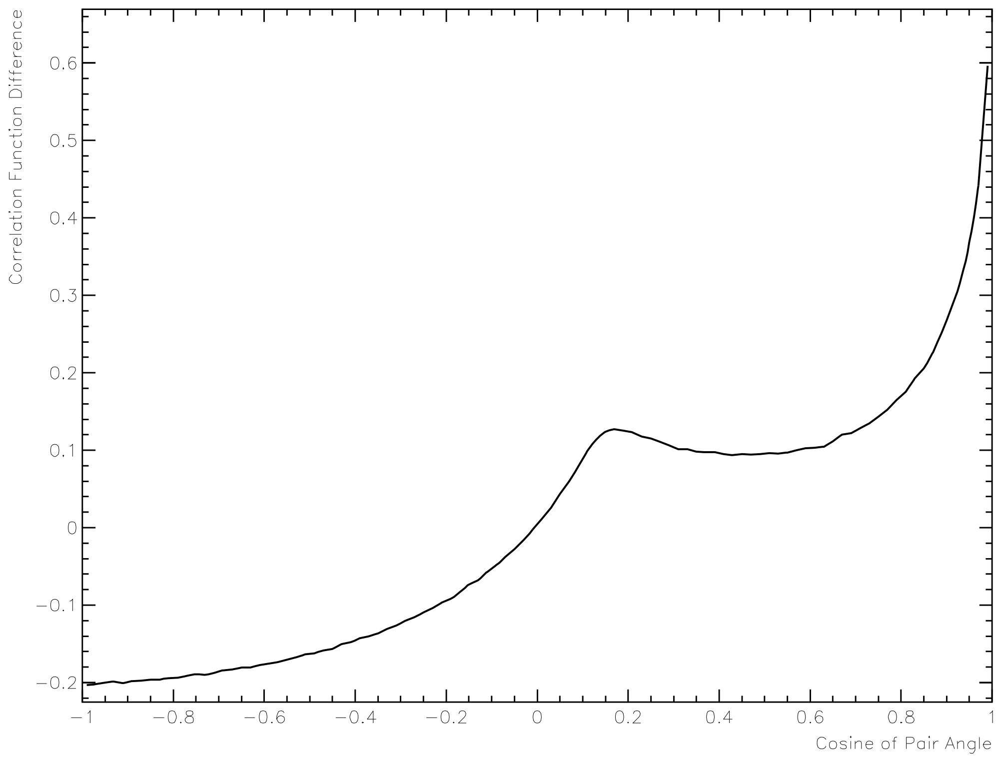


Figure 10: Difference Between CC and NC Average Correlation Functions

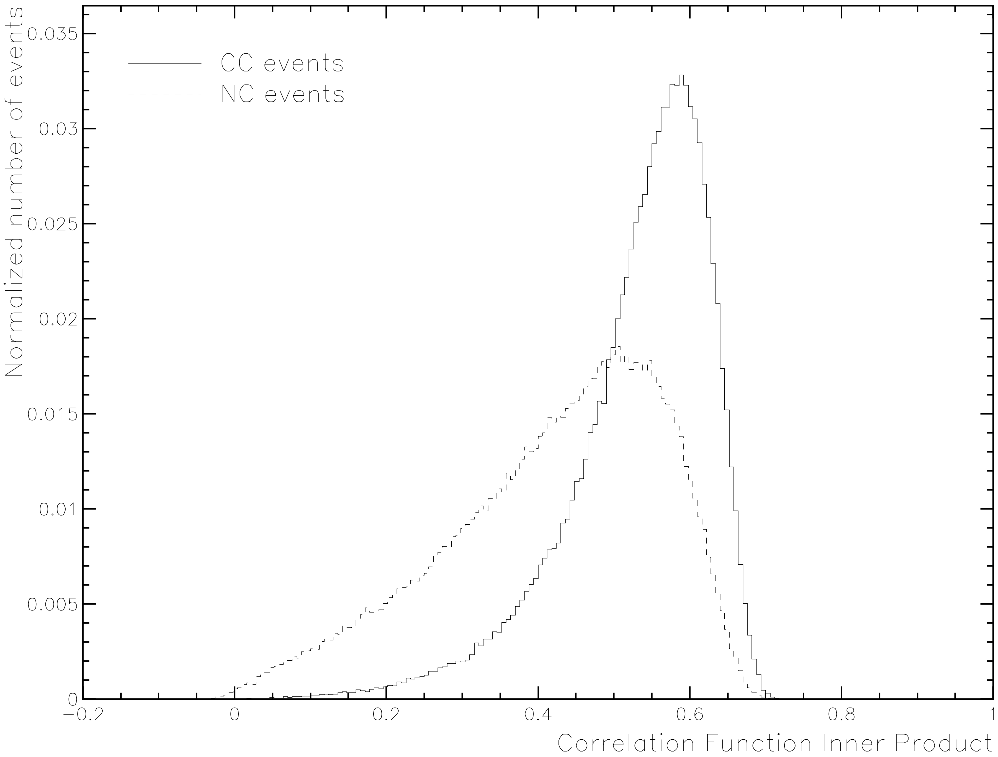


Figure 11: Inner Product with CC/NC Correlation Function Difference

### Mean Pair Angle and the $\beta$-Parameter Combination

Other statistics related to the angular distribution of hit PMTs are used within the SNO collaboration to cut background and distinguish between types of Čerenkov events. The $\langle\theta_{ij}\rangle$ statistic is the mean value of the PMT pair separation angles for an event, after early and late PMT hits have been excluded using Equation 2.

Reference [3] reports results of studying various parameters for event separation. Of the parameters derived solely from the spatial distribution of PMT hits, they found (p. 45) $\beta_1$ and $\beta_4$ to be particularly effective, and recommended that the two be used together. The particular combination $0.25\beta_1 + \beta_4$ has been used elsewhere (see [4] and [5]). The $\beta_1$ and $\beta_4$ statistics are within factors equal to the mean values of the first and fourth order Legendre polynomials of the cosines of the pair separation angles, again after application of Equation 2.

Similar statistics were calculated in this study and used for event separation, except that PMT selection was made with the RTC instead of Equation 2. The distributions of the mean value of the pair separation angle with RTC are shown in Figure 12. Figure 13 shows the distributions of the combination $0.25\beta_1 + \beta_4$.

Note that the mean pair angle, the $\beta$-parameter combination, and the two inner products discussed in the preceding sections are variations on a similar idea. They are all integrals of the product of the angular correlation function with a weighting function; they differ in the choice of weighting function.

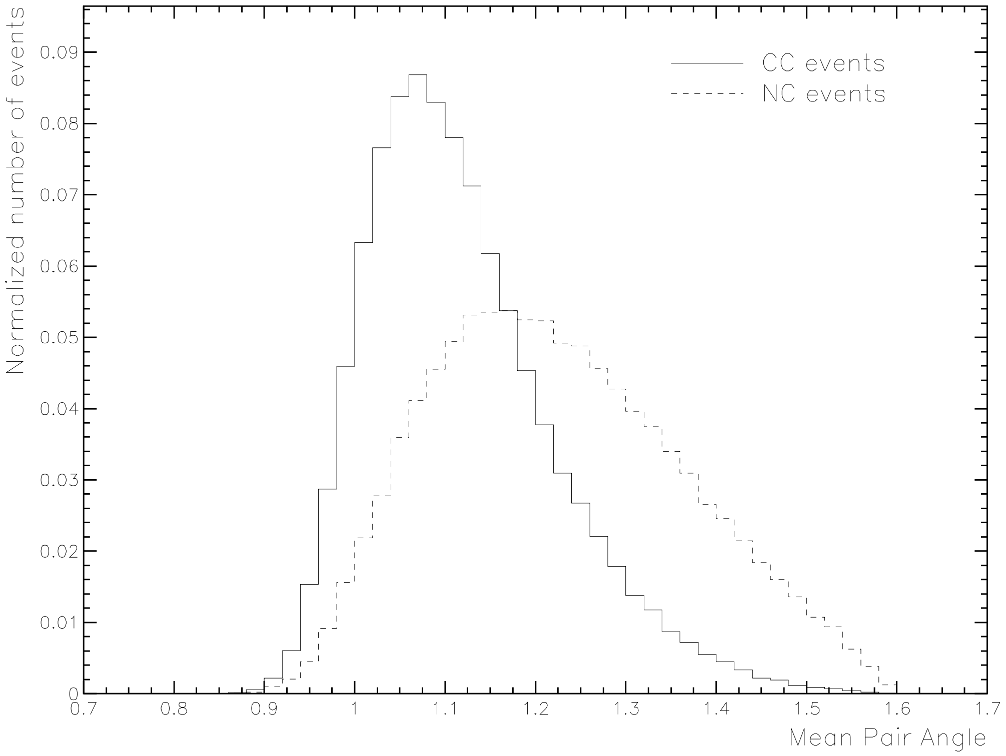


Figure 12: Mean Pair Angle

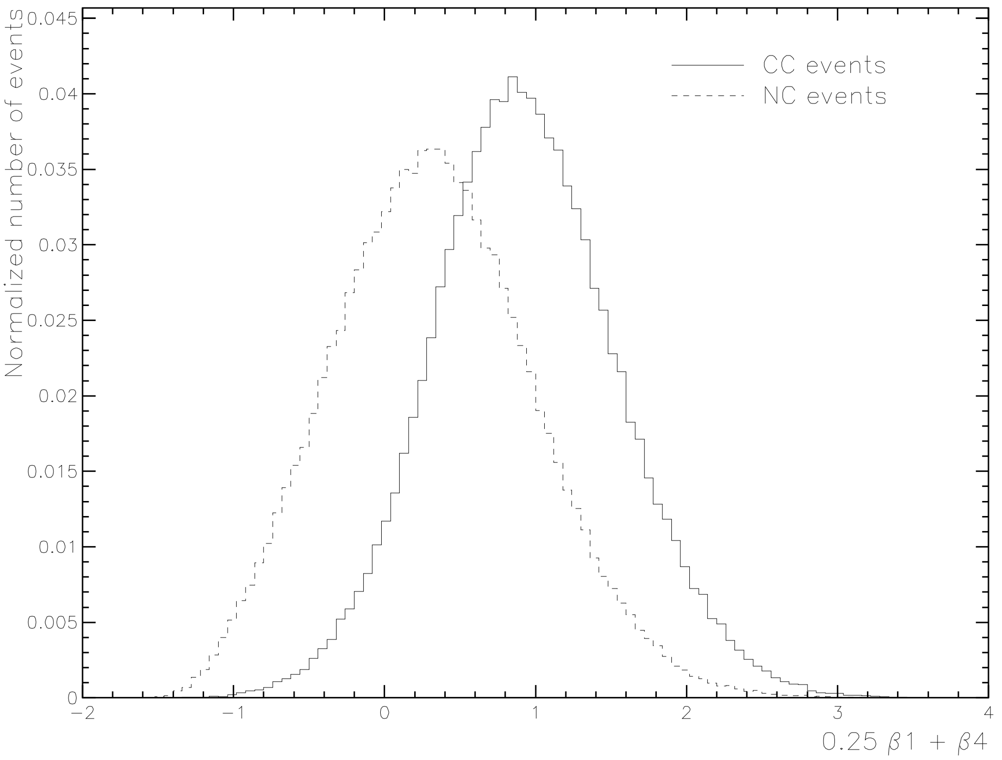


Figure 13: Beta Parameter Combination ($0.25\beta_1 + \beta_4$)

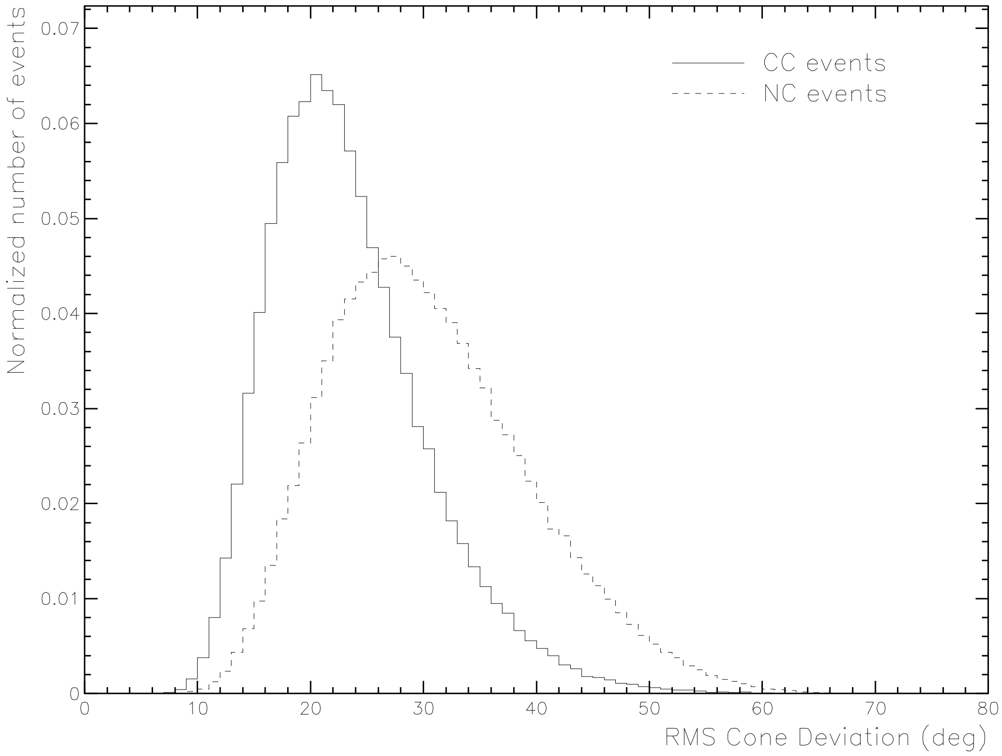


Figure 14: RMS Cone Deviation

**RMS Deviation from 42-Degree Cone**

The last statistic studied is the "RMS cone deviation". (See [5].) The vectors from the fitted event location to the hit PMTs are calculated, and the angles between those vectors and the fitted event direction are determined. The RMS cone deviation is the square root of the sum of the squares of the differences between those angles and 42 degrees. Figure 14 shows the distribution of this statistic for the study set of events.

## Event-by-Event Separation Effectiveness

It is apparent from the distributions of the various statistics that none of them provides a clean separation of CC and NC events, and by eyeball alone it is difficult to identify the best of them. The differences are small.

Event-by-event separation effectiveness was characterized in two ways. The first is simply one minus the overlap of the CC and NC distributions of the statistic. Expressed mathematically,

$$s = \frac{1}{2}\sum_{i=1}^{N} |f_{CC,i} - f_{NC,i}| \tag{10}$$

where N is the number of bins in which the two normalized distributions have been calculated, and $f_{CC,i}$ and $f_{NC,i}$ are the values of the two distributions in the i-th bin. (Others [5] have used the Kolmogorov-Smirnov statistic as a measure of separability.

If there is some value of $i$ below which one of the distributions is greater than the other and above which the inequality is reversed, the two measures are equal.) The separation effectiveness of the various statistics by this measure are listed in Table 1 in descending order.

Table 1: Separation Effectiveness

| Separation Statistic | Effectiveness |
|---|---|
| Correlation Function Ring Inner Product | 0.398 |
| $0.25\beta_1 + \beta_4$ | 0.378 |
| Correlation Function Likelihood Ratio | 0.373 |
| Correl. Fn CC/NC Difference Inner Product | 0.366 |
| RMS Cone Deviation | 0.357 |
| Mean Pair Angle | 0.342 |
| Correlation Function K-S Sum | 0.327 |

One can use any of the above statistics in a hypothesis test to categorize events. One would for example classify events as CC or NC if the value of the mean pair angle was below or above some value respectively. Figure 15 shows the separation effectiveness curves for the various statistics, which indicate how the fractions of events of each type classified correctly vary as the decision value is changed. At the midpoints of the curves, where the fractions of the two event types classified correctly are equal, the curves for the different separation statistics appear in the same order from upper right to lower left as the entries in Table 1. None of the statistics provides sufficient discrimination for event-by-event separation.

## Sample Composition Estimation (Signal Extraction)

The results of the preceding sections indicate that none of the statistics studied has sufficient discriminating power for event-by-event separation. It is probably safe to conclude that further search for a significantly more powerful separation method based on PMT hit distribution information will be fruitless. The amount of information (i.e., the number of PMT hits) is not sufficient; and in any case, significant numbers of events of the two types have Čerenkov photon distributions that are indistinguishable.

It is nevertheless desirable to use the information contained in the PMT hit distributions in SNO signal extraction. Although the hit distribution information would of course be used in conjunction with the energy, radial, and $\cos\theta_{SUN}$ distributions, the work reported here was done using spatial information only, to decide the best use of that information. The objective in all cases was to determine the fraction of CC and NC events in a mixed sample of the two types.

Three fundamentally different approaches were tried: extraction based on the distribution of a single statistic that characterizes the PMT hit distribution for each event; extraction based on comparison of the sample-average angular correlation function with what is expected for pure CC and NC samples; and maximum likelihood estimation

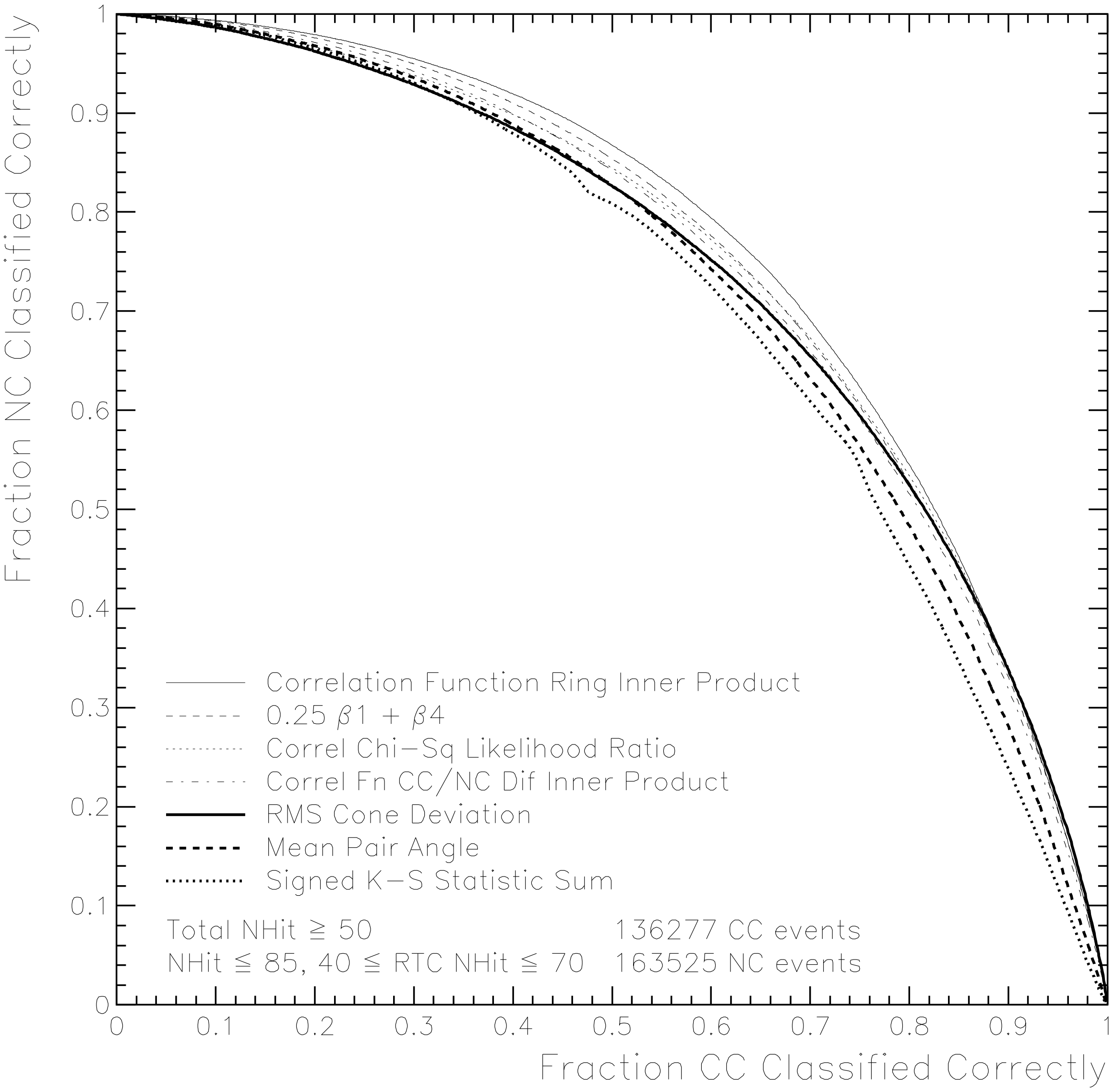


Figure 15: Separation Effectiveness Curves

using angular correlation functions. Extraction effectiveness was evaluated by generating a large number of Monte Carlo "experiments", estimating the sample composition of each by the different methods, and calculating the means and standard deviations of the results. The CC and NC statistic distributions, average correlation functions, and correlation function means and covariance matrices to which the samples were compared were derived from the standard event sets.

### Extraction Using Statistic Distributions

Signal extraction was done using each of the seven statistics that were tested in the previous section for event-by-event separability. The same maximum likelihood method was used for all. First, expected distributions, $y_{cc,i}$ and $y_{nc,i}$, of the statistic in 100 bins covering the statistic's range were derived from the events in the CC and NC standard sets and normalized to sum to one. The number of events falling in each bin was determined for the sample under consideration. The probability of obtaining an outcome with bin occupancies $k_1, k_2, ..., k_N$ is given by the multinomial probability

$$\frac{k!}{k_1!k_2!\ldots k_N!}\ p_1^{k_1}p_2^{k_2}\ldots p_3^{k_3} \tag{11}$$

where $p_i = fy_{cc,i} + (1-f)y_{nc,i}$ and $f$ is the fraction of CC events in the sample. The Fortran routine, "golden", from reference [7] was used to find the value of $f$ for which this probability is maximized.

Note that in deriving these estimates, any dependence of the statistic distributions on r, $\hat{u}\cdot\hat{r}$, or $N_{hit}$ is taken into account implicitly, because the distributions of the CC and NC events in the samples with respect to these quantities are the same as those of the standard event sets, aside from statistical variations.

### Signal Extraction by Direct Correlation Function Fit

To apply this method, the average angular correlation functions were calculated in 20 bins for the sample under consideration, and for the CC and NC standard sets. The sum of squares

$$\sum_1^N [y_i - fy_{cc,i} - (1-f)y_{nc,i}]^2 \tag{12}$$

was minimized with respect to $f$, where $y_i$, $y_{cc,i}$, and $y_{nc,i}$ are the values of the respective average correlation functions, and $N$ is equal to 20. The Fortran routine, "svdfit", from reference [7] was used to find the minimum.

### Maximum Likelihood Estimation Using Angular Correlation Functions

Suppose that event angular correlation functions are calculated at discrete points. The correlation function exhibits statistical variation, and the point values have a complex joint probability distribution. As discussed in Reference [1], the point values are highly correlated and the degree of correlation depends on $N_{hit}$. It was assumed that the distribution of values can be approximated by a multivariate normal distribution.

(Because the correlation function as used here has a fixed normalization, the number of variables in the distribution is one less than the number of points or bins in which the function is calculated.)

Suppose there is a sample of $n$ events, of which $m_x$ are of one type and $m_y$ are of a second type. The values of the correlation functions form vectors with means $\mu_x$ and $\mu_y$ and covariance matrices $V_x$ and $V_y$. The characteristic functions of the two distributions are $\phi_x(t) = \exp(i\tilde{t}\mu_x - \frac{1}{2}\tilde{t}V_x t)$ and $\phi_y(t) = \exp(i\tilde{t}\mu_y - \frac{1}{2}\tilde{t}V_y t)$, where $\tilde{t}$ is the transpose of the column vector $t$. The characteristic function of the sum of the event correlation functions is

$$\phi_S(t) = [\phi_x(t)]^{m_x}[\phi_y(t)]^{m_y} = \exp[i\tilde{t}(m_x\mu_x + m_y\mu_y) - \frac{1}{2}\tilde{t}(m_xV_x + m_yV_y)t] \quad (13)$$

and the characteristic function of the average of the event correlation functions is

$$\phi_A(t) = \phi_S(t/n) = \exp\left[i\tilde{t}(\lambda_x\mu_x + \lambda_y\mu_y) - \frac{1}{2n}\tilde{t}(\lambda_xV_x + \lambda_yV_y)t\right] \quad (14)$$

where $\lambda_x = m_x/n$ and $\lambda_y = m_y/n$. From this we see that the sample average correlation function is multivariate normal with mean $\lambda_x\mu_x + \lambda_y\mu_y$ and covariance matrix $(\lambda_xV_x + \lambda_yV_y)/n$. Given an observed (vector) value $z$ of the average correlation function the probability of obtaining that value as a function of $\lambda_x = 1 - \lambda_y$ is

$$\begin{aligned} p(\lambda_x) &= \frac{1}{(2\pi)^{\nu/2}\sqrt{\det[(\lambda_xV_x + \lambda_yV_y)/n]}} \times \\ &\quad \exp\frac{-n}{2}\left[(z - \lambda_x\mu_x - \lambda_y\mu_y)^T(\lambda_xV_x + \lambda_yV_y)^{-1}(z - \lambda_x\mu_x - \lambda_y\mu_y)\right] \end{aligned} \quad (15)$$

where $\nu$ is the number of correlation function bins minus one. The value of $\lambda_x$ that maximizes $p(\lambda_x)$ is the maximum likelihood estimate.

The situation we are dealing with here is, however, more complicated. As noted previously, the correlation function means and covariance matrices vary with $N_{hit}$, r, and $\hat{u} \cdot \hat{r}$. Perhaps this problem could be addressed by calculating mean correlation functions and covariance matrices from the standard sets as a whole, with the same cuts applied as for the sample events (the approach taken with the other extraction statistics). However, what was done was to divide the events into subsets, in increments of $N_{hit}$ of three, and further within each $N_{hit}$ interval into four subsets as shown in Equations 3 to 6. The likelihood function (Equation 15) becomes a product of likelihood functions for each of the subsets. But because the CC and NC events do not have identical distributions in $N_{hit}$, r, and $\hat{u} \cdot \hat{r}$, if $f$ is the fraction of events of type x in the sample as a whole, the expected fraction $\lambda_{x,j}$ in the j-th subset is not in general equal to $f$. Instead, the expected values of $\lambda_{x,j}$ and $\lambda_{y,j}$ are

$$\lambda_{x,j} = \frac{fh_{x,j}}{fh_{x,j} + (1-f)h_{y,j}} \quad (16)$$

and

$$\lambda_{y,j} = \frac{(1-f)h_{y,j}}{fh_{x,j} + (1-f)h_{y,j}} \quad (17)$$

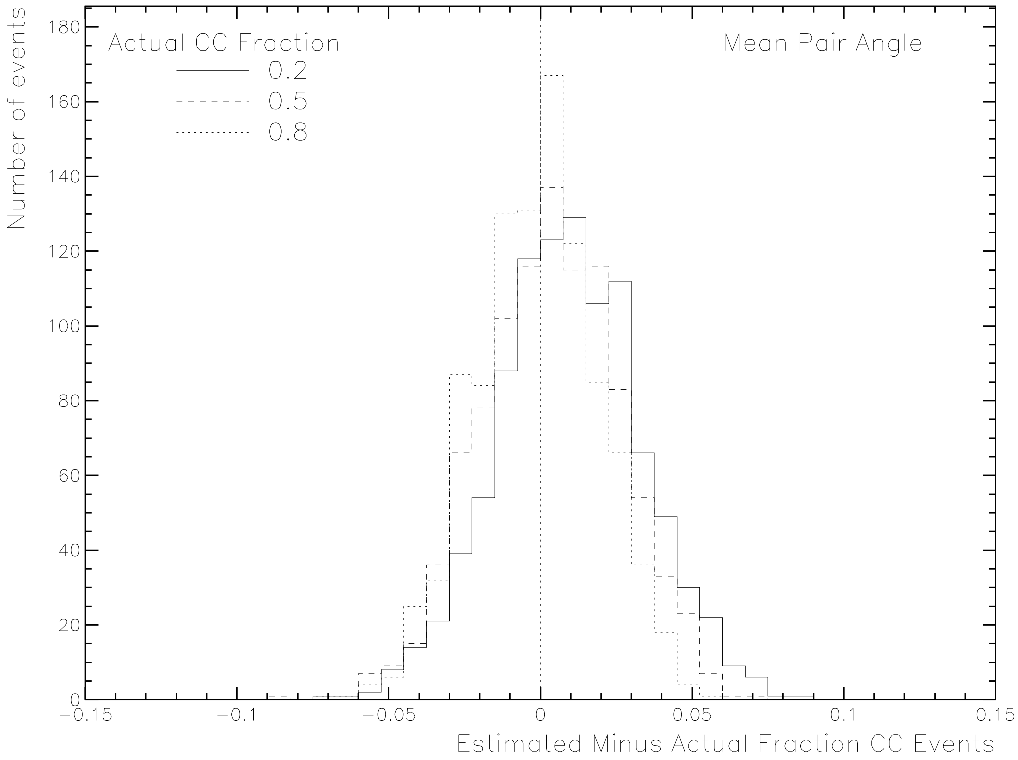


Figure 16: Estimation Errors Using the Mean Pair Angle Distribution

where $h_{x,j}$ ($h_{y,j}$) is the expected fraction of the total events of type x (y) in the j-th subset. Routines from [7] were also used in these computations: a double precision version of "choldc" for inverting the covariance matrices, and "golden" for finding the likelihood function maximum.

## Study Sample Selection and Results

Randomly chosen event samples were used to study signal extraction accuracy. For each of 11 values of $f$ (the fraction of CC events in the sample) ranging from zero to one in steps of 0.1, 1000 mixed samples of CC and NC events were drawn at random from the study event sets, with 2500 total events in each sample. The methods described above were applied to the samples to obtain estimated values of $f$, which were then compared to the actual values.

Figure 16 shows typical results, in this case for fits to the distribution of the mean pair angle. The results for all of the estimation methods are summarized in Figures 17 and 18, which show smoothed curves of the standard deviations of the sample estimates as a function of the actual fraction of CC events in the samples.

Figure 17 shows the results using six of the seven statistic distributions. Figure 18 shows the results from the direct correlation function fit, the maximum likelihood method applied to the event angular correlation functions, and the last of the statistic distribution fits. (To help in identifying the curves, the legends in Figures 17 are listed in the order of increasing standard deviation at an actual CC fraction value of 0.5.)

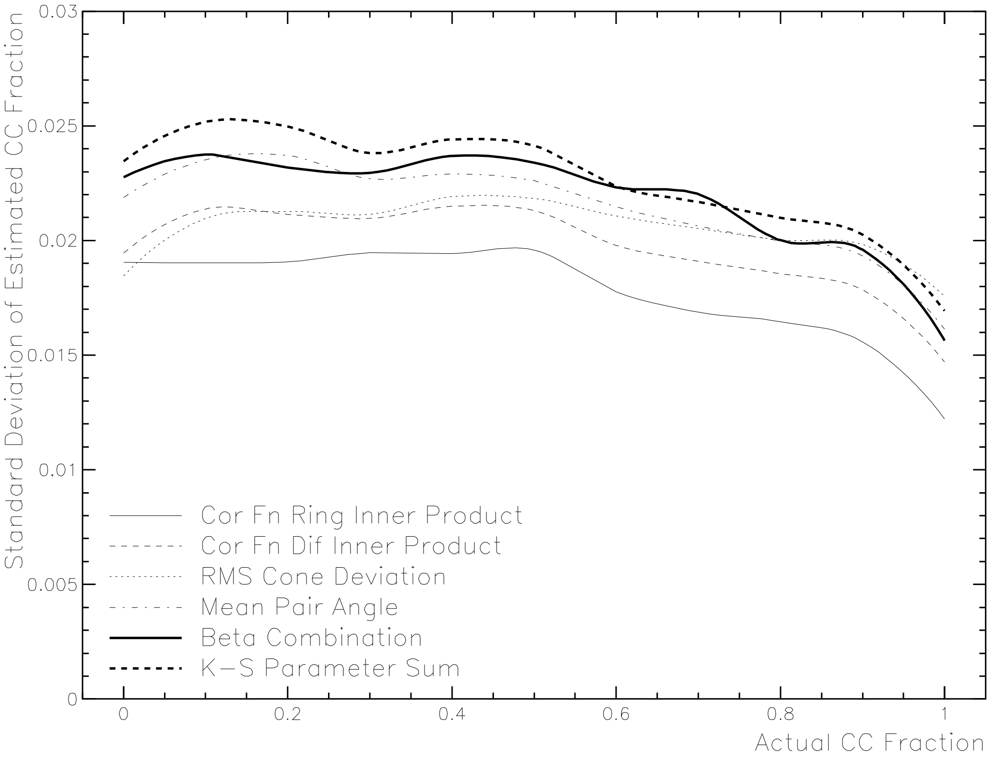


Figure 17: Standard Deviations of Estimation Errors

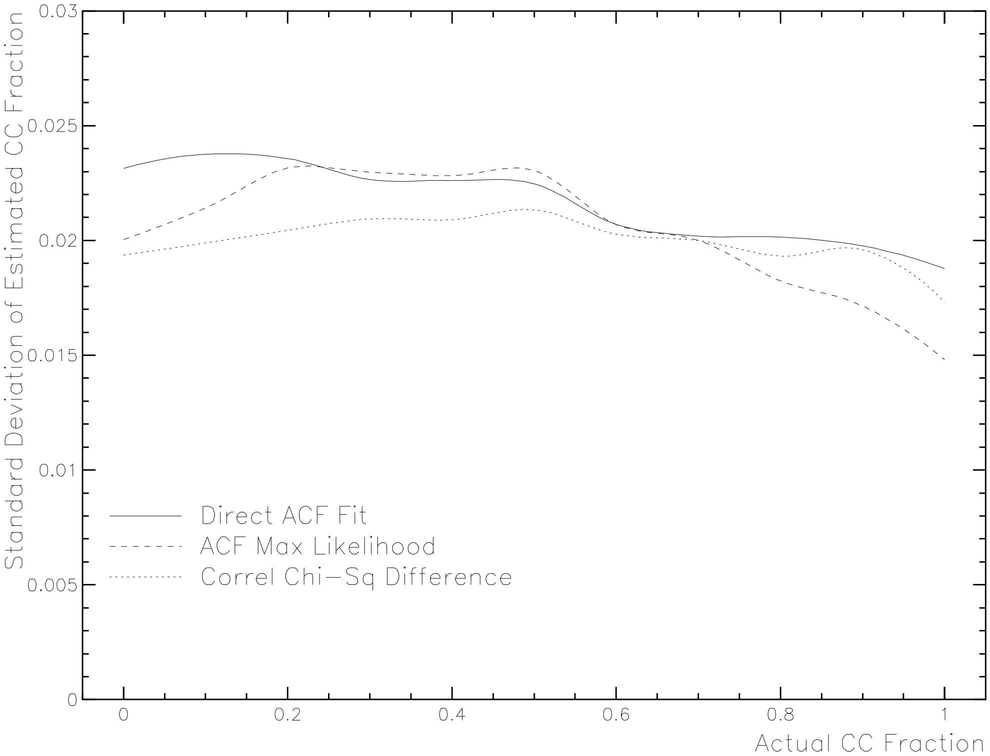


Figure 18: Standard Deviations of Estimation Errors

The mean values of the sample estimates were also calculated to ensure that the estimates are not biased. For the samples just described there were noticeable biases, but it was later concluded that these are due simply to real, albeit small, differences between the standard and study event sets. When event samples are drawn from the standard sets rather than the study sets, none of the methods produces a significant bias.

As with the event-by-event separation results, there are not large differences in the separation effectiveness as measured by the standard deviations. Also, a ranking of separation effectiveness as measured by the standard deviations is not consistent in all cases with Table 1 and Figure 15. When measured by the standard deviations, the two correlation function inner products give the best results.

Some of the methods can be eliminated from consideration for reasons other than the quality of their results. The ACF maximum likelihood method is complex, and both that method and the correlation chi-square difference require extensive preliminary computations to generate standard correlation functions and their covariance matrices. The RMS cone deviation method requires the use of fitted event direction as well as position. Finally, it's not apparent how to combine, in a statistically consistent way, the method of fitting directly to the average sample correlation function with the $N_{hit}$, r, and $\cos\theta_{SUN}$ data.

## Conclusions

The signal extraction methods studied show only small differences in their statistical power for estimating the fraction of CC events in a mixed CC/NC sample. However, when measured both by event-by-event separability and by RMS signal extraction error, the correlation function ring inner product is best by a small margin. The ring inner product signal extraction estimates have a standard deviation about 15% smaller than those of the mean pair angle.

## References


[1] W. Heintzelman. *Use of the Angular Correlation Function to Identify Background SNO Events*, SNO-STR-2000-032, Jan 2001

[2] W. Heintzelman. *Use of the Angular Correlation Function to Distinguish NC Events in Salt (Preliminary)*, SNO Analysis Note, 28 Sept 2001

[3] M.G. Bowler, S.J. Brice, X. Chen, N.A. Jelley, M.D. Lay, M.J. Lyon, M.D. Thorman, D.L. Wark, N.West, *Some Monte Carlo Studies of the Added Salt Option*. SNO Analysis Note, 4 May 1996

[4] J.A. Dunmore. *The Beta Parameters*, SNO Analysis Note, 15 Jan 2002

[5] A.D. Marino. *Studies of the Separation of Signals by Isotropy Patterns*, SNO Analysis Note, 5 Feb 2002

[6] A.D. Marino. *Studies on Leslie Events*, SNO Analysis Note, 14 Mar 2002

[7] Press, Teukolsky, Vetterling, & Flannery. *Numerical Recipes in Fortran*, Cambridge University Press, 1992