

Improvement in Fast Particle Track Reconstruction with Robust Statistics

M. G. Aartsen², R. Abbasi²⁷, Y. Abdou²², M. Ackermann⁴², J. Adams¹⁵,
J. A. Aguilar²¹, M. Ahlers²⁷, D. Altmann⁹, J. Auffenberg²⁷, X. Bai^{31,1},
M. Baker²⁷, S. W. Barwick²³, V. Baum²⁸, R. Bay⁴³, J. J. Beatty^{17,18},
S. Bechet¹², J. Becker Tjus¹⁰, K.-H. Becker⁴¹, M. L. Benabderrahmane⁴²,
S. BenZvi²⁷, P. Berghaus⁴², D. Berley¹⁶, E. Bernardini⁴², A. Bernhard³⁰,
D. Z. Besson²⁵, G. Binder^{8,43}, D. Bindig⁴¹, M. Bissok¹, E. Blaufuss¹⁶,
J. Blumenthal¹, D. J. Boersma⁴⁰, S. Bohachuk²⁰, C. Boehm³⁴, D. Bose¹³,
S. Böser¹¹, O. Botner⁴⁰, L. Brayeur¹³, H.-P. Bretz⁴², A. M. Brown¹⁵,
R. Bruijn²⁴, J. Brunner⁴², M. Carson²², J. Casey⁵, M. Casier¹³, D. Chirkin²⁷,
A. Christov²¹, B. Christy¹⁶, K. Clark³⁹, F. Clevermann¹⁹, S. Coenders¹,
S. Cohen²⁴, D. F. Cowen^{39,38}, A. H. Cruz Silva⁴², M. Danninger³⁴,
J. Daughhetee⁵, J. C. Davis¹⁷, M. Day²⁷, C. De Clercq¹³, S. De Ridder²²,
P. Desiati²⁷, K. D. de Vries¹³, M. de With⁹, T. DeYoung³⁹, J. C. Díaz-Vélez²⁷,
M. Dunkman³⁹, R. Eagan³⁹, B. Eberhardt²⁸, J. Eisch²⁷, S. Euler¹,
P. A. Evenson³¹, O. Fadiran²⁷, A. R. Fazely⁶, A. Fedynitch¹⁰, J. Feintzeig²⁷,
T. Feusels²², K. Filimonov⁴³, C. Finley³⁴, T. Fischer-Wasels⁴¹, S. Flis³⁴,
A. Franckowiak¹¹, K. Frantzen¹⁹, T. Fuchs¹⁹, T. K. Gaisser³¹, J. Gallagher²⁶,
L. Gerhardt^{8,43}, L. Gladstone²⁷, T. Glüsenkamp⁴², A. Goldschmidt⁸,
G. Golup¹³, J. G. Gonzalez³¹, J. A. Goodman¹⁶, D. Góra⁴²,
D. T. Grandmont²⁰, D. Grant²⁰, A. Groß³⁰, C. Ha^{8,43}, A. Haj Ismail²²,
P. Hallen¹, A. Hallgren⁴⁰, F. Halzen²⁷, K. Hanson¹², D. Heereman¹²,
D. Heinen¹, K. Helbing⁴¹, R. Hellauer¹⁶, S. Hickford¹⁵, G. C. Hill²,
K. D. Hoffman¹⁶, R. Hoffmann⁴¹, A. Homeier¹¹, K. Hoshina²⁷,
W. Huelsnitz^{16,2}, P. O. Hulth³⁴, K. Hultqvist³⁴, S. Hussain³¹, A. Ishihara¹⁴,
E. Jacobi⁴², J. Jacobsen²⁷, K. Jagielski¹, G. S. Japaridze⁴, K. Jero²⁷,
O. Jlelati²², B. Kaminsky⁴², A. Kappes⁹, T. Karg⁴², A. Karle²⁷, J. L. Kelley²⁷,
J. Kiryluk³⁵, J. Kläs⁴¹, S. R. Klein^{8,43}, J.-H. Köhne¹⁹, G. Kohlen²⁹,
H. Kolanoski⁹, L. Köpke²⁸, C. Kopper²⁷, S. Kopper⁴¹, D. J. Koskinen³⁹,
M. Kowalski¹¹, M. Krasberg²⁷, K. Krings¹, G. Kroll²⁸, J. Kunnen¹³,
N. Kurahashi²⁷, T. Kuwabara³¹, M. Labare²², H. Landsman²⁷, M. J. Larson³⁷,
M. Lesiak-Bzdak³⁵, M. Leuermann¹, J. Leute³⁰, J. Lünemann²⁸, O. Macías¹⁵,
J. Madsen³³, G. Maggi¹³, R. Maruyama²⁷, K. Mase¹⁴, H. S. Matis⁸,
F. McNally²⁷, K. Meagher¹⁶, M. Merck²⁷, T. Meures¹², S. Miarecki^{8,43},
E. Middell⁴², N. Milke¹⁹, J. Miller¹³, L. Mohrmann⁴², T. Montaruli^{21,3},

*Corresponding author. Email: wellons@icecube.wisc.edu, Phone: 304-542-4464, Address: Wisconsin Institutes for Discovery, 330 N. Orchard St., Madison, WI 53715

¹Physics Department, South Dakota School of Mines and Technology, Rapid City, SD 57701, USA

²Los Alamos National Laboratory, Los Alamos, NM 87545, USA

³also Sezione INFN, Dipartimento di Fisica, I-70126, Bari, Italy

⁴NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

37 R. Morse²⁷, R. Nahnauer⁴², U. Naumann⁴¹, H. Niederhausen³⁵,
 38 S. C. Nowicki²⁰, D. R. Nygren⁸, A. Obertacke⁴¹, S. Odrowski²⁰, A. Olivas¹⁶,
 39 A. Omairat⁴¹, A. O'Murchadha¹², L. Paul¹, J. A. Pepper³⁷,
 40 C. Pérez de los Heros⁴⁰, C. Pfendner¹⁷, D. Pieloth¹⁹, E. Pinat¹², J. Posselt⁴¹,
 41 P. B. Price⁴³, G. T. Przybylski⁸, L. Rädcl¹, M. Rameez²¹, K. Rawlins³,
 42 P. Redl¹⁶, R. Reimann¹, E. Resconi³⁰, W. Rhode¹⁹, M. Ribordy²⁴,
 43 M. Richman¹⁶, B. Riedel²⁷, J. P. Rodrigues²⁷, C. Rott³⁶, T. Ruhe¹⁹,
 44 B. Ruzybayev³¹, D. Ryckbosch²², S. M. Saba¹⁰, T. Salameh³⁹, H.-G. Sander²⁸,
 45 M. Santander²⁷, S. Sarkar³², K. Schatto²⁸, F. Scheriau¹⁹, T. Schmidt¹⁶,
 46 M. Schmitz¹⁹, S. Schoenen¹, S. Schöneberg¹⁰, A. Schönwald⁴², A. Schukraft¹,
 47 L. Schulte¹¹, O. Schulz³⁰, D. Seckel³¹, Y. Sestayo³⁰, S. Seunarine³³,
 48 R. Shanidze⁴², C. Sheremata²⁰, M. W. E. Smith³⁹, D. Soldin⁴¹,
 49 G. M. Spiczak³³, C. Spiering⁴², M. Stamatikos^{17,4}, T. Stanev³¹, A. Stasik¹¹,
 50 T. Stezelberger⁸, R. G. Stokstad⁸, A. Stöbl⁴², E. A. Strahler¹³, R. Ström⁴⁰,
 51 G. W. Sullivan¹⁶, H. Taavola⁴⁰, I. Taboada⁵, A. Tamburro³¹, A. Tepe⁴¹,
 52 S. Ter-Antonyan⁶, G. Tešić³⁹, S. Tilav³¹, P. A. Toale³⁷, S. Toscano²⁷,
 53 E. Unger¹⁰, M. Usner¹¹, S. Vallecorsa²¹, N. van Eijndhoven¹³,
 54 A. Van Overloop²², J. van Santen²⁷, M. Vehring¹, M. Voge¹¹, M. Vraeghe²²,
 55 C. Walck³⁴, T. Waldenmaier⁹, M. Wallraff¹, Ch. Weaver²⁷, M. Wellons²⁷,
 56 C. Wendt²⁷, S. Westerhoff²⁷, N. Whitehorn²⁷, K. Wiebe²⁸, C. H. Wiebusch¹,
 57 D. R. Williams³⁷, H. Wissing¹⁶, M. Wolf³⁴, T. R. Wood²⁰, K. Woschnagg⁴³,
 58 D. L. Xu³⁷, X. W. Xu⁶, J. P. Yanez⁴², G. Yodh²³, S. Yoshida¹⁴,
 59 P. Zarzhitsky³⁷, J. Ziemann¹⁹, S. Zierke¹, M. Zoll³⁴,
 60 and B. Recht⁴³, C. Ré⁴⁴

61 ¹*III. Physikalisches Institut, RWTH Aachen University, D-52056 Aachen, Germany*
 62 ²*School of Chemistry & Physics, University of Adelaide, Adelaide SA, 5005 Australia*
 63 ³*Dept. of Physics and Astronomy, University of Alaska Anchorage, 3211 Providence Dr.,*
 64 *Anchorage, AK 99508, USA*
 65 ⁴*CTSPS, Clark-Atlanta University, Atlanta, GA 30314, USA*
 66 ⁵*School of Physics and Center for Relativistic Astrophysics, Georgia Institute of*
 67 *Technology, Atlanta, GA 30332, USA*
 68 ⁶*Dept. of Physics, Southern University, Baton Rouge, LA 70813, USA*
 69 ⁷*Dept. of Physics, University of California, Berkeley, CA 94720, USA*
 70 ⁸*Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
 71 ⁹*Institut für Physik, Humboldt-Universität zu Berlin, D-12489 Berlin, Germany*
 72 ¹⁰*Fakultät für Physik & Astronomie, Ruhr-Universität Bochum, D-44780 Bochum,*
 73 *Germany*
 74 ¹¹*Physikalisches Institut, Universität Bonn, Nussallee 12, D-53115 Bonn, Germany*
 75 ¹²*Université Libre de Bruxelles, Science Faculty CP230, B-1050 Brussels, Belgium*
 76 ¹³*Vrije Universiteit Brussel, Dienst ELEM, B-1050 Brussels, Belgium*
 77 ¹⁴*Dept. of Physics, Chiba University, Chiba 263-8522, Japan*
 78 ¹⁵*Dept. of Physics and Astronomy, University of Canterbury, Private Bag 4800,*
 79 *Christchurch, New Zealand*
 80 ¹⁶*Dept. of Physics, University of Maryland, College Park, MD 20742, USA*
 81 ¹⁷*Dept. of Physics and Center for Cosmology and Astro-Particle Physics, Ohio State*
 82 *University, Columbus, OH 43210, USA*
 83 ¹⁸*Dept. of Astronomy, Ohio State University, Columbus, OH 43210, USA*
 84 ¹⁹*Dept. of Physics, TU Dortmund University, D-44221 Dortmund, Germany*
 85 ²⁰*Dept. of Physics, University of Alberta, Edmonton, Alberta, Canada T6G 2E1*

- 86 ²¹*Département de physique nucléaire et corpusculaire, Université de Genève, CH-1211*
87 *Genève, Switzerland*
- 88 ²²*Dept. of Physics and Astronomy, University of Gent, B-9000 Gent, Belgium*
- 89 ²³*Dept. of Physics and Astronomy, University of California, Irvine, CA 92697, USA*
- 90 ²⁴*Laboratory for High Energy Physics, École Polytechnique Fédérale, CH-1015 Lausanne,*
91 *Switzerland*
- 92 ²⁵*Dept. of Physics and Astronomy, University of Kansas, Lawrence, KS 66045, USA*
- 93 ²⁶*Dept. of Astronomy, University of Wisconsin, Madison, WI 53706, USA*
- 94 ²⁷*Dept. of Physics and Wisconsin IceCube Particle Astrophysics Center, University of*
95 *Wisconsin, Madison, WI 53706, USA*
- 96 ²⁸*Institute of Physics, University of Mainz, Staudinger Weg 7, D-55099 Mainz, Germany*
- 97 ²⁹*Université de Mons, 7000 Mons, Belgium*
- 98 ³⁰*T.U. Munich, D-85748 Garching, Germany*
- 99 ³¹*Bartol Research Institute and Department of Physics and Astronomy, University of*
100 *Delaware, Newark, DE 19716, USA*
- 101 ³²*Dept. of Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK*
- 102 ³³*Dept. of Physics, University of Wisconsin, River Falls, WI 54022, USA*
- 103 ³⁴*Oskar Klein Centre and Dept. of Physics, Stockholm University, SE-10691 Stockholm,*
104 *Sweden*
- 105 ³⁵*Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY*
106 *11794-3800, USA*
- 107 ³⁶*Department of Physics, Sungkyunkwan University, Suwon 440-746, Korea*
- 108 ³⁷*Dept. of Physics and Astronomy, University of Alabama, Tuscaloosa, AL 35487, USA*
- 109 ³⁸*Dept. of Astronomy and Astrophysics, Pennsylvania State University, University Park,*
110 *PA 16802, USA*
- 111 ³⁹*Dept. of Physics, Pennsylvania State University, University Park, PA 16802, USA*
- 112 ⁴⁰*Dept. of Physics and Astronomy, Uppsala University, Box 516, S-75120 Uppsala, Sweden*
- 113 ⁴¹*Dept. of Physics, University of Wuppertal, D-42119 Wuppertal, Germany*
- 114 ⁴²*DESY, D-15735 Zeuthen, Germany*
- 115 ⁴³*Dept. of Computer Science, University of California, Berkeley, CA 94704, USA*
- 116 ⁴⁴*Dept. of Computer Science, Stanford University, Stanford, CA 94305, USA*

117 **Abstract**

118 The IceCube project has transformed one cubic kilometer of deep natural
119 Antarctic ice into a Cherenkov detector. Muon neutrinos are detected and their
120 direction inferred by mapping the light produced by the secondary muon track
121 inside the volume instrumented with photomultipliers. Reconstructing the muon
122 track from the observed light is challenging due to noise, light scattering in the
123 ice medium, and the possibility of simultaneously having multiple muons inside
124 the detector, resulting from the large flux of cosmic ray muons.

125 This manuscript describes work on two problems: (1) the *track reconstruction*
126 problem, in which, given a set of observations, the goal is to recover the
127 track of a muon; and (2) the *coincident event* problem, which is to determine
128 how many muons are active in the detector during a time window. Rather than
129 solving these problems by developing more complex physical models that are
130 applied at later stages of the analysis, our approach is to augment the detector's
131 early reconstruction with data filters and robust statistical techniques. These
132 can be implemented at the level of on-line reconstruction and, therefore, improve
133 all subsequent reconstructions. Using the metric of median angular resolution, a

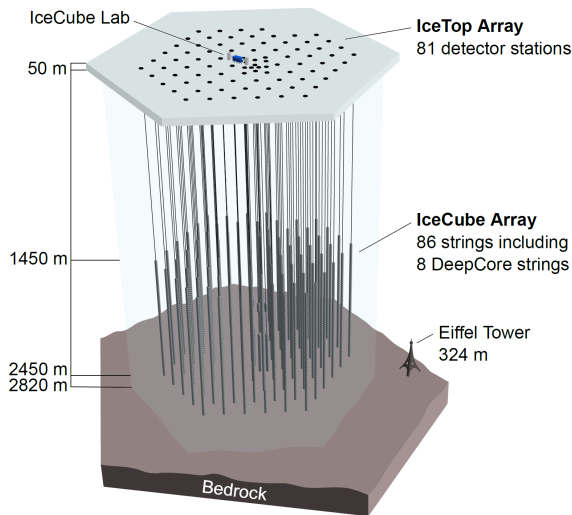


Figure 1: The IceCube neutrino detector in the Antarctic ice. A picture of the Eiffel Tower is shown for scale.

134 standard metric for track reconstruction, we improve the accuracy in the initial
 135 reconstruction direction by 13%. We also present improvements in measuring
 136 the number of muons in coincident events: we can accurately determine the
 137 number of muons 98% of the time.

138 *Keywords:* IceCube, Track reconstruction, Neutrino telescope, Neutrino
 139 astrophysics, Robust Statistics

140 **1. Introduction**

141 The IceCube neutrino detector searches for neutrinos that are generated by
 142 the universe’s most violent astrophysical events: exploding stars, gamma ray
 143 bursts, and cataclysmic phenomena involving black holes and neutron stars [1].
 144 The detector, roughly one cubic kilometer in size, is located near the geographic
 145 South Pole and is buried to a depth of about 2.5 km in the Antarctic ice [2].
 146 The detector is illustrated in Figure 1, and a more complete description is given
 147 in Section 2.

148 When a neutrino enters the telescope, it occasionally interacts in the ice and
 149 generates a muon. The neutrino direction can be inferred from a reconstruction
 150 of the muon track. Muons are also generated by cosmic rays interacting
 151 in the atmosphere, and separation of the background of cosmic ray muons and
 152 neutrino-induced muons is a necessary step for neutrino analysis. This separation
 153 is challenging, as the number of observed cosmic ray muons exceeds the
 154 number of observed neutrino muons by over five orders of magnitude [3].

155 The primary mechanism for separating the cosmic ray muons from the neu-
156 trino muons is reconstructing the muon track and determining whether the muon
157 was traveling downwards into the Earth or upwards out of the Earth. Because
158 neutrinos can penetrate through the Earth but cosmic ray muons cannot, it
159 follows that a muon traveling out of the Earth must have been generated by a
160 neutrino. Thus, by selecting only the muons that are reconstructed as up-going,
161 the cosmic ray muons can, in principle, be removed from the data. Because the
162 number of cosmic ray muons overwhelms the number of neutrino muons, high
163 accuracy is critical for preventing erroneous reconstruction of cosmic ray muons
164 as neutrino-induced.

165 Here, we examine two problems that arise in the separation of cosmic ray
166 muons from neutrino muons in the IceCube detector:

- 167 1. *Reconstruction*, in which the track of a muon is reconstructed from the
168 observed light at different positions and times in the detector.
- 169 2. *Coincident Event Detection*, in which we detect the number of muons
170 inside the detector, and assign observed photons to a muon.

171 Sophisticated reconstruction techniques have been developed that computa-
172 tionally model in detail the muon’s Cherenkov cone, as well as the scattering
173 and absorption of photons through layers of Antarctic ice with varying optical
174 properties [3–5]. Rather than further refining these techniques, the current work
175 focusses on improving the statistical techniques and optimizing data filtering in
176 the early online track reconstruction performed on the data in real time at the
177 South Pole. Besides benefiting directly any analysis that uses the online recon-
178 struction, such as the search for cosmogenic neutrinos, any later analysis will
179 benefit from improvements made at the early stages of the data collection.

180 1.1. *Related Work*

181 Track reconstruction and coincident event detection challenges are ubiqui-
182 tous in particle physics [6–8], both in particle accelerators and cosmic particle
183 detectors. While the work described in this manuscript builds on the previous
184 technique developed for the IceCube detector [3], these techniques are general
185 purpose and potentially have applications in detectors beyond IceCube.

186 1.2. *Outline*

187 We begin by describing the IceCube detector and track reconstruction chal-
188 lenges in Section 2. In Section 3, we describe the reconstruction pipeline in-
189 cluding the prior IceCube software, then we present improvements to the online
190 tracking algorithm and discuss the results. Section 4 describes improvements
191 on coincident event detection and follows a parallel structure to Section 3. We
192 conclude in Section 5.

193 2. IceCube Detector and Track Reconstruction Challenges

194 The IceCube detector is composed of 5,160 optical detectors, each containing
195 a photomultiplier tube (PMT) and an onboard digitizer [9]. The PMTs are
196 spread over 86 vertical strings arranged in a hexagonal shape, with a total
197 instrumented volume of approximately one cubic kilometer. The PMTs on a
198 given string are separated vertically by 17 m, and the string-to-string separation
199 is roughly 125 m.

200 At an abstract level, the IceCube detector operates by detecting muons
201 as they travel through the instrumented volume of ice. As the muon travels
202 through the detector, it radiates light [4], which is observed by the PMTs and
203 quantized into discrete *hits* [10]. The detector uses several trigger criteria. The
204 most commonly used trigger selects time intervals where eight PMTs (with local
205 coincidences) are fired within 5 microseconds. When a trigger occurs, all data
206 within a 10-microsecond trigger window is saved, becoming an *event*. If the
207 number of hits in an event is sufficiently large, the muon track reconstruction
208 algorithm is triggered.

209 There are several challenges for the reconstruction algorithms used in the
210 detector. Varying optical properties of the ice affect reconstruction accuracy,
211 the data may contain outlier hits due to uncorrelated noise, and there are finite
212 computational resources available to tracking code run on-site.

213 *Modeling Difficulties.* The details of the ice’s optical properties are nontrivial to
214 model. Light propagating through the ice is affected by scattering and absorp-
215 tion. These effects cannot be analytically calculated, and the optical properties
216 of the ice vary with depth [5]. In addition, the Cherenkov light originates both
217 directly from the muon and from particles showers initiated by stochastic energy
218 losses of the muon.

219 *Noise.* The noise inherent in the data is another challenge. Noise hits can
220 arise, either from the thermal background of the photocathode or from photons
221 generated by radioactive decay inside the PMT [9].

222 *Computational Constraints.* The reconstruction algorithms are also limited in
223 complexity by the computing resources available at the South Pole. The track
224 reconstruction algorithm has to process about 3,000 muons per second, so algo-
225 rithms with excessive computational demands are discouraged.

226 3. Reconstruction Improvement

227 As shown in the following, augmenting the reconstruction algorithm with
228 some basic filters and classical data analysis techniques results in significant
229 improvement in the reconstruction algorithm’s accuracy.

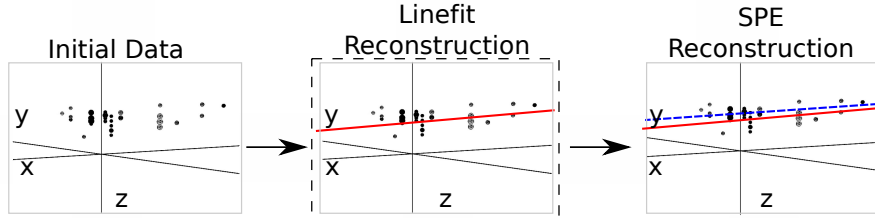


Figure 2: The reconstruction pipeline used to process data in the IceCube detector. After initial data are collected, it is then processed by some basic noise filters, which remove clear outliers. This cleaned data are processed by a basic reconstruction algorithm (solid line), which is used as the seed for the more sophisticated reconstruction algorithm (dashed line). The sophisticated reconstruction is then evaluated as a potential neutrino. The work presented in this manuscript makes changes to the basic reconstruction step (indicated by the dashed box).

230 3.1. Prior IceCube Software

231 The muon track reconstruction process (outlined in Figure 2) starts when the
 232 number of detected hits exceeds a preset threshold and initiates data collection.
 233 After the initial data are collected, the event then passes through a series of
 234 basic filters to remove obvious outliers [11].

235 This is followed by a basic reconstruction algorithm, *linefit* [12], that dis-
 236 regards the Cherenkov cone and, instead, finds the track that minimizes the
 237 sum of the squares of the distances between the track and the hits. More for-
 238 mally, assume there are N hits; denote the position and time of the i th hit as
 239 \vec{x}_i and t_i , respectively. Let the reconstructed muon track have a velocity of \vec{v} ,
 240 and let the reconstructed track pass through point \vec{x}_0 at time t_0 . Then, linefit
 241 reconstruction solves the *least-squares* optimization problem:

$$\min_{t_0, \vec{x}_0, \vec{v}} \sum_{i=1}^N \rho_i(t_0, \vec{x}_0, \vec{v})^2, \quad (1)$$

242 where

$$\rho_i(t_0, \vec{x}_0, \vec{v}) = \|\vec{v}(t_i - t_0) + \vec{x}_0 - \vec{x}_i\|_2. \quad (2)$$

243 Linefit is an approximation primarily used to generate an initial track or *seed*
 244 for a more sophisticated reconstruction.

245 The reconstruction algorithm for the sophisticated reconstruction is *Single-*
 246 *Photo-Electron-Fit (SPE fit)* [3]. SPE fit uses the least-squares reconstruction,
 247 the event data, and a parameterized probability distribution function of scat-
 248 tering in ice [3] to reconstruct the muon track. The SPE fit is the primary
 249 reconstruction algorithm used in the initial data selection and filtering run at
 250 the detector site, and the fit serves as a seed track to the more complex recon-
 251 structions used in off-site data analyses.

252 *3.2. Algorithm Improvement*

253 If angular deviations of the initial seed are large (>5-10 degrees), the simple
254 subsequent reconstruction, SPE, often does not converge to the global minimum,
255 and the efficiency is degraded. This can be resolved by more advanced but
256 time-consuming reconstruction algorithms or by improving the initial seed, as
257 described here.

258 As indicated in Equation 1, a least-squares fit models the muon as a single
259 point moving in a straight line, and hits are penalized quadratically in their
260 distance from this line. Thus, there is an implicit assumption in this model:
261 that all the hits will be near the muon. This assumption has several pitfalls:

- 262 1. It doesn't account for the distinct Cherenkov emission profile from the
263 muon.
- 264 2. It ignores the scattering effects of the ice medium. Some of the photons can
265 scatter for over a microsecond, which means that when they are recorded
266 by a PMT, the muon will have traveled over 300 m away.
- 267 3. While the noise reduction steps remove most of the outlier noise, the
268 noise hits that survive can be far from the muon. Because these outliers
269 are given a quadratic weight, they exert a huge influence over the model.

270 The first two pitfalls occur because the model is incomplete and does not
271 accurately model the data, and the third demonstrates that the model is not
272 robust to noise. The solution to this is twofold: improve the model and increase
273 the noise robustness by replacing least squares with robust statistical techniques.

274 *3.2.1. Improving the Model*

275 While disregarding the Cherenkov profile is inherent to the simplified model
276 chosen for speed reasons, removing hits generated by photons that scattered
277 for a significant length of time will mitigate the effect of ignoring the photon
278 scattering in the ice. We found that a basic filter could identify these scattered
279 hits, and improve accuracy by almost a factor of two by removing them from
280 the dataset.

281 More formally, for each hit h_i , the algorithm looks at all neighboring hits
282 within a neighborhood of r , and if there exists a neighboring hit h_j with a time
283 stamp that is t earlier than h_i , then h_i is considered a scattered hit, and is
284 not used in the basic reconstruction algorithm. Optimal values of r and t were
285 found to be 156 m and 778 ns by tuning them on simulated muon data with an
286 E^{-2} power law spectrum.

287 *3.2.2. Adding Robustness to Noise*

288 As described in Equation 1, the least squares model gives outliers quadratic
289 weight, whereas we would prefer that outliers had zero weight. There are robust
290 models in classical statistics designed to marginalize outliers. We determined
291 that replacing the least-squares model with a Huber fit [13] improves the recon-
292 struction accuracy.

293 More formally, we replace Equation 1 with the optimization problem:

$$\min_{t_0, \vec{x}_0, \vec{v}} \sum_{i=1}^N \phi(\rho_i(t_0, \vec{x}_0, \vec{v})), \quad (3)$$

294 where the Huber penalty function $\phi(\rho)$ is defined as

$$\phi(\rho) \equiv \begin{cases} \rho^2 & \text{if } \rho < \mu \\ \mu(2\rho - \mu) & \text{if } \rho \geq \mu \end{cases}. \quad (4)$$

295 Here, $\rho_i(t_0, \vec{x}, \vec{v})$ is defined in Equation 2 and μ is a constant calibrated to the
 296 data (on simulated muon events with an E^{-2} power law spectrum, the optimal
 297 value of μ is 153 m).

298 The Huber penalty function has two regimes. In the near-hit regime ($\rho < \mu$),
 299 hits are assumed to be strongly correlated with the muon's track, and the Huber
 300 penalty function behaves like least squares, giving these hits quadratic weight.
 301 In the far-hit regime ($\rho \geq \mu$), hits are given linear weights, as they are more
 302 likely to be noise.

303 In addition to its attractive robustness properties, the Huber fit's weight
 304 assignment also has the added benefit that it inherently labels points as outliers
 305 (those with $\rho \geq \mu$). Thus, once the Huber fit is computed, we can go one step
 306 further and simply remove the labeled outliers from the dataset. A better fit is
 307 then obtained by computing the least-squares fit on the data with the outliers
 308 removed. The entire algorithm has a mean runtime that is approximately six
 309 times longer than Linefit's mean runtime.

310 3.3. Results

311 The goal is to improve the accuracy of the reconstruction in order to better
 312 separate neutrinos from cosmic rays. Thus, we present three measurements: (1)
 313 the accuracy change between linefit and the new algorithm; (2) the accuracy
 314 change when SPE is seeded with the new algorithm instead of linefit; and (3)
 315 the improvement in separation between neutrinos and cosmic rays.

316 To measure the improvement generated by the changes, we use the metric
 317 of *median angular resolution* θ_{med} . The angular resolution of a reconstruction
 318 is the arc-distance between the reconstruction and the true track. The dataset
 319 is drawn from simulated neutrino data and is designed to be similar to that
 320 observed by the detector.

321 We can improve the median angular resolution of the basic reconstruction
 322 by 57.6%, as shown in Table 1. Seeding SPE with the improved basic recon-
 323 struction generates an improvement in the angular resolution of 12.9%. These
 324 improvements in the reconstruction algorithm result in 10% fewer atmospheric
 325 muons erroneously reconstructed as up-going, and 1% more muons correctly
 326 reconstructed as up-going.

Table 1: Median angular resolution (degrees) for reconstruction improvements. The first line is the accuracy of the prior least-squares model, and the subsequent lines are the accuracy measurements from cumulatively adding improvements into the basic reconstruction algorithm.

Algorithm	θ_{med}
Linefit Reconstruction (Least-Squares)	9.917
With Addition of Logical Filter	5.205
With Addition of Huber Regression	4.672
With Addition of Outlier Removal	4.211

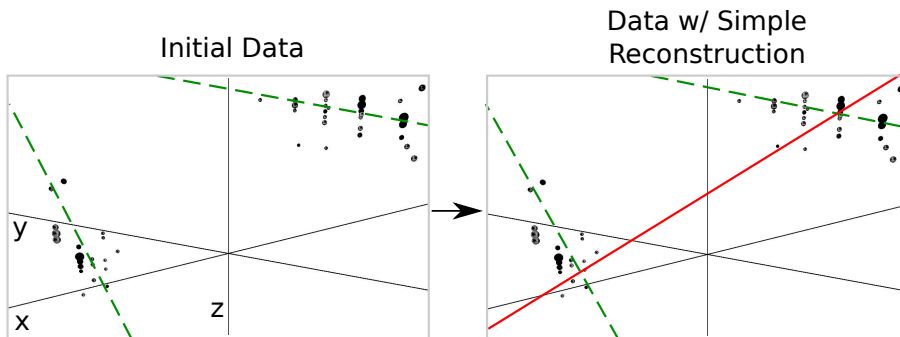


Figure 3: In this example, an event that is clearly composed of two muons (actual tracks shown as dashed lines) is treated as a single muon, and, thus, the reconstruction (solid line) is inaccurate.

327 4. Coincident Event Improvements

328 In the second study, we look at the problem of determining when more than
 329 one muon has entered the detector. In the most common case, a single muon
 330 will pass through the detector and generate an event before exiting. These events
 331 are processed by the pipeline described in Figure 2. However, for roughly 9%
 332 of the events collected by the data collection algorithm, more than one muon
 333 will be passing through the detector simultaneously, an occurrence known as a
 334 *coincident event*.

335 One of the primary sources of background noise in IceCube analyses is coincident
 336 background muons that have been erroneously reconstructed as neutrinos.
 337 To see why this occurs, consider the coincident event shown in Figure 3. There
 338 are two clear groups of hits; however, the reconstruction algorithm treats them
 339 as a single group, resulting in an erroneous reconstruction. In the ideal case,
 340 the reconstruction algorithm would identify coincident events and split them,
 341 as in Figure 4.

342 The challenge in this example is determining the number of muons in an
 343 event. Our studies show that a simple spatial clustering algorithm can solve
 344 this classification problem with less than 2% error.

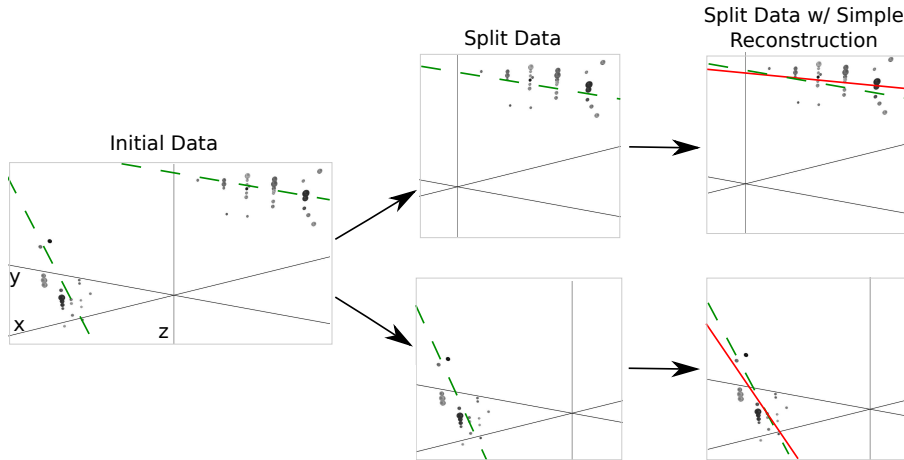


Figure 4: Ideally, the detector would split coincident events before computing the reconstruction. Splitting the event results in more accurate reconstructions (reconstructions shown as solid lines, true muon tracks shown as dashed lines). Note the difference in the reconstructions compared with Figure 3.

345 4.1. Prior IceCube Software

346 Coincident events have been a concern in the IceCube analysis [14] for years,
 347 and some software has been developed to handle coincident events. As a baseline
 348 of comparison, we use the *TTrigger* software, which is described in [15].

349 4.2. Algorithm Improvement

350 Here, we present a proximal clustering algorithm. The intuition in proximal
 351 clustering is that points local in space and time are probably from the same
 352 muon. The proximal clustering algorithm iterates through each pair of hits
 353 (i, j) and builds an adjacency matrix \mathbf{A} as

$$\mathbf{A}_{ij} = \begin{cases} 1 & \text{if } \|\Delta x^2 + \Delta y^2 + \Delta z^2 + (c\Delta t)^2\|_2 \leq \alpha, \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

354 where $\Delta x, \Delta y, \Delta z$, and Δt are the space and time differences between the pair
 355 of hits, and α is tuned to the data (in this application, the optimal value of
 356 α is 450 m). The clustering can be recovered by extracting the connected
 357 components of the graph defined by \mathbf{A} . A connected component of a graph
 358 is a subgraph such, that there exists a path between any two vertices of this
 359 subgraph.

360 4.2.1. Improving the Model

361 When implemented naively, proximal clustering succeeded for the majority
 362 of the events, but it failed if there was a gap in the muon track, which can occur
 363 when the muon travels through dusty ice layers with short scattering length. If

364 there is a significantly large gap, the algorithm erroneously separates the hits
365 into two clusters.

366 To compensate, an additional heuristic is added, *track connecting*. After the
367 data segmentation is finished, track connecting determines if separate clusters
368 should be combined. It computes the mean position and time of each cluster,
369 and connects a hypothetical muon track T between each pair of subspaces.

370 It checks if the speed s of the hypothetical track is within 25% of the speed
371 of light c , and it checks that the mean distance between hits and T in both
372 clusters is less than 60 m. If T passes both checks, the clusters are combined.

373 4.2.2. Adding Robustness to Noise

374 Proximal clustering is susceptible to noise. Noise hits close to two disjoint
375 tracks will be considered adjacent to both tracks and, thus, can connect the two
376 tracks in the adjacency matrix.

377 One heuristic that worked well at mitigating this problem was to not use all
378 of the hits in building the adjacency matrix. During data collection, some hits
379 are flagged as having a *local coincidence condition*, which indicates that both
380 they and a neighboring PMT reported a hit. These hits have a high probability
381 of not being noise hits and, thus, exclusively using them to build the adjacency
382 matrix mitigates the problem of erroneously connecting two tracks.

383 After the proximal clustering algorithm has extracted the tracks from the
384 adjacency matrix, the hits not used in the construction of the adjacency matrix
385 are simply assigned to the closest reconstructed track.

386 4.3. Results

387 There were two competing goals for coincident event detection algorithms:
388 the algorithm should be conservative enough that events containing single tracks
389 are not erroneously split, and aggressive enough that a useful fraction of coin-
390 cident events are split correctly. Our algorithm is tuned to keep almost all
391 of the single events correctly unsplit, while still correctly splitting 80% of the
392 coincident events.

393 4.3.1. Measurements

394 We modified the reconstruction pipeline shown in Figure 2, in between the
395 noise cleaning and the basic reconstruction, by adding a step for coincident event
396 detection, as shown in Figure 4. This step takes cleaned data and attempts to
397 classify the event as a single-track or multiple-track event.

398 We ran each algorithm on two datasets of simulated data. One dataset
399 comprised single-muon events, and the other dataset comprised multiple-muon
400 events. In each dataset, we measured the classification error E , which is the
401 fraction of events that were misclassified. To get a global measurement, we
402 computed the *total error* E_{tot} , defined as

$$E_{tot} = w_{\text{Single}} E_{\text{Single}} + \alpha w_{\text{Multiple}} E_{\text{Multiple}}. \quad (6)$$

Table 2: Error Rates for Classification Algorithms

Algorithm	$E_{\text{Single}}\%$	$E_{\text{Multiple}}\%$	$E_{\text{tot}}\%$
Trivial	0.0	100.0	41.5
TTrigger	11.5	31.8	23.7
Proximal clustering	0.2	18.9	8.0

403 For computing E_{tot} , we use $w_{\text{Single}} = 0.917$ and $w_{\text{Multiple}} = 0.083$, which is
404 the frequency in which single-muon and multiple-muon events appear in data
405 simulating the distribution of events that trigger the reconstruction algorithm.
406 We also include a factor of α in the weighting of the multiple-muon events. This
407 factor expresses that mischaracterizing a multiple-muon event as a single-muon
408 event degrades the quality of most higher-order analysis more than the reverse
409 mischaracterization. In our calculations, we use a value of $\alpha = 5$.

410 We present the results for the coincident event problem by measuring how
411 well each algorithm performs at determining the number of subspaces in an
412 event.

413 There are two natural comparisons for the work: the prior software TTrigger,
414 as well as the trivial algorithm, which always classifies each event as a single-
415 track event. Clearly, the latter will always get the single-track events correct,
416 and always get the multiple-track events wrong. We provide a comparison of
417 these techniques in Table 2. As shown, the new algorithm classifies the number
418 of muons in the detector 66% better than TTrigger.

419 5. Conclusions

420 We found that significant improvements can be achieved in the IceCube’s on-
421 line track reconstruction by employing some classical data analysis algorithms.
422 Optimizing data filtering and refining the least-square model have led to signif-
423 icant improvements in the accuracy of the reconstruction direction. The new
424 reconstruction software is fast enough to run on-site, and is now included in all
425 IceCube analyses.

426 We also looked at the problem of determining the number of muons in the
427 detector. We found that proximal clustering with some basic heuristics could
428 correctly determine whether an event contained a single muon or multiple muons
429 with less than 2% error, yielding a 66% improvement over the prior software.

430 References

- 431 [1] IceCube Collaboration, IceCube webpage, <http://icecube.wisc.edu/>.
- 432 [2] A. Achterberg et al., First year performance of the IceCube neutrino tele-
433 scope, *Astroparticle Physics* 26 (3) (2006) 155–173.

- 434 [3] J.Ahrens et al., Muon track reconstruction and data selection techniques in
435 AMANDA, Nuclear Instruments and Methods in Physics Research Section
436 A 524 (2004) 169–194.
- 437 [4] M. G. Aartsen et al., Measurement of South Pole ice transparency with the
438 IceCube LED calibration system IceCube Collaboration, Nuclear Instru-
439 ments and Methods in Physics Research Section A (2013) 73–89.
- 440 [5] M. Ackermann et al., Optical properties of deep glacial ice at the south
441 pole, Journal of Geophysical Research 111 (D13) (2006) D13203.
- 442 [6] F. Hirsch, on behalf of the ATLAS collaboration, Tracking and vertexing
443 with the ATLAS detector at the LHC, Nuclear Instruments and Methods
444 in Physics Research Section A: Accelerators, Spectrometers, Detectors and
445 Associated Equipment 650 (1) (2011) 218–223.
- 446 [7] R. S. Chivukulaa, M. Goldenaa, E. H. Simmons, Multi-jet physics at hadron
447 colliders, Nuclear Physics B 363 (1) (1991) 83–96.
- 448 [8] S. Ellis, J. Huston, K. Hatakeyama, P. Loch, M. Tönnesmann, Jets in
449 hadron–hadron collisions, Progress in Particle and Nuclear Physics (60)
450 (2008) 484–551.
- 451 [9] R. Abbasi et al., Calibration and characterization of the IceCube photomul-
452 tiplier tube, Nuclear Instruments and Methods in Physics Research Section
453 A 618 (2010) 139–152.
- 454 [10] R. Abbasi et al., The icecube data acquisition system: Signal capture, dig-
455 itization, and timestamping, Nuclear Instruments and Methods in Physics
456 Research Section A 601 (3) (2009) 294–316.
- 457 [11] M. Ackermann, Searches for signals from cosmic point-like sources of high
458 energy neutrinos in 5 years of AMANDA-II data, Ph.D. thesis, Humboldt-
459 Universität zu Berlin (2006).
- 460 [12] V. Stenger, Track fitting for DUMAND-II Octagon Array, Tech. rep., Uni-
461 versity of Hawaii at Manoa (1990).
- 462 [13] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University
463 Press, Pages 299, 2009.
- 464 [14] R. Abbasi et al., Measurement of the atmospheric neutrino energy spectrum
465 from 100 GeV to 400 TeV with IceCube, Physical Review D 83 (1).
- 466 [15] D. Chirkin, Measurement of the atmospheric neutrino energy spectrum
467 with IceCube, Proceedings of the 31st ICRC.