

# Study of South Pole ice transparency with IceCube flashers

Dmitry Chirkin for the IceCube collaboration

*dima@icecube.wisc.edu*

*University of Wisconsin at Madison, Madison, USA*

## Abstract

The IceCube observatory,  $1 \text{ km}^3$  in size, is now complete with 86 strings deployed in the antarctic ice. IceCube detects the Cherenkov radiation emitted by charged particles passing through or created in the ice. To realize the full potential of the detector the properties of light propagation in the ice in and around the detector must thus be known to the best achievable precision. This report presents a new method of fitting the ice model to a data set of in-situ light source events collected with IceCube. The resulting set of derived ice parameters is presented and a comparison of IceCube data with simulation based on the new model is shown.

## 1 Introduction

IceCube [1] is a kilometer-scale high energy neutrino telescope built at the geographic South Pole. The primary goal is to elucidate the mechanisms for production of high energy cosmic rays by detecting high energy neutrinos from astrophysical sources. IceCube uses the 2.8 km thick glacial ice sheet as a Cherenkov radiator for charged particles created when neutrinos collide with subatomic particles in the ice or nearby rock. Neutrino interactions can create high energy muons, electrons or tau particles, which must be distinguished from downgoing background muons based on the pattern of emitted light. The Cherenkov light from these particles is detected by an embedded array of 5160 optical sensors, most of which are deployed at depths of 1450 - 2450 m in 17 m intervals along 78 vertical cable strings, which are arranged in a triangular lattice with a horizontal spacing of approximately 125 m. The remaining 480 sensors are deployed in a more compact geometry constituting the "Deep Core".

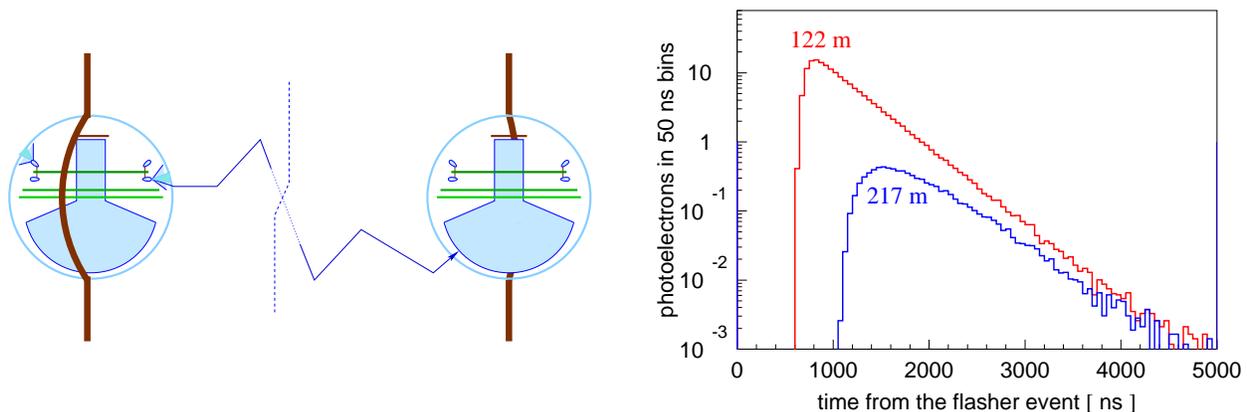


Figure 1: Left: simplified schematics of the experimental setup: the flashing sensor on the left emits photons, which propagate through ice and are detected by a receiving sensor on the right. Right: example photon arrival time distributions at a sensor on one of the nearest strings (122 m away), and on one of the next-to-nearest strings (217 m away). The goal of this work is to find the best-fit set of ice parameters, which describe these distributions as observed in data simultaneously for all pairs of emitters and receivers.

The properties of light propagation in a transparent medium can be described in terms of the average distance between successive scatters and the average distance to absorption (local scattering and absorption lengths), as well as the angular distribution of the new direction of a photon relative to old at a given scattering point. These details are used in both the simulation and reconstruction of IceCube data, thus they must be known to the best possible precision. This work presents a new, *direct fit* approach to fitting the ice properties. A global fit is performed to a set of data with in-situ light sources covering all depths of the detector, resulting in a single set of scattering and absorption parameters of ice, which describes these data best (see Figure 1). This is different from the approach of [2], where separate fits were performed to

individual pairs of emitters and receivers in the AMANDA detector<sup>1</sup>, each resulting in a measure of average properties of the surrounding ice; these were then averaged among pairs near each other to result in a table of ice parameters.

This paper introduces the data set in section 2 and simulation in section 4. The ice surrounding the detector is modeled according to the parametrization described in section 3. Data and simulation are compared with the likelihood function as discussed in sections 5 and 6. Section 7 explains how the search for the best solution was performed. Section 8 compares the result with an independent measurement [3] of the dust concentration in ice. Finally, section 9 discusses the uncertainties of the measurement (including a toy simulation study), and section 10 summarizes the result.

## 2 Flasher dataset

In 2008 IceCube consisted of 40 strings as shown in Figure 2, each containing 60 optical sensors, or digital optical modules (DOMs). Each of the DOMs consists of a 10" diameter photomultiplier tube (PMT) [4] and several electronics boards enclosed in a glass container [1]. One of the boards is the "flasher board", which has 6 horizontal and 6 tilted LEDs, each capable of emitting  $\sim 7.5 \cdot 10^9$  photons at  $\sim 405 \pm 5$  nm in a 62 ns-wide pulse.

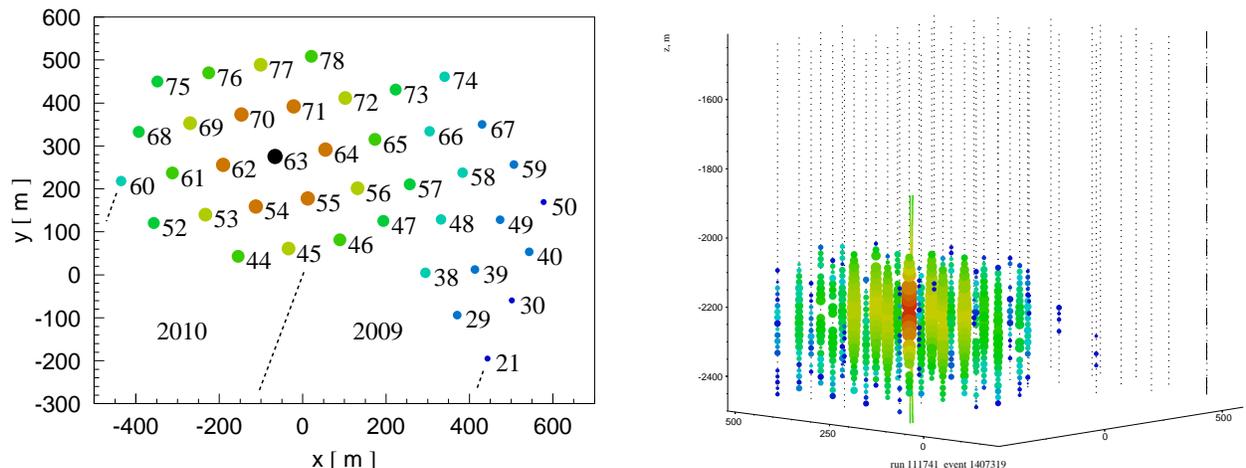


Figure 2: Left: IceCube 40-string configuration as operated in 2008. String 63 (of DOMs that were used as flashers) is shown in black; nearest 6 strings in brown. IceCube parts added in the following years lie in regions indicated approximately with dashed lines. Right: a typical flasher event, with DOM 46 on string 63 flashing. The circles are larger for DOMs that recorded greater number of photons, and the arrival time of the earliest photon in each DOM is indicated with color: from early in red to late in blue.

The PMT output signal is digitized into "waveforms" using the faster, ATWD, and slower, fADC, sampling chips [5]. The ATWD is configured to collect 128 samples with 3.3 ns sampling rate, and the fADC records 256 samples with 25 ns sampling rate. The DOMs transmit time-stamped digitized PMT signal waveforms to computers at the surface.

In a series of several special-purpose runs, IceCube took data with each of DOMs 1 – 60<sup>2</sup> of string 63 flashing in a sequence. For each of the flashing DOMs at least 250 flasher events were collected and used in this analysis. All 6 horizontal LEDs were used simultaneously at maximum brightness and pulse width settings, creating a pattern of light around string 63 that is approximately azimuthally symmetric.

As seen in Figure 3 there is a substantial variation between the charges collected in same-position DOMs on the six surrounding strings. This variation is mostly due to variations in relative orientation of the flasher LEDs with respect to the surrounding strings, relative variation of light yield between the different flasher LEDs, and some differences in distance to and depth of the six surrounding strings. The amount of variation due to these effects can be quantified with the RMS of the deviation from the mean between the six surrounding strings, shown in Figure 4. Such an estimate is conservative as a measure of our understanding of data since some of the variation can be and is correctly simulated. The

<sup>1</sup>A predecessor to IceCube, also built in the South Pole ice.

<sup>2</sup>DOMs are numbered with consecutive integers from 1 to 60 (indicating their position on a string), going from the top of each string down.

irreducible/unsimulated uncertainties include  $\sim 5 - 30$  degrees in the flasher board orientation, up to  $\sim 30\%$  in the absolute LED light output,  $\sim 1$  m in the DOM coordinates.

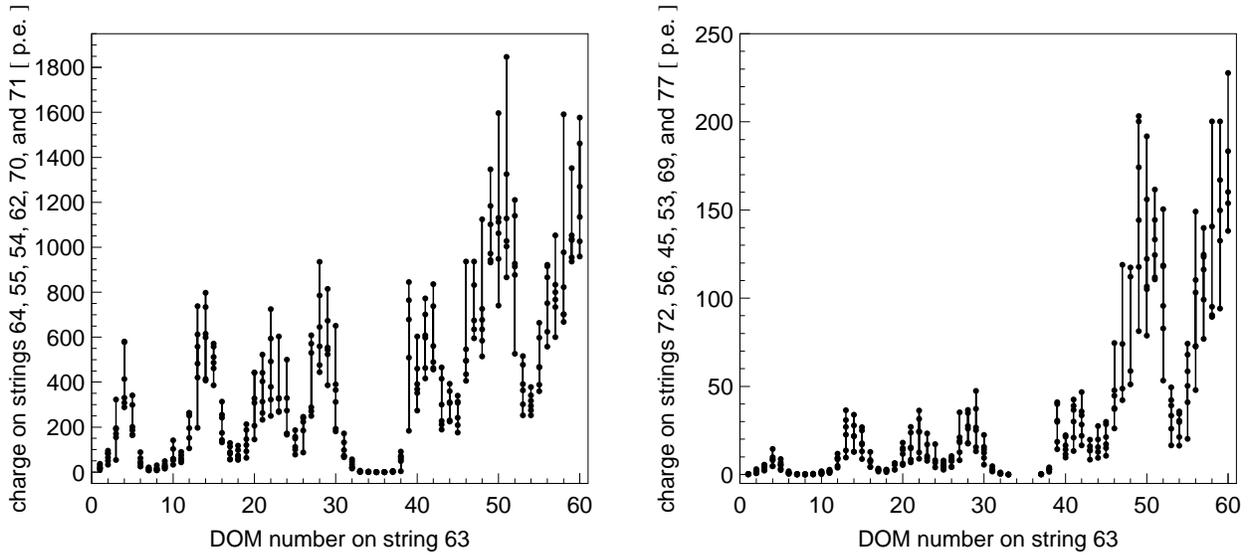


Figure 3: Charges collected by DOMs on the six nearest strings (121.8 – 126.6 m away, left) and six next-to-nearest strings (211.4 – 217.9 m away, right), observed when flashing at the same position on string 63.

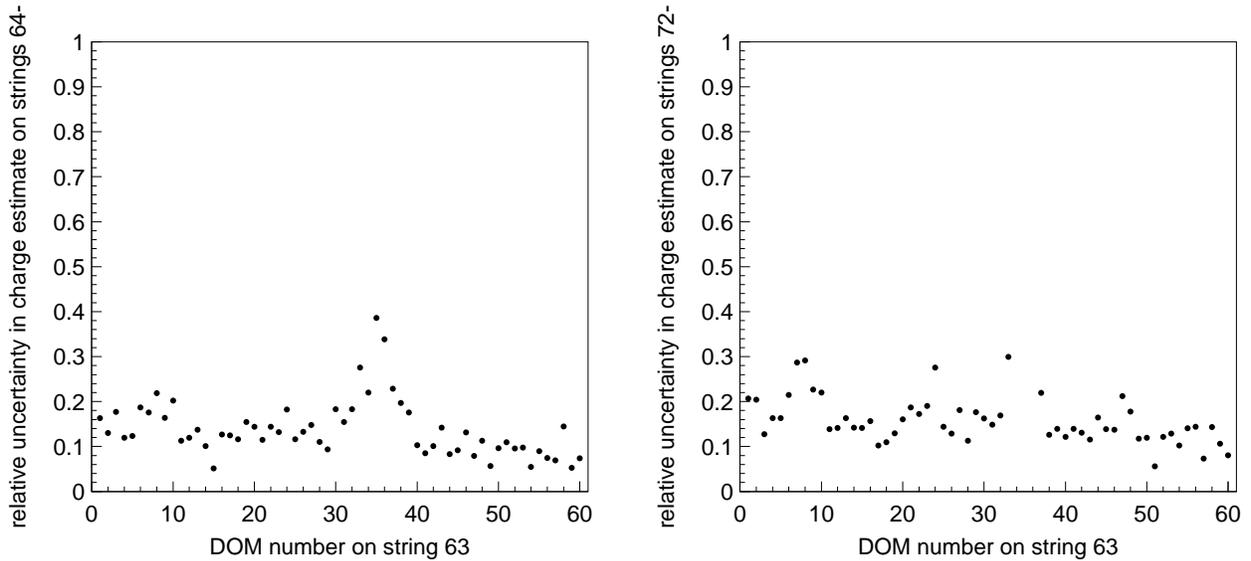


Figure 4: Relative uncertainty in the mean charge estimated from measured charges on the six nearest strings (left) and six next-to-nearest strings (right) observed when flashing at the same position on string 63.

The pulses corresponding to the arriving photons were extracted from the digitized waveforms and binned in 25 ns bins, from 0 to 5000 ns from the start of the flasher pulse (extracted from the special-purpose ATWD channel of the flashing DOM). To reduce the contribution from saturated DOMs (most of which were on string 63 near the flashing DOM) [4] and to minimize the effects of the systematic uncertainty in the simulated angular sensitivity model (of a DOM) the photon data collected on string 63 was not used in the fit.

### 3 Six-parameter ice model

This section overviews the so-called six-parameter ice model introduced in [2]. The ice is described by a table of parameters  $b_e(400)$ ,  $a_{\text{dust}}(400)$ , related to scattering and absorption at a wavelength of 400 nm, and temperature  $\delta\tau$ , given for each ice layer, and by the six parameters (fitted in [2] to AMANDA calibration data):  $\alpha$ ,  $\kappa$ , A, B, D, and E. The width of the ice layers (10 m) was chosen to be as small as possible while maintaining at least one receiving DOM in each layer. Coincidentally it is the same as the value chosen in [2].

The geometrical scattering coefficient  $b$  determines the average distance between successive scatters (as  $1/b$ ). It is often more convenient to quote the effective scattering coefficient,  $b_e = b \cdot (1 - \langle \cos \theta \rangle)$ , where  $\theta$  is the deflection angle at each scatter. The absorption coefficient  $a$  determines the average distance traveled by photon before it is absorbed (as  $1/a$ ).

The wavelength dependence of scattering and absorption coefficients within the six-parameter ice model is given by the following expressions (for wavelength  $\lambda$  in nm):

$$b_e(\lambda) = \frac{1}{\lambda_e} = b_e(400) \cdot \left(\frac{\lambda}{400}\right)^{-\alpha}$$

$$a(\lambda) = \frac{1}{\lambda_a} = a_{\text{dust}}(\lambda) + A e^{-B/\lambda} \cdot (1 + 0.01 \cdot \delta\tau), \quad \text{with} \quad a_{\text{dust}}(\lambda) = a_{\text{dust}}(400) \cdot \left(\frac{\lambda}{400}\right)^{-\kappa}.$$

The scattering and part of absorption are due to impurities (dust) embedded in the ice. The other part of absorption is due to the contribution from the pure ice, which has a mild dependence on temperature, and, thus, depth. The temperature  $T$  [K] vs. depth  $d$  [m] was parametrized in [6] as

$$T = 221.5 - 0.00045319 \cdot d + 5.822 \cdot 10^{-6} \cdot d^2, \quad \delta\tau(d) = T(d) - T(1730 \text{ m})^3.$$

The remaining two parameters  $D$  and  $E$  of the six-parameter ice model were defined in [2] in a relationship establishing a correlation  $a_{\text{dust}}(400) \cdot 400^\kappa \sim D \cdot b_e(400) + E$ , and were not used here.

This work presents the measurement of the values of  $b_e(405)$  and  $a(405)$  and relies on the six-parameter ice model to extrapolate scattering and absorption for wavelengths other than 405 nm. A number of DOMs containing flasher LEDs with four additional central wavelengths of the emitted light (340, 370, 450, and 505 nm) have been installed during the last IceCube deployment season. The future analysis of data containing light from these LEDs will verify or improve the parametrization of the wavelength dependence built into the 6-parameter model.

### 4 Simulation

Detector response to flashing each of the 60 DOMs on string 63 needs to be simulated very quickly, so that simulations based on many different sets of coefficients  $b_e(400)$  and  $a_{\text{dust}}(400)$  could be compared to the data.

A program called PPC (photon propagation code, see appendix A), was written for this purpose. It propagates photons through ice described by a selected set of parameters  $b_e(400)$  and  $a_{\text{dust}}(400)$  until they hit a DOM or get absorbed. No special weighting scheme was employed, except that the DOMs were scaled up in size (a factor 5 to 16, depending on the required timing precision<sup>4</sup>), and the number of emitted photons was scaled down by a corresponding factor ( $5^2 - 16^2$ ).

The angular sensitivity of the IceCube optical module was modeled according to the "hole ice" description<sup>5</sup> of [7], which is shown in Figure 5. The DOM acceptance (including the glass and gel transmission, and PMT quantum and collection efficiencies) was calculated according to [4] for a DOM of radius 16.51 cm. At 405 nm (flasher center wavelength) the DOM acceptance is 13.15%. The Cerenkov photons were sampled from the distribution shown in the right plot of Figure 5, which is a convolution of the DOM acceptance curve with the Cerenkov photon spectrum given by the Frank-Tamm formula:

$$\frac{dN}{d\lambda dl} = \frac{2\pi\alpha}{\lambda^2} \sin^2 \theta_c.$$

<sup>3</sup>1730 m is the depth of the center of AMANDA.

<sup>4</sup>Special care was taken to minimize biasing photon arrival times by over-sizing DOMs only in the direction perpendicular to the photon direction. Still, in the worst case a factor of 16 compared to the nominal DOM size (i.e., factor 1) introduces the maximum error of  $(16 - 1) \cdot 16.51 \text{ cm} / 22 \text{ cm/ns} = 11.3 \text{ ns}$  in the arrival time (for a DOM with radius 16.51 cm and for speed of light in ice of 22 cm/ns). An additional consideration is a small loss of DOM hit occupancy, which may occur for larger factors. However, this is largely corrected by allowing the photon to continue propagating even after it hits an oversized DOM.

<sup>5</sup>Takes into account an increased amount of scattering (with scattering length of 50 cm) in the column of ice immediately surrounding an IceCube string via a modification to the angular sensitivity curve.

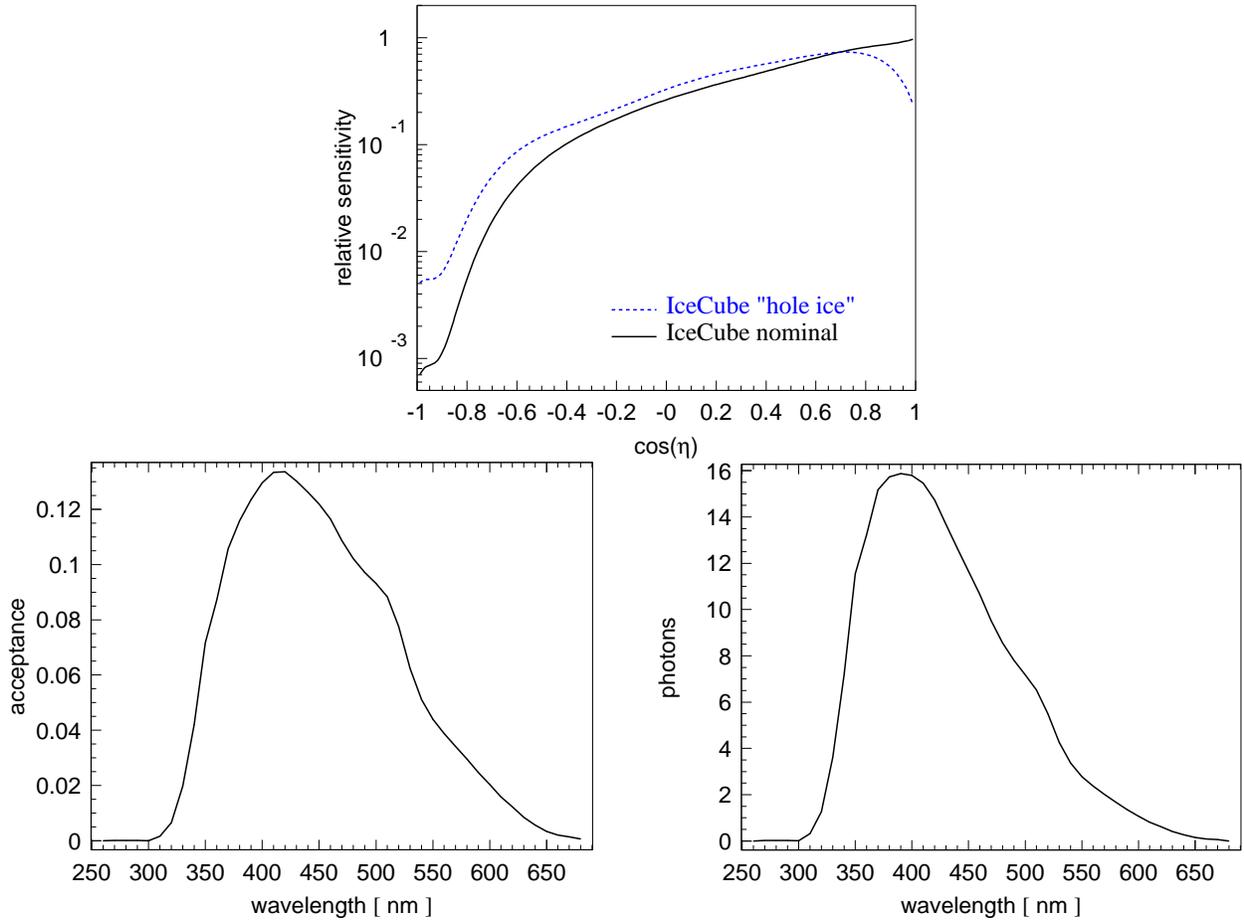


Figure 5: Top: angular sensitivity of an IceCube optical module;  $\eta$  is the photon arrival angle with respect to the PMC axis. The nominal model, based on a lab measurement, is normalized to 1.0 at  $\cos \eta = 1$ . The area under both curves is the same. Bottom left: optical module acceptance: fraction of photons arriving from a direction parallel to the PMT axis (at  $\cos \eta = 1$ ) that are recorded. Bottom right: number of Cerenkov photons (in 10 nm bins) emitted by one meter of the *bare* muon track (i.e., muon without secondary cascades), convolved with the optical module acceptance. The integral under this curve is 2450 photons.

The muon light production is treated via the use of the "effective length"  $dl$ , as described in the appendix B. The phase refractive index  $n_p$  used in the formula above (defining the Cerenkov angle  $\cos \theta_c = 1/n_p$ ) and the group refractive index  $n_g$  (used in calculation of the speed of light in medium) were estimated according to formulae from [8]:

$$n_p = 1.55749 - 1.57988 \cdot \lambda + 3.99993 \cdot \lambda^2 - 4.68271 \cdot \lambda^3 + 2.09354 \cdot \lambda^4$$

$$n_g = n_p \cdot (1 + 0.227106 - 0.954648 \cdot \lambda + 1.42568 \cdot \lambda^2 - 0.711832 \cdot \lambda^3).$$

The distribution of the photon scattering angle  $\theta$  is modeled by a linear combination of two functions commonly used to approximate scattering on impurities:  $(1 - f_{SL}) \cdot \text{HG} + f_{SL} \cdot \text{SL}$ . The first is the Henyey-Greenstein (HG) function [2]:

$$p(\cos \theta) = \frac{1}{2} \frac{1 - g^2}{[1 + g^2 - 2g \cdot \cos \theta]^{3/2}}, \quad \text{with } g = \langle \cos \theta \rangle,$$

which can be analytically integrated and inverted to yield a value of  $\cos \theta$  as a function of a random number  $\xi$  uniformly distributed on  $[0; 1]$ :

$$\cos \theta = \frac{1}{2g} \left( 1 + g^2 - \left( \frac{1 - g^2}{1 + gs} \right)^2 \right), \quad s = 2 \cdot \xi - 1.$$

The second is the simplified Liu (SL) scattering function [9]:

$$p(\cos \theta) \sim (1 + \cos \theta)^\alpha, \quad \text{with} \quad \alpha = \frac{2g}{1-g}, \quad g = \langle \cos \theta \rangle,$$

which also yields a simple expression for  $\cos \theta$  as a function of a random number  $\xi \in [0; 1]$ :

$$\cos \theta = 2 \cdot \xi^\beta - 1, \quad \text{with} \quad \beta = \frac{1-g}{1+g}.$$

Figure 6 compares these two functions with the prediction of the Mie theory with dust concentrations and radii distributions taken as described in [2]. The photon arriving timing distributions are substantially affected by the "shape" parameter  $f_{\text{SL}}$  (as shown in Figure 7), making it possible to determine this parameter from fits to data.

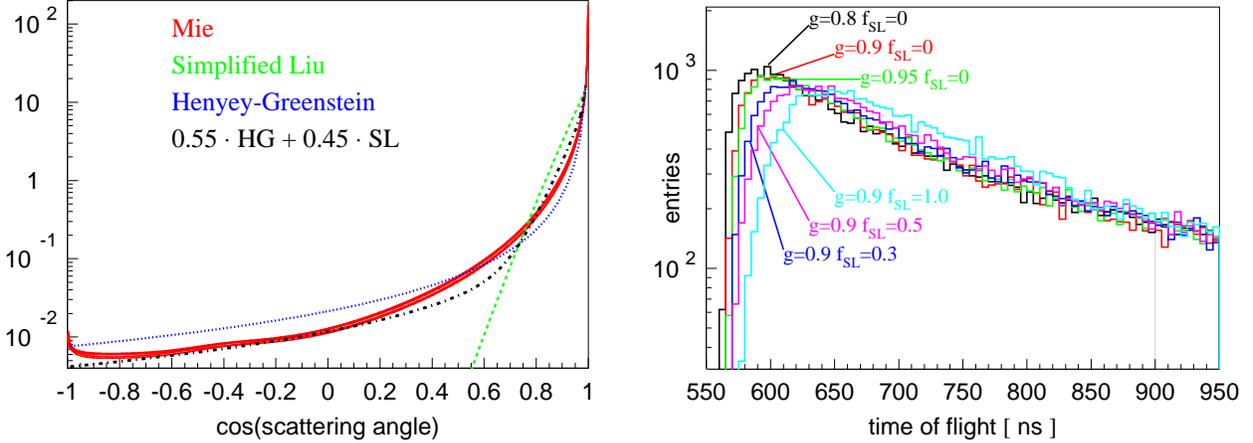


Figure 6: (left) Comparison of the Mie scattering profiles calculated at several depths of the South Pole ice with the Henyey-Greenstein (HG) [2] and simplified Liu (SL) [9] scattering functions, all with the same  $g = 0.943$ .

Figure 7: (right) Photon arriving time distributions at a DOM 125 m away from the flasher, simulated for several values of  $g = \langle \cos \theta \rangle$  and  $f_{\text{SL}}$ . The difference in peak position simulated with  $g = 0.8$  and  $g = 0.9$  is of the same order ( $\sim 10$  ns) as that between sets simulated with different values of the shape parameter  $f_{\text{SL}}$ .

The value of  $g = 0.9$  was used in this work (cf.  $g = 0.8$  in [2]). Higher values (as high as  $\sim 0.94$  [2, 10]) are predicted by the Mie scattering theory, however, these result in slower simulation, while yielding almost unchanged values of the effective scattering  $b_e$  and absorption  $a$  coefficients.

## 5 Likelihood description

Consider the amount of charge received by DOM  $i$  in time bin  $n$  when flashing DOM  $k$ . It is measured by taking data with a total photon count of  $d$  in  $n_d$  flasher events and a per-event expectation of  $\mu_d$ , and predicted by the simulation with a total photon count of  $s$  in  $n_s$  simulated events and a per-event expectation of  $\mu_s$ . Naively one expects the best approximations to  $\mu_d$  and  $\mu_s$  from data and simulated events to be  $\mu_d = d/n_d$ , and  $\mu_s = s/n_s$ .

The systematic error in describing data with simulation (i.e., describing  $\mu_d$  with  $\mu_s$ ) is  $\sigma \approx 10 - 20\%$  (estimated in section 2). One quantifies the amount of disagreement between data and simulation in the presence of such an error with a  $\chi^2_{i,n,k}$  (omitting the indices  $i$ ,  $n$ , and  $k$ ):

$$\chi^2 = \frac{(\log \mu_d - \log \mu_s)^2}{\sigma^2}.$$

The uncertainty due to this systematic error can be modeled with a probability distribution function

$$\frac{1}{\sqrt{2\pi}\sigma} \exp \frac{-(\log \mu_d - \log \mu_s)^2}{2\sigma^2}.$$

Given that  $\mu_d$  and  $\mu_s$  are not known, and the measured values are  $d$  and  $s$ , one formulates the likelihood function that describes counts measured in both data and simulation as

$$\frac{(\mu_s n_s)^s}{s!} e^{-\mu_s n_s} \cdot \frac{(\mu_d n_d)^d}{d!} e^{-\mu_d n_d} \cdot \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{-(\log \mu_d - \log \mu_s)^2}{2\sigma^2}.$$

Taking the log with a minus sign, this becomes:

$$\ln s! + \mu_s n_s - s \log(\mu_s n_s) + \ln d! + \mu_d n_d - d \log(\mu_d n_d) + \frac{1}{2\sigma^2} \log^2 \frac{\mu_d}{\mu_s} + \log(\sqrt{2\pi}\sigma) \equiv F.$$

The function  $F(\mu_s, \mu_d)$  can be easily minimized against  $\mu_s$  and  $\mu_d$ , yielding estimates of these quantities. To demonstrate this, first the derivatives of  $F$  are calculated and set to 0:

$$\mu_s \frac{\partial F}{\partial \mu_s} = \mu_s n_s - s - \frac{1}{\sigma^2} \log \frac{\mu_d}{\mu_s} = 0,$$

$$\mu_d \frac{\partial F}{\partial \mu_d} = \mu_d n_d - d + \frac{1}{\sigma^2} \log \frac{\mu_d}{\mu_s} = 0.$$

The sum of these ( $\mu_s n_s + \mu_d n_d = s + d$ ) yields an expression of  $\mu_d$  as a function of  $\mu_s$ . Plugging it into the first of the above two equations one gets

$$f = \mu_s \frac{\partial F}{\partial \mu_s}(\mu_s, \mu_d(\mu_s)) = \mu_s n_s - s - \frac{1}{\sigma^2} \log \frac{\mu_d(\mu_s)}{\mu_s} = 0.$$

This equation can be solved with a few iterations of the Newton's root finding method starting with a solution to

$$\mu_s = \mu_d(\mu_s): \quad \mu_s = \mu_d = \frac{s + d}{n_s + n_d}.$$

At each iteration the value of  $\mu_s$  is adjusted by  $-f/f'$ , where the derivative is evaluated as

$$f' = n_s \left( 1 + \frac{1}{\sigma^2} \left( \frac{1}{\mu_s n_s} + \frac{1}{\mu_d n_d} \right) \right).$$

Once the likelihood function is solved for the best values of  $\mu_s$  and  $\mu_d$ , these can be plugged into the  $\chi^2_{i,n,k}$  above. One can now write the complete  $\chi^2$  function (adding the regularization terms  $R_j$  described in the next section) as a sum over all DOMs  $i$  and time bins  $n$ , when flashing DOMs  $k$ :

$$\chi^2 = \sum_{i,n,k} \frac{(\log \mu_d - \log \mu_s)^2}{\sigma^2} + \sum_{1,2} \alpha_j R_j.$$

## 6 Regularization terms

Two regularization terms are added to the likelihood function described in the previous section. The first one is used to control the unchecked fluctuations of scattering and absorption coefficients with depth in under-constrained ice layers and is formed of terms that are numerical expressions for second derivatives of scattering and absorption with respect to the position of the ice layer:

$$R_r = \sum_{i=2}^{N-1} \left[ (\log b_e[i-1] - 2 \cdot \log b_e[i] + \log b_e[i+1])^2 + (\log a_{\text{dust}}[i-1] - 2 \cdot \log a_{\text{dust}}[i] + \log a_{\text{dust}}[i+1])^2 \right].$$

Here  $N$  is the number of ice layers in which  $b_e$  and  $a_{\text{dust}}$  are defined.

The second term is used to smooth the fluctuations in the diagram of  $a_{\text{dust}}$  vs.  $b_e$  (enforcing the notion that both are proportional to the dust concentration). It is constructed as an excess of the sum of distances between the consecutive points  $(\log b_e, \log a_{\text{dust}})$  over the shortest distance connecting the end points:

$$R_u = -D(1, N) + \sum_{j=1}^{N-1} D(j, j+1),$$

$$\text{where } D(j_1, j_2) = \sqrt{(\log b_e[j_1] - \log b_e[j_2])^2 + (\log a_{\text{dust}}[j_1] - \log a_{\text{dust}}[j_2])^2}.$$

The points  $(\log b_e, \log a_{\text{dust}})$  are sorted by the value of  $\log b_e + \log a_{\text{dust}}$  and shown in the above sum with the index  $j[i]$ .

Both of these terms affected the solution by less than  $\sim 2\%$  at detector depths at their chosen strengths  $\alpha_{r,u}$ , which was verified by re-running the fits without them. The regularization terms would likely become more important if the width of ice layers (10 m in this work) were chosen to be much smaller than the spacing between DOMs on a string (17 m).

## 7 Fitting the flasher data

The six horizontal LEDs within a single DOM flashing at maximum brightness and width emit  $\sim 4.5 \cdot 10^{10} \pm 30\%$  photons [1] (at room temperature in the lab, without the surrounding DOM glass sphere). Only 13.15% of these, i.e.,  $5.9 \cdot 10^9$ , remain after accounting for the DOM acceptance (as explained in section 4). Using a DOM size scaling factor of 16 only  $2.3 \cdot 10^7$  photons need to be simulated ( $16^2 = 256$  times fewer).

A number of  $9765625 \approx 10^7$  of photons simulated with a scaling factor of 16 corresponds to  $2.5 \cdot 10^9$  photons without scaling (i.e., DOM size scaling factor of 1.0), or  $1.9 \cdot 10^{10}$  real photons leaving the flasher DOM (after accounting for the receiving DOM acceptance). This is a "unit bunch" of photons, which is simulated in  $\sim 1$  second on a single GPU (see appendix A).

In the following a "photon yield factor"  $p_y$  is the number of unit bunches that correspond to a given number of real photons. E.g.,  $4.5 \cdot 10^{10}$  photons emitted by a flasher board correspond to a photon yield factor of  $p_y = 2.37$ . Additional considerations (such as partial shadowing of the DOM surface by the supporting cables) lower this estimate by 10% to  $2.12 \pm 0.66$ . This represents an upper limit on the photon yield factor since a fraction of photons ( $\sim 6\%$ ) is absorbed by the glass sphere or reflected back into the flasher DOM.

Data from all pairs of emitter-receiver DOMs (located in the same or different ice layers, altogether  $\sim 38700$  pairs) contributed to the fit to  $\sim 200$  ice parameters (scattering and absorption in 10 m layers at detector depths of 1450 to 2450 m). Two  $\chi^2$  functions were used in fitting the data: the  $\chi_q^2$  constructed with one term from each emitter-receiver pair (using the total recorded charge), and the  $\chi_t^2$  constructed with recorded charge split in 25 ns bins. Although  $\chi_t^2$  used the available information more fully,  $\chi_q^2$  turned out to be somewhat more robust with respect to fluctuations (between the simulated sets) and also faster to compute. Thus,  $\chi_q^2$  was used in an initial search for a solution, with  $\chi_t^2$  applied in the final fits.

Both  $b_e(400)$  and  $a_{\text{dust}}(400)$  are roughly proportional to the concentration of dust (this would be precise if the dust composition in the ice were the same at all depths). This motivates the following simplification in the initial search for the minimum of  $\chi_q^2$ : in each layer both  $b_e(400)$  and  $a_{\text{dust}}(400)$  are scaled up or down by the same relative amount (in the range 1 – 40%), preserving their ratio to each other.

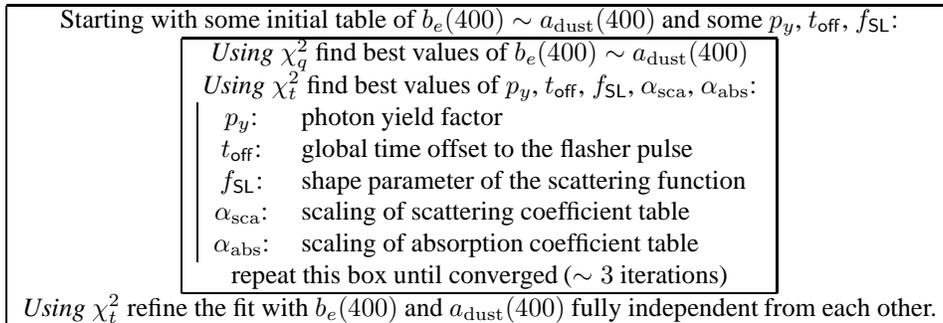


Table 1: Flow chart of the global fit procedure to ice/flasher parameters.

Starting with the homogeneous ice described with  $b_e(400) = 0.042 \text{ m}^{-1}$  and  $a_{\text{dust}}(400) = 8.0 \text{ km}^{-1}$  (average of [2] at detector depths) the minimum of  $\chi_q^2$  is found in  $\sim 20$  steps. At each iteration step the values of  $b_e(400)$  and  $a_{\text{dust}}(400)$  are varied in consecutive ice layers, one layer at a time. Five flashing DOMs closest to the layer, which properties are varied, are used to estimate the variation of the  $\chi^2$ . Figure 8 shows ice properties after each of 20 steps of the minimizer.

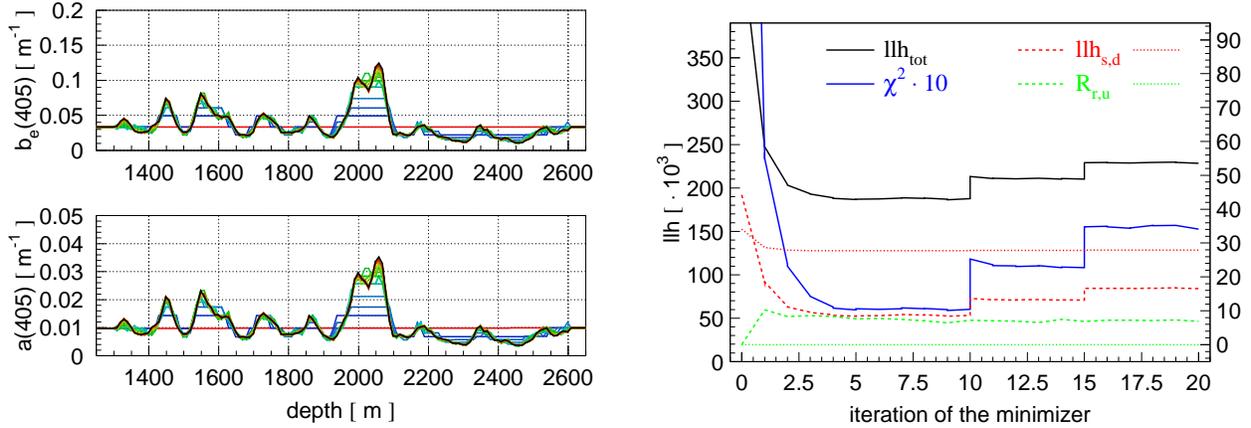


Figure 8: Left: values of  $b_e(405)$  and  $a(405)$  vs. depth after 20 steps of the minimizer. The black curve shows fitted values after the last step of the minimizer. Right:  $\chi^2$  values achieved after each step of the minimizer. The starting "homogeneous ice" value is  $1.34 \cdot 10^5$ . Regularization terms  $R_{r,u}$  use the scale on the right. Also shown are the Poisson terms for simulation and data ( $\text{llh}_{\text{s,d}}$ ) and the full likelihood including all terms ( $\text{llh}_{\text{tot}}$ ). The  $\chi^2$  may jump when the number of simulated flasher events is increased, but then goes down as the minimizer steps through the iterations. For the iteration steps 1-10: 1 flasher event is simulated; for 11-15: 4; for 16-20: 10.

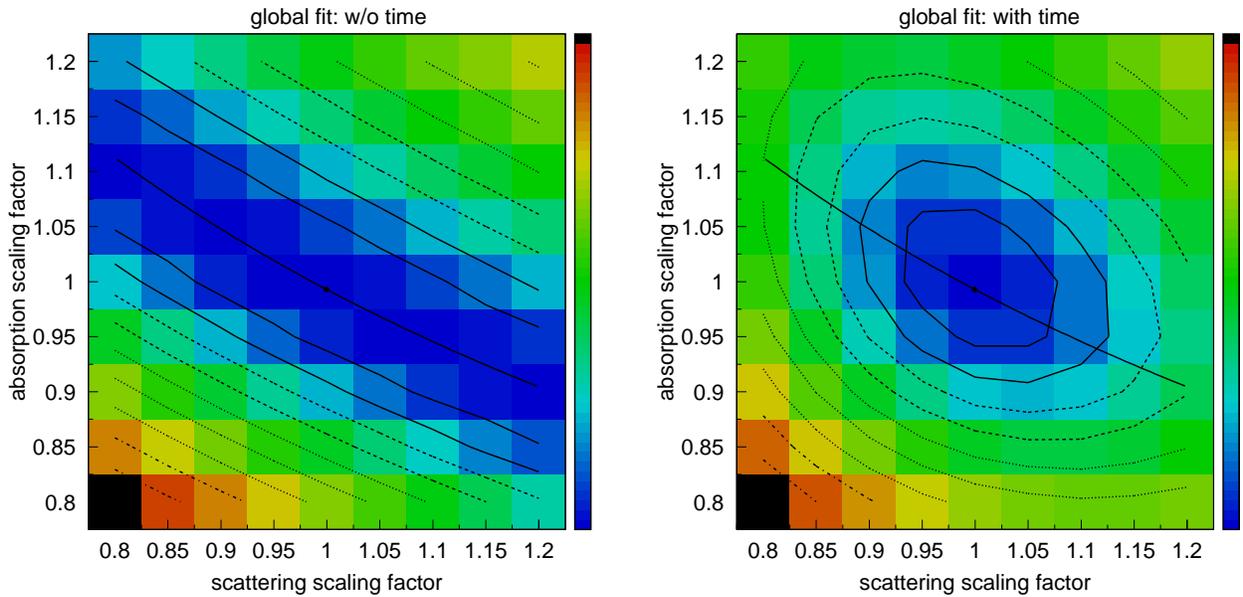


Figure 9: Likelihood functions in the vicinity of their minima: constructed using only charge information (left), and using full timing information (right). The values are shown on a log scale (with colors and contours). The ranges of values shown are:  $\chi_q^2 = 1.43 \cdot 10^4 \dots 1.51 \cdot 10^5$  (left) and  $\chi_t^2 = 1.05 \cdot 10^5 \dots 4.01 \cdot 10^5$  (right).

Next the search for minimum of  $\chi_t^2$  is performed in the space of the overall time offset from the flasher start time  $t_{\text{off}}$ , photon yield factor  $p_y$ , shape parameter  $f_{\text{SL}}$  of the scattering function (see section 4), and scaling coefficients applied to the

depth tables of  $b_e(400)$  and  $a_{\text{dust}}(400)$ .

The  $b_e(400)$  and  $a_{\text{dust}}(400)$  of the solution are scaled to produce the likelihood profiles shown in Figure 9. The minimum of  $\chi_q^2$  has an extremely oblong shape, and the direction of its longest extension is determined. The point along the line drawn in this direction is chosen to minimize the  $\chi_t^2$ . Using the timing information appears necessary to resolve both  $b_e(400)$  and  $a_{\text{dust}}(400)$ . At this point the starting "homogeneous ice" values of  $b_e(400)$  and  $a_{\text{dust}}(400)$  are adjusted and the entire procedure is repeated.

Finally, the solution is refined by minimizing  $\chi_t^2$ , varying  $b_e(400)$  and  $a_{\text{dust}}(400)$  at each step of the minimizer 4 times (combinations of  $b_e \pm \delta b_e$  and  $a_{\text{dust}} \pm \delta a_{\text{dust}}$ ). The entire procedure described above is also outlined in Table 1.

The best fit is achieved for  $p_y = 2.40$ , which is near the  $p_y$  value of the average photon yield measured in the lab. Since the best value of  $p_y$  is calculated by the method itself, the resulting table of  $b_e(400)$  and  $a_{\text{dust}}(400)$  is independent of a possible constant scaling factor in the charge estimate or the absolute sensitivity of a DOM. The best fit values of the other parameters are  $t_{\text{off}} = 13 \text{ ns}^6$ ,  $f_{\text{SL}} = 0.45$  (see Figure 10).

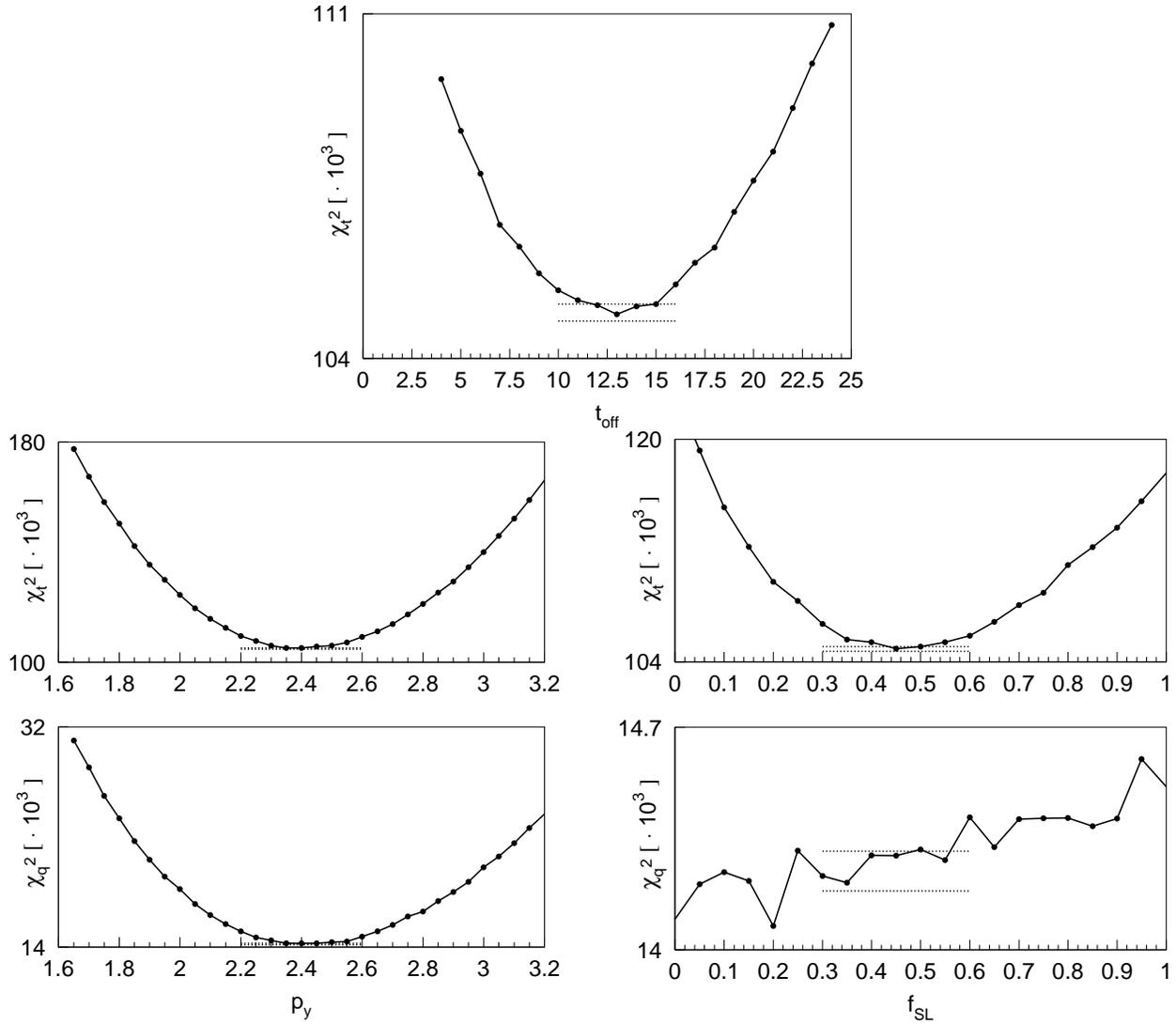


Figure 10: Behavior of  $\chi_t^2$  and  $\chi_q^2$  in the vicinity of the found minimum in  $t_{\text{off}}$ ,  $p_y$ , and  $f_{\text{SL}}$ . All plots are shown on a linear scale. Horizontal dotted lines show the  $\pm 1\sigma$  range due to purely statistical fluctuations in the simulation estimated for the best-fit model. The minimum in  $t_{\text{off}}$  and  $f_{\text{SL}}$  is only a feature of  $\chi_t^2$  but not  $\chi_q^2$ .

<sup>6</sup>The fitted value of  $t_{\text{off}}$  is of the same order as the rise time of the flasher pulse, which was ignored in the simulation.

## 8 Dust logger data

Several dust loggers [3] were used during the deployment of 7 of the IceCube strings to result in a survey of the structure of ice dust layers with extreme detail (with the effective resolution of  $\sim 2$  millimeters). These were then matched up across the detector to result in a *tilt map* of the South Pole ice, as well as a high-detail *average dust log*, a record of a quantity proportional to the dust concentration vs. depth. Additionally, the EDML<sup>7</sup> ice core data was used to extend the dust record to below the lowest dust-logger-acquired point.

The table of dust layer elevation (the *tilt map*) provides the layer shift (relief) from its position at the location of a reference string at distance  $r$  from this string along the average gradient direction (225 degrees SW), see Figure 11. Thus, the  $z$ -coordinate of a given layer at  $r$  is given by  $z_r = z_0 + \text{relief}(z_0, r)$ . Between the grid points  $z_r$  was calculated by linear interpolation in  $z_0$  and  $r$ . The equation was solved by simple iteration resulting in a table of  $z_0(z_r) - z_r$  vs.  $z_r$  given at several points along the gradient direction. Combined with the dust depth record at the location of the reference string (at  $r = 0$ ) this yields a complete description of the dust profile in and around the detector (assuming that the concentration of dust is maintained along the layers).

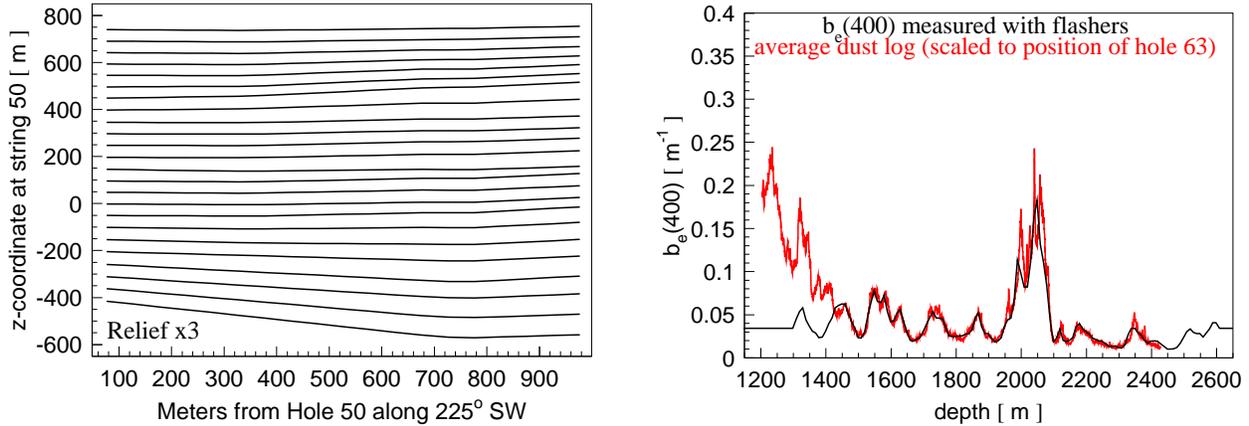


Figure 11: Left: extension of ice layers along the average gradient direction. Relief is amplified by a factor of 3 to enhance the clarity of the layer structure. Right: comparison of the average dust log with the effective scattering coefficient  $b_e(400)$  measured with the flasher data.

The correlation between the effective scattering coefficient measured with the IceCube flasher data and the average dust log (scaled to the location of string 63) is excellent, as shown in Figure 11. All major features match, have the right rise and falloff behavior, and are of the same magnitude. Some minor features are washed out in the flasher measurement.

Having established the correlation with the average dust log, the EDML-extended version of the log was used to build an initial approximation to the fitting algorithm described in the previous section. This resulted in a recovery and enhancement of several features in the scattering and absorption vs. depth that were previously washed out. Additionally, the solution is now biased towards the scaled values of the extended log (instead of to the somewhat arbitrary values of the initial homogeneous ice approximation) in the regions where the flasher fitting method has no resolving power, i.e., above and below the detector.

## 9 Uncertainties of the measurement

To study the precision of the reconstruction method a set of flasher data was simulated with PPC (250 events for each of the 60 flashing DOMs on single string). The agreement between the simulated and reconstructed ice properties is to within  $\sim 5\%$ , see Figure 12. Due to the dramatically lower number of recorded photon hits in the layer of dirtiest ice at  $\sim 2000$  m (the *dust peak*), more simulation was necessary to reconstruct its ice properties: 250 events per flasher were used within the *dust peak*, whereas only 10 events per flasher were used everywhere else. The proper amount of simulation was used to achieve the best possible precision of the final result of section 10.

<sup>7</sup>East Dronning Maud Land, see [3].

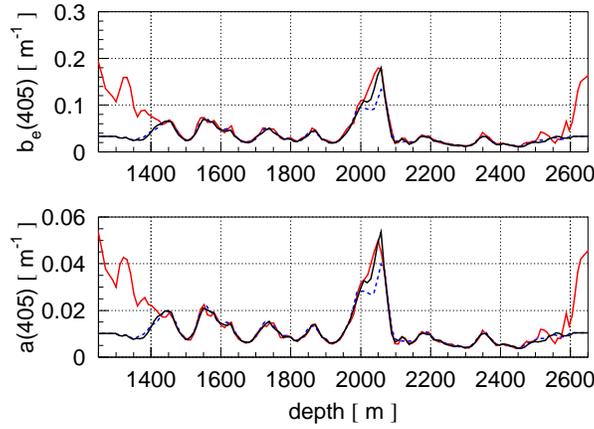


Figure 12: Reconstructed ice properties in black for simulated flasher events with input ice properties in red. The ice properties in the *dust peak* are reconstructed correctly with 250 simulated events per flasher. The blue dashed curve shows the result achieved with only 10 simulated events per flasher.

This verification approach was used to quantify the uncertainty in the measured values of  $b_e(405)$  and  $a(405)$  due to the lack of knowledge of the precise flasher output timing profile. Reconstructing the simulation which used the 62 ns-wide rectangular shape of the flasher pulse with a hypothesis that all photons are emitted simultaneously at flasher start time leads to systematic shifts in the measured ice properties of at most  $\sim 6.5\%$ .

Several pulse extraction methods with and without correcting for PMT saturation (using the saturation model of [4]) were tried in extracting photon hits from the flasher data, and the ice properties were reconstructed for each and compared. This led to an estimate of the uncertainty due to detector calibration and pulse extraction (in waveforms of up to  $\sim 1000$  photoelectrons) of  $\sim 3 - 5\%$ .

Reconstructing the data with the azimuthally-symmetrical vs. 6-fold "star" pattern of flasher LED light leads to no discernible difference in the resulting ice properties. Also, if the DOMs on the flashing string are not used in the fits, the difference between the ice properties reconstructed for nominal or hole ice angular sensitivity models is negligible.

Finally, the uncertainty due to statistical fluctuations in the sets simulated during the reconstruction procedure are estimated at  $\sim 5 - 7\%$ . This uncertainty could be reduced with more simulated events per flasher (at least 10 were simulated for each configuration, cf. 250 in data). However, the entire fitting procedure already takes  $\sim 10$  days of calculation to produce a result, so the number of simulated events cannot be increased much beyond the used value.

## 10 Results

The effective scattering and absorption parameters of ice measured in this work are shown in Figure 13 with the  $\pm 10\%$  gray band corresponding to  $\pm 1\sigma$  uncertainty at most depths. The uncertainty grows beyond the shown band at depths above and below the detector.

Figure 13 also shows the AHA (Additionally Heterogeneous Absorption) model, which is based on the ice description of [2] extrapolated to cover the range of depths of IceCube and updated with a procedure enhancing the depth structure of the ice layers. The AHA model provided the ice description of IceCube prior to this work.

The amplitude and timing distributions in the flasher data agree well with the new model, as shown in Figures 14 and 15. Figure 16 shows data and simulation of the tilted-LED flashers (which were not used in the fit of this work). Figure 17 shows the per-event number of hit DOMs, or  $N_{ch}$ , for IceCube flasher events. Figures 18 and 19 compare the photon arrival time-derived distributions in muon data<sup>8</sup> with those in detector simulation based on the ice model of this work, and Figure 20 shows the average event depth and size.

<sup>8</sup>Data collected by the IceCube detector in the normal operating mode.

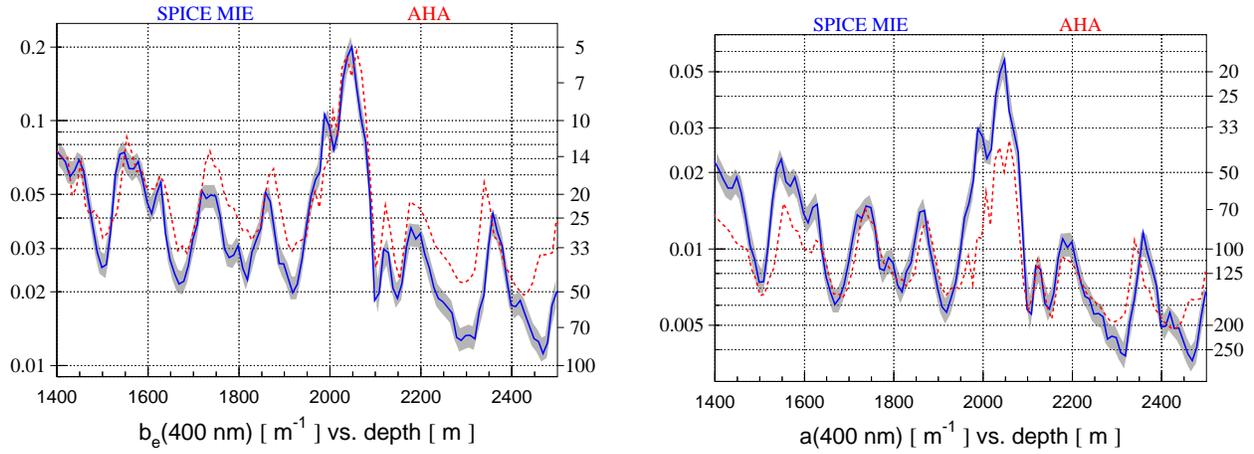


Figure 13: Values of  $b_e(400)$  and  $a(400)$  vs. depth for converged solution in solid blue. The updated model of [2] (AHA) is in dashed red. The uncertainties of the AHA model at the AMANDA depths of  $1730 \pm 225$  m are  $\sim 5\%$  in  $b_e$  and  $\sim 14\%$  in  $a$ . The scale and numbers to the right of each plot indicate the corresponding effective scattering  $1/b_e$  and absorption  $1/a$  lengths in [m].

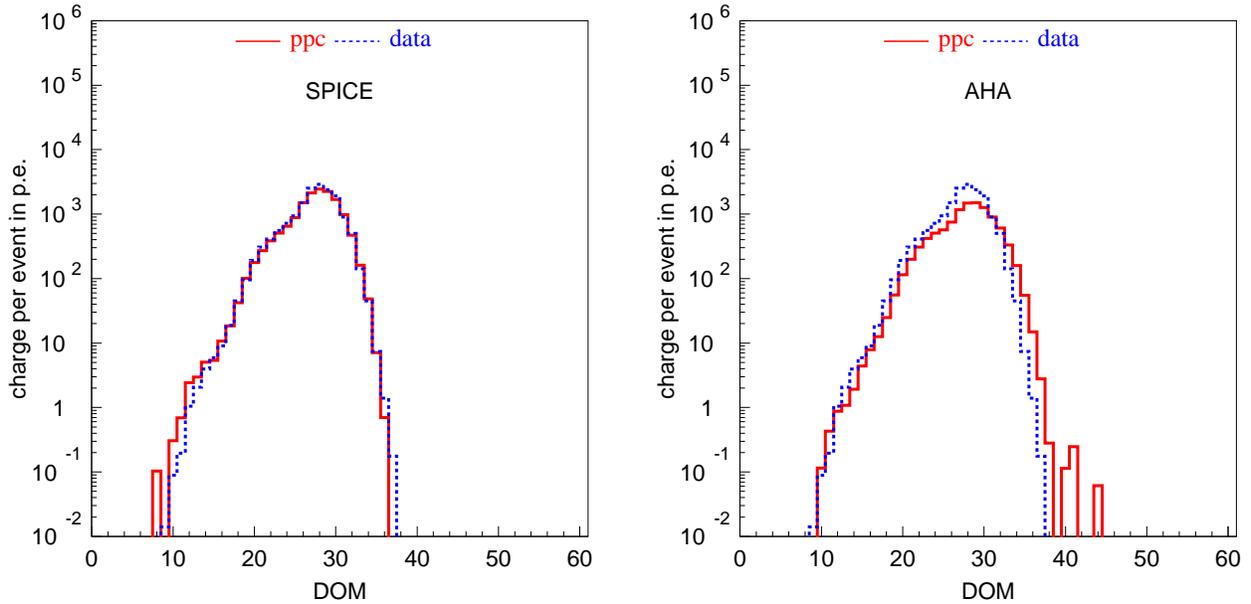


Figure 14: Left plot: ice model of this work (SPICE), right plot: AHA. Charges received on 6 nearest strings when DOM 63-27 was flashing.

## 11 Conclusions

The precise knowledge of the optical properties of ice employed by the IceCube detector is crucial in the analysis of the IceCube data. The scattering and absorption of ice (averaged in 10 m depth bins) were obtained in a fit to the special-purpose in-situ light source data collected in 2008.

Figures 14-20 demonstrate the remarkable improvement in the precision of simulation based on the ice description obtained with the *direct fit* method of this work over that based on the previously used AHA ice model. The disagreement of simulation with muon data has been reduced from a factor in some cases as high as 1.8 to less than 10% in most observables.

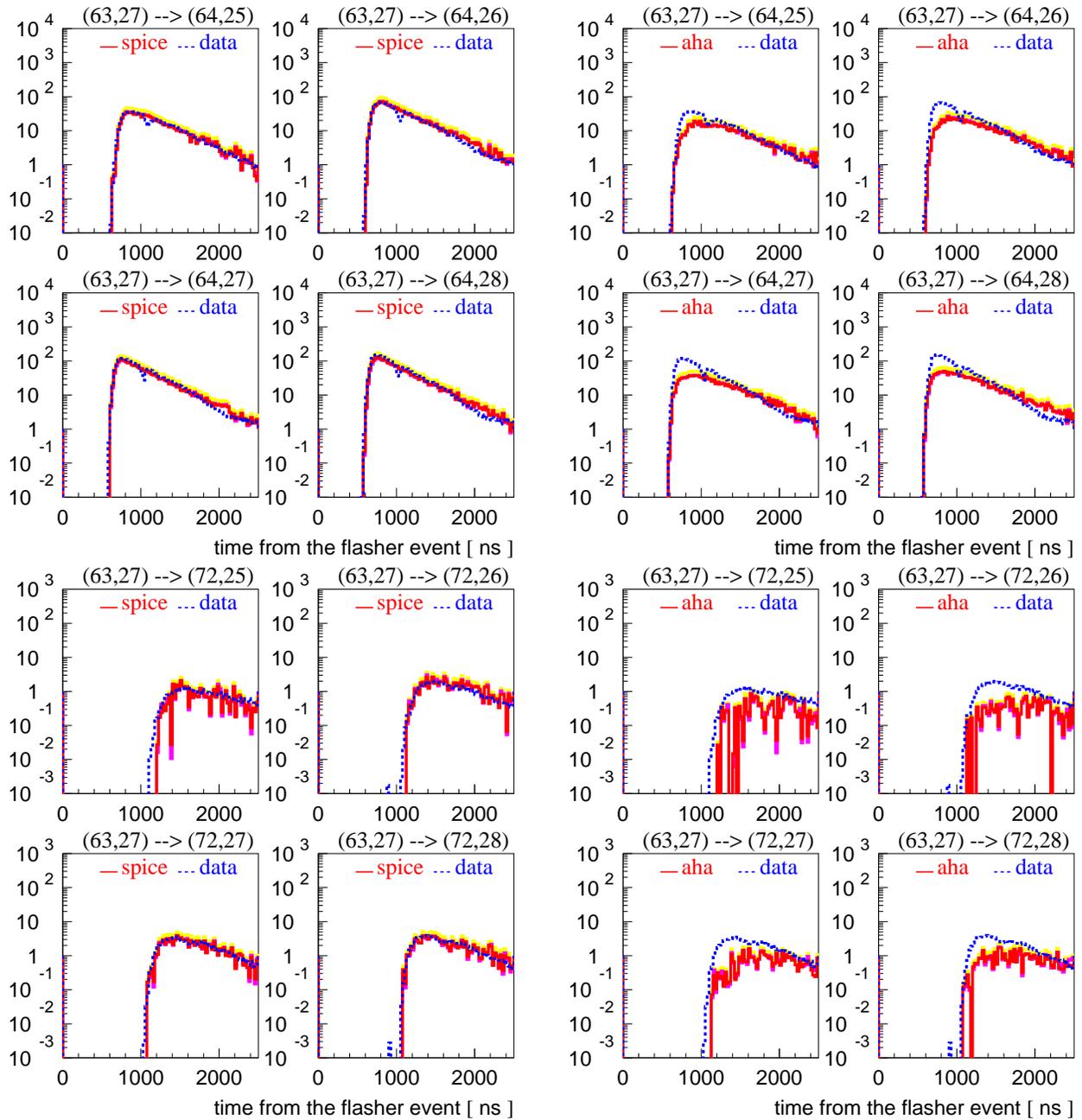


Figure 15: Left plots: ice model of this work (SPICE), right plots: AHA. Timing distributions on 4 DOM positions closest in depth to the flashing DOM 63-27 on 6 nearest strings (top) and 6 next-to-nearest strings (bottom). The dip in the timing distributions is visible at high received charges and corresponds to the transition region between the part of the waveform captured with ATWD (first  $\sim 450$  ns) and fADC. The simplified flasher simulation used in this work does not exhibit such a feature.

## A Photon Propagation Code

Three different versions of the program (available from [11]) were written: one in C++, another in Assembly (for the 32-bit i686 with SSE2 architecture), and a version that employs the NVIDIA GPUs (graphics processing units) via the CUDA programming interface [12]. The relative performance of these different implementations (for simulating both flashers and

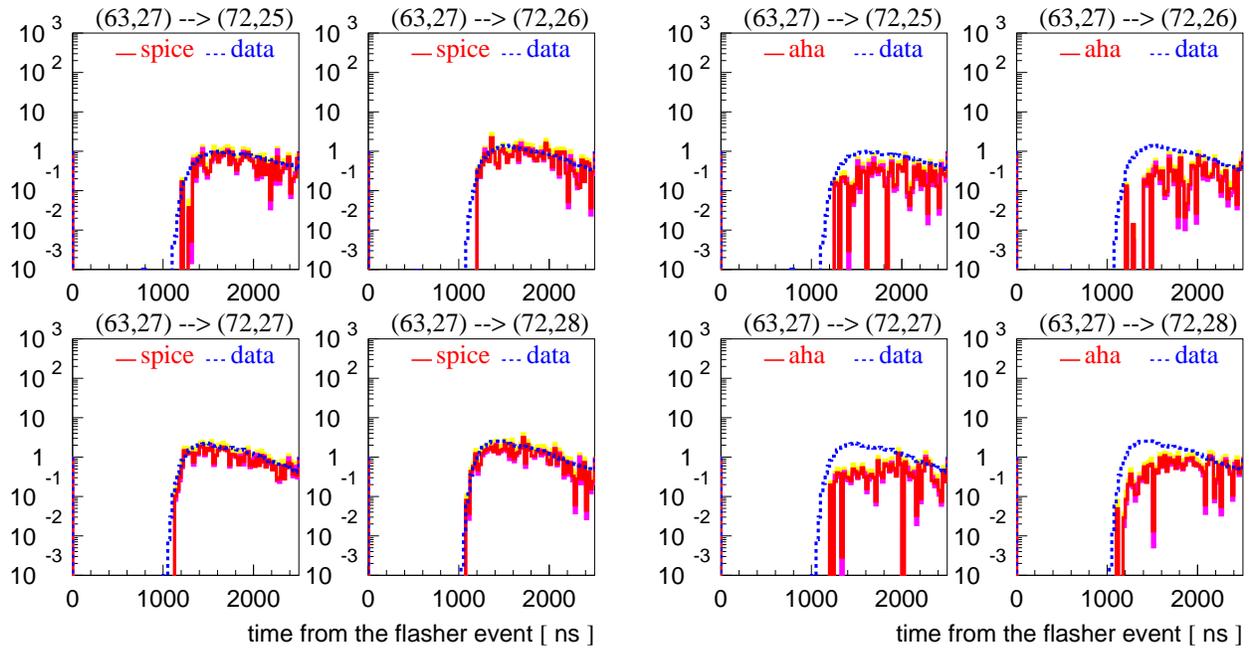


Figure 16: Left plot: ice model of this work (SPICE), right plot: AHA. Timing distributions on 4 DOM positions closest in depth to the DOM 63-27 on 6 next-to-nearest strings when flashing tilted LEDs of DOM 63-27.

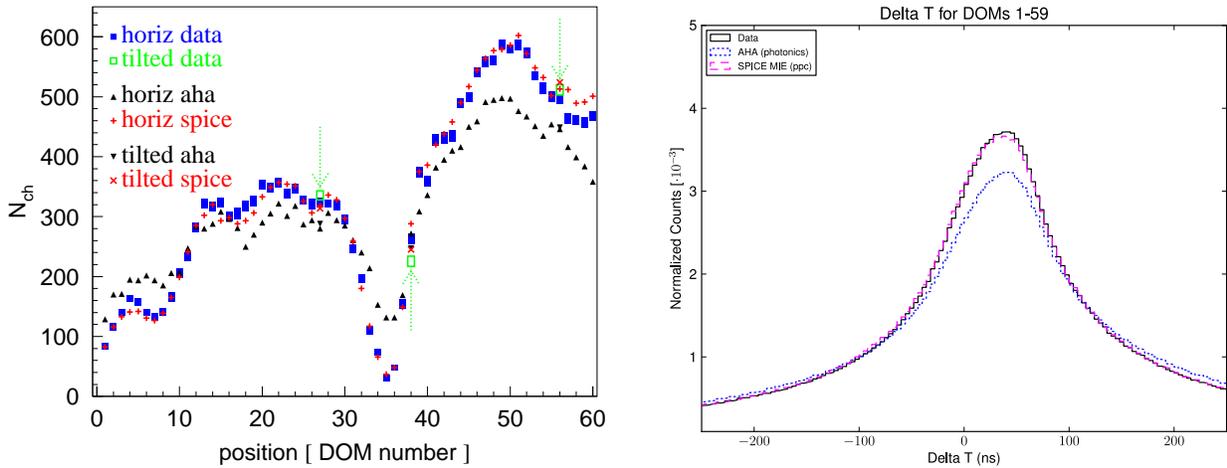


Figure 17: (left) Comparison of the average  $N_{ch}$  of events in data and simulation for horizontal and tilted-LED flashers. The arrows mark 3 positions of the tilted-LED flashers. The "spice" denotes simulation based on the ice model of this work, "aha" denotes simulation based on the AHA model.

Figure 18: (right) Distribution of time differences (all curves are normalized to 1 event per histogram) between hits on DOMs and their immediate lower neighbors in muon data and simulation. The peak is shifted towards positive values since most hits are caused by photons left by muons traveling down.

Cerenkov light from muons) is compared in Table 2.

The writing of the GPU version of PPC was prompted by a similar project called i3mcm1 [13], which showed that acceleration factors  $\sim 100$  compared to the CPU-only version were possible. After demonstrating the impeccable agreement between test simulation sets made with the C++, Assembly, and GPU implementations of PPC, and with i3mcm1, the GPU

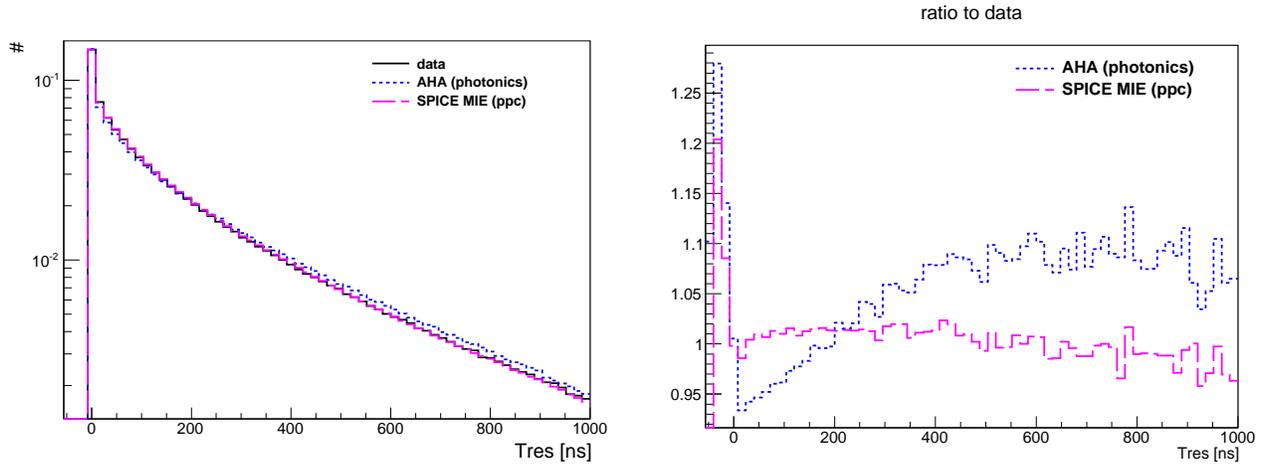


Figure 19: Left: distribution of muon time residuals (all curves are normalized to 1 event per histogram): time delay due to scattering of photons arriving from the reconstructed muon track in data and simulation. Right: ratio of simulation to data.

	C++	Assembly	GTX 295 GPU
flasher	1.00	1.25	147
muon	1.00	1.37	157

Table 2: Speedup factor of different implementations of PPC compared to the C++ version.

version of PPC was chosen for the analysis of this work on a GPU-enabled computer with i7 920 (2.67 GHz) CPU and 3 GTX 295 NVIDIA cards (6 GPUs).

## B Muon and cascade light production

The light yield of the muon and all of its secondaries (ionization losses and delta electrons, bremsstrahlung, electron pair production, and photonuclear interaction [14]) with energies below 500 MeV is parametrized in [15]<sup>9</sup> by substituting the length  $dl$  of the Cerenkov light-emitting segment of a *bare* muon of energy  $E$  with the "effective length"

$$dl_{\text{eff}} = dl \cdot (1.172 + 0.0324 \cdot \log_e(E \text{ [GeV]})) .$$

The light yield of cascades is also parametrized in [15] via the use of the "effective length":

$$\begin{aligned} dl_{\text{eff}} &= 0.894 \cdot 4.889/\rho \text{ m/GeV} \cdot E \text{ [GeV]} && \text{for electromagnetic cascades} \\ dl_{\text{eff}} &= 0.860 \cdot 4.076/\rho \text{ m/GeV} \cdot E \text{ [GeV]} && \text{for hadronic cascades.} \end{aligned}$$

These formulae were derived for muons in water, but are given here for propagation in ice ( $\rho = 0.9216$  is the ratio of the densities of ice<sup>10</sup> and water). This work relies on newer parametrization of the cascade light yield of [16]<sup>11</sup>:

$$\begin{aligned} dl_{\text{eff}} &= 5.21 \text{ m/GeV} \cdot 0.924/\rho \cdot E \text{ [GeV]} && \text{for electromagnetic cascades} \\ dl_{\text{eff}} &= F \cdot 5.21 \text{ m/GeV} \cdot 0.924/\rho \cdot E \text{ [GeV]} && \text{for hadronic cascades.} \end{aligned}$$

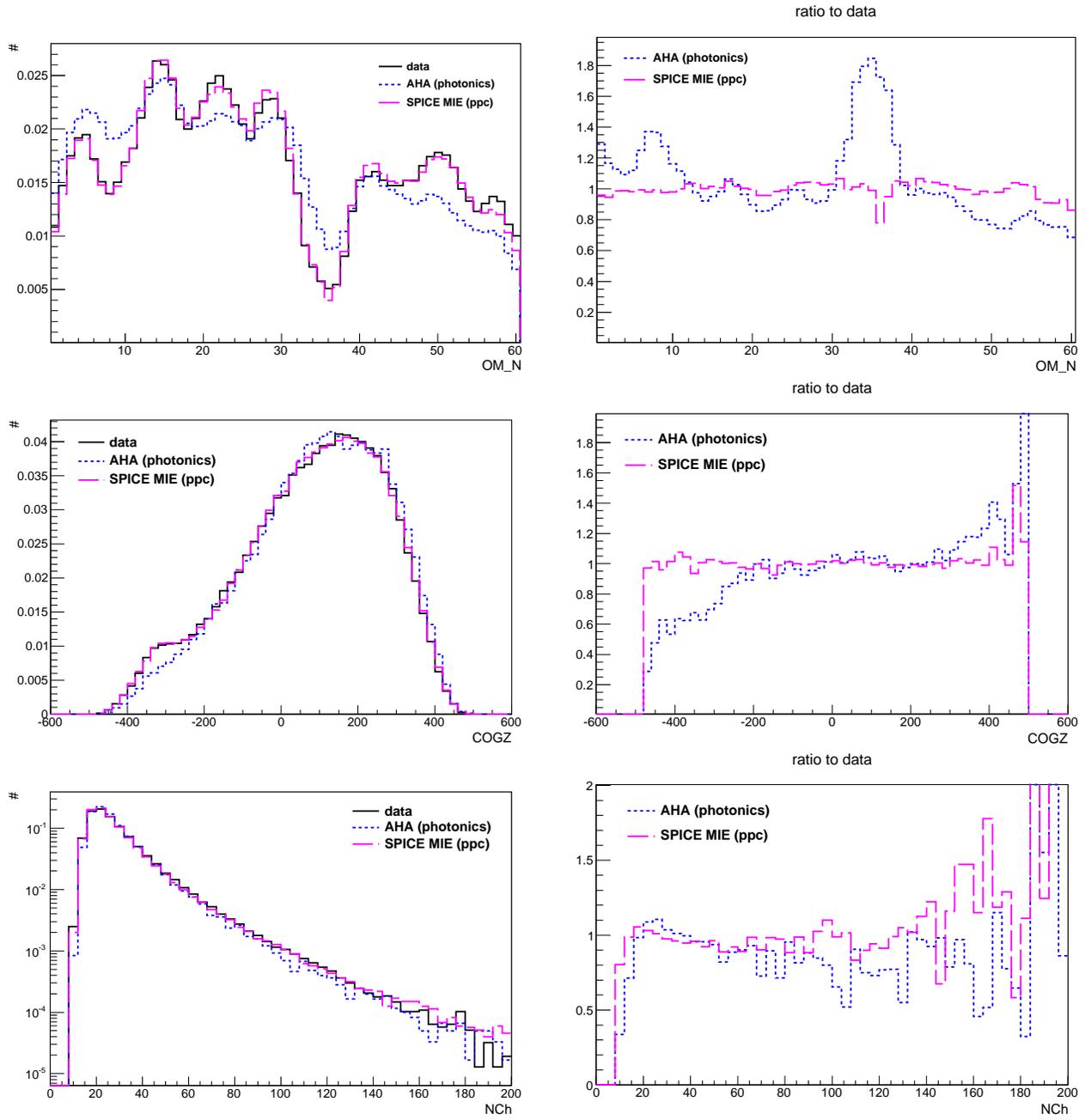
Here  $F$  is a ratio of the effective track length of the hadronic to electromagnetic cascades of the same energy  $E$ . It is approximated with a gaussian distribution with the mean  $\bar{F}$  and width  $\sigma_F$ :

$$\bar{F} = 1 - (E \text{ [GeV]}/E_0)^{-m} \cdot (1 - f_0), \quad E_0 = 0.399, \quad m = 0.130, \quad f_0 = 0.467,$$

<sup>9</sup>The formula 7.97 contains a typo; however, the caption within Figure 7.56 (B) is correct, with  $\text{LOG}(E)$  understood as  $\ln(E) \equiv \log_e(E)$ .

<sup>10</sup>Taken at the center of IceCube (depth of 1950 m, temperature  $-30.4^\circ \text{C}$ ); cf.  $\rho = 0.9167$  at  $0^\circ \text{C}$ .

<sup>11</sup>The axis labels in Figure 3.2 are correct; formula 3.4 needs to be corrected as in this text.



$$\sigma_F = F \cdot \text{rms}_0 \cdot \log_{10}(E [\text{GeV}])^{-\gamma}, \quad \text{rms}_0 = 0.379, \quad \gamma = 1.160.$$

The longitudinal development of cascades is described by sampling the displacement  $l$  from the start of the cascade to the photon emission point from the following distribution (ignoring the LPM elongation) [15]:

$$l = L_{\text{rad}} \cdot \Gamma(a)/b, \quad L_{\text{rad}} = 35.8 \text{ cm}/\rho,$$

where  $\Gamma(a)$  is a gamma distribution with the shape parameter  $a$ . Parameters  $a$  and  $b$  are given by:

$$\begin{aligned} a &= 2.03 + 0.604 \cdot \log_e(E), & b &= 0.633 & \text{for electromagnetic cascades} \\ a &= 1.49 + 0.359 \cdot \log_e(E), & b &= 0.772 & \text{for hadronic cascades.} \end{aligned}$$

All photons are emitted strictly at the Cerenkov angle with respect to the emitting track segment. These, except for the *bare* muon itself, are assumed to be distributed according to

$$dl/dx \sim \exp(-b \cdot x^a) \cdot x^{a-1}, \quad \text{with } x = 1 - \cos(\theta).$$

The coefficients  $a = 0.39$  and  $b = 2.61$  were fitted to a distribution of 100 GeV electron cascades from [15] (see Figure 21) and are fairly constant with energy and are used to describe the hadronic cascades as well.

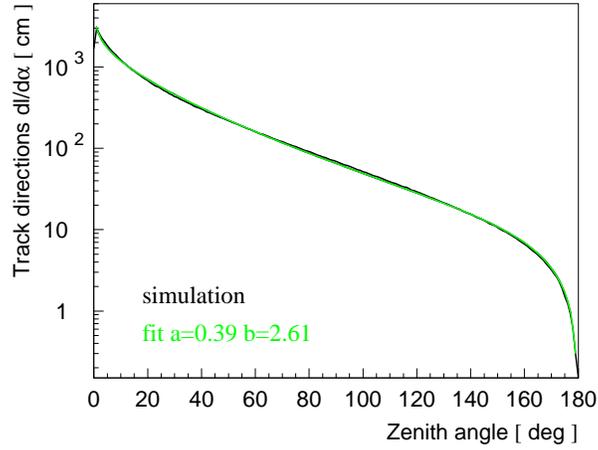


Figure 21: Fit to the angular track-length distribution for 100 GeV electron cascades. The simulation line in black is taken from figure 7.44 of [15], the fit in green is to the function given in this text.

## References

- [1] A. Achterberg et al. First year performance of the icecube neutrino telescope. *Astroparticle Physics*, 26:155, 2006.
- [2] M. Ackermann et al. Optical properties of deep glacial ice at the south pole. *J. Geophys. Res.*, 111:D13203, 2006.
- [3] R. C. Bay et al. South pole paleowind from automated synthesis of ice core records. *J. Geophys. Res.*, 115:D14126, 2010.
- [4] R. Abbasi et al. Calibration and characterization of the icecube photomultiplier tube. *Nucl. Instrum. Methods A*, 618:139–152, 2010.
- [5] R. Abbasi et al. The icecube data acquisition system: Signal capture, digitization, and timestamping. *Nucl. Instrum. Methods A*, 601:294–316, 2009.
- [6] K. Woschnagg and P. B. Price. Temperature dependence of absorption in ice at 532 nm. *Applied Optics*, 40(15):2496, 2001.
- [7] J. Lundberg et al. Light tracking through ice and water - scattering and absorption in heterogeneous media with photonics. *Nucl. Instrum. Methods A*, 581:619, 2007.
- [8] P. B. Price and K. Woschnagg. Role of group and phase velocity in high-energy neutrino observatories. *Astroparticle Physics*, 15:97, 2001.
- [9] Pingyu Liu. A new phase function approximating to Mie scattering for radiative transport equations. *Phys. Med. Biol.*, 39:1025, 1994.
- [10] D. Chirkin. Mie scattering code and plots: <http://icecube.wisc.edu/~dima/work/ICESCA>. 2002.
- [11] D. Chirkin. photon propagation code: <http://icecube.wisc.edu/~dima/work/WISC/ppc>. 2010.
- [12] NVIDIA CUDA: <http://www.nvidia.com/cuda>. 2010.
- [13] Tareq Abuzayyad. i3mcml: <http://wiki.icecube.wisc.edu/index.php/i3mcml>. 2010.
- [14] D. Chirkin and W. Rhode. Propagating leptons through matter with muon monte carlo (mmc). *hep-ph/0407075*, 2008.
- [15] C. H. Wiebusch. *Ph.D. thesis, Physikalische Institute, RWTH Aachen*, pages 85–96, 1995.
- [16] M. Kowalski. On the Cherenkov light emission of hadronic and electro-magnetic cascades. *AMANDA internal report 20020803*, <http://internal.icecube.wisc.edu/reports/amanda/>, 2002.