# Multiple-Hit Effect in the SNO Photomultiplier Tube Array

A.W.P. Poon
*Nuclear Physics Laboratory, University of Washington, Seattle, WA 98195, USA*

**SNO-STR-97-038**

**Abstract**

One of the most favourable parameters in analysing the signals from the SNO photomultiplier tube (PMT) array is the number of fired PMTs — $N_{hits}$. However, this parameter is intrinsically non-linear because of the multiple-hit effect. In this report, we investigated how this affects the linearity of SNO detector's energy response. We have also developed an algorithm to correct for this non-linear effect. We shall then apply the gain correction technique which we have developed previously to gain-correct this multiple-hit corrected $N_{hits}$. Finally, we investigated how multiple hits affect the detector resolution, and reached the conclusion that $N_{hits}$ distributions do not give the maximum amount of statistical information.

# Contents

# 1 Introduction

In our previous report [1], we have developed an algorithm to correct for the positional dependence of SNO detector's energy response. This algorithm uses photomultiplier tube (PMT) array information—the number of fired PMTs ($N_{hits}$) or electronically calibrated charge ($N_{pe}$), and the parameter $R_{PMT}$, which is the distance between the event vertex and the PMT array in the electron's direction of propagation.

However, a shortcoming of this gain correction technique is its inability to correct for the multiple-hit effect in each event. The gain corrected energy calibration line extracted from $N_{hits}$ information has a $\sim$3% systematic uncertainty in the calibrated energy near 10 MeV. This uncertainty arises from multiple hits not being correctly accounted for in extracting the gain correction function. This problem can be eliminated by using $N_{pe}$, which is linear in energy.

The calibrated charge spectrum $N_{pe}$ is a broader distribution than the real generated photoelectron spectrum ($N_{gen}$) at a fixed energy. To first approximation, one can envision the energy spectrum reconstructed using the gain correction technique based on $N_{pe}$ information as the real distribution of generated photoelectrons convolved with the single photoelectron spectrum of the PMT. The single photoelectron spectrum that is used in SNOMAN is shown in Figure 1. As a result, any energy spectrum reconstructed using $N_{pe}$ information will be broader than the corresponding one reconstructed using $N_{hits}$ or $N_{gen}$.

To understand the systematic effects involving multiple hits, we investigate its dependence on energy and on position in both the $D_2O$ and neutral current detector (NCD) array installed scenarios. The Monte Carlo data sets we used in this study was generated using the same SNOMAN configuration as in our last report [1]. In this report, we shall summarise the findings of this investigation.

# 2 Energy and Positional Dependence of the Multiple-Hit Effect

## 2.1 Energy Dependence

To understand the energy dependence of the multiple-hit effect, we generated electrons in the energy range of 2 to 25 MeV with an uniform distribution. The data set contained 30,000 isotropic electrons distributed evenly within the $D_2O$ volume. In Figure 2, a plot of the real number of photoelectrons $N_{gen}$ is plotted against the number of fired PMTs $N_{hits}$ for this data set. In both running scenarios, it is clear that as the energy of the event increases, the bigger the spread in the deviation between $N_{gen}$ and $N_{hits}$ from linearity. The three-dimensional plots in Figure 3 demonstrate that the difference between the number of photoelectrons generated and the number of fired PMTs," $\Delta$,

$$\Delta = N_{gen} - N_{hits}, \tag{1}$$

has a much broader distribution at higher energy.

Looking at Figures 2 and 3 more closely, one would be astonished by the fact that even at $N_{hits}\sim$100 (about 10 MeV), the magnitude of $\Delta$ can be as much as 10% of $N_{hits}$. This percentage increases substantially as energy increases.

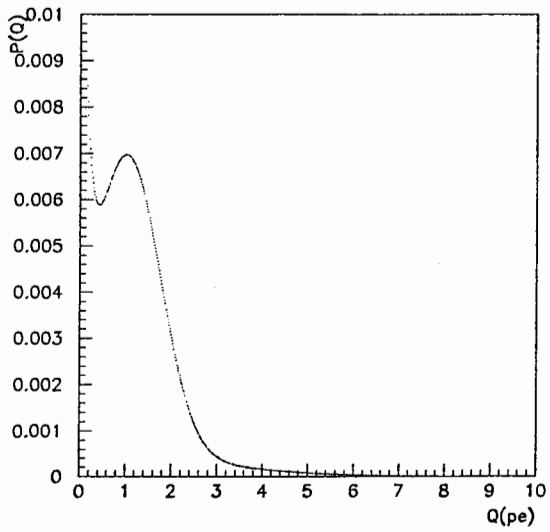


Figure 1: Single photoelectron spectrum used in SNOMAN.

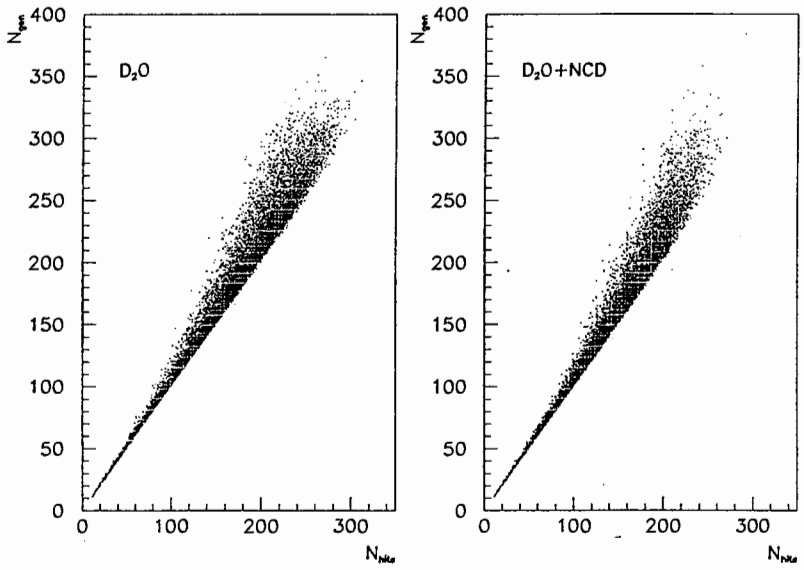


Figure 2: Real number of photoelectrons $N_{gen}$ and $N_{hits}$. The data shown here represents 30,000 electrons generated with energy drawn uniformly in the range of 2 to 25 MeV. The electrons are evenly distributed in the $D_2O$ volume, and have an isotropic angular distribution. It is clear that at higher $N_{hits}$ the spread of $N_{gen}$ becomes greater. This is an indication of an increasing number of PMTs receiving multiple hits.

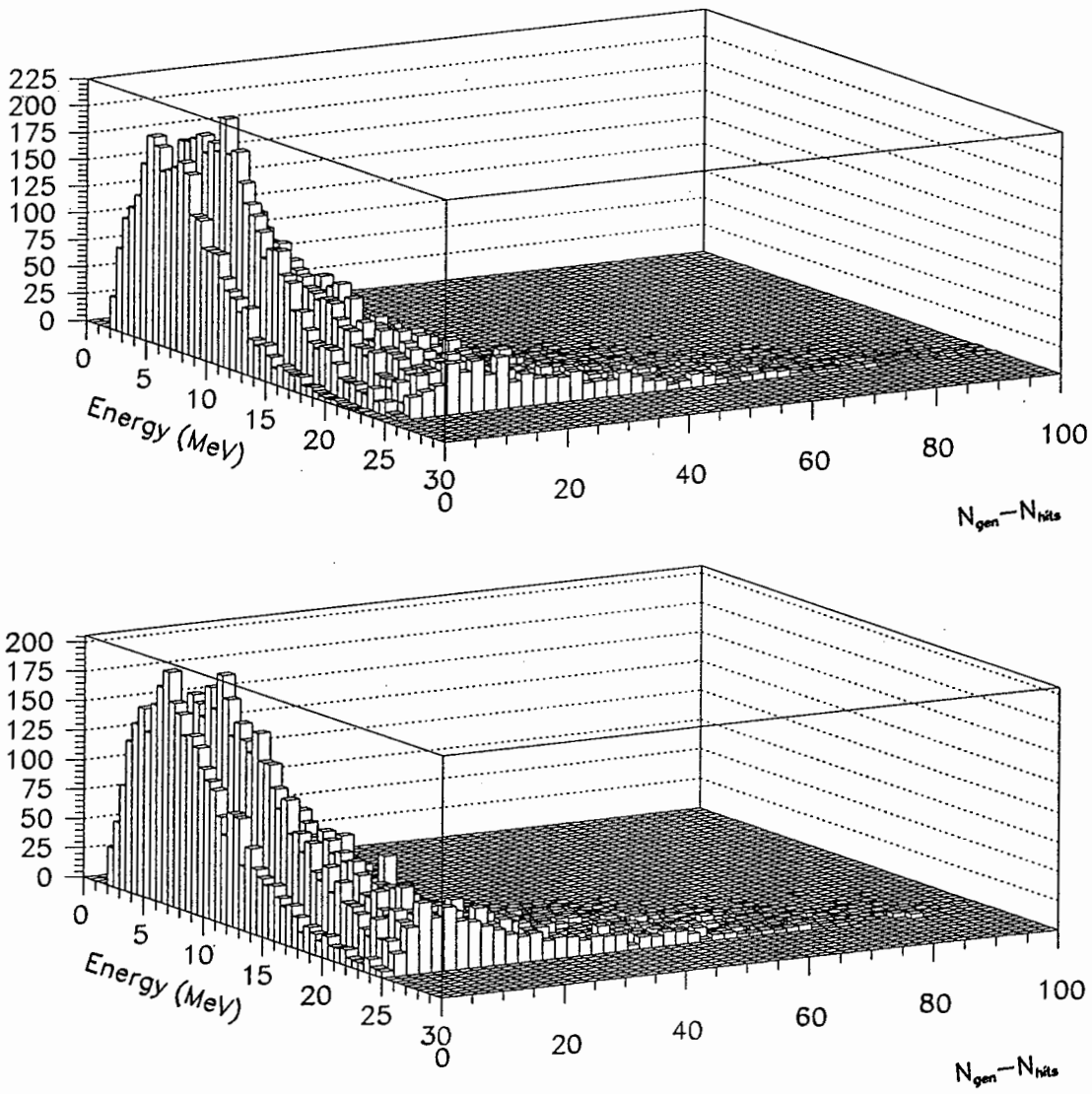


Figure 3: Energy dependence of the multiple-hit effect. The parameter, $\Delta$, defined as $N_{gen}$-$N_{hits}$ has a broader distribution as energy increases. The top panel shows the $\Delta$ distribution in the $D_2O$ running scenario, whilst the bottom panel shows the NCD installed scenario. The data shown here represent 30,000 electrons generated with energy drawn uniformly in the range of 2 to 25 MeV.

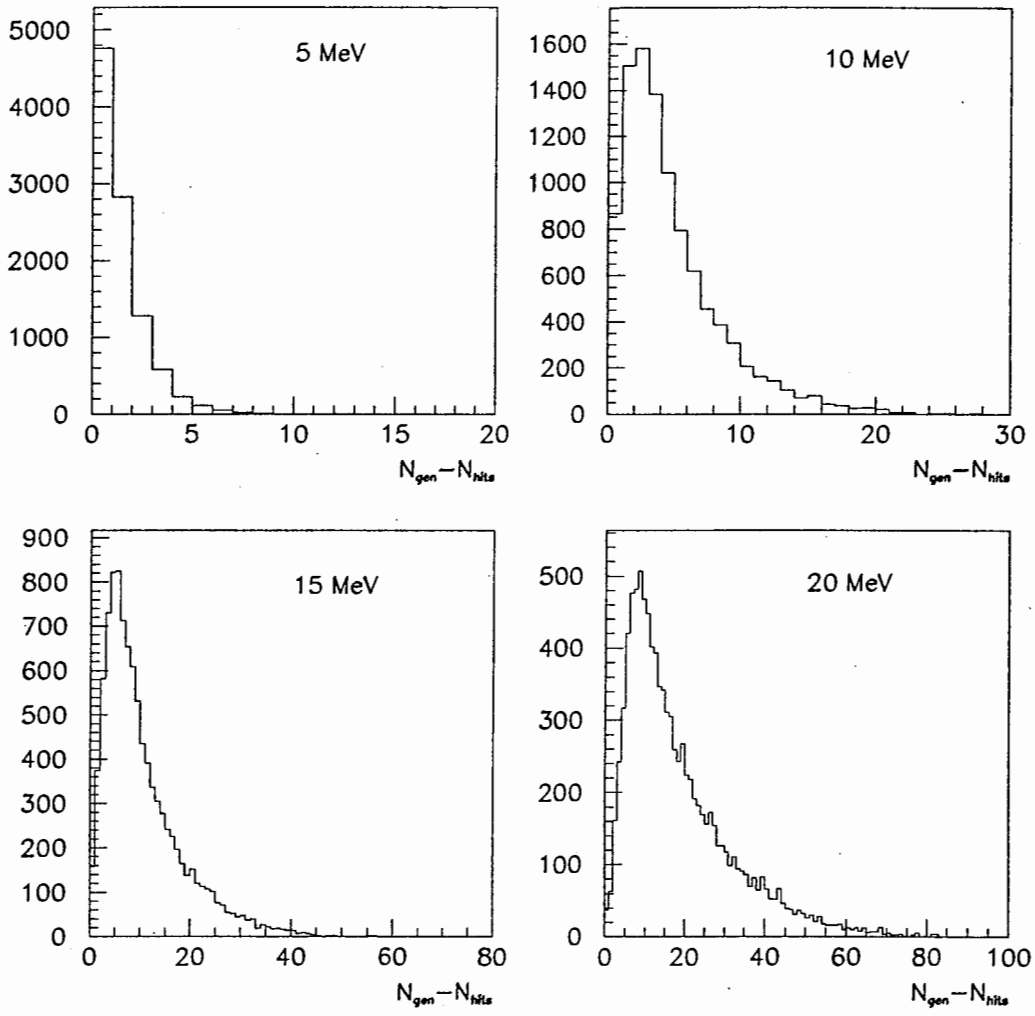


Figure 4: Distribution of $\Delta$ for 5, 10, 15 and 20-MeV electrons.

| Energy (MeV) | $D_2O$ | | $D_2O$+NCD | |
|---|---|---|---|---|
| | $\delta_{>0}$ | $\langle\Delta\rangle$ | $\delta_{>0}$ | $\langle\Delta\rangle$ |
| 5 | 0.517±0.009 | 1.43±0.01 | 0.475±0.008 | 1.34±0.01 |
| 10 | 0.912±0.013 | 4.76±0.04 | 0.863±0.013 | 4.30±0.04 |
| 15 | 0.984±0.014 | 10.42±0.08 | 0.967±0.014 | 9.37±0.09 |
| 20 | 0.996±0.014 | 18.38±0.08 | 0.992±0.014 | 16.27±0.13 |

Table 1: Energy dependence of the proportion of multiple-hit events and the mean $\Delta$ for electrons distributed evenly within the $D_2O$ volume. The uncertainties quoted in this table are statistical uncertainties.

For reasons that will become apparent later in Section 3, we have also generated monoenergetic electron data sets at various energies. In Figure 4, the distribution of $\Delta$ is shown for 5, 10, 15, 20-MeV isotropic electrons generated evenly within the $D_2O$ volume. The $\Delta$-histograms clearly demonstrate why using only the raw $N_{hits}$ information to extract energy of an event is susceptible to a major systematic uncertainty. For instance, over 50% of the events have $\Delta > 0$ at an energy of 5 MeV. We define this proportion of events having non-zero $\Delta$ as $\delta_{>0}$:

$$\delta_{>0} = \frac{N_{\Delta>0}}{N_{trigger}} \tag{2}$$

where $N_{\Delta>0}$ is the number of events that registered a $\Delta$ greater than zero, and $N_{trigger}$ is the number of triggered events. In Table 1, we have summarised the energy dependence of $\delta_{>0}$ and the mean $\Delta$ for the monoenergetic electrons distributed evenly within the $D_2O$ volume. We have verified that the $N_{gen}$-energy relation is linear; therefore, we can view $\langle\Delta\rangle$ as an estimate of $N_{hits}$'s deviation from linearity. Given that $N_{gen}\sim100$ at 10 MeV in the $D_2O$ running scenario, average $N_{hits}$ deviates from linearity by almost 5% at this energy.

## 2.2 Positional Dependence

Instead of using $R_{PMT}$ as in [1], we use a refined parameter in this study of the positional dependence of the multiple hit effect. We define this parameter $D_{PMT}$ as:

$$D_{PMT} = \frac{1}{N_{hits}} \sum_{i=1}^{N_{hits}} \left|\vec{r}_i - \vec{R}_v\right| \tag{3}$$

where $\vec{R}_v$ is the position of the event vertex generated by the Monte Carlo, the set $\{\vec{r}_i\}$ is the position of the $i^{\text{th}}$ hit PMT. Therefore, $D_{PMT}$ is the average distance between the event vertex and the hit PMTs.

The positional dependence of $\Delta$ is shown in Figure 5 for electrons in both run scenarios. The three-dimensional plots show that $\Delta$ has a much longer tail at small $D_{PMT}$. These events are the ones originated close to the acrylic vessel and are directed towards the PMT array in the outward direction. This general observation is in agreement with a similar analysis performed with the parameter $R_{PMT}$ in our previous report [1].

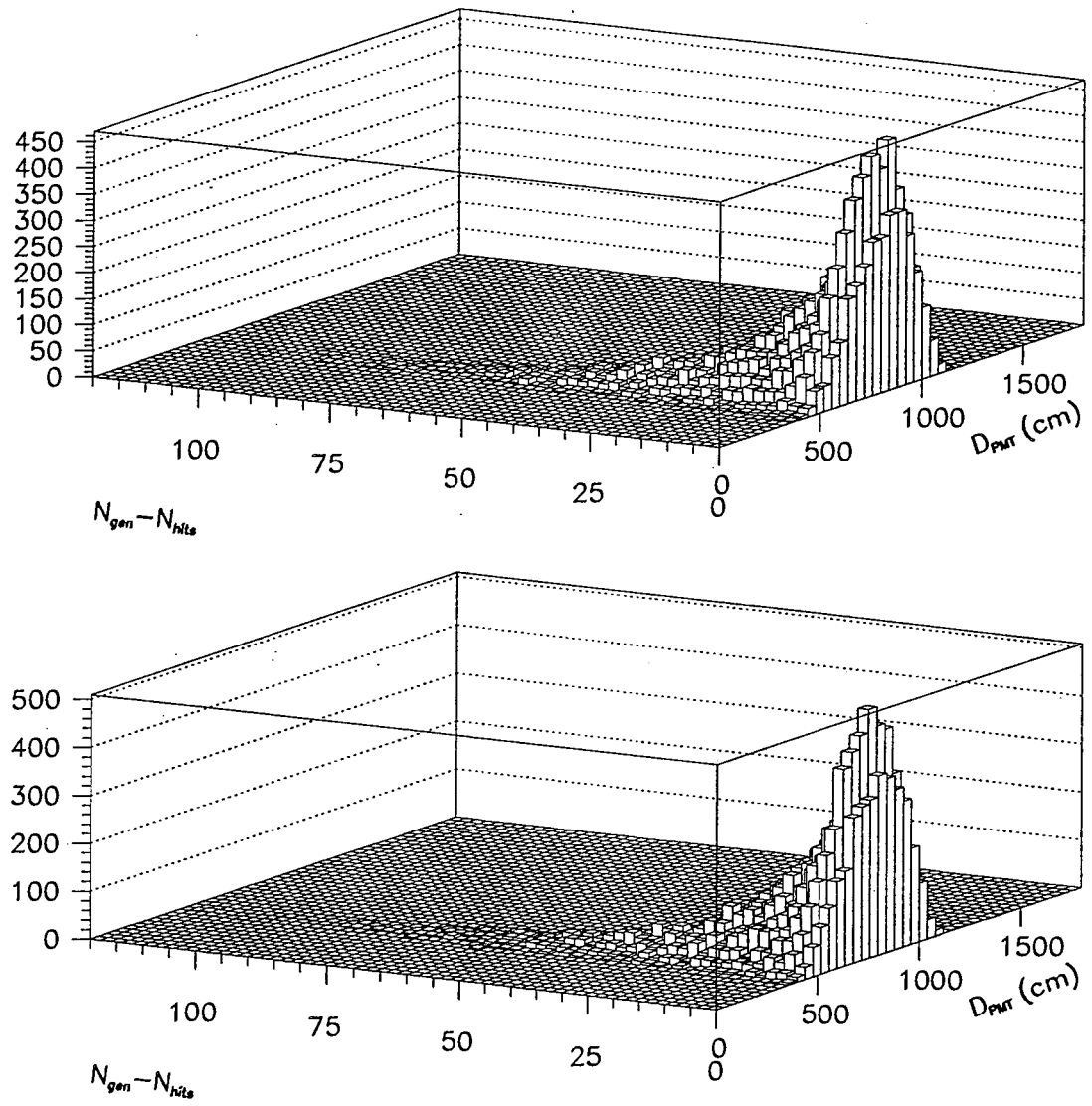


Figure 5: Positional dependence of the multiple-hit effect. The top panel shows the $\Delta$ distribution in the $D_2O$ running scenario, whilst the bottom panel shows the NCD installed scenario. $\Delta$ has a much longer tail at small $D_{PMT}$. These are the events originated close to the acrylic vessel and directed in the outward direction. The data shown here represent 30,000 electrons generated with energy drawn uniformly in the range of 2 to 25 MeV.

| Energy (MeV) | $D_2O$ | |
|---|---|---|
| | $\delta_{>0}$ | $\langle\Delta\rangle$ |
| 5 | 0.355±0.007 | 0.95±0.01 |
| 10 | 0.852±0.013 | 2.57±0.02 |
| 15 | 0.987±0.014 | 5.42±0.03 |

Table 2: Energy dependence of the proportion of multiple-hit events and the mean $\Delta$ for standard electron candles in $D_2O$. The uncertainties quoted in this table are statistical uncertainties.

We want to compare $\langle\Delta\rangle$ in Table 1 to standard electron calibration "candles" in order to understand the significance of positional dependence. For our standard electron candles, we used 5, 10, and 15-MeV electrons generated at the centre of the SNO detector in the pure $D_2O$ running scenario. In Table 2, we show the results of our analysis. Given the huge difference in $\langle\Delta\rangle$ between these standard candles and the isotropic cases in Table 1, it is apparent that one must handle the positional dependence of the signal with great care.

## 3 An Algorithm to Correct for the Multiple-Hit Effect

After a somewhat qualitative study in the previous sections on the energy and positional dependence of the multiple-hit effect, we shall describe in this section an algorithm which we have developed to correct for this effect.

Since it has been demonstrated [1, 2] that a quadratic fit of the energy-$N_{hits}$ calibration curve is better than a linear fit in the "zeroth order approximation,"[1] it is natural to assume that there is a second-order component in the $N_{hits}$ to $N_{gen}$ conversion:

$$N_{gen} = N_{hits} + a_2 N_{hits}^2 \tag{4}$$

where $a_2$ is the second-order contribution which is dependent on the position of the event. We parameterise $a_2$ as a function of $D_{PMT}$:

$$a_2(D_{PMT}) = \beta_0 + \frac{\beta_1}{(D_{PMT} - \beta_2)^2} \tag{5}$$

where $\beta_i$ are parameters to be fitted. Naively speaking, this $a_2$ contribution should vary as $\sim D_{PMT}^{-2}$ because the number of PMTs that lie on a "ring" subtended by the Čerenkov cone goes as $\sim D_{PMT}^2$.

To generate the $D_{PMT}$ dependence of $a_2$, we divided events in the 30,000-electron data set into 20-cm $D_{PMT}$ bins. For each of these bins, we extracted the corresponding $a_2$ by fitting $N_{gen}$ to the relationship in Eqn.(4). After extracting the $a_2$ function from all the bins, it was then fitted to Eqn.(5). This last fit for the two running scenarios is shown in Figure 6.

[1] We introduced this term in our previous report [1]. This means that the average energy response of the SNO detector is extracted by using monoenergetic events distributed isotropically and evenly in the $D_2O$ volume *without* any correction. That is, the raw $N_{hits}$ output is used.

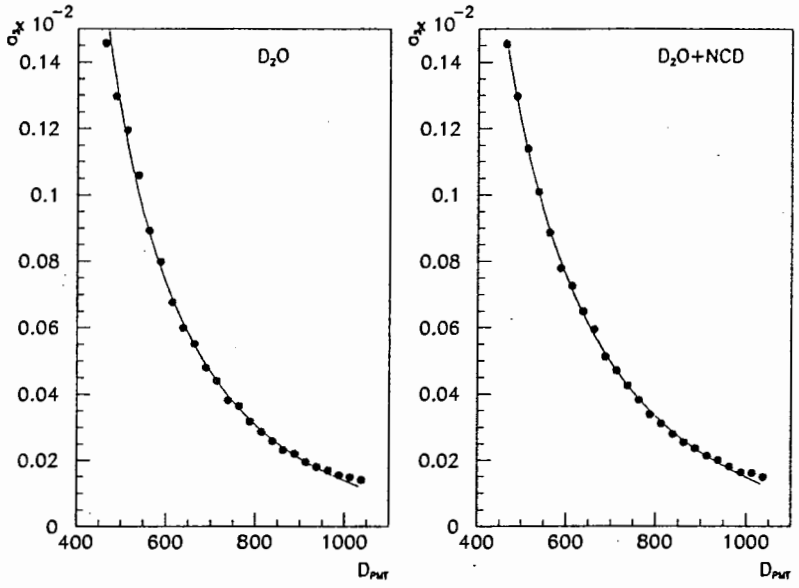


Figure 6: Fitting the second order contribution in $N_{hits}$ non-linearity as a function of $D_{PMT}$. The function $a_2$ was fitted according to Eqn.(5).

With the function $a_2$ determined for both the $D_2O$ and the NCD-installed running scenarios, we can then correct for the multiple-hit effect on an event-by-event basis using the available $N_{hits}$ and vertex information. We first ran the events in the 30,000-electron data sets through this correction mechanism. The corrected $N_{hits}$ distributions are shown in Figure 7. The agreement between the multiple-hit corrected $N_{hits}$ spectra and the corresponding Monte Carlo generated $N_{gen}$ spectra is good. To further demonstrate the power of this correction algorithm, the distributions of the difference between the corrected $N_{hits}$ and Monte Carlo $N_{gen}$ for both running scenarios are shown in Figure 8. The sharp peaks at zero in this latter figure convince us further that this algorithm performs this multiple-hit correction to a high degree of accuracy.

So far, we have been using in our analysis the vertex position $\vec{R}_v$ generated by SNOMAN to avoid introducing any event fitter dependent effect. In the real SNO detector, this position is not known and can only be estimated by an event fitter. In our previous report [1], we demonstrated that the simplest fitter of all—the time fitter—can reconstruct event vertex and electron direction reliably in both running scenarios. We wanted to verify that the good agreement obtained in the multiple-hit correction analyses above will also hold for time fitter analysed events. We defined the time fitter analysed position parameter $D_{PMT}^{fit}$ as

$$D_{PMT}^{fit} = \frac{1}{N_{hits}} \sum_{i=1}^{N_{hits}} \left| \vec{r}_i - \vec{R}_v^{fit} \right| \tag{6}$$

where $\vec{R}_v^{fit}$ is the event vertex fitted by the time fitter. We repeated the analysis by first extracting the corresponding $a_2^{fit}$ in this scenario. The shape of the resulting $a_2^{fit}$ functions looks similar to

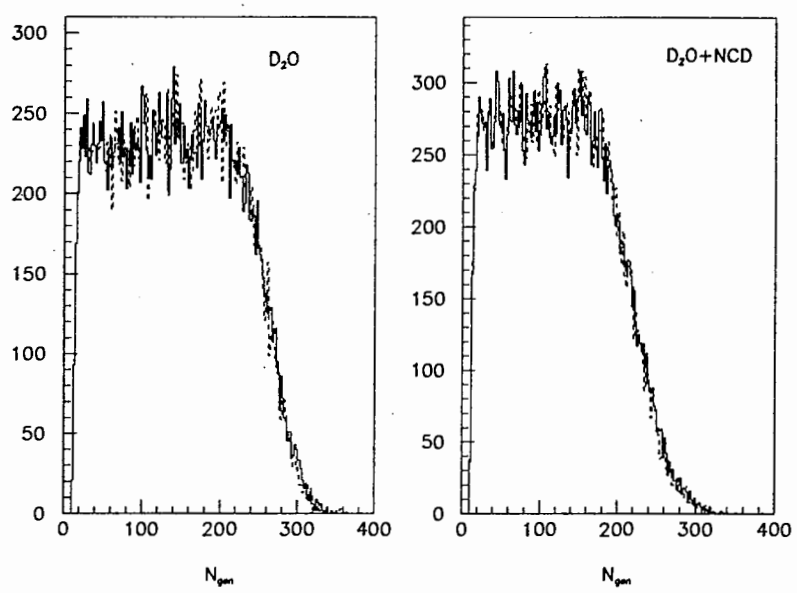


Figure 7: Comparing the multiple-hit corrected $N_{hits}$ distributions to Monte Carlo generated $N_{gen}$ distributions. $N_{hits}$ from the 30,000-electron data set was corrected on an event-by-event basis. The Monte Carlo generated $N_{gen}$ distributions are shown as the solid histograms, whilst the ones corrected from $N_{hits}$ using the multiple-hit correction scheme are shown as dashed histograms.

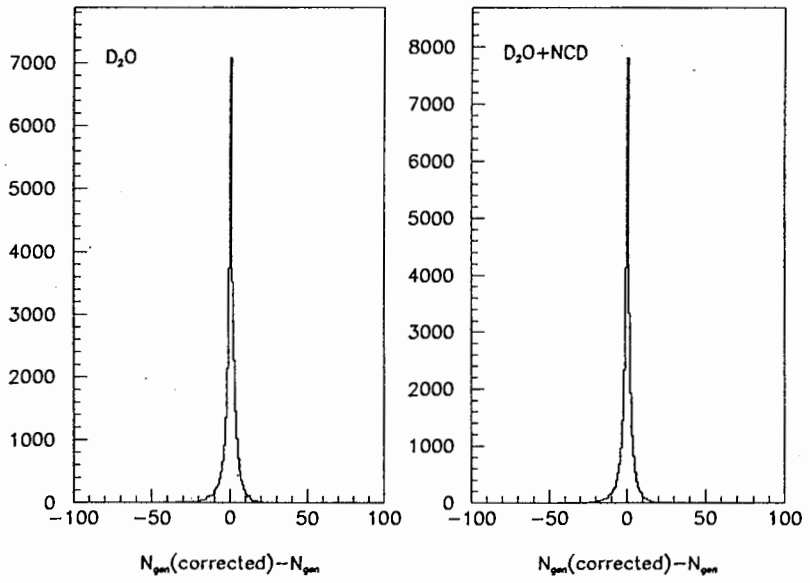


Figure 8: Difference between multiple-hit corrected $N_{hits}$ distributions and Monte Carlo generated $N_{gen}$ distributions.

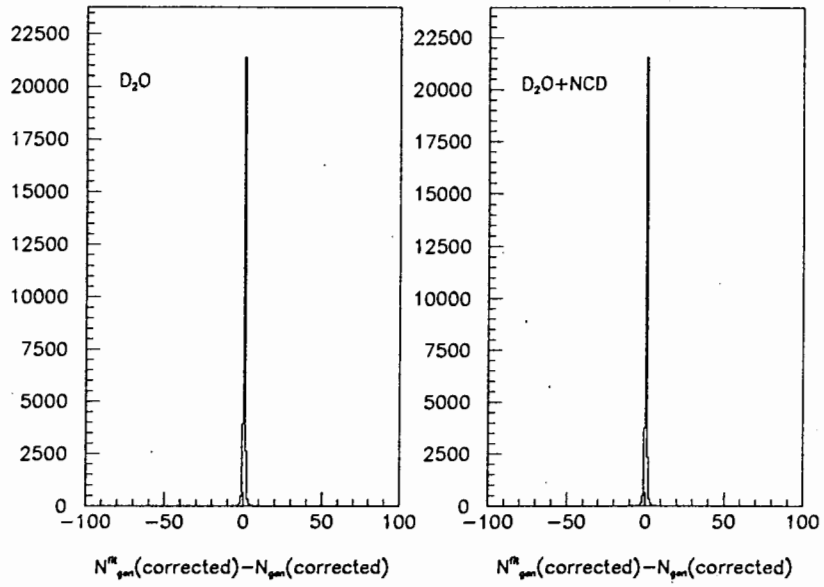


Figure 9: Comparing the multiple-hit corrected $N_{hits}$ distributions using time-fitter fitted vertices to that corrected using Monte Carlo generated vertices.

those in Figure 6, although the numerical values for the $\beta_i$ parameters differ somewhat. Finally, we found that there is not any degradation in the multiple-hit correction algorithm when the time-fitter analysed parameters were used. In Figure 9, the difference between the multiple-hit corrected $N_{hits}$ distributions using time-fitter fitted vertices and that corrected using Monte Carlo generated vertices is shown for both the $D_2O$ and the NCD-installed running scenarios.

Astute readers would realise that the good agreement between the multiple-hit corrected $N_{hits}$ distributions and the $N_{gen}$ distributions might be biassed. This is because we corrected the $N_{hits}$ distributions using a correction function extracted from the same data set. To verify the robustness of this correction scheme, we used the correction functions extracted from the 30,000-electron data sets to correct for monoenergetic electron $N_{hits}$ spectra at 5, 10, 15, and 20 MeV. Each of these monoenergetic sets contains approximately 10,000 triggered events. The $N_{hits}$ corrected spectra in the $D_2O$ running scenario are shown in Figure 10, whilst those for the NCD-installed scenario are shown in Figure 11. It is clear that the multiple-hit correction scheme we have developed did an excellent job in correcting these monoenergetic electron data sets as well.

When one inspects the $N_{gen}$ spectra for the both the pure $D_2O$ and the NCD-installed scenarios in Figure 11, it is clear that the spectra have a slight asymmetry. In the pure $D_2O$ scenario, this asymmetry arises mainly from light attenuation in the $D_2O$, the acrylic vessel and the light water. This asymmetry is worse in the NCD-installed case because of extra light scattering and absorption by the NCD array.

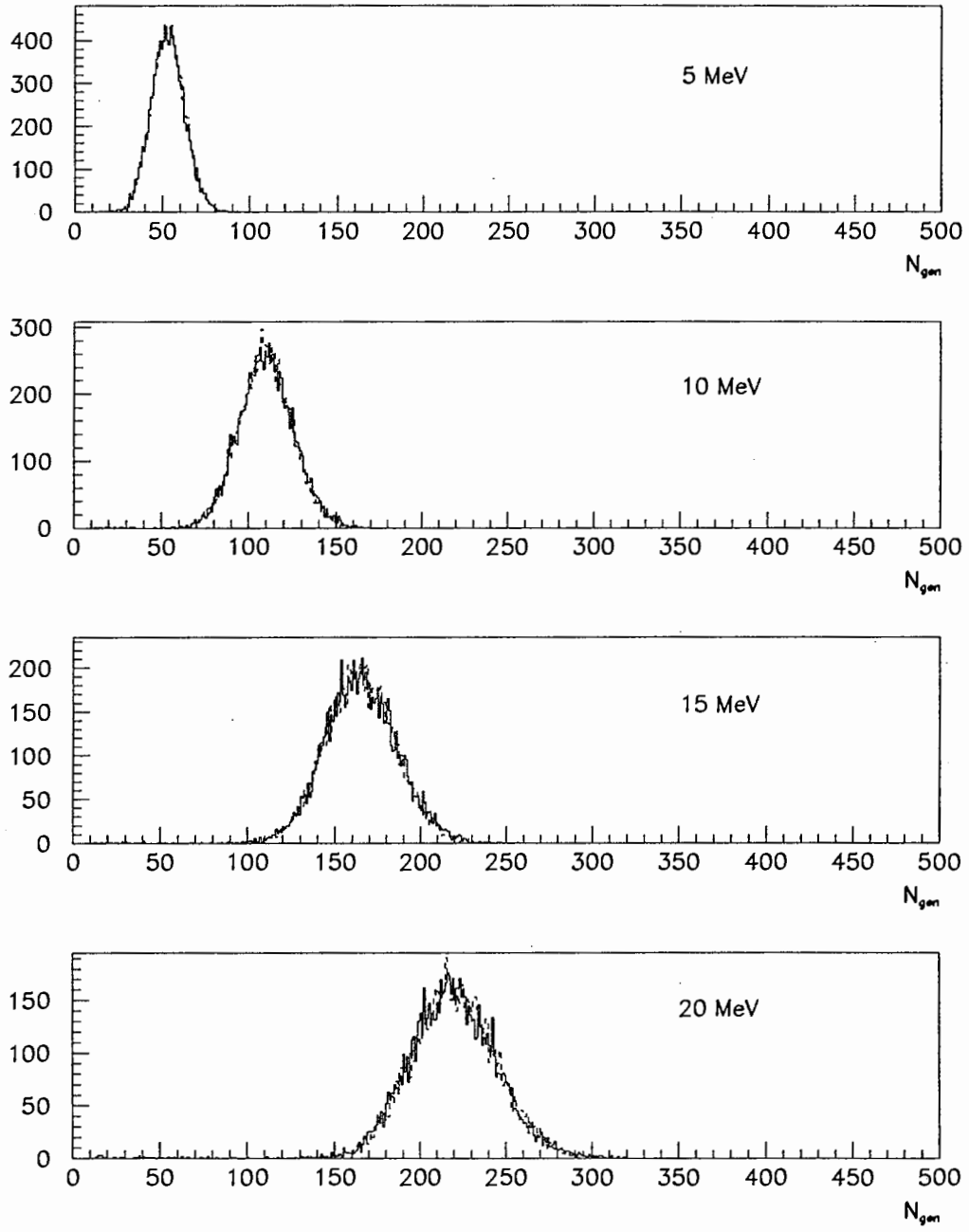


Figure 10: Reconstructing monoenergetic electron $N_{gen}$ distribution in the $D_2O$ running scenario. This is a consistency check to make sure that the good agreement obtained in Figure 7 is not biassed. The solid histograms are the Monte Carlo generated $N_{gen}$ distributions, whilst the dashed histograms are the multiple-hit corrected ones based on $N_{hits}$ information.

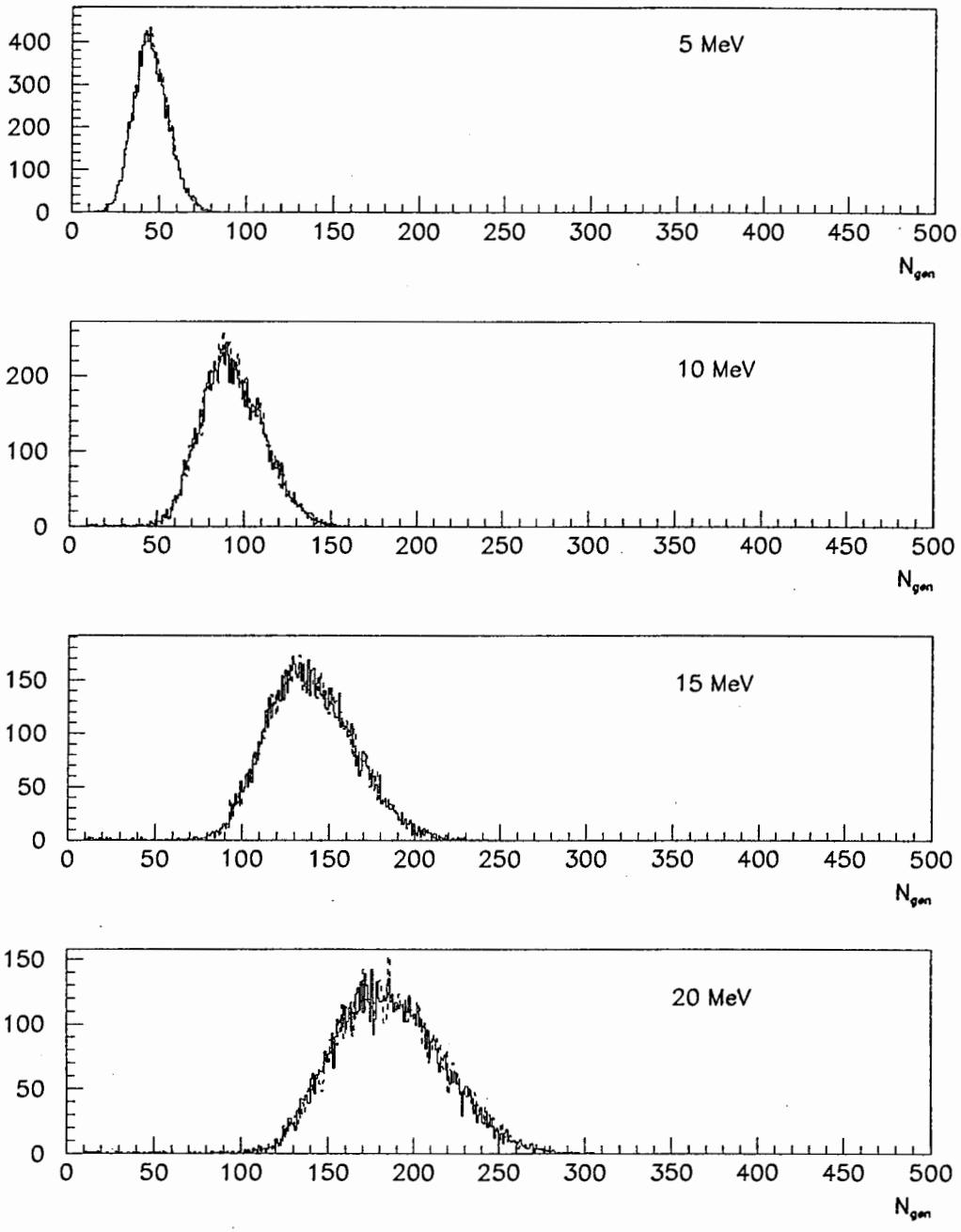


Figure 11: Reconstructing monoenergetic electron $N_{gen}$ distribution in the NCD-installed running scenario. This is a consistency check to make sure that the good agreement obtained in Figure 7 is not biassed. The solid histograms are the Monte Carlo generated $N_{gen}$ distributions, whilst the dashed histograms are the multiple-hit corrected ones based on $N_{hits}$ information.

| Energy (MeV) | $D_2O$ | |
|---|---|---|
| | $\delta_{>0}$ | $\langle\Delta\rangle$ |
| 5 | 0.374±0.007 | 1.16±0.01 |
| 10 | 0.836±0.013 | 4.00±0.03 |
| 15 | 0.970±0.014 | 9.94±0.08 |
| 20 | 0.994±0.014 | 17.16±0.08 |

Table 3: Energy dependence of the proportion of multiple-hit events and the mean $\Delta$ for gamma rays distributed evenly within the $D_2O$ volume. The uncertainties quoted in this table are statistical uncertainties.

# 4 Multiple-Hit Effect in Gamma-Ray Events

Whilst relativistic electrons emit Cĕrenkov photons readily in the $D_2O$ volume, gamma rays have to be converted into electrons through multiple Compton scattering and pair production before their detection. As we have noted previously [1], the gamma-ray signals are smaller than the electron signals of the same energy. This is primarily because of multiple Compton scattering as each of the Compton electrons would stop radiating Cĕrenkov photons once their kinetic energy has dropped below the Cĕrenkov threshold.

As the energy of the gamma ray increases, the probability that it would be converted to electrons through Compton scattering decreases whilst that for pair production increases. Therefore, we would expect different Cĕrenkov photon hit pattern on the PMT array for electrons and gamma rays. The difference between the electron and the gamma-ray hit patterns is small in the solar neutrino energy regime. But we should still ask ourselves whether the multiple-hit correction technique we developed for electrons above would work for gamma-ray events as well.

In Table 3, $\delta_{>0}$ and $\langle\Delta\rangle$ for 5, 10, 15 and 20-MeV isotropic gamma rays distributed evenly within the $D_2O$ volume are shown. Comparing these values with the corresponding ones for electrons in Table 1, we can see that multiple-hit effect appears to affect gamma-ray events to a slightly smaller extent than electron events at the same energy. In Figure 12, we show this comparison.

However, this difference in the influence of the multiple-hit effect on electron and gamma-ray events might attribute to either the difference in the PMT hit pattern or simply the reduction in Cĕrenkov light output in gamma-ray events. However, one would expect the latter case to have a bigger influence. Hence, we normalised $\langle\Delta\rangle$ by the average number of generated photoelectrons $\langle N_{gen}\rangle$ for both electron and gamma-ray events. In Figure 13, we show this quantity $\langle\Delta\rangle/\langle N_{gen}\rangle$ as a function of energy. The close agreement between the electron and the gamma-ray events indicates that the influence of multiple hits differs primarily by the difference in their Cĕrenkov light output.

In our previous report [1], we noted that gamma rays converted in the acrylic vessel have an enhanced Cĕrenkov light output because of acrylic's higher index of refraction. Given that these events are originated closer to the PMT array, the probability that each hit PMT would receive multiple photons increases if the gamma rays are directed towards the array. In Figure 14(a), $N_{gen}$ is plotted against $D_{PMT}$ for 10,000 isotropic 20-MeV gamma ray events distributed evenly within

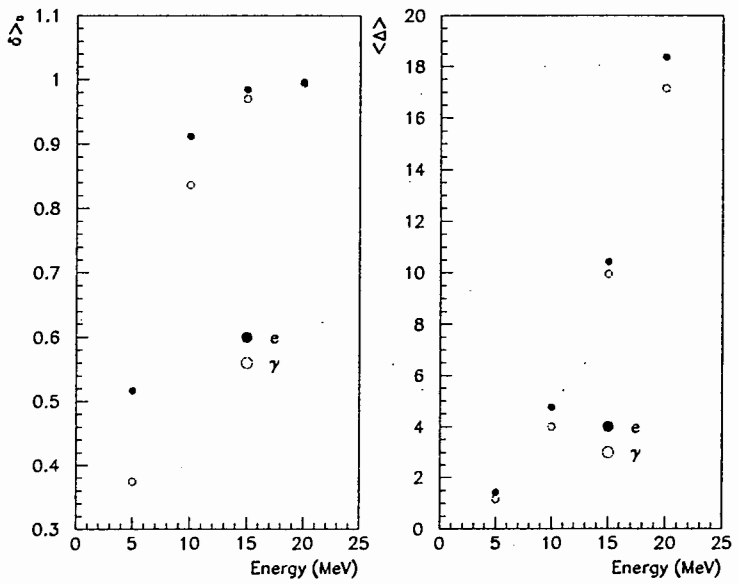


Figure 12: Comparing multiple-hit effect for electron and gamma-ray events. We can see that multiple-hit effect appears to affect gamma-ray events to a slightly smaller extent than electron events at the same energy

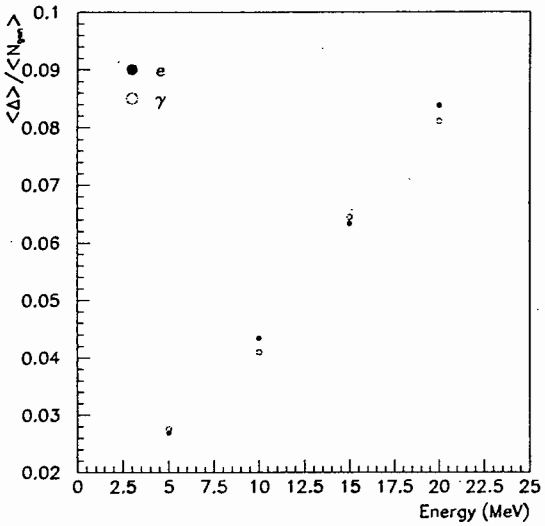


Figure 13: $\langle\Delta\rangle$ normalised by the average number of generated photoelectrons $N_{gen}$ for monoenergetic electrons and gamma rays. It is clear from this plot that the influence of multiple hits on electron and gamma-ray events differs primarily by the difference in their Cĕrenkov light output.

the $D_2O$ volume. The long tail at small $D_{PMT}$ illustrates this enhancement. We placed a cut on this tail as shown in the figure. In Figure 14(b), the $\Delta$ distribution for events within this cut and for the whole data set are shown. It is clear that events within this cut dominate the histogram at large $\Delta$. Finally, in Figure 14(c), we show how well the time fitter can reconstruct the events within this cut. Almost all these events were reconstructed outside the vessel. This is expected since the vertices of the converted electrons are in the acrylic vessel. Moreover, the time fitter systematically pulls the fitted vertex towards the PMT array.

We should apply the multiple-hit correction scheme we developed in the last section to the gamma-ray events to check for its robustness. The time fitter was used to reconstruct these gamma-ray events. We then used the reconstructed information to calculate $D^{fit}_{PMT}$, and to multiple-hit correct the $N_{hits}$ information using the correction function generated for the electron events in the last section. This usage of the electron correction function is to ensure that it is not necessary to perform a particle identification prior to this correction. In Figure 15, we show the multiple-hit corrected $N_{hits}$ spectra along with Monte Carlo generated $N_{gen}$ spectra. The agreement between them is good.

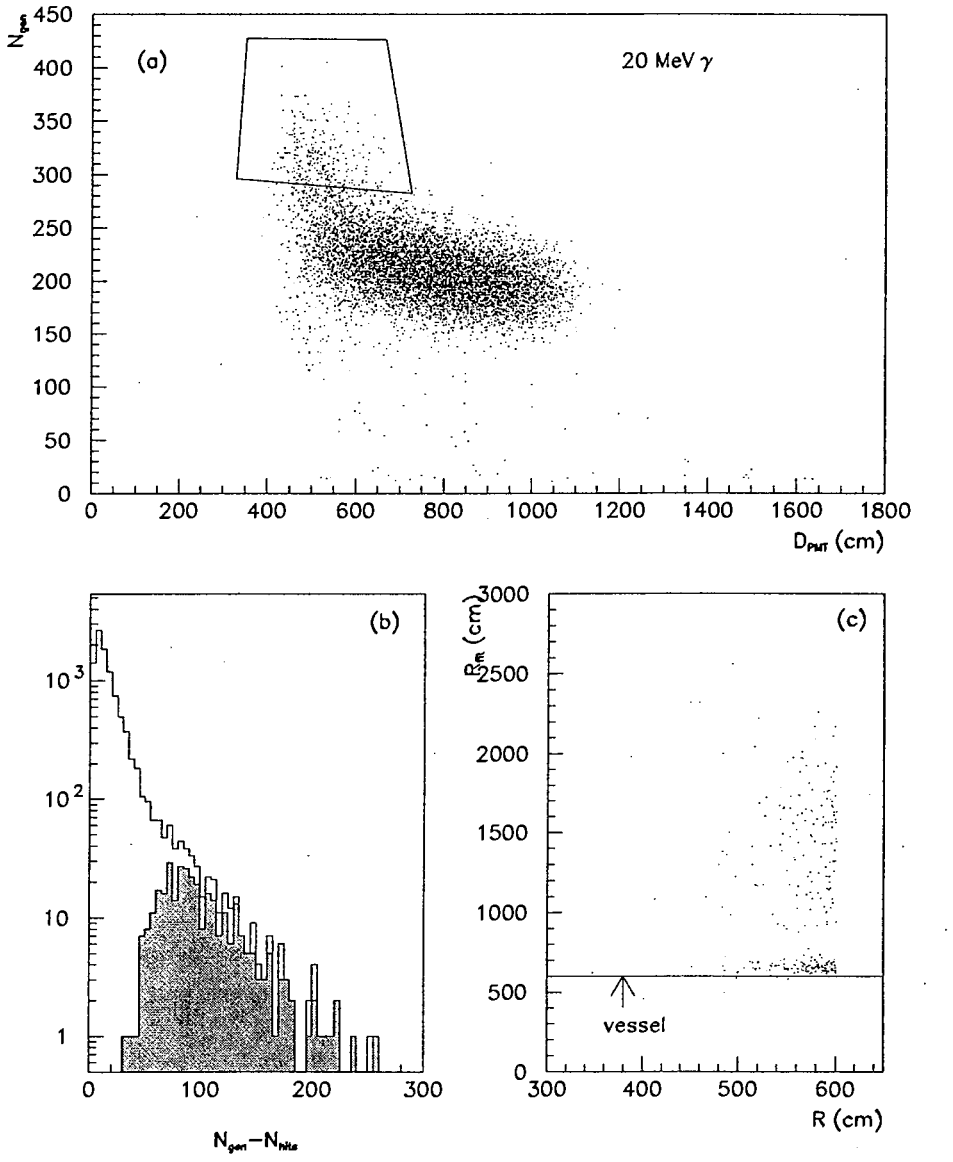


Figure 14: Positional dependence of $N_{gen}$ for 20-MeV gamma rays. (a) $N_{gen}$ is plotted against $D_{PMT}$ for 10,000 isotropic 20-MeV gamma rays distributed evenly within the $D_2O$ volume. The tail at small $D_{PMT}$ indicates enhanced Cĕrenkov light output due to gamma-ray conversion in the acrylic vessel; (b) $\Delta = N_{gen} - N_{hits}$ is plotted for the whole data set and for the events within the cut shown in (a). The shaded histogram shows the distribution of $\Delta$ after this cut had been applied. It is clear that events within this cut dominate the histogram at large $\Delta$; (c) The time fitter fitted radial distance ($R_{fit}$) is plotted against the origin of the gamma rays ($R$) for events within the cut. Almost all of the events within this cut were reconstructed outside the $D_2O$ volume.

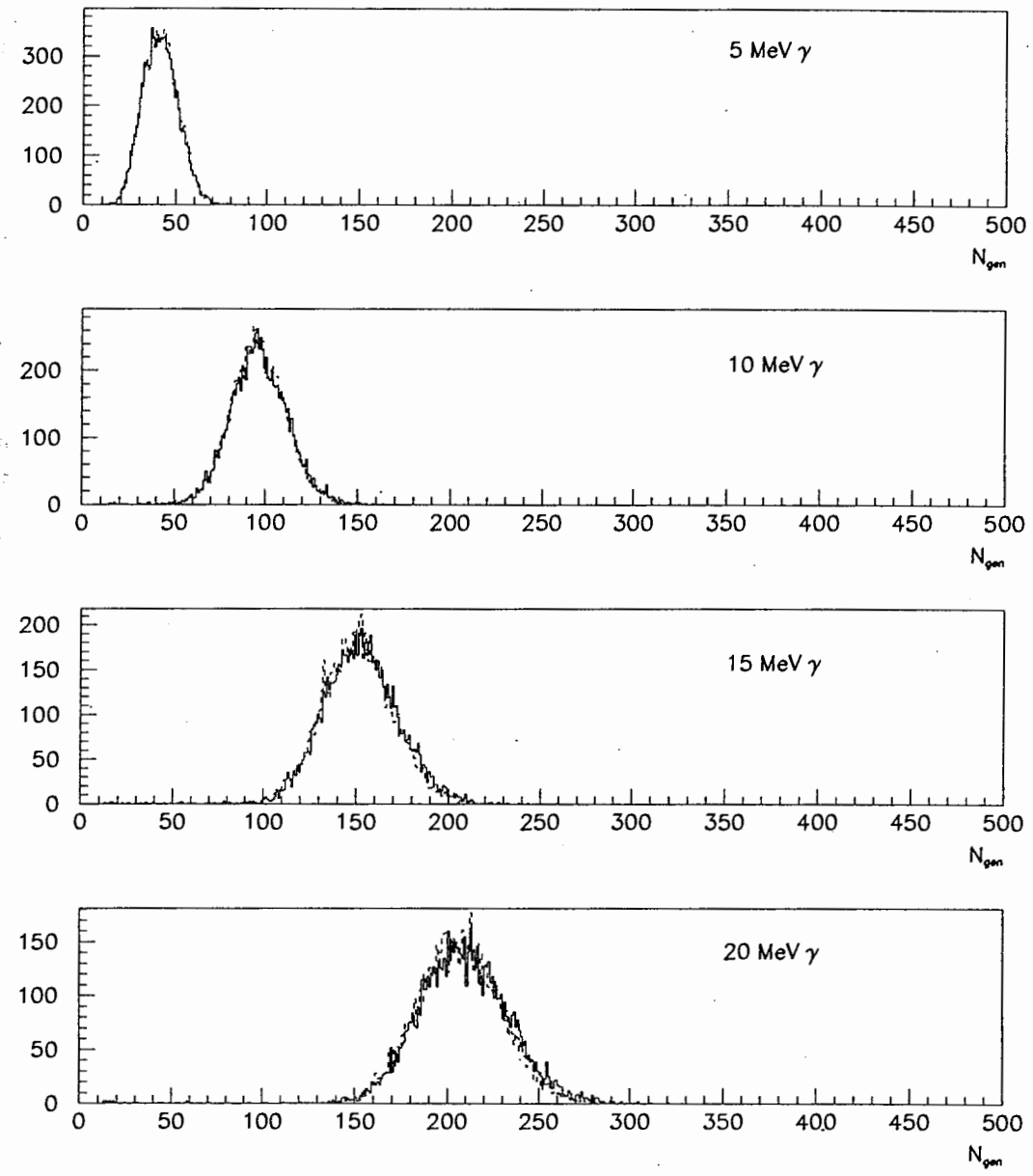


Figure 15: Reconstructing monoenergetic gamma-ray $N_{gen}$ distribution in the $D_2O$ running scenario. We used time fitter reconstructed $D_{PMT}^{fit}$ to multiple-hit correct the events. The correction function used was the one generated with the electron data. The solid histograms are the Monte Carlo generated $N_{gen}$ distributions, whilst the dashed histograms are the multiple-hit corrected ones based on $N_{hits}$ information.

# Multiple-Hit Effect in the SNO Photomultiplier Tube Array

A.W.P. Poon
*Nuclear Physics Laboratory, University of Washington, Seattle, WA 98195, USA*

**SNO-STR-97-038**

**Abstract**

One of the most favourable parameters in analysing the signals from the SNO photomultiplier tube (PMT) array is the number of fired PMTs — $N_{hits}$. However, this parameter is intrinsically non-linear because of the multiple-hit effect. In this report, we investigated how this affects the linearity of SNO detector's energy response. We have also developed an algorithm to correct for this non-linear effect. We shall then apply the gain correction technique which we have developed previously to gain-correct this multiple-hit corrected $N_{hits}$. Finally, we investigated how multiple hits affect the detector resolution, and reached the conclusion that $N_{hits}$ distributions do not give the maximum amount of statistical information.

| Energy (MeV) | $\mu_{N_{gen}}$ | $\sigma_{N_{gen}}$ |
|---|---|---|
| 5 | 50.3 | 8.4 |
| 10 | 103.4 | 13.1 |
| 15 | 155.5 | 16.4 |

Table 4: $N_{gen}$ statistics for 5, 10, and 15-MeV monoenergetic electron standard candle sets. The electrons were generated at the centre of the SNO detector. The resulting $N_{gen}$ distributions were fitted to a normal distribution. The statistical uncertainty of the numbers is less than 0.2%.

# 5 Applying the Gain Correction Technique to Multiple-Hit Corrected Events

As we mentioned earlier, the gain correction technique we developed in our previous study [1] broke down when trying to correct for the intrinsically non-linear entity $N_{hits}$. With the multiple-hit correction scheme above, we can then linearise the $N_{hits}$ response. In this section, we shall concentrate on applying the gain correction technique to the multiple-hit corrected $N_{hits}$ distributions.

First we need to generate the "standard candle" for the gain correction mechanism. We generated monoenergetic electron events from the centre of SNO with energies of 5, 10 and 15 MeV, and extracted the mean $N_{gen}$ from these runs by fitting the spectra to a normal distribution. In Table 4, we summarised the extracted information.

Following the same procedure as we outlined in [1], we generated the gain correction function $G^i_{N^{fit}_{gen}}(D^{fit}_{PMT})$ where $N^{fit}_{gen}$ represents the multiple-hit corrected $N_{hits}$ using the time fitter fitted $D^{fit}_{PMT}$ for $i$=$D_2O$ or $D_2O$+NCD. Only events that were reconstructed with a vertex within the $D_2O$ volume were used in extracting the gain correction function. In Figure 16, a plot showing the gain correction functions extracted using different candle sets demonstrates that the non-linearity associated with the gain correction functions extracted with only $N_{hits}$ information that we observed in [1] have been corrected for.

As we argued in [1], we would expect

$$G^i_{N^{fit}_{gen}}(D^{fit}_{PMT}) \propto \exp\left(\lambda D^{fit}_{PMT}\right) \qquad i = \mathrm{D_2O},\ \mathrm{NCD} \tag{7}$$

where $\lambda$ is an average attenuation length for optical photons propagating through the $D_2O$, acrylic vessel and the light water. Hence, we fitted the gain correction functions in Figure 16 to a linear relationship. In Figure 17, linear fits to the gain correction functions are shown.

Good agreement amongst the fits in each of the two subsets ($D_2O$, and $D_2O$+NCD) prompted us to use a weighed average of the fits to generate an average gain correction function $\langle G^i_{N^{fit}_{gen}}(D^{fit}_{PMT})\rangle$ for each running scenario.

To combine the multiple-hit correction and the gain correction, $N_{hits}$ information was first processed by the multiple-hit algorithm on an event-by-event basis. The multiple-hit corrected output for each event was then processed by the gain correction mechanism. In Figures 18 and 19, the two-tier

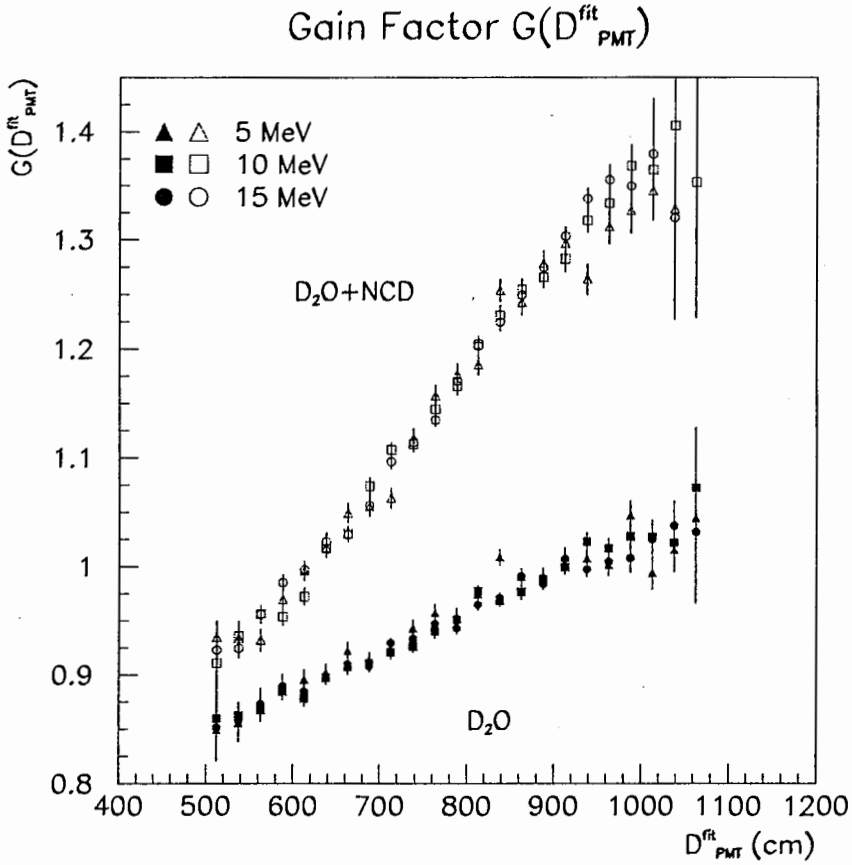


Figure 16: Gain correction functions extracted using multiple-hit corrected $N_{hits}$ using 5, 10 and 15 MeV standard electron candles. Only events that were reconstructed with a vertex within the $D_2O$ volume were used in extracting the gain correction functions. The good agreement amongst the gain correction functions extracted using different standard candles convinced us that the non-linearity associated with $N_{hits}$ that we observed previously have been correctly dealt with.

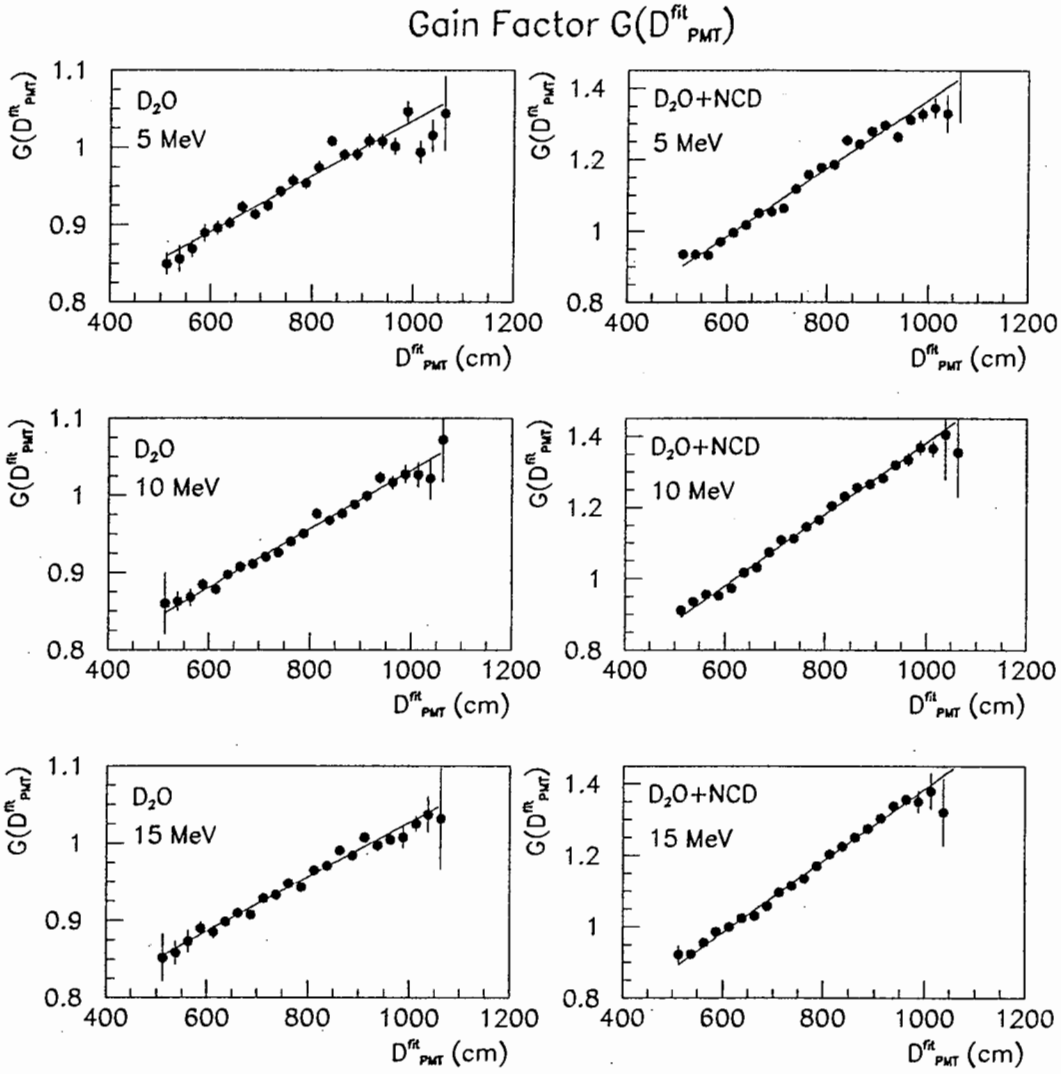


Figure 17: Linear fit of $G^i_{N^{fit}_{gen}}(D^{fit}_{PMT})$ to $D^{fit}_{PMT}$

corrected energy spectra are shown along with the corresponding real energy spectra convoluted with the detector resolution function.

We extracted the energy calibration line and the energy resolution of the detector by running monoenergetic electron events through the two-tier correction system. These events were generated with an isotropic angular distribution and an uniform distribution in position within the $D_2O$ volume. The energy calibration and the energy resolution functions are shown in Figure 20. In the figure, the energy calibration lines are linear. One would notice a significant improvement in the energy resolution in the NCD installed scenario when comparing the two-tier corrected spectra with the "zeroth order" $N_{hits}$ resolution. In Figure 21, we compare these two resolution functions in the NCD installed scenario. As a comparison, the "zeroth order" $N_{hits}$ and the two-tier corrected resolution in the $D_2O$ running scenario are also shown. When comparing the "zeroth order" $N_{hits}$ resolution and the two-tier corrected resolution in the $D_2O$ running scenario, it is clear that the former has a marginally better resolution. This is counter-intuitive as more information was used in processing the signals in the latter case. We shall investigate this "anomaly" in the next section.

An improvement in the energy resolution after the events have been processed by the two-tier correction scheme has a significant impact on SNO data analysis. A previous analysis by Skensved [3] showed that the SNO analysis threshold might be as high as 5.6 MeV when the NCD array is in place. With this two-tier correction scheme we developed, the energy resolution has significantly improved in the NCD running scenario. We estimated that the analysis threshold might be lowered to $\sim$5.3 MeV when this analysis technique is employed under the NCD running scenario.

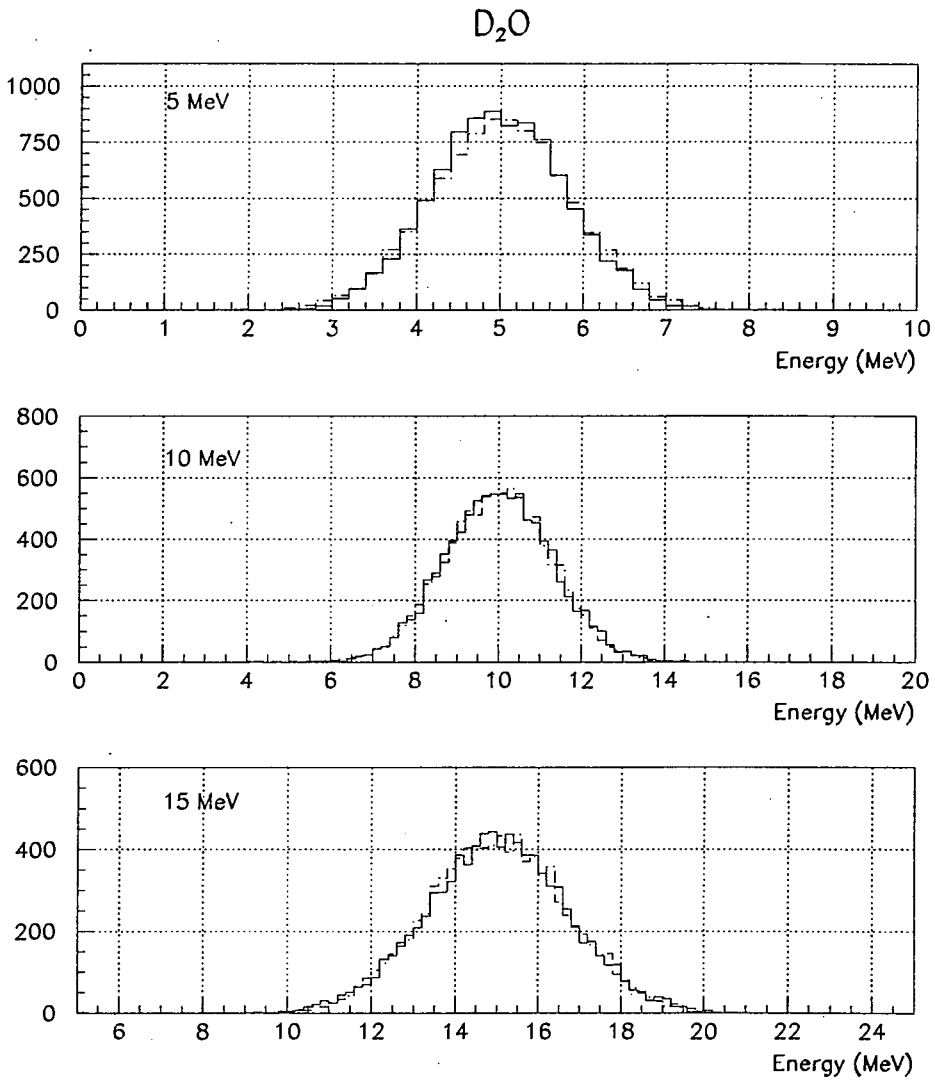


Figure 18: Comparing the two-tier correction scheme with theoretical expectation in the $D_2O$ running scenario. The solid histograms are the two-tier corrected distribution, and the dashed histograms are generated by convoluting the detector resolution function with a theoretical energy line.

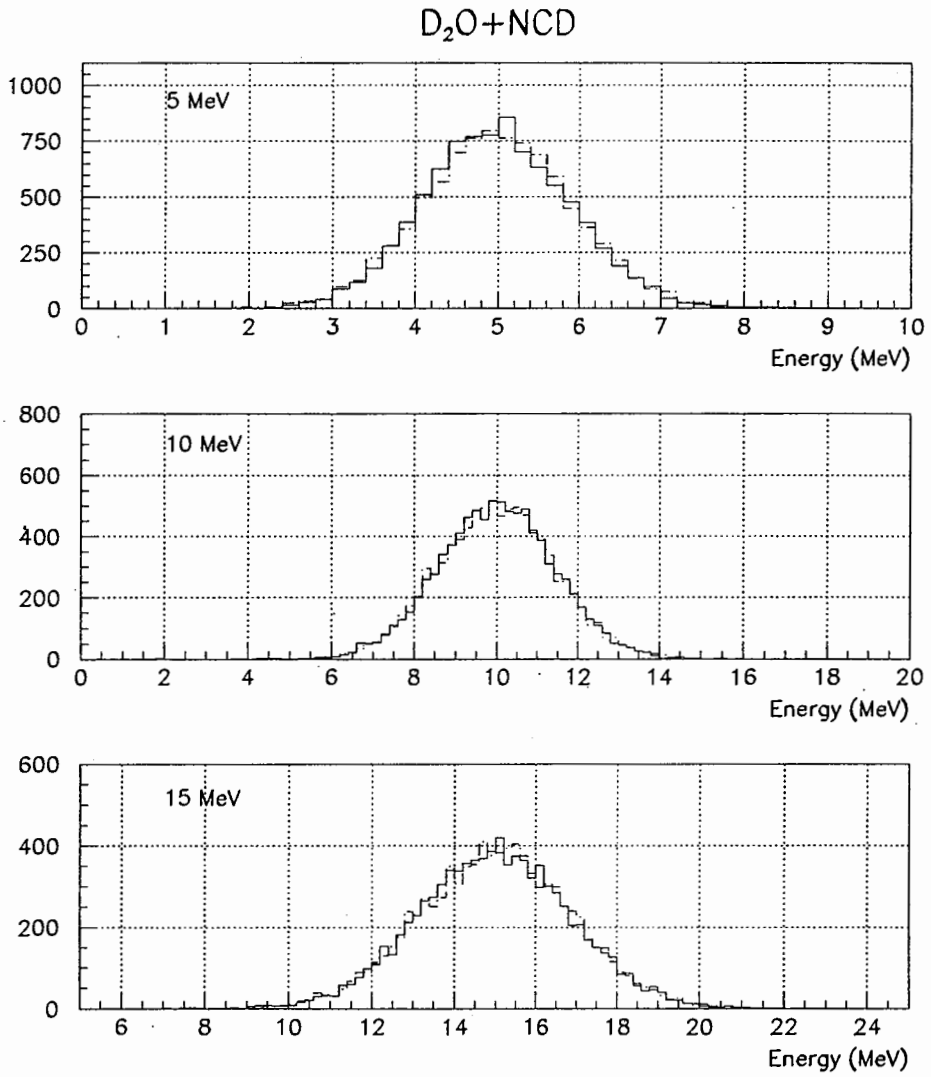


Figure 19: Comparing the two-tier correction scheme with theoretical expectation in the NCD installed running scenario. The solid histograms are the two-tier corrected distribution, and the dashed histograms are generated by convoluting the detector resolution function with a theoretical energy line.

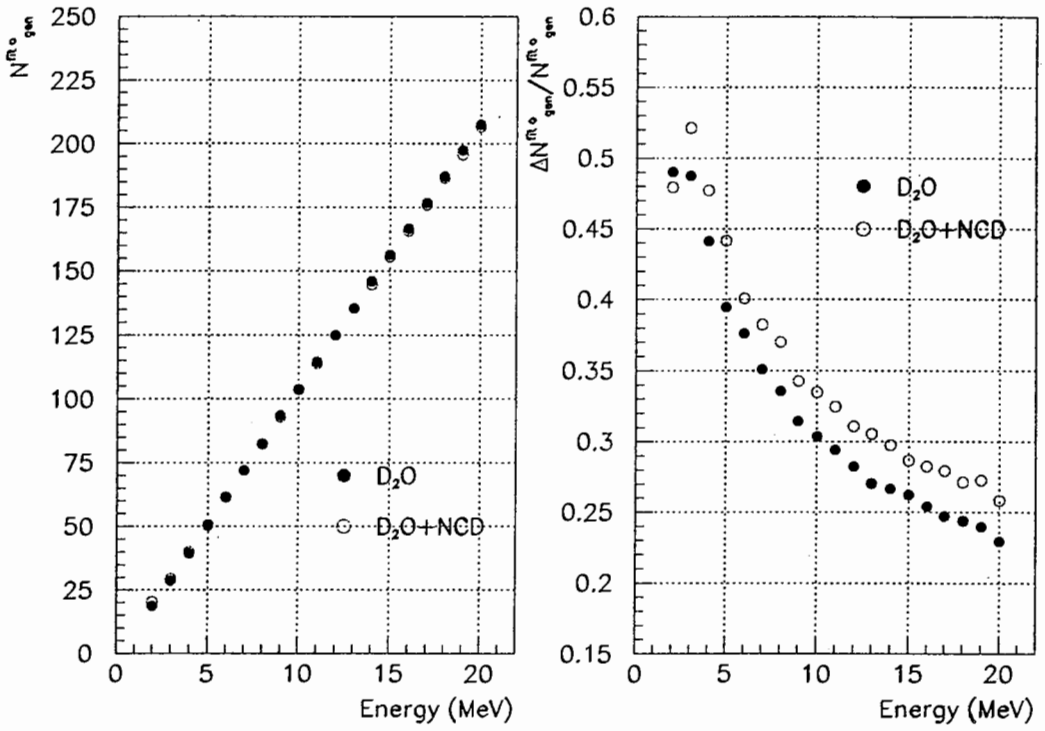


Figure 20: Energy calibration curve and detector resolution using multiple-hit and gain-corrected information. The energy calibration line is now linear, and there is a significant improvement for the energy resolution in the NCD installed scenario.

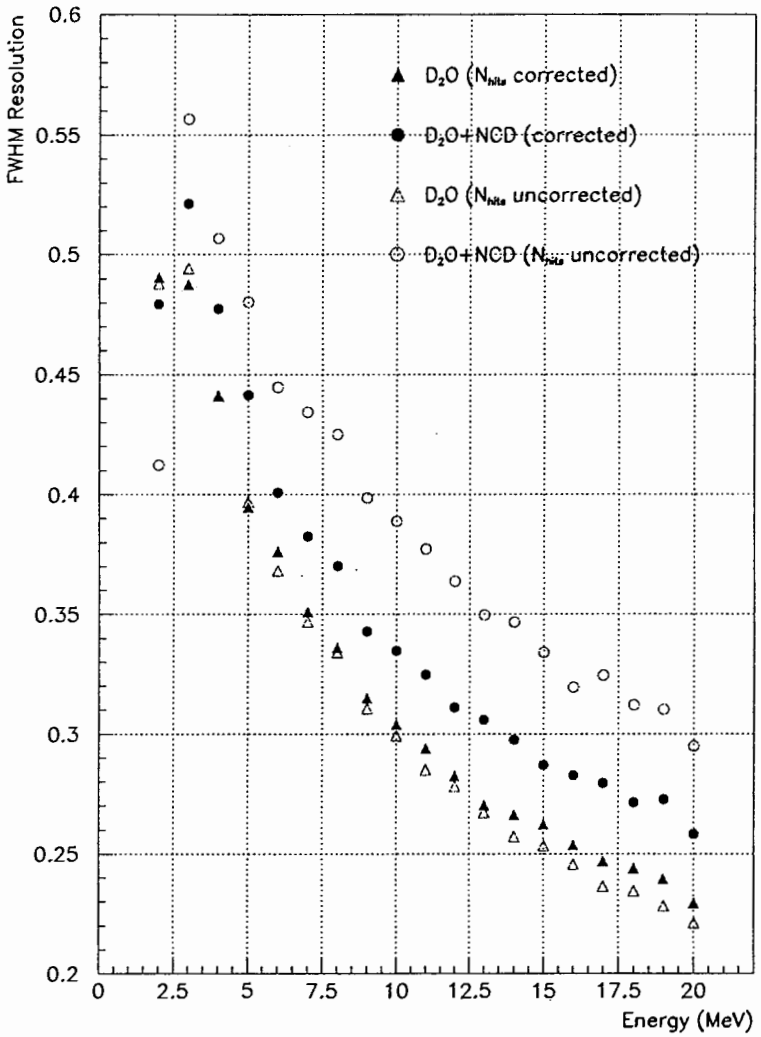


Figure 21: Comparing the energy resolution in the "zeroth order approximation" and the two-tier corrected electron events.

Figure 22: A simple Monte Carlo analysis to understand the first two moments in $N_{hits}$ distribution. In the configuration above, the number of available bins $N_{PMT}$=12, the number of filled bins $N_{hits}$=7, and the total number of detected objects $N_{gen}$=9.

## 6 The $N_{hits}$ Resolution Anomaly

In the last section, we found that the $N_{hits}$ resolution in the "zeroth order" approximation is *better* than the two-tier corrected resolution. This is somewhat counter-intuitive as the two-tier correction scheme linearise the energy response. Therefore, photon counting is done more correctly in the two-tier corrected quantities. We shall try to understand this "anomaly" in this section.

We shall study the results of a very simple Monte Carlo experiment. In this experiment, we have a certain number of bins (number of available PMT $N_{PMT}$) to accept objects (Cĕrenkov photons) randomly thrown at them. Not all of these objects will land inside a bin, and there is a fixed probability (quantum efficiency $p_{qe}$) that an object will land inside a bin. We set $p_{qe}$ to $\frac{1}{3}$. What we want to determine from this experiment is the dependence of the mean and the standard deviation of the number of occupied bins ($N_{hits}$) and the number of detected objects ($N_{gen}$) on the number of available bins $N_{PMT}$. The diagram in Figure 22 shows the configuration of this Monte Carlo experiment. This simple Monte Carlo configuration can be compared to a "ring" of PMTs that lie right at Cĕrenkov cone angle, and these PMTs are ready to accept Cĕrenkov photons in the absence of any light scattering in the $D_2O$.

The distribution of the number of detected objects $N_{gen}$ follows a binomial distribution, with the mean $\langle N_{gen} \rangle$ and the standard deviation $\sigma_{N_{gen}}$ given by

$$\langle N_{gen} \rangle = p_{qe} N^{\circ}_{gen} = \frac{N^{\circ}_{gen}}{3} \tag{8}$$

$$\sigma_{N_{gen}} = \sqrt{N^{\circ}_{gen} p_{qe} (1 - p_{qe})} = \frac{\sqrt{2 N^{\circ}_{gen}}}{3}. \tag{9}$$

In Table 5, we have summarised the values of $\langle N_{gen} \rangle$ and $\sigma_{N_{gen}}$ for the cases of $N^{\circ}_{gen}$ =100, 200 and 300. We have verified that the results of our Monte Carlo experiment agree with these expressions for the $N^{\circ}_{gen}$ cases we considered.

In Figure 23, we show the mean and the standard deviation of $N_{hits}$ as a function of $N_{PMT}$ with different total number of generated objects $N^{\circ}_{gen}$ for this experiment. These values were extracted by fitting the $N_{hits}$ distributions to a normal distribution. In Figure 24, the fractional width of the $N_{hits}$ distribution for different $N^{\circ}_{gen}$ is shown. The non-linearity of $N_{hits}$ is clearly shown in these figures. It can also be seen that these extracted values ($\langle N_{hits} \rangle$, $\sigma_{N_{hits}}$, and $\sigma_{N_{hits}}/\langle N_{hits} \rangle$) are all less than the corresponding theoretical expectations for $N_{gen}$ in Table 5.

To better understand the $N_{hits}$ anomaly, we plotted the two ratios, $\langle N_{hits} \rangle / \langle N_{gen} \rangle$ and $\sigma_{N_{hits}}/\sigma_{N_{gen}}$, against the number of available bins $N_{PMT}$. Given that $\langle N_{gen} \rangle$ is a linear quantity, we want to

| $N^{\circ}_{gen}$ | $\langle N_{gen} \rangle$ | $\sigma_{N_{gen}}$ | $\sigma_{N_{gen}}/\langle N_{gen} \rangle$ |
|---|---|---|---|
| 100 | 33.33 | 4.71 | 0.141 |
| 200 | 66.66 | 6.67 | 0.100 |
| 300 | 100.0 | 8.17 | 0.082 |

Table 5: Expected mean, standard deviation and fractional width for $N_{gen}$ in the simple Monte Carlo experiment to understand the $N_{hits}$ resolution anomaly. We have assume a quantum efficiency $p_{qe}$ of $\frac{1}{3}$.

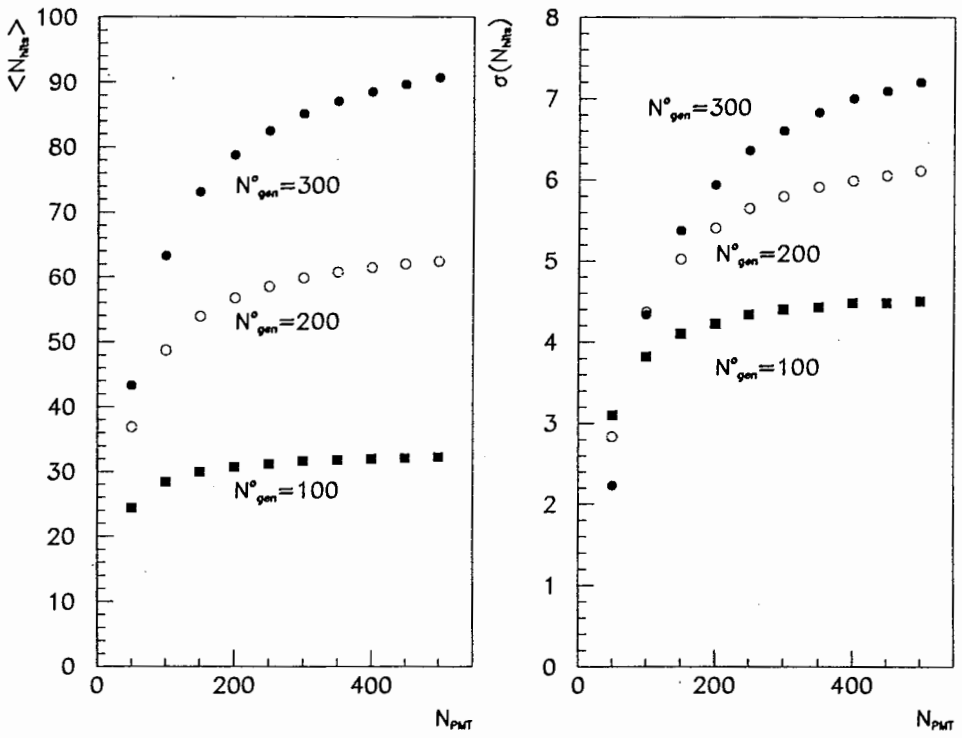


Figure 23: Mean and standard deviation of $N_{hits}$ in the simple Monte Carlo experiment to understand the $N_{hits}$ resolution anomaly.

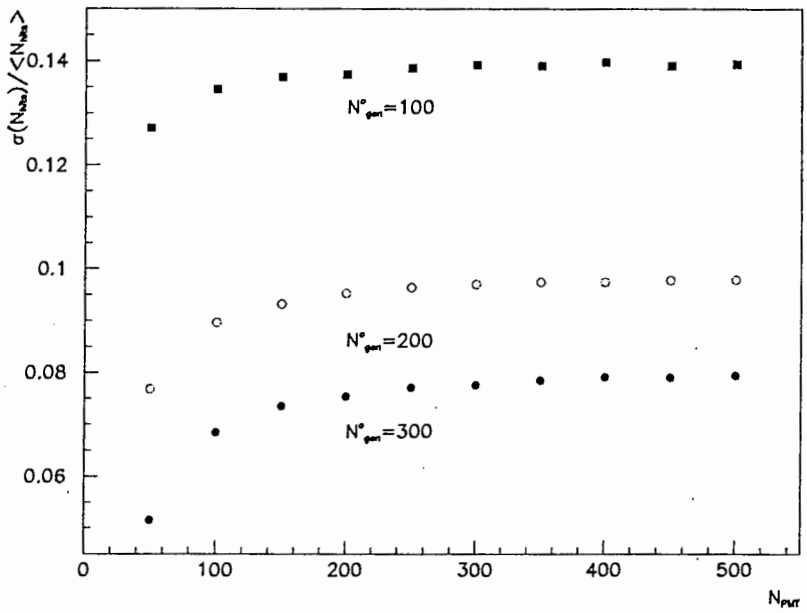


Figure 24: Fractional width of $N_{hits}$ in the simple Monte Carlo experiment to understand the $N_{hits}$ resolution anomaly.

determine the relative deviation from linearity for $\langle N_{hits} \rangle$ and $\sigma_{N_{hits}}$ by inspecting these ratios. This is shown in Figure 25.

It should be clear that as $N_{PMT} \rightarrow \infty$, these ratios should approach unity. Of course, this theoretical limit could not be reached for signals with a strong directionality, such as the Čerenkov light cone, in a finite-size detector. However, if one looks at these ratios closely, it is surprising to learn that even for the case of $N^{\circ}_{gen} = 100$ with $N_{PMT}$=500, i.e. a factor of 5 difference in the number of generated photons to the number of available bins, the ratios $\langle N_{hits} \rangle / \langle N_{gen} \rangle$ and $\sigma_{N_{hits}} / \sigma_{N_{gen}}$ are about 4% and 5% from unity.

Another interesting fact that is obvious from this figure is that the relative deviation from linearity for $\langle N_{hits} \rangle$ and $\sigma_{N_{hits}}$ are different for the same $N_{PMT}$. In fact,

$$\frac{\langle N_{hits} \rangle}{\langle N_{gen} \rangle} > \frac{\sigma_{N_{hits}}}{\sigma_{N_{gen}}} \tag{10}$$

in all the cases we have considered. What this means is that the "resolution" for $N_{hits}$ will be better than $N_{gen}$ even though we have done all the counting correctly in the latter case. This is the primary reason why the $N_{hits}$ resolution is "better" than the two-tier corrected resolution in the $D_2O$ running scenario as we found in the last section. The two-tier correction scheme did improve the resolution in the NCD-installed scenario because the severe degradation in resolution in this scenario is primarily caused by the positional and directional dependence of the detector response. The two-tier correction corrects for these dependencies and significantly improves the resolution.

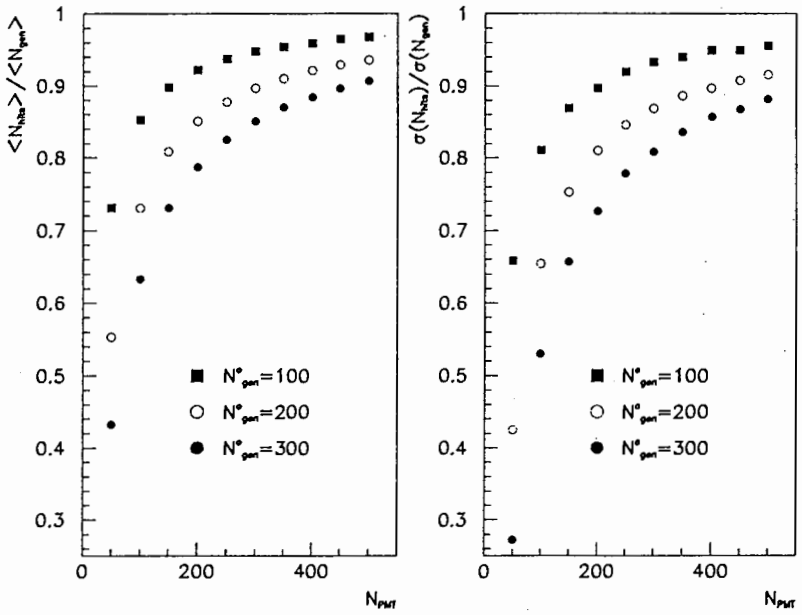


Figure 25: Relative deviation from linearity for $\langle N_{hits}\rangle$ and $\sigma_{N_{hits}}$ in the simple Monte Carlo experiment to understand the $N_{hits}$ resolution anomaly.

To further illustrate this last point, we have plotted the two-tier corrected full-width-at-half-max (FWHM) resolution $R_{2-tier}(E)$, the Monte-Carlo generated $N_{gen}$ FWHM resolution $R_{Ngen}(E)$ and the $N_{hits}$ FWHM resolution $R_{Nhits}(E)$ as a function of energy for isotropic, monoenergetic electrons evenly distributed in a pure $D_2O$ run scenario in Figure 26. When one inspects this plot, one would find that

$$R_{Ngen}(E) > R_{2-tier}(E) > R_{Nhits}(E). \tag{11}$$

Since we have demonstrated in our simple Monte Carlo experiment above that $R_{Ngen}(E) > R_{Nhits}(E)$, this would imply that the two-tier correction scheme does improve the $N_{gen}$ resolution $R_{Ngen}$.

One can conclude from this simple Monte Carlo analysis that even though the $N_{hits}$ resolution appears to be "better" than the $N_{gen}$ resolution, one should be really careful in interpreting its true significance. Whilst $N_{hits}$ has a narrower distribution, it is $N_{gen}$ which represents the true photon counting statistics and contains maximum amount of statistical information. It is the incorrect photon counting statistics inherent to $N_{hits}$ which contributes to the "degradation" in detector resolution in the two-tier corrected resolution $R_{2-tier}$ the pure $D_2O$ run scenario in Figure 21. To further illustrate this last point[2], we can imagine a situation in which there is only one available bin, i.e. $N_{PMT}$=1. In this scenario, $N_{hits}$ is either 0 and 1 regardless what $N_{gen}$ is. Either count in $N_{hits}$ does not contain much information. It is clear that $N_{gen}$ contains more statistical information than $N_{hits}$ in this hypothetical case.

[2]This argument originated from a discussion with S.R. Elliott

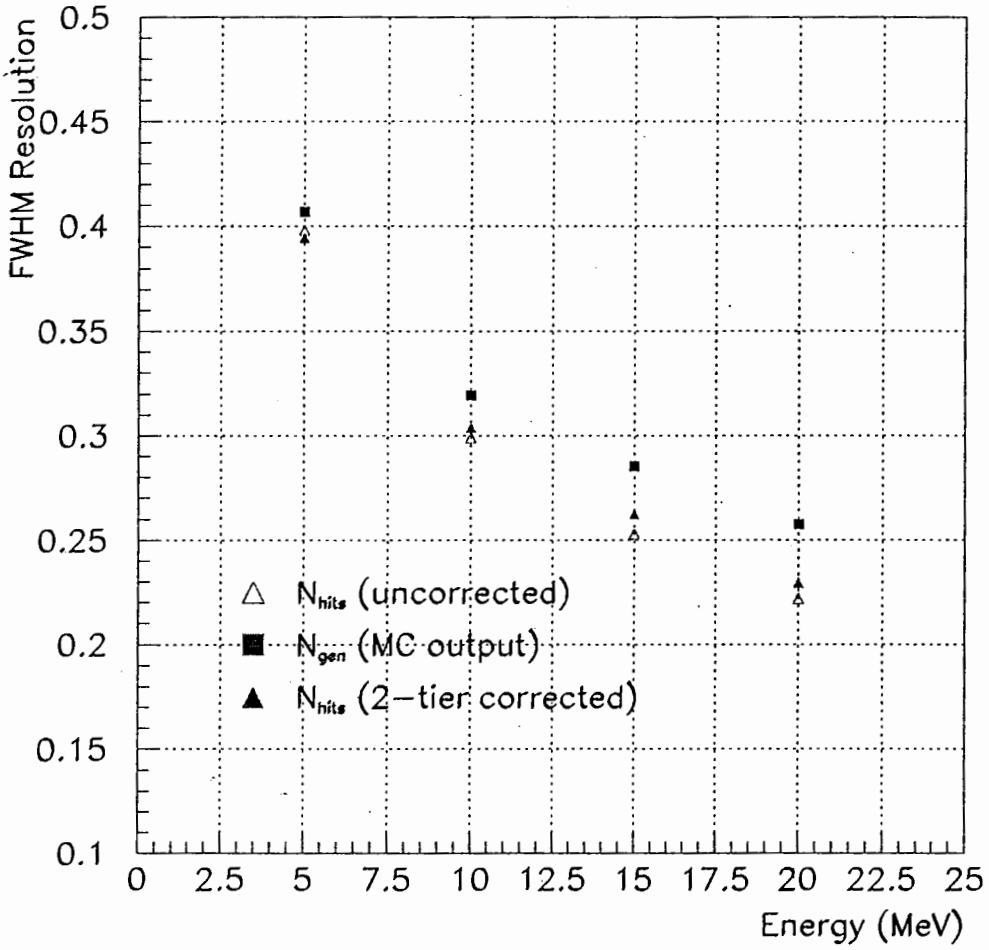


Figure 26: Comparing the two-tier corrected detector resolution to Monte Carlo generated $N_{gen}$ resolution.

# 7 Conclusions

In this report, we have demonstrated that using only the raw $N_{hits}$ information from the data stream to perform the energy calibration may lead to a significant systematic uncertainty. This is primarily a direct consequence of the multiple-hit effect. We have shown that by correcting the $N_{hits}$ signal with the function in Eqn.(4), the energy non-linearity in $N_{hits}$ can be removed.

Previous attempts [1] in gain correcting the energy spectra using the raw $N_{hits}$ information failed because of this multiple-hit effect. We repeated the gain correction analysis using the multiple-hit corrected $N_{hits}$ information, and demonstrated that the gain correction functions extracted from standard candles of different energies are consistent with one another. We further demonstrated that this two-tier correction scheme can reconstruct energy spectra in good agreement with theoretical spectra convoluted with the detector resolution function.

We have demonstrated that the parameter $D_{PMT}$ can be used to correct for the energy and positional dependencies of the multiple-hit effect. We further demonstrated that this parameter can be used in place of $R_{PMT}$ in the gain correction algorithm.

The energy resolution calculated from the two-tier corrected events shows dramatic improvement in the NCD installed scenario. This will reduce the solar neutrino data analysis threshold to ~5.3 MeV in the NCD installed scenario. A previous analysis by Skensved [3] showed that without any $N_{hits}$ signal correction this threshold might be as high as 5.6 MeV. A reduction in the analysis threshold will allow a more sensitive search for charged-current spectral distortion as predicted by the MSW mechanism.

We also performed a simple Monte Carlo experiment trying to understand the $N_{hits}$ resolution anomaly. We found that multiple-hit effect is the origin for this anomaly, in which the $N_{hits}$ resolution appears to be better than the two-tier corrected resolution. It was concluded that $N_{hits}$ *does not represent the true photon counting statistics* and it *does not contain the maximum amount of statistical information.*

# References


[1] A.W.P. Poon, S.R. Elliott, and R.G.H. Robertson, *Energy Response of the SNO Detector in the Presence of the Neutral Current Detector Array*, SNO internal report SNO-STR-97-010, 1997.

[2] M. Lay, *Energy Response and PMT Backgrounds in SNO*, SNO internal report STR-96-015.

[3] T.J. Bowles, P.J. Doe, A. Hime, R.G.H. Robertson, T.C. Spencer, P.M. Thornewell, J.B. Wilhelmy, and J.F. Wilkerson, *Neutral-Current Detection in the Sudbury Neutrino Observatory*, page 60, Los Alamos National Laboratory FIN-94-ER-E324, January 1992.