SNO-STR-2000-022

Using Time and Angle Information in Reconstruction and the SNOMAN Path Fitter

Josh Klein, Mark Neubauer

University of Pennsylvania

March 22, 2001

Abstract

We describe here the design and performance of the Path Fitter, a maximum likelihood reconstruction algorithm which incorporates time and angular information in the fit for both position and direction. The Path Fitter includes this information via a 'first principles' approach to the handling of tube-by-tube hit probabilities, and the resultant performance is demonstrated to be superior to other standard SNO fitters. We also describe some of the studies of angular fitting alone, uncovering some of the subtleties of using an average (rather than event-by-event) angular PDF. Suggestions for upgrades to the Path Fitter are also presented, and it is hoped that with the Path Fitter as a template others will be able to test and incorporate better ideas into the fitting procedure.

Contents

List of Figures

List of Tables

1 Introduction

Many reconstruction methods have been developed for SNO over the years. These make up a diverse set of algorithms which all try come up with clever (and sometimes rather elaborate) ways to best estimate the position and direction of events occurring in the detector. Reconstruction is one of the most important parts of neutrino analysis in SNO because of the importance of accurately determining the position and direction of events for perserving distributions that distiguish the signals (CC, ES, NC) and for rejecting backgrounds. It may seem peculiar that there are so many reconstruction methods in SNO given that the information one gets out of events in so simple - time (relative to the trigger) and integrated charge for each hit PMT. It is, however, this relatively small amount event of information which makes reconstruction challenging in SNO, as fitters need to determine the best way to use all of the existing information in events.

This report provides documentation for a relatively new reconstruction method - The Path Fitter - that is part of the standard SNOMAN distribution[1]. It also includes a general treatment of the effects of including PMT angular information along with time information in vertex reconstruction, as this is the primary approach taken by the Path Fitter which largely sets it apart from other standard fitters.

The organization of this reports is as follows. Reconstruction using time and angular information is discussed in Section 2. Speed of light consideration relevant for fitters is described in Section 3. The Path Fitter reconstruction method - both general approach and the fitter implemented in SNOMAN - is presented in Section 4.1, with some performance on signal MC detailed in Section 5. Finally, some notes on late calculations are placed in Appendix A, which may prove useful for future upgrades to the Path Fitter.

[1] versions 4.0083 and later.

2 Time and Angle Information in Reconstruction

More information is good. —*Anonymous*

Well, more information can't hurt. —*Anonymous, two years later*

In any fitting process, we expect that the inclusion of additional information—more data points, for example—will generally act to improve the fit. Even in the case where the information is of poor quality, such as an additional data point with a large uncertainty, the proper weighting of the information in the fit should ensure that the fitted parameters are not much changed. Of course, this is only true if the additional information is not systematically in error.

In the reconstruction of SNO events, we have several opportunities for including 'more information.' A simple time fitter will throw out light which it feels is 'too late' or 'too early' to be used in the fit, but which in principle could be included if one understood how the arrival times of the late photons were distributed in the detector. Similarly, the angular distribution of phototube hits in the detector also provide information about the event position (and, more obviously, about the event direction), and thus the inclusion of the angular distribution of PMT hits can in principle improve the accuracy of the reconstructed position.

There are further motivations, however, for including the angular information in reconstruction of the event position. One of the well-known systematics associated with the reconstruction of Cerenkov light is the so-called 'drive': the positions tend to reconstruct downstream of the true event position, along the event direction. The cause of this effect is the angular symmetry of the Cerenkov light—the narrow angular separation of the tubes provides a poor constraint on the fit position along the axis of symmetry, and any amount of late light not properly accounted for will tend to push the event closer to the tubes in order to make the expected time difference larger, thus better matching the measured time differences. Given that the drive is caused by the angular symmetry of the tubes, it seems natural to imagine that including the angular symmetry explicitly in the fit will provide some correction for the effect. The worse the late-light problem (for example, because of hardware problems or poor calibrations) the more important the angular information may be.

While the principle is simple, in practice demonstrating that the angular information aids in the position reconstruction is not an easy task. Earlier work [1] showed that simply including the angular distribution of phototube hits (averaged over positions and directions in the D_2O) actually over-corrected the drive. Events were pushed backwards because this allowed more tubes to be crowded in to the peak of the $dN/d\cos\theta$ distribution. The limiting case, of course, was to drive the event far outside the PSUP, where essentially all tubes could be placed into this peak. Work by Steve Brice [2] showed that the angular information itself was too weak to improve on the position determined by the time information. Bill Frati, on the other hand, uses the 'backward drive' created by the angular fit to cancel the forward drive of the time fit, by explicitly tuning weighting factors for each.

At least part of the problems found previously were caused by the fact that the angular information was not being included properly. Averaging over all positions and directions removes a lot of information—in particular the fact that the solid angle of the tubes increases the probability of a tube being hit as one moves closer. The Path Fitter does a more proper accounting of the angular information by recalculating the the angular PDF at each hypothesized position and direction by convolving the average distribution of *photons* produced by a Cerenking electron with the solid angle of the hit tubes seen from that hypothesized position and direction.

2.1 2D 'Perfect World' Fitting

Attempting to study the effects of angular fit in the full Monte Carlo with its associated complex geometries and light patterns is not the best first-order approach. Instead, we created a very simple 2D world in which the coverage was 100%, there was no loss of light to attenuation, the phototubes had perfect quantum efficiency, and the generation of events used the identical angular and time distributions as was used to fit them.

What we discovered here was somewhat surprising, and is shown in Figure 1, which plots the fit x and y positions for 10000 events, when only angular information is used in the fit. The crescent shape of the fit event positions (actually a segment of an ellipse) is characteristic of what we call the 'swirl'—the fact that fitting an event using only angular information appears to drive the fit positions off axis and forwards.

The explanation of the swirl is simple: hit tubes that are near the Cerenkov angle can be kept at the Cerenkov angle if the event position is moved forward and off-axis as long as

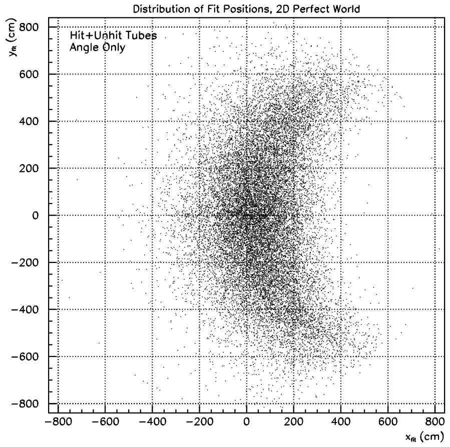

Figure 1: Scatterplot of x and y fit positions for 2D world with event generated at (0,0) with direction along the x axis.

the event direction is altered to compensate for the shift. Figure 2 shows an illustration of how this works. Figures 3 and 4 show a simple event display for an individual fit compared to the original position and direction.

Moving *backwards* in the case of angular fitting is much harder, because the angle between the hit tubes becomes smaller than the Cerenkov angle as the fit position moves further away from the tubes than the true position. In addition, by moving forward, the increase in PMT solid angle means that the fit can make up some loss in likelihood due to the tubes moving a little off the Cerenkov cone. The increased solid angle also allows some slop in the fit direction which can place the tubes closer to the Cerenkov angle when the fit is closer.

Based on these plots, we might expect to gain something by including in the likelihood the probability of tubes to *not* be hit. After all, as we move closer we not only increase the solid angle of the tubes which are hit, but also those which are not hit. Shouldn't unhit tubes near the Cerenkov angle penalize the fit as one gets closer to the tube? The Path Fitter implements an 'unhit tube probability' if the user selects this option, but we found that while there was some improvement it did not become significant until many tubes were

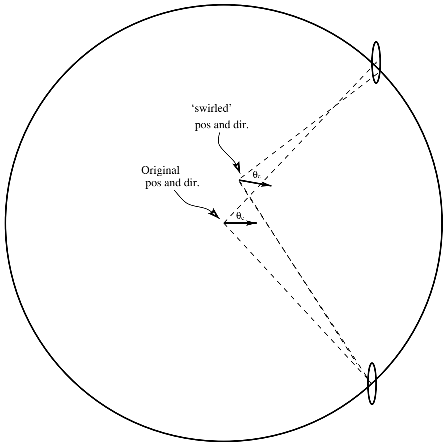

Figure 2: Schematic explanation of 'swirl' effect in angular fitting.

hit. In the case of high energy events, such an option might be practical and useful. Were the Path Fitter to also fit for event energy (as a generalized maximum likelihood, perhaps), one would have to reduce the energy of an event as one got close to tubes where only a few were hit, and this would reduce the probability that other tubes in the event were hit and perhaps push the events back, away from the PSUP.

Additionally, one could imagine using not only hit tubes, but charge, which would further push an event back if it were near tubes with only small charges. The Path Fitter as it stands does allow more than one hit on a tube through its poisson probability, but in the real world where charge resolution is poor, this is not used. In the 2D world this *is* implemented but again does not appear to have a large effect.

The first question we must ask is then, does this make the inclusion of angular fitting worse than not using it at all? The answer is that in this case the angular fitting in combination with time fitting still does better than time fitting only. The reason is that the while the time fit has a poor constraint along the direction of the event, the angular information has its best contraint in that direction. Figure 5 shows a contour plot of the likelihood space for fits using time only, and Figure 6 using angular information only. The elongation of the

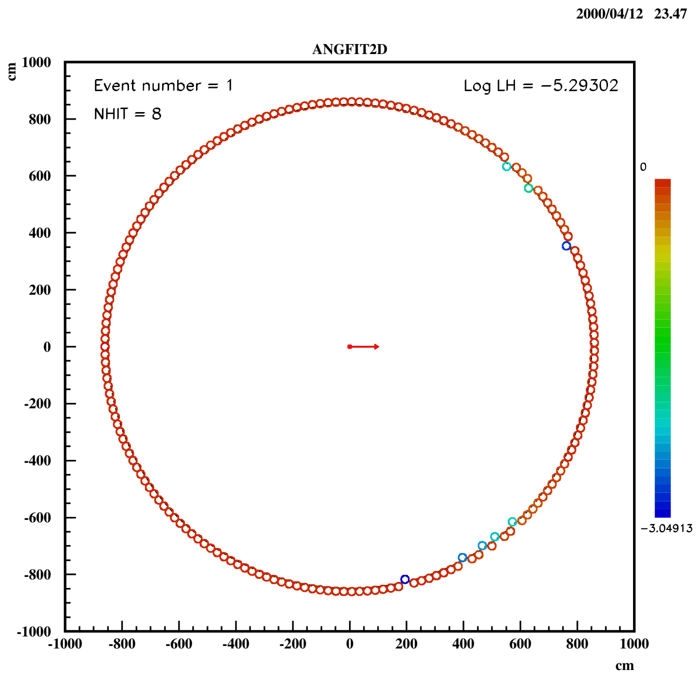

Figure 3: 2D event display showing generated event position and the hit tubes in the event. The colors correspond to the log probability contribution to the likelihood of each PMT (hit and unhit).

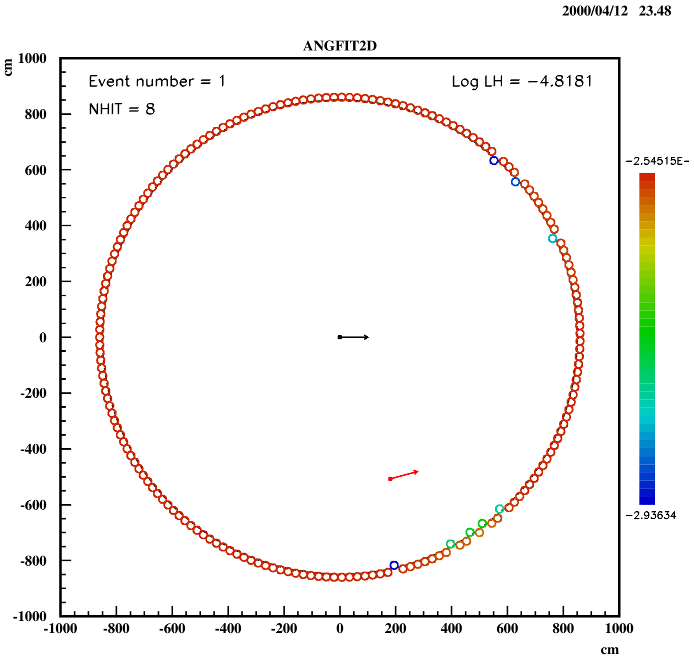

Figure 4: 2D event display showing original and 'swirled' event position. The colors correspond to the log probability contribution to the likelihood of each PMT (hit and unhit).

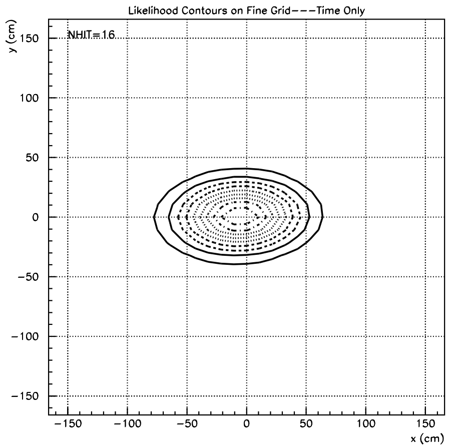

Figure 5: Likelihood contour for 2D fits using time only.

time fit contours along the direction of the event shows the characteristic effects of geometry on time fitting, though in the perfect 2D world in which there is no late light there is no systematic offset of the position. The shape of the angular contour, however, shows that right near the fit position the angular contour is steepest in the forward direction and overall has its best constraint *along* the direction of the time fit. While advantageous, in the 2D perfect world the angular information is not likely to buy us much on position resolution because the timing information is already so good. However, as the timing gets worse—due to scattering, TSLH effects, bad calibrations, late pulsing, variations in c, etc.—the angular information will become more important.

Figure 7 compares the resolution in the 2D world for time only and time+angle fits. For a narrow tube time spread ($\sim$ 1.5 ns) there is a small improvement in the position resolution. Figure 8 shows the effects of the angular fit on events whose tube time spread is made unnaturally large. Here the angular information significantly narrows the position resolution. These plots are the first time we have seen that the proper inclusion of the angular information improves the fit for position.

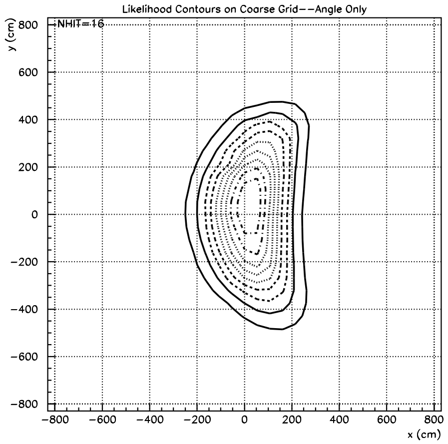

Figure 6: Likelihood contour for 2D fits using angular information only.

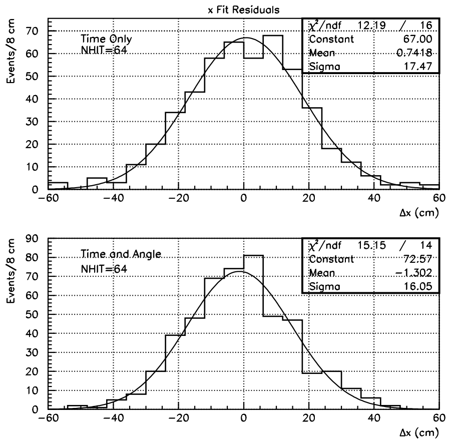

Figure 7: Comparison of x residuals in 2D world using time only and time+angle information, for 1.5 ns tube timing.

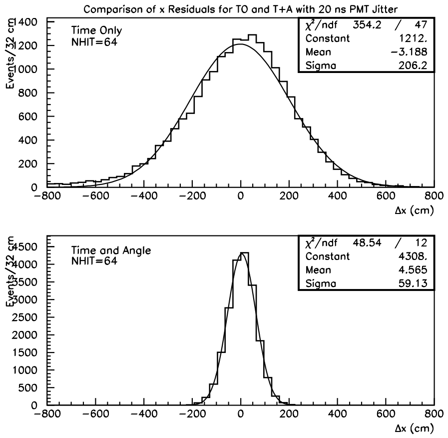

Figure 8: Comparison of x residuals in 2D world using time only and time+angle information, for 20 ns tube timing.

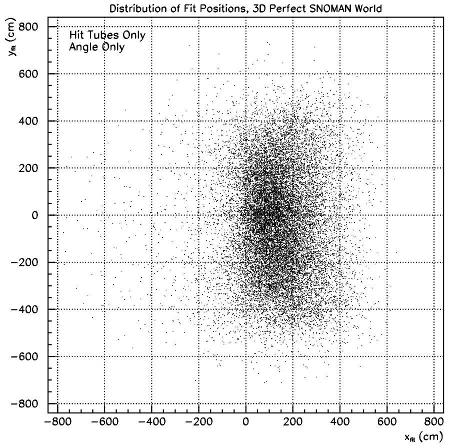

Figure 9: Scatterplot of x and y fit positions in 'perfect' 3D world.

2.2 3D Fitting

The next question we need ask is, do these conclusions apply in a more realistic world than the simple 2D case? To explore this, we went back to SNOMAN and turned off all the things that made SNOMAN different in timing than our 2D perfect world: scattering, reflections, different speeds of light (we used a fixed value for all photons), late pulsing in the tubes (we replaced the complete distribution with a gaussian), and anything else which might make the timing distribution asymmetric or non-gaussian. Figure 9 shows a scatterplot of the fit positions of an event at the center, projected onto two dimensions.

It is clear from Figure 9 that, even accounting for the increased density due to the projection from three onto two dimensions, the distribution of fit positions is not the same in 3D as it is in 2D. While there is some evidence of the swirl, this lies beneath another distribution which looks more like the forward drive of a time-only fitter.

What is the source of the difference between 3D and 2D? Doug Cowen pointed out that the generation of events in 2D and 3D are different: in 2D, we used our angular PDF to both generate the angular distribution of hit tubes and to fit them, while in 3D SNOMAN (and EGS) generates the hit tube positions. Bill Heintzelman further pointed out that our angular PDF assumes that the distribution of hit tubes is ϕ symmetric which is not true for most real events—the multiple scatters of the electron are not guaranteed to be azimuthally symmetric around the initial (or even average) direction. So in fitting in 2D, we were fitting a correct ϕ symmetric hypothesis to our ϕ symmetric event distributions (of course, in 2D, there really isn't a 'ϕ' but just the mirror symmetry about the x axis) while in 3D we were fitting the ϕ symmetric PDF to an inherently (event-by-event) non-symmetric ϕ distribution of hit tubes.

How could this be proven? We did this by replacing SNOMAN's calls to EGS with a call to an angular photon distribution identical to the one we used in 2D but with the addition of explicit ϕ symmetry to make it three dimensional. The projected scatterplot of fit positions is now shown in Figure 10, which looks much more like the 2D case (but even narrower). The apparent forward shift of the points on the scatterplot relative to (0,0) is due to the projection of the three dimensional distribution onto two dimensions—the most probable fit postion is actually at the origin.

We thus conclude that our fit hypothesis is not yet ideal for 3D fitting, and ultimately some improvement may come by creating a PDF which more correctly represents the distribution of tube hits for an individual event. This is not particularly easy, as we must in principle deal with the probabilities of individual electron scattering trajectories, including the scatter-to-scatter correlations in direction. However, we believe that with such a correct PDF, we will gain even more from the angular information.

However, as the rest of this memo describes, even our slightly incorrect PDF gives us an improvement on many fronts over time-only fitters.

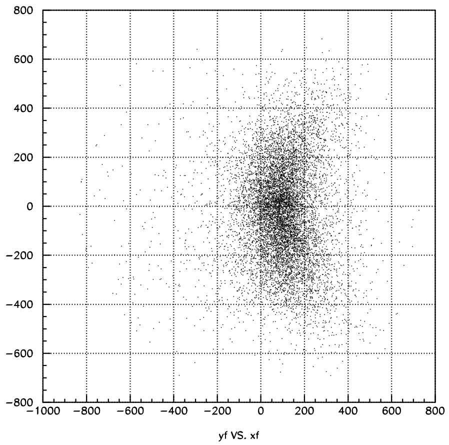

Figure 10: Scatterplot of x and y fit positions in 'perfect' 3D world, using ϕ symmetric PDF to generate the angular distribution of photons.

3 Speed of Light Considerations

We might imagine that the speed of light that should be used for reconstruction would be an easily determined number. As it turns out, however, SNO's multiple media make this a far more complex problem than that of a single-media Cerenkov detector.

On the one hand, there is the obvious problem of the varying indices of refraction as photons travel from D_2O, through acrylic, and H_2O. This in itself is a small effect—the heavy and light water have nearly identical indices while the acrylic is neither far off nor (in most cases) very thick.

A far more complex (and important) issue is interplay between the broad spectrum of Cerenkov light and the responses of each of the media to different wavelengths. The steep dependence of the indices of each media with wavelength mean that photons drawn from different spectra can have significantly different average (or most probable) speeds. Since the absorption of each medium alters the spectrum of detected photons slightly, the path that a photon takes from its origin point to its detection point may have a large effect on its most probable speed.

Figure 11 shows the spectrum of detected photons for events in the D_2O, illustrating how the flat Cerenkov spectrum is affected by passing through heavy water, acrylic, light water, and finally by the phototube's wavelength response (of course, in the last case, the PMT is not changing the spectrum but just detecting a portion of it).

Comparing the photon speed distribution for events in the heavy water to those in the light water illustrates the overall problem nicely. Figure 12 is the distribution of speeds for detected photons for events originating in the D_2O while Figure 14 is that for events in the H_2O. Figure 13 and compare the distribution of speeds for D_2O to the actual spread in arrival times at the PMT's caused by the spread in speeds. Two things are apparent from the figures: the average speed of photons for events in the light water is slightly lower than that for photons in the heavy water, and spread of speeds in the light water is much broader than that for the heavy water. The decreased average H_2O speed results from the fact that prompt light originating in the light water may travel directly to the phototubes without going through any acrylic. These photons will have more short wavelengths than those which pass through the acrylic, and hence they will on average travel slower. The increased spread for photons originating in the light water is due to the fact that these photons can travel to

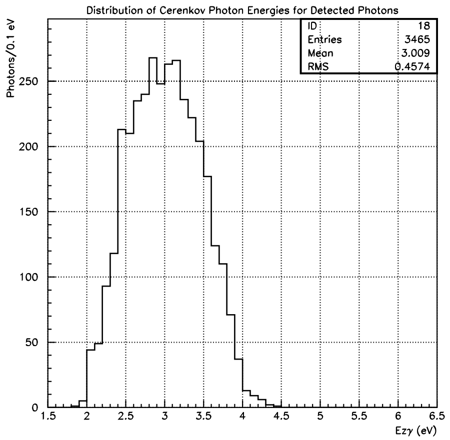

Figure 11: Distribution of energies for detected photons produced by events in the D_2O.

the PMT's through very different paths—from the light water only paths to paths passing through two thicknesses of acrylic or even through the neck twice. In the cases where the photons do travel through the acrylic twice, even more of the high frequency photons are lost, making the average speed for these paths higher than that for the photons originating in the D_2O.

In principle, the correct way of handling these varying speed distributions would be to include the differences in the actual time PDF by convolving the prompt light (and late light in a more complex way) with the possible speed distribution for each photon path. A more simple approach would be to use the most probable speed for each photon path or, even a speed which represents the most probable for all positions and paths, which is what we have chosen for the Path Fitter for now by using 22.75 cm/ns, which works reasonably well for both light and heavy water.

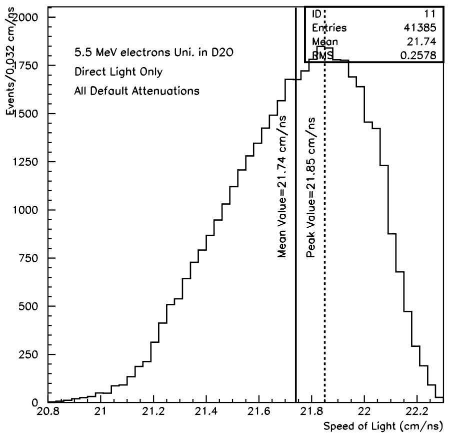

Figure 12: Distribution of speeds for detected photons produced by events in the D_2O.

4 Path Fitter Reconstruction Method

The Path Fitter reconstruction method has grown out of attempts to improve time-only based fitters by simultaneously including PMT angular information with hit PMT time information to determine the most likely vertex of an event. The motivation and success of this approach under ideal MC situations was discussed in Section 2. Another motivation was to create a general maximum likelihood fitter which uses a different PDF for each hypothesized electron position and direction in the detector that is a properly weighted sum of photon path probabilities (direct, scattered, reflected, etc). This is an ambitious task indeed, both in terms of CPU resources and the ability to accurately calculate (or model) the late light as there is currently no calibration late light in SNO. As a result, the fitter installed in SNOMAN is a scaled-down version of original Path Fitter principles which attempts to incorporate the best parts of the method with simplifying assumptions about the late light component to events.

Section 4.1 describes the general approach of maximum likelihood fitting using both time

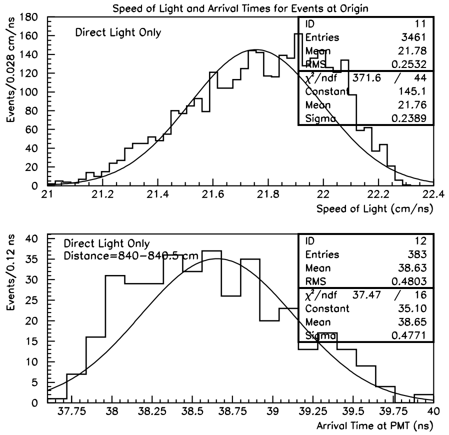

Figure 13: Distribution of speeds for detected photons produced by events in the D_2O and resulting spread in PMT arrival times.

and angular information. It is argued that this approach provides a natural framework within which to refine reconstruction in SNO by including more detailed information about the detector into the fitting procedure. Section 4.2 details the Path Fitter as implemented in SNOMAN (development version 4.0083 and later).

4.1 General Considerations

4.1.1 Fitter Development - A Brief History of Time and Angle

Aside from the initial work [1] by Josh Klein on maximum likelihood fitting using time and angle, our work on reconstruction started with the idea of using Monte Carlo on a event-by-event basis to determine the PDF used in the process of fitting events. The main point was that an ideal fitter in terms of using all available information would be a maximum likelihood fitter that employs the full MC to determine the probability density for a given PMT to record a given charge and time. In practice, the PDF for a given hypothesized electron position,

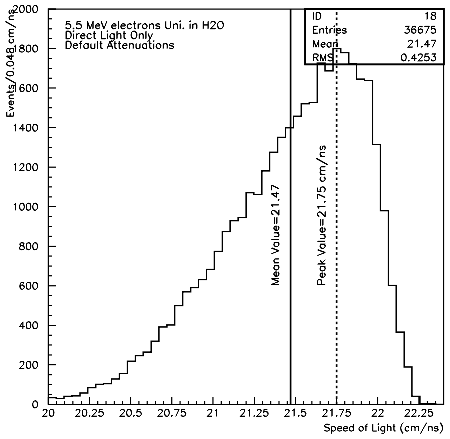

Figure 14: Distribution of speeds for detected photons produced by events in the H_2O.

direction, time, and energy would be determined by simulating this electron many times (an infinite number, in principle) and noting the relative number of times that each PMT records a given time and charge.

In practice, there are several problems with such a fitter. The MC does not perfectly simulate the physics nor the detector response. A more serious difficulty is that simulating the large number of events required to get a statistically significant PDF and re-doing this many times as the minimization routine searches the phase space of possibilities requires an impracticable amount of CPU power. An attempt was made to speed up the MC PDF determination by preprocessing the Cerenkov photon production in terms of the angular distribution of light produced by the electron. This is possible because while PDF determination by the previously described method involves summing over many events, one can just use the angular distribution of Cerenkov light to get the PDF by just generating one event having a very large number of photons. The MC is then used to simply propagate the photons and simulate the detector response. Such a fitter - called the "bomb" fitter since a large photon bomb with Cerenkov angular distribution is used to estimate the PDF at

each minimization step - was attempted and abandoned because of the impractible amount of CPU time required to fit events.

In the Path Fitter, a different approach is taken to estimating the PDF. In principle, the SNOMAN MC is simply using Monte Carlo techniques to do a set of integrals - albeit a very complicated set of integrals containing many different parts that reflect the richness and complexity of the detector. It may be possible to simply calculate the most important of these integrals to get PDFs which are reasonable approximations of the "true" PDFs resulting from the full MC calculations.

The challenge in calculating a reasonably accurate PDF is to properly reproduce correlations between timing and angular distributions of hit PMTs. The principle approach of the Path Fitter is to break the PDF into a sum of terms, each of which represents the probability for a given PMT to be hit at a certain time if the Cerenkov photon causing the hit follows a particular type of path (e.g. direct light, scattering, reflections). Each of the terms then has the proper correlation between time and angle that can, in principle, be calculated. Therefore, we have the following estimate for the PDF,

$$\mathcal{P}_i(n_i, t_i; \vec{r_e}, \hat{d}_e, t_e, E) = \sum_{jpaths} P_{ij}(n_i, t_i; \vec{r_e}, \hat{d}_e, t_e, E)$$

where P_{ij} is the probability density corresponding to the j^{th} path type for the i^{th} PMT to be hit n_i times and record a time t_i, given an electron initially at $\vec{r_e}$, t_e in the direction $\hat{d}_e$. Note that higher order terms (e.g. paths corresponding to scatter + reflection or scatter + scatter) are left out of the PDF calculation.

The path-dependent density functions P_{ij} separate into the product of time and angle parts,

$$P_{ij} \sim P_{ij}^{TIM}(n_i, t_i) P_{ij}^{ANG}(n_i)$$

The reason this is not shown as an equality is because the calculation for each of the path types actually involves integration of this functional form. For example, the scattering term involves integration over all scatterers in the detector, and there are several paths that a reflected photon can take into a given phototube. Note that n_i is included in the time function to suggest the possibility of including multi-photon effects in the time PDF. The time function P_{ij}^{TIM} would then have the same form (e.g. Gaussian PMT jitter) for each of the paths, differing only in the photon path distance. The angular part is the Poisson probability for a given PMT to be hit n_i times,

$$P_{ij}^{ANG}(n_i) = \frac{1}{n_i!}\lambda_{ij}^{n_i} e^{-\lambda_{ij}^{n_i}} \qquad \lambda_{ij}^{n_i} \equiv N_\gamma \rho_{ij}$$

where N_γ is the number of Cerenkov photons generated by the electron (can be roughly estimated from NHIT and the overall PMT efficiency + coverage), and ρ_{ij} is the probability for a Cerenkov photon generated by the electron to be detected by the i^{th} PMT. The probability ρ_{ij} depends on the angular distribution of Cerenkov photons from the electron, the solid angle subtended by the i^{th} PMT from the electron position, and optical effects such as attenuation and reflectivity.

The main disadvantage of using calculated rather than Monte Carlo generated PDFs in a fitter is the difficultly of properly estimating contributions from reflections and scattering. In terms of reflections, only the dominant PSUP reflections were included in PDF calculations and the PSUP was approximated as a spherical mirror with uniform reflectivity. Scattering was considered by assuming a uniform distribution of scatterers in the detector volume, which scatter light isotropically with a probability related to the scattering length of the medium.

While PDFs generated from modeling reflections and scattering in this way qualitatively match the Monte Carlo results, there are several problems with this approach. From a practical standpoint, a typical fit requires many such calculations to be performed, since the PDF needs to be recalculated at each electron phase space point. The scattering component to this calculation is particularly CPU intensive, since it involves mult-dimensional integration for each hit PMT to add up contributions from all scatterers in the detector volume. In addition, modeling of the PSUP as a spherical mirror is too simplistic as there are substantial non-specular PSUP reflections which need to be included. Appendix A contains some notes on simple-minded calculations of reflections and scattering for the interested reader.

In the interest of implementing a fitter which incorporates the basic path fitter approach but avoids the difficulties of calculating reflections and scattering, a simplified path fitter has been installed in SNOMAN. This fitter is described in Section 4.2.

4.1.2 PDF Normalization

Since a different PDF is used for each electron position $\vec{r_e}$, time t_e and direction $\hat{d}_e$, one needs to worry about the relative normalization of the PDFs[2]. If one ignores multi-photon effects and considers a signal photon path in the sum, the normalization is trivial

$$\sum_{n_i=0}^{\infty} \int_{-\infty}^{\infty} P_{ij}(n_i, t_i; \vec{r_e}, \hat{d}_e, t_e) dt_i = \sum_{n_i=0}^{\infty} \frac{1}{n_i!} \lambda_{ij}^{n_i} e^{-\lambda_{ij}^{n_i}} \int_{-\infty}^{\infty} P_{ij}^{TIM}(t_i) dt_i = 1 \qquad (1)$$

if P_{ij}^{TIM} is itself normalized. The fact that the normalization is independent of $\vec{r_e}$, t_e, and $\hat{d}_e$ greatly simplifies the fitting procedure. If one now considers multiple photon paths, the are additional multiplicative terms that reflect the relative probabilities of photon paths. For example, if there were only direct and reflected light in the detector, the reflection term would be multiplied by the fractional amount of reflected light (related to the mean reflection coefficient) and the direct term would be preceded 1 minus this probability to normalize the overall PDF.

4.1.3 Background vs. Signal PDFs

The explicit approach used to fit events with the Path Fitter is to use as much information about the signals one is trying reconstruct as possible to estimate the most likely position, direction, and time of the event. The reconstruction is performed under the hypothesis that the event was initiated by a single relativistic electron. For CC and ES signals, this is valid hypothesis. A given neutron signal event may involve several Compton-scattered electrons from the capture γ, however, we are fortunate enough to have copious amounts of calibration data (^{16}N) very near the capture γ energy to test how well these reconstruct under the single electron hypothesis. After a particular fit is completed, the consistency of this hypothesis with the observed time and angular distribution of hits is the event is tested. Reconstructed events which are inconsistent with this hypothesis (at some level) are rejected. Aside from rejecting good events (e.g. external $\beta-\gamma$s) that are poorly reconstructed in either position or direction, this approach has the added advantage of rejecting non-Cerenkov backgrounds such as flashers.

[2]Technically, each PDF also needs to be normalized to one for proper interpretation of the likelihood.

It is often stated that a better approach to reconstructing events in SNO is to fit events under the hypothesis that they are initiated by backgrounds. The idea is that a fitter's most important role is to reject external backgrounds that dominate the data sample, and therefore one should use as much information about the backgrounds as possible to accurately locate these events in the detector so that they can be eliminated. One then hopes that the fitter also performs well on the signals of interest, but this is less important since the signal events represent a small fraction of the data.

There are several problems which this approach which make such a fitter inferior, in principle, to ML fitters operating under the signal hypothesis. The signals of interest are simple final states like single e^-'s and γ's generated within the inner region of the detector. These events are far easier to model than background events such as PMT $\beta-\gamma$s or the myriad of instrumental backgrounds that exist in real data taking. As such, one expects that signal PDFs will always be better determined (in MC, at least) than background PDFs. A more serious shortcoming of the background PDF approach is just how to produce a reliable background PDF. One may try to use a PDF which represents a weighted sum of all known backgrounds to the signals we are trying to measure. Even if one sets aside the very serious difficulty of determining the background event topologies (how well are PMT $\beta-\gamma$ tails modeled? how do I get instrumental background PDFs?), determining the proper weighting of all the backgrounds relevant to a particular measurement is a difficult problem. Even if one can determine through external means that a particular (small) set of backgrounds dominate a given signal of interest and then generate a PDF properly weighted over these backgrounds, the flexibility of such fitter is called into question. This is because the background makeup is such a strong function of analysis threshold. For example, if one is attempting to analyze data above 4 MeV, the dominate backgrounds are likely to be light water and PMT $\beta-\gamma$s. However, why would one care about $\beta-\gamma$ backgrounds if he or she is analyzing above 8 MeV? In this case, the background PDF should be heavily weighted with high energy external γ-rays and instrumental backgrounds (how to I determine the latter?). It is therefore argued that the best approach to reconstructing events in SNO to fit under the signal hypothesis and reject events that are inconsistent with this hypothesis based on the reconstructed position/direction, the observed event time and angular distribution, and known properties of Cerenkov light.

4.1.4 Discarding Hits During the Fit

Most fitters use some sort of algorithm to discard PMTs during the fit in an effort to improve vertex reconstruction. The Elastic Fitter uses a particularly elegant method of de-weighting hits that is in close analogy to the method of simulated annealing (described in Section 4.2). The idea is that PMT hits not due to direct Cerenkov light are messy and uncertain, and since they add no useful information to the fit an attempt should be made to ignore these hits. A simple example of a fitter that follows this procedure is the SNOMAN Time Fitter. An implicit assumption in the Time Fitter algorithm is that hits are due to direct light - photons propagating directly from a Cerenkov source to the PMTs. This assumption is actually quite a poor one, since a sizable fraction of hits are not from direct Cerenkov light but rather reflections, scattering, or random PMT noise coincidences. These hits produce significant tails in the χ^2 distribution and also break down the approximation of using PMT jitter as the time error for each hit since this error is now non-Gaussian. The Time Fitter attempts to circumvent this problem by selectively throwing out hits in the events and re-doing the χ^2 minimization.

Any method of selecting direct light hits is non-ideal and can lead to vertex reconstruction tails when non-direct light hit remain in the fit. This may be particularly serious at low energies where it is more difficult to pick out direct light from other types of PMT hits. There is also an unavoidable sacrifice of direct light hits in the process, which degrades the vertex resolution by throwing out information that is valid under the fitter hypothesis. In the Path Fitter, all PMTs which pass the full timing calibration as used in the fit[3]. Maximum likelihood allows one to incorporate any type of information into the PDF. For example, a hit way out of time with the prompt peak (determined by the hypothesized electron position) has a certain probability of being random noise, which is incorporated as a flat probability in time and angle with relative magnitude related to the average PMT noise in the detector and occupancy of the event. Of course, a PDF is only as good as the information that goes into it. One might worry that it is better to simply throw out late hits rather than assign them probabilities because of the difficulty of determining an accurate PDF. This may be true, although in the end it is a trade-off between the dangers of properly selecting non-direct

[3]Eventually, channel property information from ANXX banks should be used to ignore hits from bad channels.

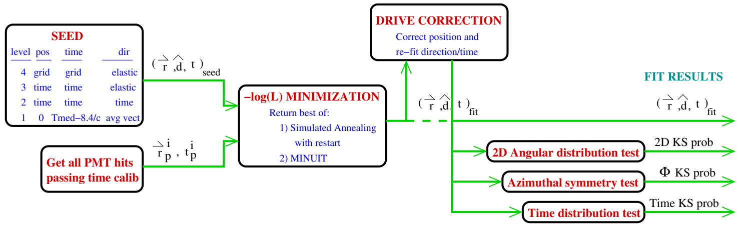

Figure 15: Diagram showing the general flow of the path fitter algorithm.

hits and the dangers of assigning non-direct hits the wrong probability. In the Path Fitter, the choice of using all the hits in the event is made and the late light component is modeled in a very simple way as a step function at the prompt time ($\Theta(t_i - t_{prompt})$).

4.2 The SNOMAN Path Fitter

The path fitter implemented in SNOMAN fits events by maximizing the likelihood function

$$\mathcal{L} = \prod_{i=1}^{NHIT} \mathcal{P}_i(n_i = 1, t_i; \vec{r_e}, \hat{d_e}, t_e)$$

where $\mathcal{P}_i$ is the probability density function for the i^{th} PMT which is described in Section 4.2.1. In practice, the $-log\mathcal{L}$ function is minimized to get the fit position, direction, and time. Note that only hit PMTs are included in the likelihood function. Unhit PMTs add information into the fit (and fit naturally into the maximum likelihood framework) and their exclusion is only because of the extra CPU time required at the occupancies relevant for solar neutrino events.

The general flow of the path fitter algorithm is shown in Figure 15. The initial ingredients of the fit are the set of PMTs used in the fit and a seed electron phase space point for the minimization routines. All hit PMTs which pass the full timing (ECA + PCA) calibrations are used in the fit. The seed point for the fit depends on the user specified seed level and the results of the other fitters (grid, elastic, and time fitters). For example, if the user specifies a seed level of 4 (the default), then the initial position/time are from the grid fitter and the initial direction is the elastic fitter direction, unless either of these fitters fail or the position

fit is outside the PSUP region. In the latter scenarios, the fitter enters the next lowest seed level and continues, unless this lower seed level is below the minimum seed level specified by the user, in which case the fitter fails. Details regarding the other seed levels are shown in Figure 15.

The $-log\mathcal{L}$ function minimization proceeds through several steps. First the function is minimized using the method of simulated annealing (specifically using the Numerical Recipes routine *ameobsa*). This method has proven effective in global minimization problems when local function minima exist. The basic idea is that an n+1 dimensional simplex randomly samples the function at very large values of a parameter T (called the "temperature"), but at very low temperatures the simplex follows the downhill simplex method of Nedler and Mead. At intermediate temperatures, the downhill simplex method is followed except that each vertex of the simplex has a non-zero probability of taking an *uphill* step from function value F_1 to F_2 ($F_2 > F_1$) given by

$$P_{up} \sim e^{\frac{-(F_2 - F_1)}{T}}$$

where F is the value of the function at the given simplex point. The minimization proceeds by lowering the temperature from an initial value (T_o) which is large compared to typical changes in the log-likelihood to a smaller temperature according to the "scheduling function"

$$T_i = T_o(1 - \tfrac{i}{I})^{\alpha}$$

where i is the step index, I is the total number of temperature steps, and $\alpha = 4$ (default). The phase space point with the lowest function value ($-log\mathcal{L}$) is then returned.

The overall minimization strategy is to first get the best point (i.e. electron phase space point with the lowest $-log\mathcal{L}$) from simulated annealing using the initial seed point from the seed levels. The simulated annealing is then performed again using the previous best point as a seed. Finally, MIUNIT (using "MINIMIZE" strategy) is called using the current best point as the seed. The best point from the three minimization stages is returned as the fit electron position, direction, and time.

After the fit is complete, the user has the option of applying a drive correction to the fit point and calculating several goodness-of-fit (GOF) criteria. The drive correction and GOF criteria and described in Section 4.2.2 and Section 4.2.3, respectively.

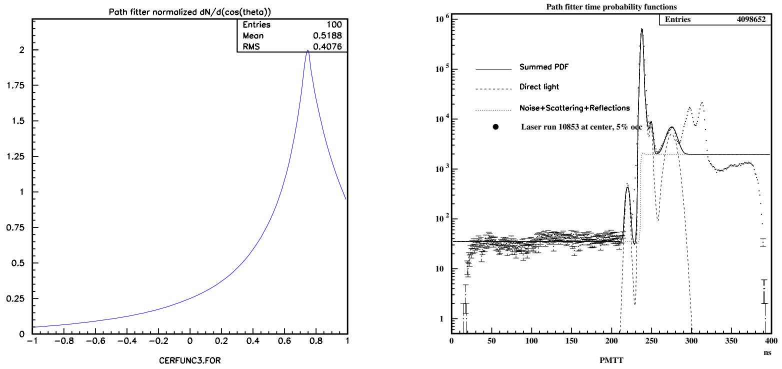

Figure 16: (a) Angular distribution of Cerenkov photons relative to the initial direction of a Monte Carlo 5 MeV electron. (b) Hit PMT time distributions which make up the direct light (dashed) and "other" light (dotted) time probability density functions. The solid dots correspond to hit PMT times from laserball run 10853 at the center of the detector.

4.2.1 The Path Fitter PDF

In the simplified path fitter implemented in SNOMAN, all PMT hits are assumed to be from two separate photon paths - direct light and "other light", which represents all non-direct types such as reflections, scattering, and random PMT noise. Specifically, the probability density function for the ith PMT to be hit once at time t_i given an electron initially at $\vec{r_e}$, t_e in direction $\hat{d}_e$ is given by

$$\mathcal{P}_i(n_i = 1, t_i; \vec{r_e}, \hat{d}_e, t_e) = f_{direct}\mathcal{P}^{TIM}_{direct}(t_i)\mathcal{P}^{ANG}_{direct}(1) + f_{other}\mathcal{P}^{TIM}_{other}(t_i)\mathcal{P}^{ANG}_{other}(1)$$

where $\mathcal{P}^{TIM}_{direct}(\mathcal{P}^{ANG}_{direct})$ and $\mathcal{P}^{TIM}_{other}(\mathcal{P}^{ANG}_{other})$ are the time (angular) probability density functions for direct and other light, respectively, and $f_{direct} = 0.879$ and $f_{other} = 0.121$ are the respective light fractions based on fits to laserball data.

The time probability density functions are derived from fits to laserball data at the center of the detector (see Figure 16(a)). For direct light, the time PDF is a sum of four Gaussian functions corresponding to prompt, pre-pulse, late-pulse, and after-pulse PMT hits, where the time offset is determined by the direct time-of-flight to the PMT. The time PDF for hits

from other sources is modeled as the sum of a flat PMT noise distribution and $\Theta(t_i - t_{prompt})$ for late light from reflections and scattering.

The angular PDF for non-direct light is a normalized uniform distribution. Therefore, no attempt is made to use angular information in scattered or reflected light as was the case in the original path fitter, and the only time information which is used is the simple fact that the majority of hits outside the prompt peak are late. The direct light angular PDF used in the path fitter includes the angular distribution of Cerenkov photons from a multiply scattered electron and the solid angle subtended by the PMT in question. Specifically, the angular PDF is the Poisson probability for a single hit of the i^{th} PMT is given by

$$\mathcal{P}^{ANG}_{direct} = N_\gamma \rho_i e^{-N_\gamma \rho_i}$$

where $N_\gamma = \frac{NHIT}{0.55}$ is used as a rough estimate of the number of emitted Cerenkov photons and

$$\rho_i = \frac{1}{2\pi} g(\cos\alpha_i)\Omega_i$$

where $g(\cos\alpha)$ is the angular distribution of Cerenkov photons relative to the initial direction of a 5 MeV electron according to the SNOMAN implementation of EGS4 code (see Figure 16(b)), α_i is the angle of the i^{th} PMT relative to the hypothesized electron direction from the electron position, and Ω_i is the solid angle of the i^{th} PMT viewed from the electron position given by

$$\Omega_i = \frac{\pi r_c^2}{d_i^2}\hat{d}_i \cdot \hat{r}_i$$

where r_c is the concentrator radius, $\hat{r}_i$ is the unit position vector of the i^{th} PMT, and $\vec{d_i}$ is the vector connecting the electron position and the position of the i^{th} PMT. The angular distribution, $g(\cos\alpha)$, is fit to the SNOMAN results as piece-wise continuous function consisting of two 10^{th}-degree polynomials (one for each side of the Cerenkov angle) and Gaussian centered on the (cosine) of the Cerenkov angle.

4.2.2 Drive Correction

It was determined through Monte Carlo studies that the path fitter, much like other fitters in SNOMAN, contains a systematic "drive" in the vertex fitting which causes the mean fit residual along the initial electron direction to be positive. The drive was discussed in Section 2 in the context of time and angle fitting. While many interesting things were learned about the drive (e.g. seems to be related to the assumption of ϕ symmetry in each event), no analytical fix is currently implemented in the Path Fitter. The approach taken to improve the Path Fitter was to empirically study the effect in Monte Carlo and allow the user to apply a correction to the vertex based on the results of these studies.

It was determined that the magnitude of the drive is related to the distance, M_{gen}, along the generated direction from the generated position to the PSUP, but the correlation with M_{fit} is small inside the D_2O (see Figure 17). Therefore, when the drive correction option is enabled by the user (default), the position returned by the fitter is pulled back in a direction *opposite* the fit direction by an amount determined from empirical studies of the average amount of drive (15.27 cm). After this correction is applied, the direction and time are re-fit at the drive corrected position.

4.2.3 Goodness-of-Fit Criteria

The Path Fitter can be requested by the user to return GOF information which assesses the quality of the fit. The GOF information is in terms of a KS probability for each of three quality tests - two angular tests and one time test - each of which are related to the consistency between the event expectations based on properties of Cerenkov light and the detector response.

Two of the GOF tests examine the angular distribution hit PMTs to determine the fit quality. For these tests, $(\cos\alpha, \psi)$ of each PMT is first determined, where α and ψ are the polar and azimuthal angles of a given hit PMT's position vector in a coordinate system that is centered on the fit position with the z-axis along the fit direction and an x-y orientation chosen at random. One of the angular tests - the "azimuthal symmetry test" - returns the KS probability for the ψ distribution to be flat, using the fact that Cerenkov light is produced azimuthally symmetric about the (instantaneous) electron direction. The other angular test compares the 2-D $\cos\alpha, \psi$ distribution with the hit PMT angular probability density function one expects for Cerenkov light. The angular PDF that is used in the 2-D angular test is based on the angular function used for fitting 2-D KS probability is used to assess fit quality.

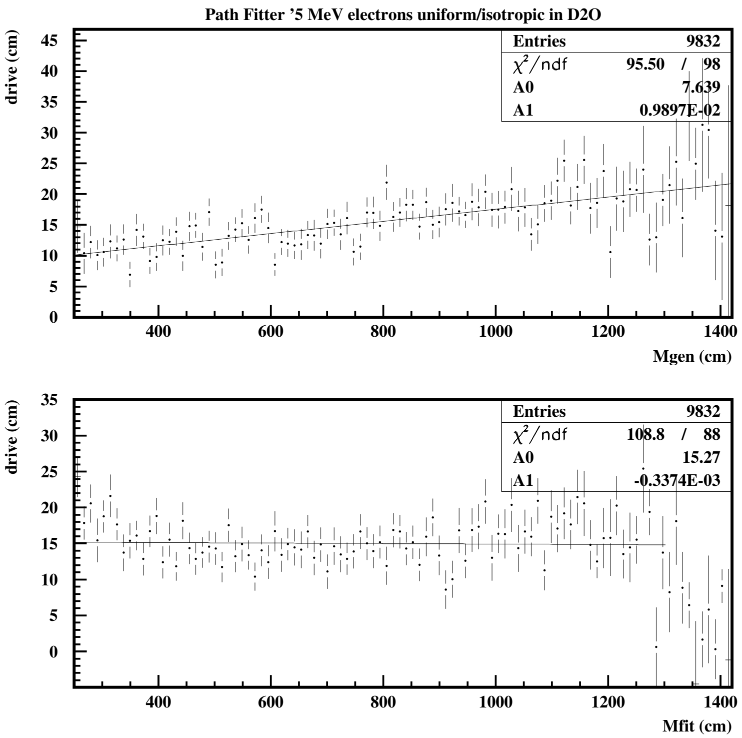

Figure 17: Path Fitter drive (cm) versus (a) M_{gen} and (b) M_{fit}. $M_{gen(fit)}$ is the distance along the generated (fit) direction from the generated (fit) position to the PSUP.

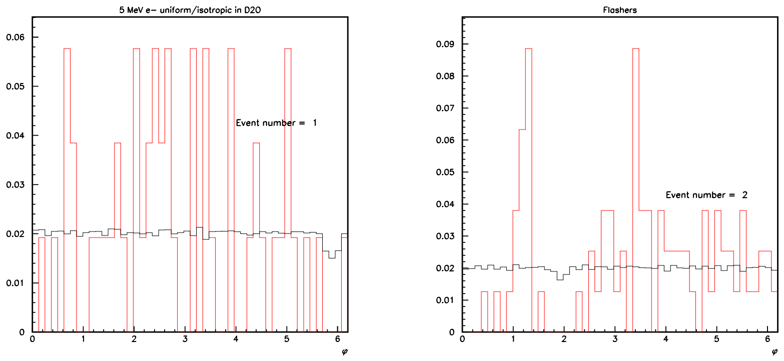

Figure 18: Example of azimuthal symmetry KS test for fit quality applied to (a) a 5 MeV MC electron and (b) a PMT flasher event.

As as example of the angular tests, consider two events which have been fit well inside the D_2O - a Monte Carlo 5 MeV electron fitting near its generated position/direction and a PMT flasher event. Figure 18 shows the ψ distribution for both events and the calculated probability density function taking into account the Cerenkov photon angular distribution, PMT solid angle, etc. Note the dip in each distribution caused by missing PMT in the neck region. As previously mentioned, the PMT ψ distribution is simply compared to a uniform distribution, and the KS probability is calculated. For this particular example, the MC and flasher events have azimuthal symmetry KS probabilities of 0.7 and 0.03, respectively. Note that while the flasher event looks like it has a less uniform ψ distribution and has a low KS probability value, it is probably not enough to reject this event based on the azimuthal symmetry test alone. However, when one considers both the $\cos\alpha$ and ψ together as in the 2D angular test, the two events are much more distinct (see Figure 19). The MC and flasher events have 2D angular KS probabilities of 0.4 and 10^{-5}, respectively.

The angular GOF tests have different strengths and weaknesses. The azimuthal symmetry test hinges on a simple assumption about the symmetry of Cerenkov light and should be relatively independent of electron energy. However, it does not take into account PMT solid angle effects (although this could be put in) which breaks the azimuthal symmetry of the

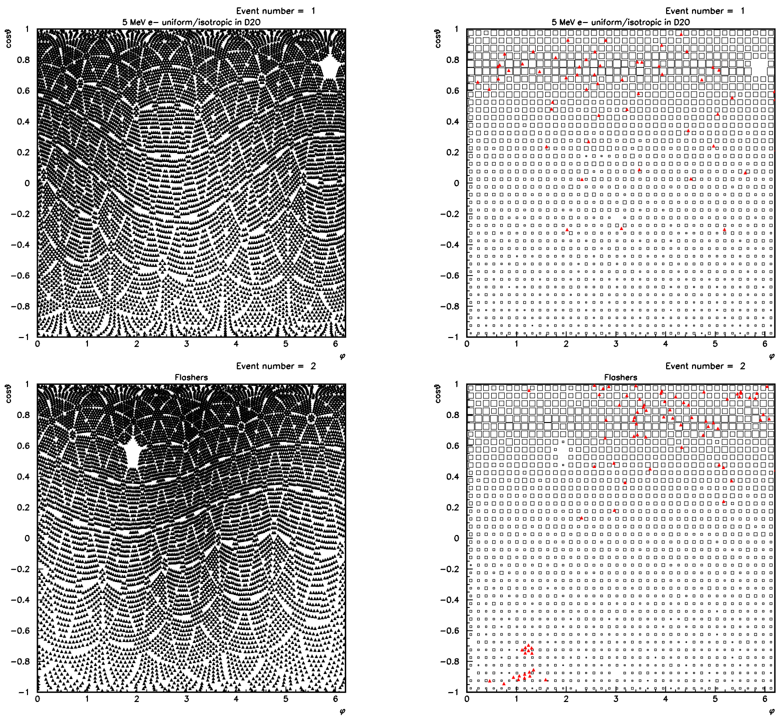

Figure 19: Example of 2D angular distribution KS test for fit quality applied to (top) a 5 MeV MC electron and (bottom) a PMT flasher event.

PMT hits even from a symmetric photon distribution. This only becomes a problem near the PSUP, as the effects of PMT solid angle on the symmetry of the hit PMT distribution are small inside the D_2O. On the other hand, the 2-D angular test includes PMT solid angle and in some sense is a generalization of the azimuthal symmetry test, so one might expect this test to have a smaller radial bias. The downside of the 2-D angular test, aside from the longer time it takes to perform the test due to the extra dimension in the KS test, is that the polar distribution depends on energy due differences in the electron scattering.

An elaborate time KS test is implemented in SNOMAN along with the two angular tests just described. The time test involves calculating the expected PMT time spectrum after a fit and comparing this to the observed hit PMT time in the event. This calculation involves the direct light PDF and the reflected times computed by the specular reflection code discussed in Appendix A.1 (scattering could be included as well). The performance of the time test were disappointing at best. This is not really unexpected - it is one thing to fit under simplifying assumptions about late light, but it is another thing altogether to test the quality of a fit based on these assumptions if they are not a particularly good model. Its possible that when further knowledge of late light is determined in SNO in the future such a test might be made useful, but for now use of this test is not recommended. In practice, one is better off using a more "blunt" test for consistency with Cerenkov light such as the fractional number of hits inside a prompt time window, like the ITR variable[3] developed by Vadim Rusu.

4.2.4 Using the Fitter

The Path Fitter code is included in SNOMAN development version 4.0083 and later releases. The fitter is run with a call to the FTP processor which should be preceded by all other fitters used to seed the Path Fitter. An example PEL used to generate and fit Monte Carlo is shown below:

```
define event_loop
   call MCO
   call FTT
   call FTG
   call FTE
   call FTP
   call ANL
```



```
 quit_event
end_def
```

In terms of user control, all Path Fitter parameters can be adjusted by changing the relevant entry in **fitter.dat**, included below:

```
*DO  TFTP   1   -i(30I 1F 3I 2F 3I 6F 1I 1F 7I)
#.  Bank for bank for path fitter parameters
#.
#.       Standard Database Header
#.
19750101        0 20380517 03331900  #.  1..4   Intrinsic validity
       0        0        0           #.  5..7   Data type, Task type, Format no.
       0        0        0           #.  8..10  Creation Date, Time, Source Id.
19750101        0 20380517 03331900  #. 11..14  Effective validity
       0        0                    #. 15..16  Entry Date Time
4*0                                  #. 17..20  Spare
10*0                                 #. 21..30  Temporary data (not in database)
#.
#.    End of Standard Database Header
#.
#.    User Data.
21.8                         #. C_AVG
150                          #. Iterations within amebsa at a given temp
10                           #. Total amebsa calls (number of temp steps)
4                            #. Parameter used in temperature scheduling
1.E-4                        #. Frational tolerance for simulated annealing
200.                         #. Starting temperature
150                          #. Restart1: Iterations within amebsa at a given temp
10                           #. Restart1: Total amebsa calls (number of temp steps)
4                            #. Restart1: Parameter used in temperature scheduling
1.E-4                        #. Restart1: Frational tolerance for simulated annealing
10.                          #. Restart1: Starting temperature
50.                          #. X,Y, and Z width of starting simplex
0.1                          #. COSTHETA width of starting simplex
0.3                          #. PHI width of starting simplex
30.                          #. PMTT width of starting simplex
100                          #. Maximum number of roots requested from ZBRAK
1.E-2                        #. Tolerance on root found by ZBRENT
```

```
0                     #. PRINT_MODE = 0 for no output
                      #.             1 for debug
4                     #. SEED LEVEL
                      #.    level position   time     direction
                      #.    ----- -------- ---------- ----------
                      #.      4     grid   grid       elastic
                      #.      3     time   time       elastic
                      #.      2     time   time       time
                      #.      1     center tmed-8.4/c avg vector
2                     #. minimum allowed seed level
#x6                   #. Goodness-of-fit mask
                      #.    Mask bit          GOF type
                      #.    --------          --------
                      #.       0     Time distribution KS test
                      #.       1     2-D angular distribution KS test
                      #.       2     Azimuthal symmetry KS test
1                     #. Fitting mode = 1 for full time and angle
                      #.             = 2 for time only
1                     #. Drive correction flag
                      #.    = 0 for no correction
                      #.    = 1 apply correction and refit dir,time
3                     #. Minimization strategy (1-3)
                      #.    = 1 for MINUIT only
                      #.    = 2 for Simulated Annealing + MINUIT
                      #.    = 3 for Simulated Annealing w/restart + MINUIT
```

While the TFTP section of **fitter.dat** contains all Path Fitter parameters that the user can control, some of these parameters are historical artifacts of the fitter development process and not relevant for typical user control. A list of Path Fitter mnemonics for parameters which the typical user may wish to change along with a brief description of each are included in the file **run_path_fitter.cmd** in the SNOMAN **prod** directory.

5 Path Fitter Performance

In this section, performance of the Path Fitter vertex and angular reconstruction on signal MC is presented, which some comparison made to the Grid, Elastic and Time Fitters. Much

recent work on Path Fitter characterization has been done, looking at performance on MC, calibration sources, and various types of instrumental backgrounds. A comprehensive presentation of these studies is an STR in itself and is outside the scope of this STR on the description of the Path Fitter as a reconstruction method. The interested reader should see [4] for more details on Path Fitter performance.

5.1 Vertex Reconstruction

In this section, vertex reconstruction performance of the Path Fitter is compared to the Time and Grid Fitters. The comparisons were done for CC electrons simulated over the TOXENN AB+10% livetime. Other types of MC should be compared in a systematic way, particularly backgrounds, but this represents a first pass at MC signal studies which are relevant for signal extraction.

Figure 20 shows the x-position residual distribution for the three fitters. Residual distributions for y and z are essentially identical to Figure 20. Notice that the Path Fitter has smaller tails than either of the other fitters and is also narrower in width (resolution). These facts are also easily seen in the rdev distribution shown in Figure 21, which is a plot of the fit distance from the true MC position for the different fitters.

Recall from the discussion in Sections 2 and 4.2.2 about fitter drive that most fitters - including the Path Fitter - possess a systematic pull of fit position in the direction of the initial (or average) electron direction. Section 4.2.2, in particular, discussed how the Path Fitter can be instructed to apply an empirical MC-based correction for this drive. Figure 22 shows drive distributions for the three different fitters, where it is evident that while the other fitters possess this systematic pull, the correction has removed this systematic effect from the Path Fitter(according to MC). Also, notice that in the Path Fitter the drive distribution is narrower overall.

Since the drive correction was determined based on MC studies, one may wonder how sensitive the Path Fitter drive is to changes in the optical model in the MC. For example, is the correction fine-tuned to the particular amount of scattering in the MC at the time that the drive was studied? The Path Fitter drive is actually quite robust to large changes in scattering, as Figure 23 demonstrates with the TOXENN extreme optical parameter sets. Optics 1 and 3 are best and worst case in terms of scattering in the detector (see [4] for a

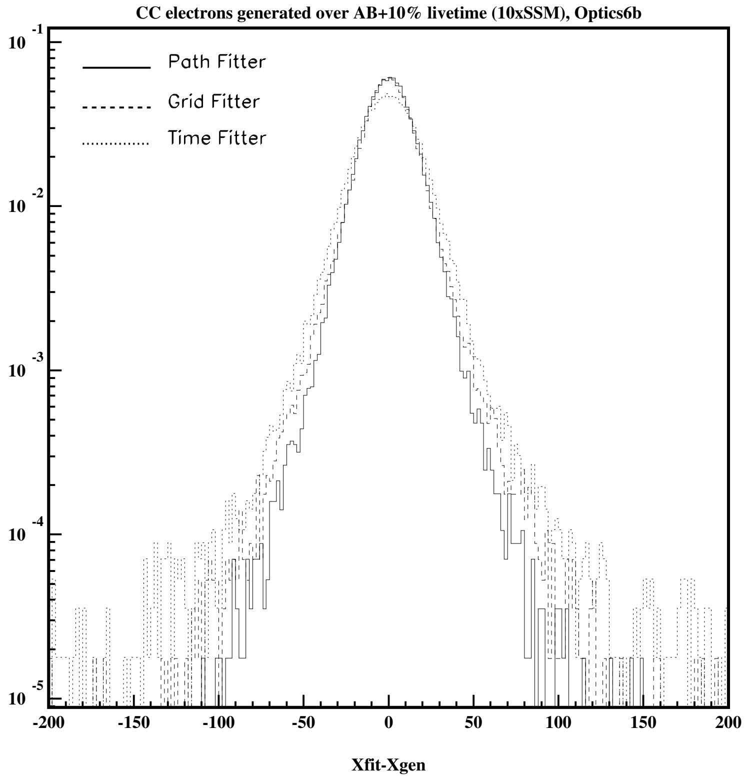

Figure 20: Distribution of x residuals for the three different fitters.

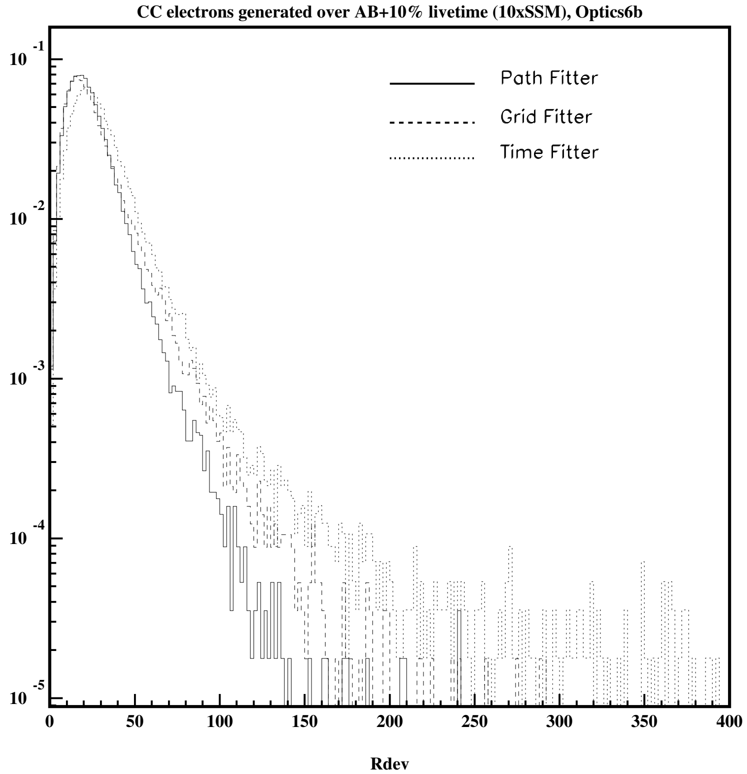

Figure 21: Rdev distribution for the three different fitters. Rdev is the fit distance from the true MC position.

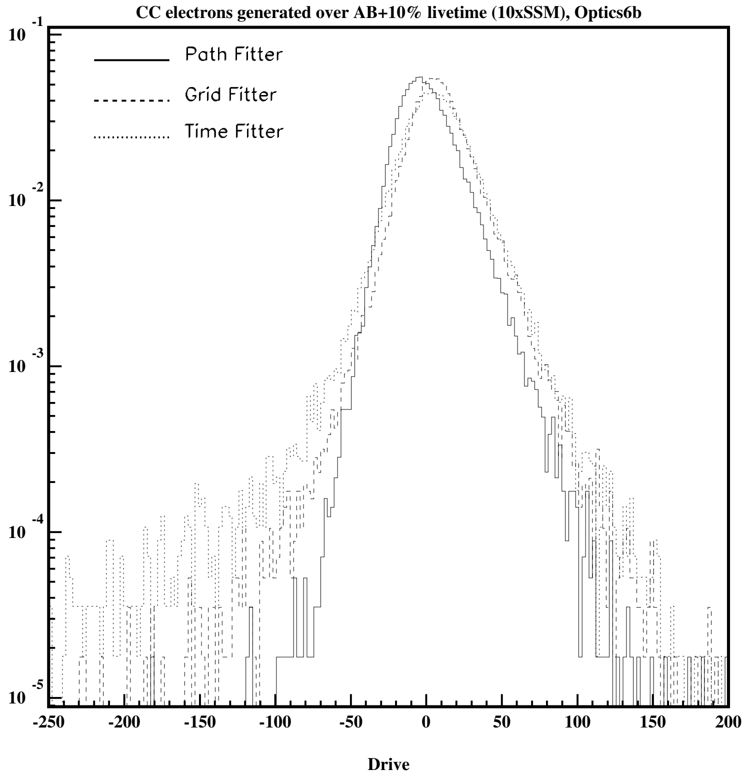

Figure 22: Comparison of drive on MC CC electrons for the three different fitters.

full description).

Table 1 shows are summary of vertex reconstruction performance comparisons for MC CC electrons.

	Time Fitter	Grid Fitter	Path Fitter
X resolution (cm, Gaussian Fit)	17.59±0.07	14.19±0.06	13.87±0.05
X Raw σ (cm)	21.06	17.00	15.17
Y resolution (cm, Gaussian Fit)	17.64±0.07	14.22±0.06	13.87±0.05
Y Raw σ (cm)	21.18	17.12	15.22
Z resolution (cm, Gaussian Fit)	17.50±0.07	14.17±0.06	13.96±0.05
Z Raw σ (cm)	20.81	16.90	15.20
R_{dev} peak position (cm)	21	15	15
R_{dev} mean (cm)	30.22	24.43	22.85
R_{dev} σ (cm)	22.34	16.59	13.11
Drive mean (cm)	6.41	8.50	0.45
Drive sigma (cm)	27.51	22.37	19.85

Table 1: Summary of vertex reconstruction performance comparisons for MC CC electrons.

5.2 Angular Reconstruction

In this section, the angular reconstruction performance of the Path Fitter on MC CC and ES electrons is presented, with comparisons made to the Elastic Fitter. The direction "fit" returned by the Time Fitter is simply the mean PMT vector relative to the fit position. As a result, the angular reconstruction quality is poor and therefore comparisons are not all that useful. The "Improved" Grid Fitter[5] has the capability of doing a separate fit for direction after the vertex is determined, but this feature has not been tested in the analysis contained in this memo. Our understanding is that the angular reconstruction performance is quite similar to the Elastic Fitter, so comparison to the Elastic Fitter should suffice until direct comparisons can be made.

Figure 24 shows the a comparison between the Path and Elastic Fitters for reconstruction

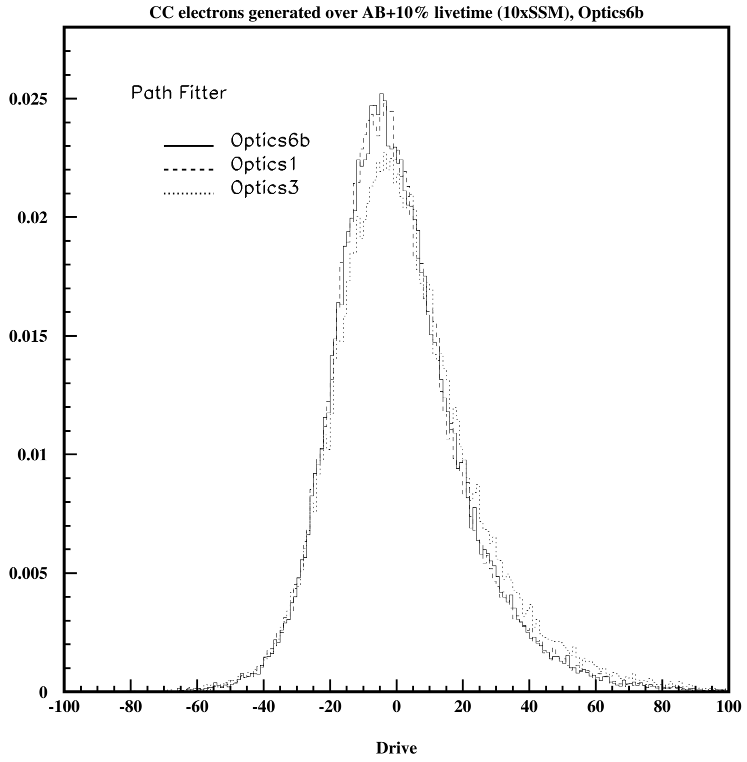

Figure 23: Sensitivity of Path Fitter drive to scattering in the MC, via the TOXENN extreme optical parameter sets.

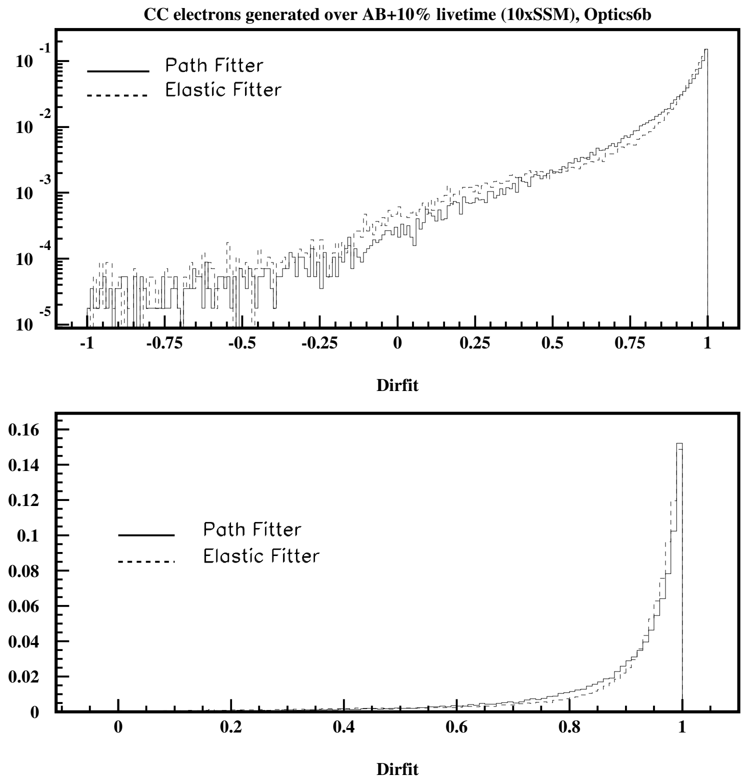

Figure 24: Distribution of the cosine of the angle between the fit direction and initial direction of the MC CC electron for the different fitters.

quality of the initial CC electron direction. Specifically, this shows distributions of the cosine of the angle between the fit direction and initial direction of the MC CC electron for the different fitters, shown on both a linear and log scale so one can see the tails. Note that angular reconstruction is quite similar for the two fitters, both in terms of the forward peak and the tails. However, one might argue from these plots that the Path Fitter is a bit sharper in the extreme forward part of the fit and that it also has a smaller tail of poorly reconstructed directions.

Figure 25 shows a similar plot to Figure 24, except this time for the $\cos\theta_{sun}$ distribution of MC ES electrons. In this case, it looks like the Elastic Fitter puts slightly more events into the ES peak, while the Path Fitter shows a smaller tail on the ES $\cos\theta_{sun}$ distribution.

6 Summary

As described here, the Path Fitter provides a 'first principles' approach to incorporating the angular information into fitting SNO events. The motivation for including the angular information is both to help reduce the drive which is associated with time fitters and to improve the fits overall due to the non-ideal timing of the tubes. Residual TSLH effects, late pulsing, scattering, poor calibrations—all can degrade the quality of a time-only fitter but which have little effect on the angular part of a time and angle fitter. We have found that the Path Fitter's overall performance is excellent and that simultaneously fitting for both position and direction is not only achievable but preferable to reconstructing them with separate fitters.

Perhaps the most critical point, however, is not how well the Path Fitter performs but how much better it can yet be. Upgrading the Path Fitter is as simple as altering a PDF or adding a multiplicative factor for a path-dependent probability. Improvements such as accounting for photon path-dependent speeds of light, the attenuation lengths of the various media, a more detailed late-light PDF or an even more proper angular PDF can all be added easily once they have been calculated. We therefore feel strongly that the Path Fitter should continue to be improved, hopefully allowing an even better rejection of backgrounds, or the possibility of moving further out in fiducial volume and thus increasing the statistics we can use for the ES measurements and spectral measurements.

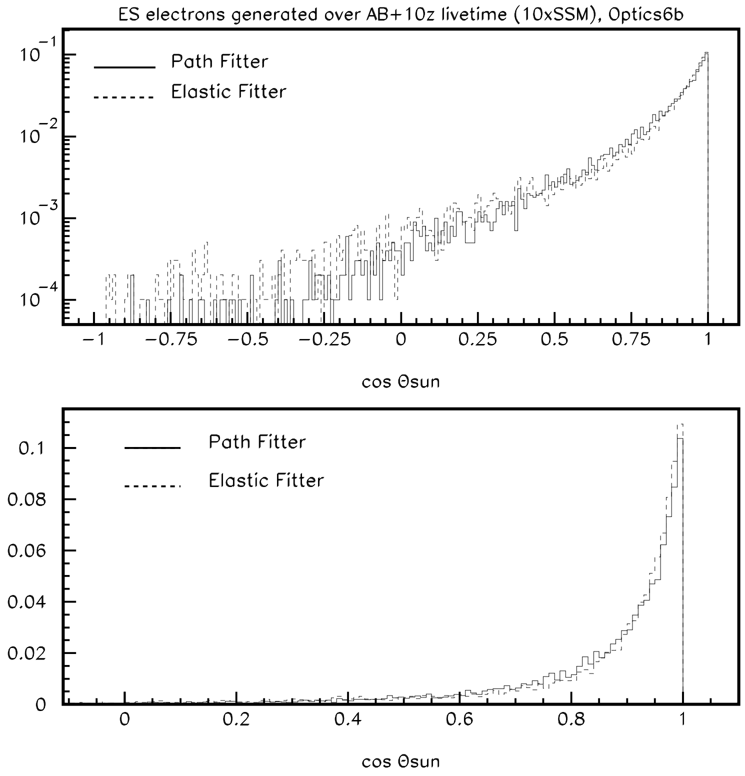

Figure 25: $\cos\theta_{sun}$ distribution of MC ES events for the different fitters.

We hope that with the information in this document the Path Fitter will not only be used by others but that improvements such as the ones suggested here will also receive attention. We therefore urge others interested in reconstruction to look on the Path Fitter not just as an excellent way of determining position and direction, but as a template for testing out and fine tuning even more ideas on reconstruction.

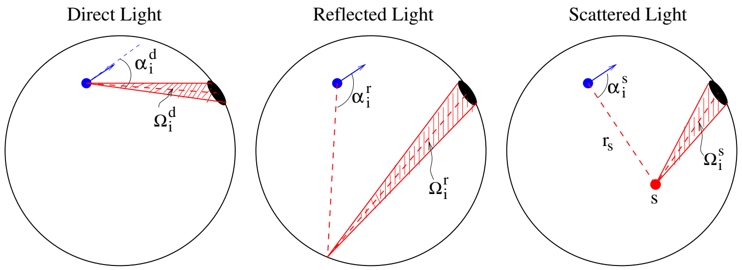

Figure 26: Schematic pictures of the different photon path considered in the Path Fitter PDF calculations.

A Notes on Calculation of Late Light

As described in Section 4.1, the original Path Fitter approach was to construct a PDF for each electron position, direction, and time that is a sum of all possible photon paths from the electron to a given PMT, weighted by the relative probability of each path (See Figure 26). That is

$$\mathcal{P}_i(n_i, t_i; \vec{r_e}, \hat{d}_e, t_e) = \sum_{jpaths} P_{ij}(n_i, t_i; \vec{r_e}, \hat{d}_e, t_e) \tag{2}$$

Grouping the overall PDF into the sum of direct light, scattering, and reflections, we get

$$\begin{aligned}\mathcal{P}_i(n_i, t_i; \vec{r_e}, \hat{d}_e, t_e) = f_{direct} P^{direct}(n_i, t_i; \vec{r_e}, \hat{d}_e, t_e) \\ + f_{refl} \sum_{jpaths} P_j^{refl}(n_i, t_i; \vec{r_e}, \hat{d}_e, t_e) \\ + f_{scat} \int_{s\epsilon V} P_s^{scat}(n_i, t_i; \vec{r_e}, \hat{d}_e, t_e)\end{aligned} \tag{3}$$

The point of this appendix is to document the calculations behind some of the attempts that were made to include, in a simple way, late light in the Path Fitter PDFs. The calculational toy model used was to assume spherical specular reflections off the PSUP and uniform, isotropic scattering in the water volume, with both assumed to be wavelength independent. Reflections and scattering are described in Section A.1 and A.2, respectively.

In Equation 3, the direct light term represents a single path, which if one only considers PMTs hit once, we get

$$P^{direct}(t_i) = P^{TIM}(t_i) N_\gamma \rho_i e^{-N_\gamma \rho_i} \tag{4}$$

where

$$\rho_i = \frac{1}{2\pi} g(\cos\alpha_i^d)\Omega_i^d \tag{5}$$

Recall that $g(\cos\alpha)$ is the angular distribution of Cerenkov photons and ω_i^d is the solid angle subtended by the i^{th} PMT viewed from the hypothesized electron position.

The next two section outline how one could get PDFs for reflected and scattered light in this simplified calculational toy model.

A.1 Spherical Specular Reflections

The dominant reflections in SNO events are reflection off of the PMT glass and concentrators. The actual reflection angles that result in the SNO geometry are complicated, particularly when one considers arbitrary source positions within the detector. However, one might expect there to be a near specular component to the reflections off the PMT glass, and so it is worthwhile considering how one would incorporate a simple specular reflection model into reconstruction.

The form of the PDF for reflections is same as Equation 4, except now there are (as will be demonstrated in Section A.1.3) discrete specular reflection paths, j, that connect the hypothesized source position and i^{th} PMT,

$$P_j^{refl}(t_i) = P_j^{TIM}(t_i) N_\gamma \rho_{ij} e^{-N_\gamma \rho_{ij}} \tag{6}$$

where

$$\rho_{ij} = \frac{1}{2\pi} g(\cos\alpha_{ij}^r)\Omega_{ij}^r \tag{7}$$

In the case of reflections, α_{ij}^r is the angle between the hypothesized electron direction and the vector from the electron position to the point on the PSUP where the j^{th} reflection takes place, and Ω_{ij}^r is the solid angle subtended by the i^{th} PMT when viewed from the reflection point (see Figure 26). Therefore, the entire problem is reduced to solving the general spherical mirror problem for all possible specular reflection path connecting a point inside the sphere and a point on the sphere. This general solution derived in Section A.1.3 and code implementing this solution is installed as part of the Path Fitter in SNOMAN (see the routine **ftp_refl_solve.for**).

A.1.1 Introduction

There are two distinct approaches to spherical reflections presented in this appendix, corresponding to two differently stated problems. In Section A.1.2, the problem of generating spherical reflections given a source point and the direction of the generated photon is discussed. In this case, a solution of the 1^{st} order reflection problem involves finding the distance to the sphere, the reflection point on the sphere, and the direction of the reflected photon. This is more along the lines of a Monte Carlo approach where one can generate the n^{th} order reflection path of the photon through an iterative solution of the 1^{st} order problem. The consideration of this formulation of the reflection problem is largely historical, as it was used in the "bomb" fitter briefly discussed in Section 4.1.1. It is inclusion in this memo is mostly for completeness of the discussion on spherical reflections.

Incorporating reflections into the Path Fitter requires a different formulation of the reflection problem, in terms of a source and sink point rather than a source point and direction. Section A.1.3 deals with the problem of generating all possible 1^{st} order specular reflection paths between an arbitrary source point (e.g a photon) within a sphere and a sink point on the sphere (e.g. a phototube). This involves finding the generated photon direction from the source point, the reflection point on the sphere, and the total path length for each reflection path solution. This approach is potentially useful for a fitter which considers Cerenkov photon reflection paths from a relativistic electron to a hit phototube (e.g. the Path Fitter), however it lacks the ability to iteratively determine multiply reflected photon paths in a simple way like the previous approach.

A.1.2 Reflection Paths Given Source Point and Direction

The problem of generating all possible reflection paths (up to n^{th} order) reduces to iterative calls to a routine which, given a photon generation point G and a directional unit vector $\hat{d}_\gamma$, can find (see Figure 27):

- The distance, m, from the photon generation point to the sphere in the direction of $\hat{d}_\gamma$.
- The point on the sphere P where the reflection will take place.
- The directional unit vector $\hat{d}_r$ of the reflected photon.

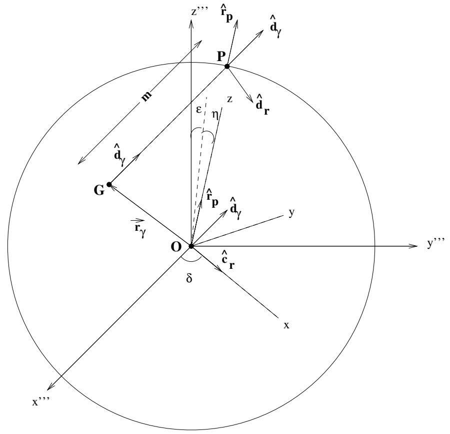

Figure 27: Reflection problem defined by source point G and photon direction $\hat{d}_\gamma$

Let $\vec{r_\gamma} \equiv \vec{OG}$ represent the generated photon position vector and $\vec{r_p} \equiv \vec{OP}$ represent the reflected photon position vector on the sphere so that $\vec{r_p} = R\hat{r}_p$. Then we have the vector equation

$$m\hat{d}_\gamma + \vec{r_\gamma} = \vec{r_p}$$

relating the photon direction to the generated and reflected points. Squaring both sides yields an equation quadratic in m

$$m^2 + 2m\vec{r_\gamma} \cdot \hat{d}_\gamma + r_\gamma^2 - R^2 = 0$$

which has solutions solutions given by

$$m = -\vec{r_\gamma} \cdot \hat{d}_\gamma \pm \sqrt{\left(\vec{r_\gamma} \cdot \hat{d}_\gamma\right)^2 - (r_\gamma^2 - R^2)}$$

Reflection solutions only exist for m nonnegative and real, so we require

$$m > 0 \qquad \text{and} \qquad \left(\vec{r_\gamma} \cdot \hat{d}_\gamma\right)^2 - (r_\gamma^2 - R^2) \geq 0$$

There are three specific cases to consider:

- *G inside sphere ($r_\gamma < R$):* Positive root ensures there is one and only one solution:

$$m = -\vec{r_\gamma} \cdot \hat{d}_\gamma + \sqrt{\left(\vec{r_\gamma} \cdot \hat{d}_\gamma\right)^2 - (r_\gamma^2 - R^2)}$$

- *G on sphere ($r_\gamma = R$):* For $m > 0$, we get one solution:

$$m = -\vec{r_\gamma} \cdot \hat{d}_\gamma + \sqrt{\left(\vec{r_\gamma} \cdot \hat{d}_\gamma\right)^2}$$

- *G outside sphere ($r_\gamma > R$):* For $m > 0$, there are two cases:
  - *Case 1:* $(\vec{r_\gamma} \cdot \hat{d}_\gamma)^2 = r_\gamma^2 - R^2$

    Photon tangent to sphere $\Rightarrow$ one solution:

$$m = -\vec{r_\gamma} \cdot \hat{d}_\gamma$$

  - *Case 2:* $(\vec{r_\gamma} \cdot \hat{d}_\gamma)^2 > r_\gamma^2 - R^2$

    Photon within tangents to sphere $\Rightarrow$ two solutions:

$$m = \begin{cases} -\vec{r_\gamma} \cdot \hat{d}_\gamma - \sqrt{\left(\vec{r_\gamma} \cdot \hat{d}_\gamma\right)^2 - (r_\gamma^2 - R^2)} & \text{(near point)} \\ -\vec{r_\gamma} \cdot \hat{d}_\gamma + \sqrt{\left(\vec{r_\gamma} \cdot \hat{d}_\gamma\right)^2 - (r_\gamma^2 - R^2)} & \text{(far point)} \end{cases}$$

Once m is known, the vector to the reflection point on the sphere is simply given by

$$\vec{r_p} = m\hat{d}_\gamma + \vec{r_\gamma}$$

Now that the reflection point has been determined, the reflected directional unit vector $\hat{d}_r$ can be computed assuming specular reflection. $\hat{d}_r$ is related to $\hat{r}_p$ and $\hat{d}_\gamma$ by the following three relations:

$$\begin{array}{cc} \hat{r}_p \cdot \hat{d}_\gamma = -\hat{r}_p \cdot \hat{d}_r & (\theta_i = \theta_r) \\ (\hat{r}_p \times \hat{d}_\gamma) \cdot \hat{d}_r = 0 & \text{(coplanar)} \\ \hat{d}_r \cdot \hat{d}_r = 1 & \text{(unit vector)} \end{array}$$

Direct algebraic solution of this system of equations is intractable in the original (SNO) coordinate system for arbitrary $\hat{r}_p$ and $\hat{d}_\gamma$. The strategy is then to rotate into a system in which the previous set of equations is easily solved for $\hat{d}_r$ and then rotate back into the original system. The rotated system is chosen such that the z-axis lines up with $\hat{r}_p$ and $\hat{d}_\gamma$ lies in the xz-plane. So in the final rotated system (unprimed), the reflected directional unit vector is $\hat{d}_r = d_{r_x}\hat{x} + d_{r_y}\hat{y} + d_{r_z}\hat{z}$ and

$$\begin{aligned} \hat{r}_p \cdot \hat{d}_\gamma = -\hat{r}_p \cdot \hat{d}_r &\Rightarrow d_{r_z} = -d_{\gamma_z} \\ (\hat{r}_p \times \hat{d}_\gamma) \cdot \hat{d}_r = 0 &\Rightarrow d_{r_y} = 0 \\ \hat{d}_r \cdot \hat{d}_r = 1 &\Rightarrow d_{r_x} = \pm\sqrt{1 - d_{\gamma_z}^2} \end{aligned}$$

Therefore, in the rotated system,

$$\hat{d}_r = \pm\sqrt{1 - d_{\gamma_z}^2}\,\hat{x} - d_{\gamma_z}\hat{z}$$

where + is for $d_{\gamma_x} > 0$ and - is for $d_{\gamma_x} \leq 0$.

So the problem is now reduced to computing the rotation matrix $\mathcal{R}$ (and $\mathcal{R}^{-1}$ through matrix inversion or inverse rotation) between the original (triple-primed) and final rotated (unprimed) coordinate systems. This is accomplished by a (ZXY) Euler rotation about three Euler angles δ, ϵ, η. In words, this is a rotation around the z'''-axis by angle δ, then rotation around the x''-axis by angle ϵ, and finally a rotation around the y'-axis by angle η. Suitably chosen δ, ϵ, and η then aligns the z-axis with $\hat{r}_p$ and $\hat{d}_\gamma$ will lie in the xz-plane, allowing use of the previously derived equation for $\hat{d}_r$.

The angle δ is related to the x'''-component of a unit vector $\hat{c}_r$ lying in the $x'''y'''$-plane which is coplanar to the vectors $\hat{r}_p$ and $\hat{d}_\gamma$. Specifically, $\delta = \cos^{-1}(c_{r_{x'''}})$, and so we now need to calculate $\hat{c}_r$ in the triple-primed system. Leaving the tedious algebraic details out, the solution for $\hat{c}_r = c_{r_{x'''}}\hat{x} + c_{r_{y'''}}\hat{y}$ of the equations

$$(\hat{r}_p \times \hat{d}_\gamma) \cdot \hat{c}_r = 0 \qquad \hat{c}_r \cdot \hat{c}_r = 1$$

yields

$$c_{r_{x'''}} = \frac{r_{p_{x'''}} d_{\gamma_{z'''}} - r_{p_{z'''}} d_{\gamma_{x'''}}}{n_c} \qquad \text{and} \qquad c_{r_{y'''}} = \frac{r_{p_{y'''}} d_{\gamma_{z'''}} - r_{p_{z'''}} d_{\gamma_{y'''}}}{n_c}$$

where

$$n_c = \sqrt{(r_{p_{x'''}} d_{\gamma_{z'''}} - r_{p_{z'''}} d_{\gamma_{x'''}})^2 + (r_{p_{y'''}} d_{\gamma_{z'''}} - r_{p_{z'''}} d_{\gamma_{y'''}})^2}$$

So then,

$$\delta = \cos^{-1}\left(\frac{r_{p_{x'''}} d_{\gamma_{z'''}} - r_{p_{z'''}} d_{\gamma_{x'''}}}{\sqrt{(r_{p_{x'''}} d_{\gamma_{z'''}} - r_{p_{z'''}} d_{\gamma_{x'''}})^2 + (r_{p_{y'''}} d_{\gamma_{z'''}} - r_{p_{z'''}} d_{\gamma_{y'''}})^2}}\right)$$

The angle ϵ is related to the z''-component of the unit vector $\hat{r}_p$ projected in the $y''z''$-plane. Therefore, we need to do the previous z''' rotation by angle δ and transform $\hat{r}_p$ into the double-primed system (the system after the δ rotation). Then,

$$\epsilon = \cos^{-1}\left(\frac{r_{p_{z''}}}{\sqrt{r_{p_{y''}}^2 + r_{p_{z''}}^2}}\right)$$

Now when we perform a rotation by angle ϵ around the x''-axis, the $\hat{z}'$, $\hat{x}'$, $\hat{r}_p$, and $\hat{d}_\gamma$ unit vectors are all coplanar. This leaves one more rotation around the y'-axis to generate the desired final rotated (unprimed) system (and the rotation matrix $\mathcal{R}$) by angle η given by

$$\eta = \cos^{-1}(r_{p_{z'}})$$

Now that we have $\mathcal{R}$, we can translate any vector $\vec{V}'''$ in the triple-primed (original) system to the unprimed (rotated) system by

$$\vec{V} = \mathcal{R}\vec{V}'''$$

In order to transform back, we need to compute $\mathcal{R}^{-1}$ through either matrix inversion or by doing the inverse rotation (i.e. (YXZ) Euler rotation with Euler angles $-\eta$, $-\epsilon$, $-\delta$) and then

$$\vec{V}''' = \mathcal{R}^{-1}\vec{V}$$

Therefore, with $\mathcal{R}$, $\mathcal{R}^{-1}$, and the previous relations, the reflected direction of the photon in the SNO frame is determined.

As mentioned earlier, one of the nice things about setting up and solving the spherical reflection problem in terms of a generation point and an initial photon direction is that multiple reflections are handled simply by iteration of the above algorithm. As an example, Figure 28 shows a few 5^{th}-order reflection paths off a two sphere system generated using a fortran routine which implements the algorithm in this section.

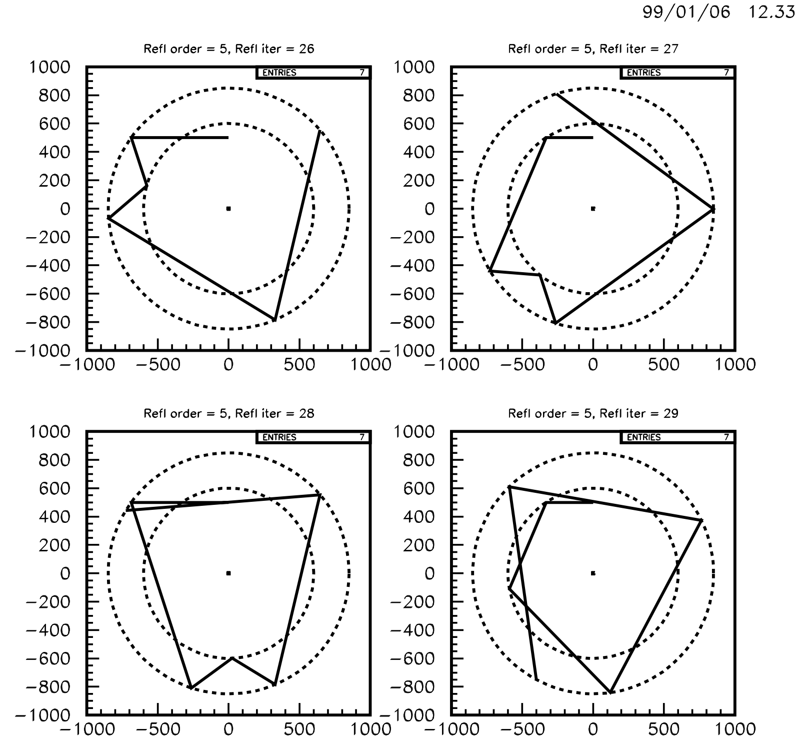

Figure 28: Reflection paths off concentric spheres of radius 850 and 600cm for an initial photon at (0,500) in the (-1,0) direction

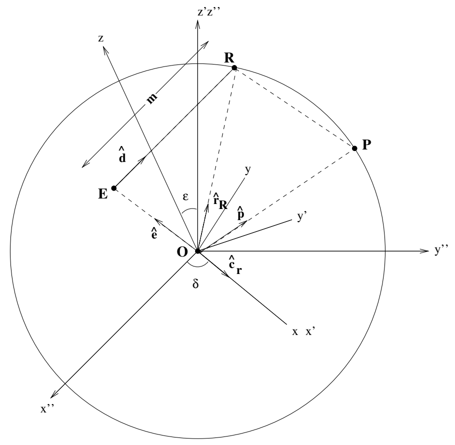

Figure 29: Reflection problem defined by source point E and sink point P

A.1.3 Reflection Paths Given Source and Sink Points

In this case, we are given a photon source point E within a spherical region of radius R centered at point O and a photon detection point P on the surface of the sphere (see Figure 29). For all possible specular reflection paths from E to P, we wish to find:

- The point on the sphere R where the reflection will take place.
- The directional unit vector $\hat{d}$ of the generated photon from the source point E.
- The total path length from the photon generation point E to the sink point P (i.e. $\vec{ER} + \vec{RP}$) in the direction of $\hat{d}$.

It should be noted that the labeling of source and sink points is merely for conceptual purposes - we are just trying to find a point R on the sphere connecting points E and P which is consistent will the constraints of specular reflection.

The intersection of a plane defined by the points E, P, and O and a sphere of radius R centered at O is a great circle of radius R. It on this great circle in which the reflection

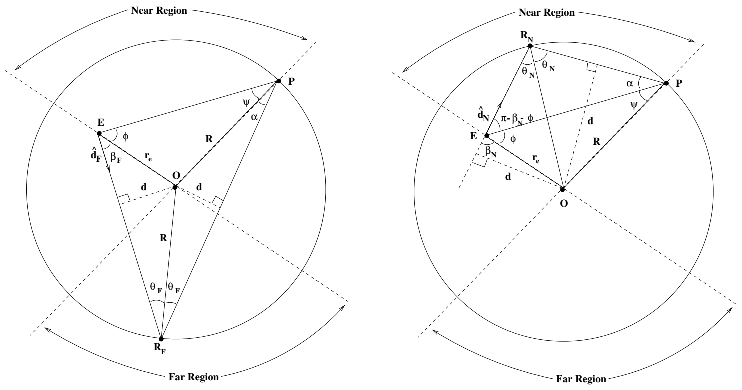

Figure 30: Geometry involved in near and far reflection solutions.

problem is solved in the analysis that follows. It may or may not be intuitive that for non-antipodal points E and P (i.e. E, P, and O are *not* colinear) all solutions *must* lie on this great circle, so the degree to which other specular reflection solutions exist will be later investigated. For E and P antipodal, there is clear degeneracy in the choice of a plane though E, P, and O cutting the sphere and these intersections define a set of great circles on which the reflection problem needs to be solved. However, this solution set is equivalent to the ring defined by the intersection of the plane passing through the two solution points, obtained on any one great circle in the set, that is perpendicular to line $\vec{OE}$ (or equivalently $\vec{OP}$ or $\vec{PE}$).

The geometry involved in solving the problem on the great circle is shown in figure 30 for reflections in the far and near regions of the circle. It should be obvious that any specular reflection point R must lie in one of these regions, otherwise a vector from the origin cannot bisect the reflection angle. After lines of (necessarily equal) length d are drawn perpendicular to lines $\vec{ER}$ and $\vec{PR}$, elementary trigonometry yields the following relations for the far and near reflection solutions:

$$\frac{d}{R} = sin\theta_F \qquad \frac{d}{r_e} = sin\beta_F \qquad \frac{d}{R} = sin\alpha \qquad 2\theta_F + \alpha + \beta + \phi + \psi = \pi \qquad \text{(far region)}$$

$$\frac{d}{R} = sin\theta_N \qquad \frac{d}{r_e} = sin\beta_F \qquad \frac{d}{R} = sin(\alpha + \psi) \qquad 2\theta_N + \alpha - \beta - \phi = 0 \qquad \text{(near region)}$$

Combining the four equations for the far reflection as well as for the near reflection and incorporating the dot product relation

$$\phi + \psi = \pi - cos^{-1}(\hat{e} \cdot \hat{p})$$

we get a transcendental equation in β_F and β_N for the near and far regions, respectively:

$$\frac{R}{r_e} sin\{\tfrac{1}{3}[cos^{-1}(\hat{e} \cdot \hat{p}) - \beta_F]\} - sin\beta_F = 0 \qquad \text{(far region)}$$

$$\frac{R}{r_e} sin\{\tfrac{1}{3}[\pi - cos^{-1}(\hat{e} \cdot \hat{p}) + \beta_F]\} - sin\beta_F = 0 \qquad \text{(near region)}$$

These transcendental equations can be solved for β numerically using root finding algorithms such as bisection or Brent's Method.

We now want to generate the rotation matrix $\mathcal{R}$ relating the original (double-primed) coordinate system to a system (unprimed) where $\hat{e}$ and $\hat{p}$ are in the xz-plane so that we can specify the reflection solution points. This is accomplished by a (ZX) Euler rotation about two Euler angles δ and ϵ. With appropriate replacement of variables, the angles δ and ϵ are the same as in Section 27, namely,

$$\delta = \cos^{-1}\left(\frac{p_{x''}e_{z''} - p_{z''}e_{x''}}{\sqrt{(p_{x''}e_{z''} - p_{z''}e_{x''})^2 + (p_{y''}e_{z''} - p_{z''}e_{y''})^2}}\right) \qquad \epsilon = \cos^{-1}\left(\frac{e_{z'}}{\sqrt{e_{y'}^2 + e_{z'}^2}}\right)$$

This set of rotations yield the rotation matrix $\mathcal{R}$, and an (XZ) Euler rotation about angles $-\delta$ and $-\epsilon$ generates the inverse rotation matrix $\mathcal{R}^{-1}$ so that we can transform any vector from one system to another by

$$\vec{V} = \mathcal{R}\vec{\mathcal{V}}'' \qquad \text{and} \qquad \vec{V}'' = \mathcal{R}^{-1}\vec{V}$$

Now that we have the set of solutions for near and far reflection angles β_i and the rotation matrices $\mathcal{R}$ and $\mathcal{R}^{-1}$, its just a matter of calculating the vectors $\hat{d}$ and $\vec{R} \equiv \vec{OR}$ in the unprimed system and rotating them into the original (double-primed) system. The unit vector $\hat{d}$ specifying the photon direction relative to the source point vector $\vec{OE}$ is easily calculated in the unprimed frame again using rotation of axes. Figure 31 shows how the solution to this problem is equivalent to rotating a unit vector $\hat{x}''$ by angles $-\beta$ and $-\xi$. Note that $sin\xi = e_z$ and $cos\xi = \pm e_x$, where $+$ is for $e_x > 0$ and $-$ is for $e_x < 0$. Using the rotation matrix $\mathcal{R}'$ (not to be confused with the previously defined matrix $\mathcal{R}$!) for a rotation by angle θ in the plane,

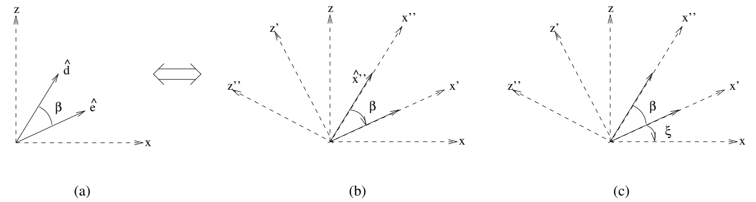

Figure 31: (a) $\hat{d}$ referenced by angle β off arbitrary vector $\hat{e}$ (b) rotation of $\hat{x}''$ by angle $-\beta$ (c) then by angle $-\xi$.

$$\mathcal{R}'(\theta) = \begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix}$$

then for $e_x > 0$,

$$\hat{d} = \mathcal{R}'(-\xi)\mathcal{R}'(-\beta)\ \hat{x}''$$

$$\begin{pmatrix} d_x \\ d_z \end{pmatrix} = \begin{pmatrix} e_x & -e_z \\ e_z & e_x \end{pmatrix} \begin{pmatrix} cos(-\beta) & sin(-\beta) \\ -sin(-\beta) & cos(-\beta) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e_x cos\beta - e_z sin\beta \\ e_x sin\beta + e_z cos\beta \end{pmatrix}$$

and for $e_x < 0$,

$$\hat{d} = \mathcal{R}'(-\xi)\mathcal{R}'(-\beta)\ (-\hat{x}'')$$

$$\begin{pmatrix} d_x \\ d_z \end{pmatrix} = \begin{pmatrix} -e_x & -e_z \\ e_z & -e_x \end{pmatrix} \begin{pmatrix} cos(-\beta) & sin(-\beta) \\ -sin(-\beta) & cos(-\beta) \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} e_x cos\beta + e_z sin\beta \\ e_x sin\beta - e_z cos\beta \end{pmatrix}$$

It should be clear from Figure 30 that while these relations for $\hat{d}$ are correct in the near region, for the far region the directional unit vector gets an extra minus sign. So in terms of β and $\hat{e}$, the photon directional unit vector in the far region is

$$\hat{d}_F = \begin{cases} (e_z sin\beta_F - e_x cos\beta_F)\hat{x} - (e_x sin\beta_F + e_z cos\beta_F)\hat{z}, & \text{for } e_x > 0 \\ -(e_x cos\beta_F + e_z sin\beta_F)\hat{x} + (e_z cos\beta_F - e_x sin\beta_F)\hat{z}, & \text{for } e_x < 0 \end{cases}$$

and in the near region

$$\hat{d}_N = \begin{cases} (e_x cos\beta_N - e_z sin\beta_N)\hat{x} - (e_x sin\beta_N + e_z cos\beta_N)\hat{z}, & \text{for } e_x > 0 \\ (e_x cos\beta_N + e_z sin\beta_N)\hat{x} + (e_x sin\beta_N - e_z cos\beta_N)\hat{z}, & \text{for } e_x < 0 \end{cases}$$

Now that we have $\hat{d}$, we can calculate the vector $\vec{R}$ specifying the point in the sphere in which the reflection takes place and the total path length $\vec{ER} + \vec{OR}$. For both the near and far solutions, we have the vector relation

$$m\hat{d} + r_e\hat{e} = R\hat{r}_R$$

Squaring both sides of this equation we get

$$m^2 + {r_e}^2 + 2mr_e(\hat{e}\cdot\hat{d}) = R^2$$

Substituting $\hat{e}\cdot\hat{d} = -cos\beta$ for the far reflection and $\hat{e}\cdot\hat{d} = cos\beta$ for the near solution, we get

$$\vec{R} = \begin{cases} m\hat{d}_F + r_e\hat{e}, & \text{in far region} \\ m\hat{d}_N + r_e\hat{e}, & \text{in near region} \end{cases}$$

where m is the length of $\vec{ER}$ given by

$$m = \begin{cases} r_e cos\beta_F + \sqrt{R^2 - {r_e}^2 sin^2\beta_F}, & \text{in far region} \\ -r_e cos\beta_N + \sqrt{R^2 - {r_e}^2 sin^2\beta_N}, & \text{in near region} \end{cases}$$

Therefore, the total path length $m_{tot} \equiv \mid \vec{ER} + \vec{RP} \mid = m + \mid R\hat{p} - \vec{R} \mid$ or

$$m_{tot} = \begin{cases} r_e cos\beta_F + \sqrt{R^2 - {r_e}^2 sin^2\beta_F} + \sqrt{2R(R - \hat{p}\cdot\vec{R}_F)}, & \text{in far region} \\ -r_e cos\beta_N + \sqrt{R^2 - {r_e}^2 sin^2\beta_N} + \sqrt{2R(R - \hat{p}\cdot\vec{R}_N)}, & \text{in near region} \end{cases}$$

where $\vec{R}_F$ and $\vec{R}_N$ are the vectors to the reflection points in the far and near region respectively which were previously determined.

Using the inverse rotation matrix $\mathcal{R}^{-1}$, we can rotate the derived vectors $\hat{d}$ and $\vec{R}$ into the original (double-primed) coordinate system. Therefore, the direction $\hat{d}$ of the photon from source point E, the point R on the sphere where the reflection will take place, and the total photon path length m_{tot} for all possible specular reflection paths between E and P have been determined.

Earlier in the section it was stated that all specular reflection points for non-antipodal points lie on the unique great circle defined by these points. The computer routine which implements the algorithm described in section A.1.2 can be used to test this conjecture. In fact, this routine can be used to find all 1^{st} order reflection paths between two points by generating photons in all directions from the source point and recording the direction of any

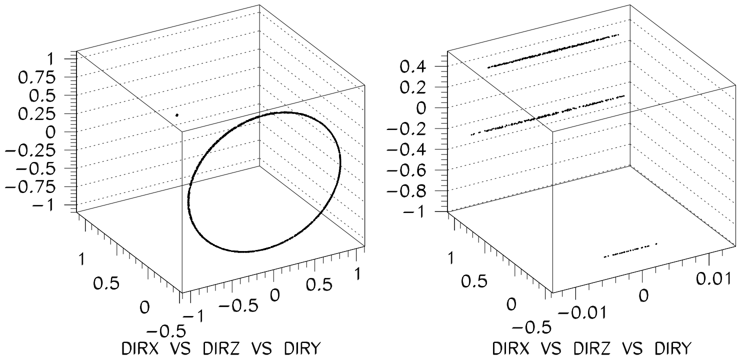

Figure 32: Photon direction cosines for (a) antipodal and (b) non-antipodal source and sink points.

photons ending up within some small tolerance of the sink point. The mean of the distribution of those photon directions should be a good estimator of the true photon direction solution between the source and sink points. Figure 32 shows the results of such an approach for one set of (a) antipodal and (b) non-antipodal source and sink points, which backs up the previous conjecture about the non-antipodal pair solution set. Figure 32(a) corresponds to the source point at $(0, 0, \frac{2}{3}R)$ and the sink point at $(0, 0, -1)$, in which the ring of direction solutions is evident. Figure 32(b) corresponds to the source point at $(0.905R, 0, 0.095R)$ and sink point at $(0, 0, 1)$, in which all three solutions lie in the xz-plane. What do we get from the computer routine implementing the algorithm described in this section for the non-antipodal pair example above? Figure 33 shows the transcendental equations for this source/sink pair in which three solutions - one far reflection and two near reflections - are clearly evident. When these β solutions are properly transformed to give the allowed photon directions, they correspond precisely to the means of the dir_x, dir_y, and dir_z distributions (see Figure 32(b)) generated by the computer routine from section A.1.2. The results from analysis of the non-antipodal pair example are summarized in Table 2:

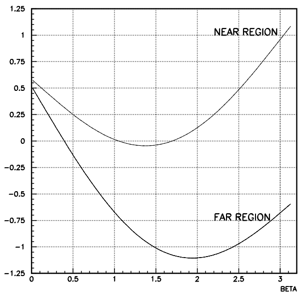

Figure 33: Plots of the transcendental equations in the near and far regions for non-antipodal source and sink points.

A.2 Uniform Scattering

In this section, the problem of including a simple scattering model into the Path Fitter PDF is discussed. In this case, we outline how one would include uniform, isotropic scattering, with photons scattered exactly once on their way from the hypothesized electron position to a given PMT.

The form of the PDF for scattering is same as Equation 4, except now there are a continuum of paths - one for each scatterer at S, that connect the hypothesized source position and i^{th} PMT,

$$P_s^{scat}(t_i) = P_s^{TIM}(t_i) N_\gamma \rho_{is} e^{-N_\gamma \rho_{is}} \tag{8}$$

where ρ_{is} is the probability for a Cerenkov photon generated by the electron and scattered at S to be detected by the i^{th} PMT.

The scattered light PDF is constructed by considering the probability that a PMT at $\vec{r_i}$ is hit at time t_i by an photon scattered isotropically within the volume constrained by the PSUP. Incorporating scattering into the PDF is complicated in general, so the following simplifying assumptions are made:

- A photon generated at E scatters exactly once along its path to P

Point to point algorithm				Point and dir algorithm		
$\beta(rad)$	dir_x	dir_y	dir_z	dir_x	dir_y	dir_z
0.3944	-0.8781	0.0000	-0.4785	-0.8780±0.0001	-0.0006±0.0004	-0.4787±0.0002
1.047	0.4073	0.0000	0.9133	0.4074±0.0002	0.0005±0.0003	0.9132±0.0001
1.700	-0.2320	0.0000	0.9727	-0.2320±0.0002	0.0004±0.0004	0.9727±0.0001

Table 2: Results for the non-antipodal pair: source pt at $(0.905R, 0, 0.095R)$ and sink pt at = $(0, 0, 1)$ implementing the algorithms described in section A.1.2 and section A.1.3. A tolerance of 10 cm was used in the "point and dir" method of section A.1.2 and 1σ errors are shown.

- The angular distribution of scattered photons at any scatter point S is uniform (i.e. scattering is isotropic)
- The scattering volume within the PSUP is uniform with a scattering length λ_s

The scattered light PDF is then constructed by calculating the probability P_s for a photon to hit a PMT at P with time t_i after scattering at an arbitrary point S within the PSUP volume V, and then integrating over all possible S in V (see Figure A.2),

$$P^{scat}(t_i) = \int_{S\epsilon V} P_s^{scat}(t_i)dV$$

The angular probability, ρ_{is}, in Equation 8 can be thought of as the product of several probabilities,

$$\rho_{is} = P_1 P_2 P_3 P_4$$

where the probabilities P_i are shown below

- P_1 = probability that a relativistic electron moving in a direction $\hat{d}_e$ will generate a Cerenkov photon in a direction $\hat{d}_\gamma(\theta, \phi)$
  $= g[\alpha_s(\theta, \phi)]$, where $\alpha_s(\theta, \phi) = cos^{-1}[\hat{d}_e \cdot \hat{d}_\gamma(\theta, \phi)]$
- P_2 = probability of scattering after a distance r, which is the interaction probability times the photon survival probability over a distance r
  $= \frac{1}{\lambda_s} e^{-\frac{r}{\lambda_s}}$

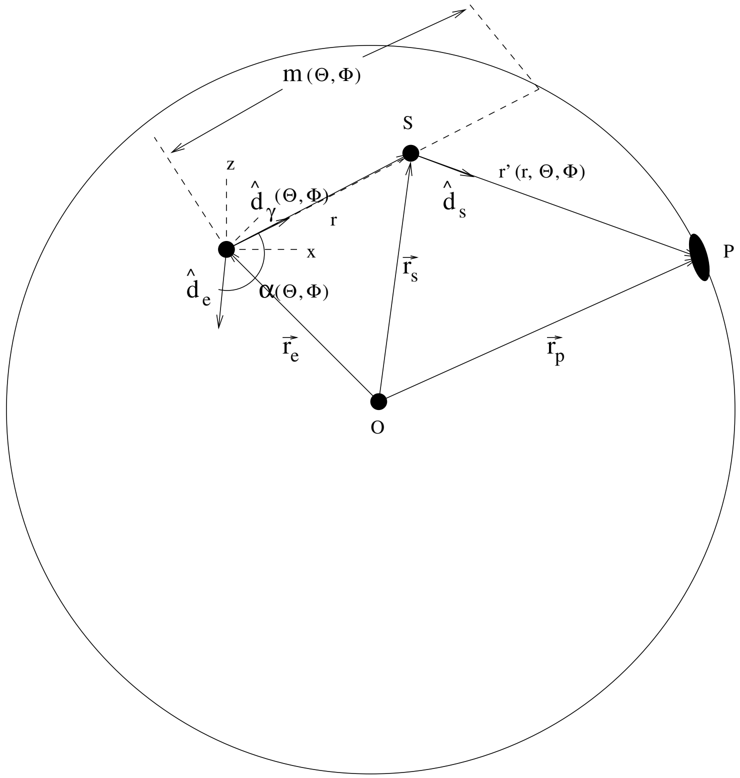

Figure 34: Figure showing the vectors invloved in the scattering problem.

- P_3 = probability of scattering at S into the solid angle $\Omega(r,\theta,\phi)$ subtended by a PMT at P times the probability of scattering in a direction $\hat{SP} \equiv \hat{r}'(r,\theta,\phi)$
  $= \Omega(r,\theta,\phi)\frac{1}{4\pi} = [\hat{r}'(r,\theta,\phi)\cdot\hat{r}_p)]\frac{\pi r_c{}^2}{[r'(r,\theta,\phi)]^2}\frac{1}{4\pi}$,
  where $\vec{r}'(r,\theta,\phi) = \vec{r}_p - \vec{r}_s$ is the distance from the scatterer at S to the PMT at P.

- P_4 = probability of traveling a distance $r'(r,\theta,\phi)$ without scattering, which under our assumption of exactly one scatter,
  $= 1$ (otherwise it would be $e^{-\frac{r'(r,\theta,\phi)}{\lambda_s}}$)

The total scattering probability is then calculated by integrating over all $S\epsilon V$, and is given by

$$P^{scat}(t_i) = \int_0^{2\pi}\int_0^{\pi}\int_0^{m(\theta,\phi)} P^{TIM}(t_i,r,\theta,\phi)N_\gamma\rho_i(r,\theta,\phi)e^{-N_\gamma\rho_i(r,\theta,\phi)}\ sin\theta\ r^2\ drd\theta d\phi \tag{9}$$

where

$$\rho_i(r,\theta,\phi) = \frac{1}{4\pi\lambda_s}g[\alpha_s(\theta,\phi)]\Omega(r,\theta,\phi)e^{-\frac{r}{\lambda_s}} \tag{10}$$

and

$$m(\theta,\phi) = -\vec{r_e}\cdot\hat{d}_\gamma(\theta,\phi) + \sqrt{\left(\vec{r_e}\cdot\hat{d}_\gamma(\theta,\phi)\right)^2 - (r_e^2 - R^2)} \tag{11}$$

is the distance from the electron to the PSUP along in the direction intersecting the scattering point S (i.e. $\hat{d}_\gamma(\theta,\phi)$)

Recall that $P^{TIM}(t_i,r,\theta,\phi)$ is the time function that represents the probability of recording a PMT time t_i given a photon TOF, $\frac{r+r'(r,\theta,\phi)}{c}$. One could imagine a using a Gaussian time jitter for the form of $P_{TIM}(r,\theta,\phi)$,

$$P^{TIM}(r,\theta,\phi) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}\{t_i - [\frac{r+r'(r,\theta,\phi)}{c}+t_e]\}^2}. \tag{12}$$

Note that the scattering probability shown in Equation 9 involves 3D integration at each hypothesized electron position, direction, and time. This is very computationally intensive. One could make the simplifying approximation of no PMT jitter, handwaving that the time smearing from the scattering dominates over PMT jitter, which reduces the dimensionality of the integration. This effectively replaces $P^{TIM}(t_i,r,\theta,\phi)$ with a δ-function at the scattered photon TOF, and is given by

$$P^{TIM}(r,\theta,\phi) = \delta(r - r_{TOF}(\theta,\phi)),$$

where $r_{TOF}(\theta,\phi)$ is the root (if one exists) of the equation

$$t_p - [\frac{r+r'(r,\theta,\phi)}{c} + t_e] = 0$$

such that $0 \leq r_{TOF}(\theta,\phi) \leq m(\theta,\phi)$

Inserting the δ-function form of $P^{TIM}(r,\theta,\phi)$ and performing the radial integration yields

$$P^{scat}(t_i) = \int_0^{2\pi}\int_0^{\pi} N_\gamma\rho_i(r_{TOF},\theta,\phi)e^{-N_\gamma\rho_i(r_{TOF},\theta,\phi)}\ sin\theta\ r_{TOF}^2\ d\theta d\phi \tag{13}$$

This form of probability could then be used as the scattering PDF in fitting events, once the scattering length is known.

References

[1] J. R. Klein, *Maximum Likelihood Fitting to Both Time and Angle*, SNO technical Report SNO-STR-94-055, University of Pennsylvania (1995).

[2] S. J. Brice, *The Elastic Fitter*, SNO technical report SNO-STR-95-040, Oxford University (1995).

[3] V. Rusu, *Post Reconstruction Cuts*, Tech. Rep., University of Pennsylvania (2001).

[4] TOXENN, *Analysis for First Solar Neutrino Paper*, Tech. Rep., Triumf, Oxford Univeristy, Rutherford Appleton Laboratory, University of Pennsylvania (2001).

[5] X. Chen, *An Improved Grid Fitter*, SNO technical report SNO-STR-00-004, Lawrence Berkeley National Laboratoy (2000).