APPLICATION OF BAYES' THEOREM TO MUON TRACK RECONSTRUCTION IN AMANDA

T. DeYoung[1] and G. C. Hill[2] for the AMANDA Collaboration
[1] Santa Cruz Institute for Particle Physics, University of California, Santa Cruz, CA 95064, USA
[2] Department of Physics, University of Wisconsin, Madison, WI 53706, USA

Abstract
The AMANDA neutrino telescope detects neutrinos by observing Cherenkov light from secondary leptons produced in charged current neutrino interactions. At lower energies, a background of penetrating muons approximately 10^5 times as numerous as neutrino-induced muon events is rejected by looking for muons travelling upward through the earth. Because the detector is designed to maximize effective volume rather than for precision measurements, incorporating this information into the muon track reconstruction algorithm via Bayes' theorem leads to a substantial reduction in the misreconstruction rate.

1 THE AMANDA DETECTOR

The Antarctic Muon and Neutrino Detector Array (AMANDA) is designed to detect secondary leptons produced in high energy neutrino charged-current interactions with the South Polar ice cap or the bedrock below. The leptons are detected through the Cherenkov light emitted as they pass through the ice, which is extremely transparent. The detector presently consists of a three-dimensional array of 677 optical modules (OMs — essentially photomultiplier tubes inside glass pressure vessels) buried deep in ice cap, as shown in Fig. 1.

Because the high energy neutrinos the instrument is intended to detect are rare and the neutrino interaction cross section is low, AMANDA was designed to maximize effective volume at the expense of precision reconstruction of events. The limiting scale of the OM spacing is set by the optical parameters of the ice, which has a typical absorption length of around 95 m and an effective scattering length of 20–30 m [1, 2]. The typical spacing of OMs is therefore 10–20 m vertically and 20–50 m horizontally.

Most AMANDA triggers are produced by muons of relatively low energy (10's to 100's of GeV at the detector), nearly all of which are produced in air showers above the South Pole and penetrate to the detector. However, muons that decay in flight before reaching the surface produce high energy neutrinos, which will occasionally interact near the detector to regenerate the muons. For air showers above Antarctica these neutrino-induced muons are buried in the direct muon flux, but regenerated muons can be observed from air showers in the Northern Hemisphere as well. The angular distributions of penetrating and regenerated muons are shown in Fig. 2.

2 TRACK RECONSTRUCTION

The sparsity of the detector array means that most low energy muons produce signals ("hits") in relatively few OMs. Typical events consist of 15–30 hits over the several hundred meters of contained track. Furthermore, the low absorption of ice relative to the scattering length means that many of the photons collected are highly scattered, arriving well after the Cherenkov time (i.e., the time at which photons emitted at the Cherenkov angle and travelling in a straight line would have been observed). The relative fraction of "direct" (unscattered) photons decreases with distance from the track, as shown in Fig. 3. Because the arrival time distributions are highly asymmetric, AMANDA has historically used a maximum likelihood method to reconstruct the muon track parameters from the observed time delays of the hits.

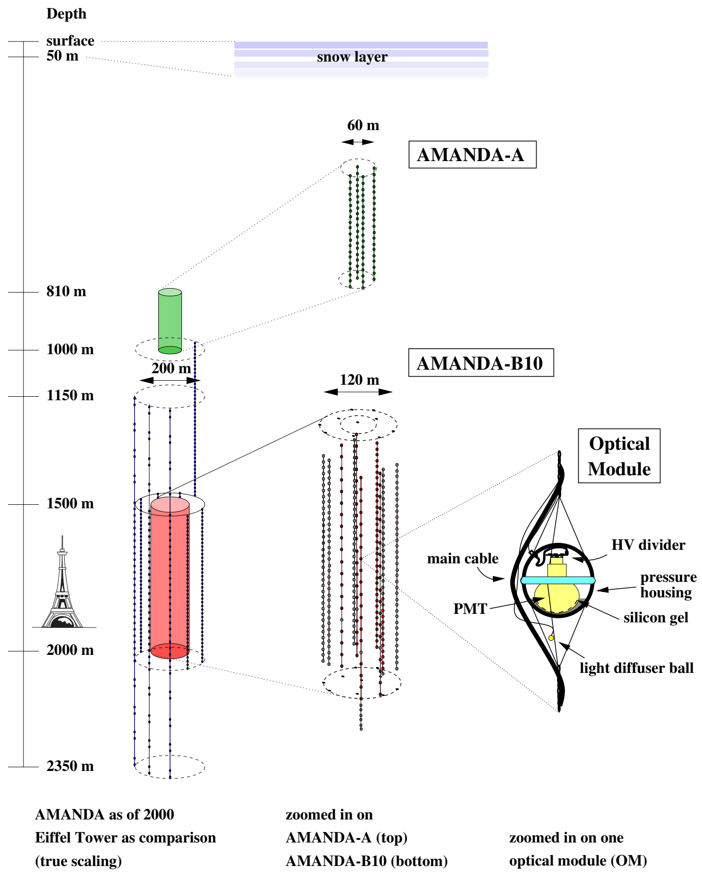

Fig. 1: The present AMANDA detector, with the Eiffel Tower shown to illustrate the scale. This paper describes data taken in 1997, using the ten inner strings labeled AMANDA-B10, shown in expanded view in the center.

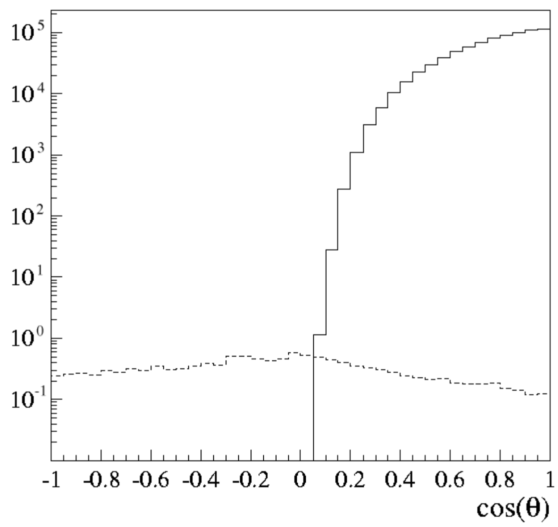

Fig. 2: Zenith angle distribution of simulated AMANDA muon events. The solid line represents penetrating downgoing muons from cosmic ray air showers [3] over Antarctica, simulated by *CORSIKA* [4] and propagated by *mudedx* [5, 6]. The dashed line indicates muons regenerated from neutrinos produced in air showers [7], simulated by *nusim* [8, 9].

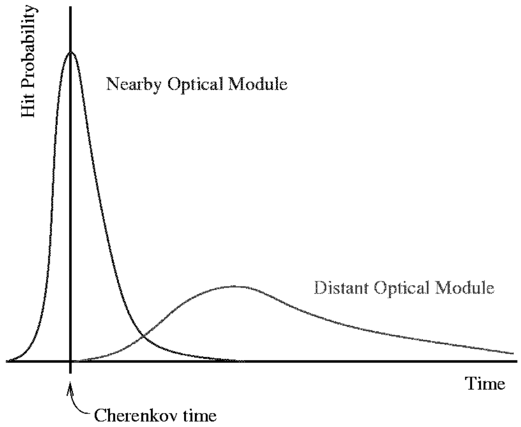

Fig. 3: Cartoon sketch of time residual (delay) distributions of photons arriving at an optical module near the muon track, compared to a distant OM. Times are relative to the arrival time expected for photons in the Cherenkov cone; negative times are possible because the distributions are smeared by the approximately Gaussian jitter in the response of the PMT and electronics. As the distance from the muon track increases, scattering becomes more important.

The time delay pdf's illustrated in Fig. 3 are calculated by simulating the propagation of photon fields produced by muons in ice, including the effects of scattering and absorption [10] and of the hardware response [11]. A timing likelihood

$$\mathcal{L} = \prod_{i=1}^{N_{\rm hit}} p(t_{\rm res}^{(i)} \,|\, d_{\perp}^{(i)}, \theta_{\rm ori}^{(i)}) \tag{1}$$

is then constructed, where $t_{\rm res}$ is the time residual of the first photon (multiple photons in rapid succession are typically not resolved), and $d_{\perp}$ and $\theta_{\rm ori}$ are the distance and orientation of the OM relative to the muon track, which are dependent on the parameters (x, y, z, θ, ϕ) of the muon track. For computational simplicity the probability that an OM be hit is normally not included, nor is any correction made for the bias toward earlier times introduced if multiple photons arrive, but only the time of the first is recorded. The track is reconstructed by varying the muon track parameters until the minimum of the likelihood function is found. Fig. 4 shows a reconstructed muon track from the 1997 data. This event, with 41 hits, is slightly brighter than the typical muon event.

3 BACKGROUND REJECTION

Examining the events reconstructed with the likelihood method described above, about $2 \cdot 10^3$ times too many upgoing events are found. Clearly, a large fraction of the apparently upgoing tracks are in fact misreconstructed downgoing muons. These misreconstructed events constitute a background that must be removed before neutrino events can be identified. Because of the complexity of the task, and to increase the robustness of the results, two separate analyses of the data using different methods were undertaken by the AMANDA collaboration. The analyses will be only sketched out here; details are available in [12].

3.1 Improved Maximum Likelihood

In the first analysis, cuts were applied in several parameters indicative of accurately reconstructed muon tracks, reducing the data set by a factor of approximately 10^5. At this point, a more complete likelihood function

$$\mathcal{L}^{\rm MPE} = \mathcal{L}_{\rm time}^{\rm MPE} \cdot \mathcal{L}_{\rm hit} \tag{2}$$

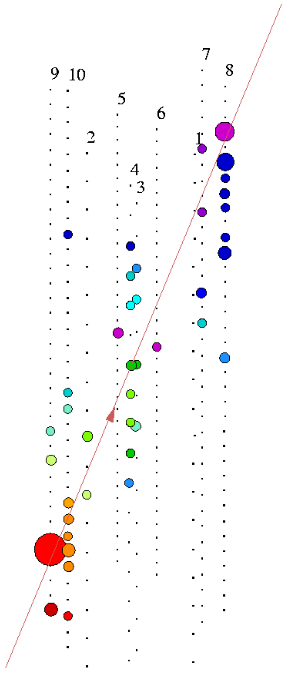

Fig. 4: An upgoing muon track recorded in 1997. The dots indicate OMs in the 10-string AMANDA detector; larger colored circles show OMs that detected light from the muon track. The size of the circles indicates the amplitude of the PMT signal, and the shading indicates the relative timing, early hits at the bottom left and later ones at top right. The line marks the reconstructed muon track.

was constructed and used to refit the surviving tracks. Here $\mathcal{L}_{\text{time}}^{\text{MPE}}$ is the timing likelihood of Eq. 1, save that the probability function is corrected for sampling bias in the case of multiple photoelectrons:

$$p(t_{\text{res}} \,|\, d_\perp, \theta_{\text{ori}}) \;\longrightarrow\; N \cdot p(t_{\text{res}} \,|\, d_\perp, \theta_{\text{ori}}) \cdot \left(\int_{t_{\text{res}}}^{\infty} p(t' \,|\, d_\perp, \theta_{\text{ori}})\, dt' \right)^{(N-1)}$$

where N is the total number of photoelectrons detected. The hit likelihood

$$\mathcal{L}_{\text{hit}} = \prod_{i=1}^{N_{\text{hit}}} P(d_\perp^{(i)}, \theta_{ori}^{(i)}) \;\cdot \prod_{i=N_{\text{hit}}+1}^{N_{\text{OM}}} \left(1 - P(d_\perp^{(i)}, \theta_{ori}^{(i)})\right)$$

is just the probability that the hit tubes be hit, and that quiet tubes not be hit (regardless of time delay), in the given event [13]. This improved likelihood fit further reduced the number of misreconstructed tracks, and with the application of some additional track-quality cuts a nearly pure sample of atmospheric neutrino events was obtained, as seen in Fig. 5 [12, 14].

3.2 Bayesian Reconstruction

The second analysis used Bayes' theorem to incorporate prior knowledge of the muon flux into the reconstruction algorithm [15, 16]. This prior knowledge is based on observations of downgoing atmospheric muons, which comprise the vast majority of AMANDA triggers, and on simulations of AMANDA response to the atmospheric neutrino flux measured by previous experiments. For a particular muon track hypothesis $\mu = \mu(x, y, z, \theta, \phi)$, Bayes' theorem states that

$$P(\mu \,|\, \text{data}) = \mathcal{L}(\text{data} \,|\, \mu)\ \mathrm{P}(\mu),$$

ignoring a normalization factor $P(\text{data})$, which is a constant for the observed event. Here $\mathcal{L}$ is the likelihood of Eq. 1 (or, better, Eq. 2), and $P(\mu)$ represents our prior knowledge of the muon flux to

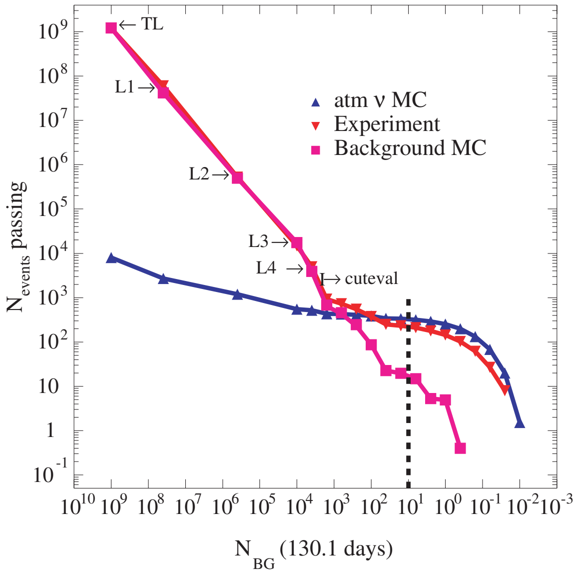

Fig. 5: The number of events surviving the first analysis, for progressively tighter sets of cuts. Note that the background Monte Carlo (simulated misreconstructed downgoing muons) and the data are in good agreement at all levels, until the background are reduced below the level of the signal at a rejection level of 10^{-6}.

which AMANDA is exposed. The most important feature of that flux is its zenith dependence, shown in Fig. 2, so for this analysis we refit the data, multiplying the timing likelihood from Eq. 1 by a one-dimensional prior $P(\theta)$ which is essentially a piecewise polynomial fit to the sum of the curves shown in Fig. 2. Including this prior information reduced the rate of misreconstructions in the 1997 data set by a factor of approximately 410 compared to the timing likelihood fit. At this point, track-quality cuts were imposed on parameters similar to those used in the other analysis, although as shown in Fig. 6 the background rejection factor required was much lower, allowing fewer and looser cuts to be applied.

Subsequent to this analysis, the authors have learned that the same technique was developed by the NEVOD experiment, who were thus able to separate a single atmospheric neutrino candidate from a background of 10^{10} atmospheric muons in a $6 \times 6 \times 7.5$ m^3 surface detector [17].

4 DISCUSSION

AMANDA is a rather unusual instrument in that it was designed to barely work; in the interest of maximizing effective volume, the amount of information collected for each event is purposely kept to the minimum level at which reconstruction is possible. At the same time we are searching for a small signal in a relatively very large background; misreconstructions are thus both common and critically important. In this regime it is absolutely essential to make good use of the relatively sparse information collected by the detector, and this has resulted in a great deal of work on reconstruction and background rejection techniques.

The two analyses described here achieved comparable results, identifying similar numbers of neutrino muon events with similar levels of background contamination. Neither approach is thus clearly superior, and in future work a combination of the two approaches should produce even better results. It is noteworthy, however, that the addition of very simple prior information, a polynomial function of a single variable, yielded improvement that in the more traditional likelihood analysis required the development of several cuts and major improvements of the likelihood function. The same technique could also be used in AMANDA for reconstructing lepton energy, for example, and it seems likely that it could be applied profitably in other experiments when reconstructing parameters that follow a non-uniform distribution.

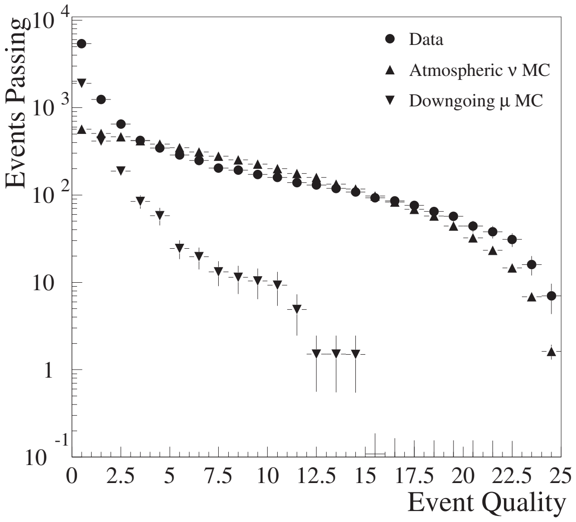

Fig. 6: The number of events surviving the second analysis as cuts are tightened, analogous to Fig. 5. Quality of $Q = 0$ corresponds to events reconstructed as upgoing after cross talk and other instrumental effects were removed from the data stream; without such cleaning approximately 10 times as many events would be reconstructed as upgoing before cuts.

Another interesting fact is that the most tenacious backgrounds in the two analyses were very different. With the Bayesian analysis, "normal" downgoing muons were rejected very effectively, and the vast majority (90%) of the misreconstructed muons were contaminated with cross talk between OM channels [16]. Cross talk in AMANDA can result in apparently upgoing time sequences between the real and induced hits, and the reconstruction algorithm based on the simple timing likelihood was misled by the cross talk. It was primarily to remove these contaminated events that quality cuts had to be applied.

In the traditional analysis, by contrast, the dominant component of the background at every stage of data reduction was regular misreconstructed downgoing muons, in particular nearly horizontal muons for which the likelihood maximum was very broad and extended across the horizon. For these muons, the best maximum-likelihood fit is highly dependent on the stochastics of photon propagation and detection, and the misreconstruction rate is large. Quality cuts imposed to reject these events had the side effect of removing most of the events contaminated by cross talk, as the cuts were much less sensitive to cross talk than the likelihood.

Finally, there is a more theoretical point to address. The second analysis has been referred to here as 'Bayesian,' and its development was guided by Bayesian statistical theory. It is not quite a properly Bayesian reconstruction, however: to keep the computational load to a reasonable level, marginalization of the track position parameters (x, y, z) was not performed; the posterior was simply maximized. Nevertheless, it is a reasonable approximation to a Bayesian reconstruction, subject to operational constraints, and the authors generally prefer to think about the algorithm in the framework of Bayesian statistics.

On the other hand, Bayes' theorem holds also in frequentist statistics. From the frequentist standpoint, all of the prior information used in the 'Bayesian' reconstruction is perfectly acceptable: it is derived from prior measurements, and does not contain any subjective element (or at least, no more than does any scientific result). It is therefore perfectly valid to see this algorithm as a case in which Bayes' theorem can be fruitfully applied in the context of frequentist statistics.

Acknowledgments

T. DeYoung is grateful to R. Cousins, M. Goldstein, F. James, and H. Prosper for excellent discussions of the Bayesian reconstruction from the standpoints of frequentist and Bayesian statistics. G. C. Hill thanks A. A. Petrukhin for helpful comments on this work and for discussions on the NEVOD reconstruction.

AMANDA is supported by the following agencies: U.S. National Science Foundation, Office of Polar Programs; U.S. National Science Foundation, Physics Division; University of Wisconsin Alumni Research Foundation; U.S. Department of Energy; Swedish Natural Science Research Council; Swedish Polar Research Secretariat; Knut and Alice Wallenberg Foundation, Sweden; German Ministry for Education and Research; U.S. National Energy Research Scientific Computing Center (supported by the Office of Energy Research of the U.S. Department of Energy); UC-Irvine AENEAS Supercomputer Facility; Deutsche Forschungsgemeinschaft (DFG). D.F. Cowen acknowledges the support of the NSF CAREER program and C. Pérez de los Heros acknowledges support from the EU 4th framework of Training and Mobility of Researchers.

References

[1] P. B. Price, K. Woschnagg, and D. Chrikin, Geophys. Res. Lett. **27** (2000) 2129.

[2] K. Woschnagg (for the AMANDA Collaboration), in *Proc. of the 26th Intl. Cosmic Ray Conference* (Salt Lake City, Utah, 1999), Vol. 2, 200–203.

[3] B. Wiebel-Sooth and P. L. Biermann, *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*, Vol. VI 3C of *New Series* (Springer-Verlag, 1998), 37–90.

[4] D. Heck *et al.*, Report FZKA 6019, Forschungszentrum Karlsruhe, Karlsruhe, Germany.

[5] W. Lohmann, R. Kopp, and R. Voss, Preprint CERN–85–03, Geneva.

[6] R. Kopp, W. Lohmann, and R. Voss, MUDEDX: Parametrization and Monte Carlo generation of energy loss of high energy muons (1–10000 GeV); for astrophysicists up to 100000 GeV, Version 2.02, 1995, private communication.

[7] P. Lipari, Astrop. Phys. **1** (1993) 195.

[8] G. C. Hill, Astrop. Phys. **6** (1997) 215.

[9] G. C. Hill, Ph.D. thesis, University of Adelaide, Adelaide, Australia, 1996.

[10] A. Karle, in *Simulation and Analysis Methods for Large Neutrino Telescopes* (DESY, Zeuthen, Germany, 1998), pp. 174–185, DESY-PROC–1999–01.

[11] S. Hundertmark, Ph.D. thesis, Humboldt Universität zu Berlin, Berlin, Germany, 1999, available at URL http://amanda.berkeley.edu/manuscripts/.

[12] X. Bai *et al.*, *Observation of High Energy Atmospheric Neutrinos with the Antarctic Muon and Neutrino Detector Array*, submitted to Phys. Rev. D.

[13] C. Wiebusch, in *Simulation and Analysis Methods for Large Neutrino Telescopes* (DESY, Zeuthen, Germany, 1998), pp. 302–316, DESY-PROC–1999–01.

[14] M. Gaug, Diploma thesis, Humboldt Universität zu Berlin, Berlin, Germany, 2001, DESY-THESIS–2001–022, available at URL http://amanda.berkeley.edu/manuscripts/.

[15] G. C. Hill, in *Proc. of the 27th Intl. Cosmic Ray Conference* (Hamburg, Germany, 2001), Vol. 3, 1279–1282.

[16] T. R. DeYoung, Ph.D. thesis, University of Wisconsin, Madison, 2001, available at URL http://amanda.berkeley.edu/manuscripts/.

[17] V. M. Aynutdinov *et al.*, Part. and Nucl. Lett. **109** (2001) 43.