

Improvement in Muon Track Reconstruction with Robust Statistics

M. G. Aartsen², R. Abbasi²⁷, Y. Abdou²², M. Ackermann⁴¹, J. Adams¹⁵,
J. A. Aguilar²¹, M. Ahlers²⁷, D. Altmann⁹, J. Auffenberg²⁷, X. Bai^{31,1},
M. Baker²⁷, S. W. Barwick²³, V. Baum²⁸, R. Bay⁷, J. J. Beatty^{17,18},
S. Bechet¹², J. Becker Tjus¹⁰, K.-H. Becker⁴⁰, M. Bell³⁸,
M. L. Benabderrahmane⁴¹, S. BenZvi²⁷, J. Berdermann⁴¹, P. Berghaus⁴¹,
D. Berley¹⁶, E. Bernardini⁴¹, A. Bernhard³⁰, D. Bertrand¹², D. Z. Besson²⁵,
G. Binder^{8,7}, D. Bindig⁴⁰, M. Bissok¹, E. Blaufuss¹⁶, J. Blumenthal¹,
D. J. Boersma³⁹, S. Bohaichuk²⁰, C. Boehm³⁴, D. Bose¹³, S. Böser¹¹,
O. Botner³⁹, L. Brayeur¹³, H.-P. Bretz⁴¹, A. M. Brown¹⁵, R. Bruijn²⁴,
J. Brunner⁴¹, M. Carson²², J. Casey⁵, M. Casier¹³, D. Chirkin²⁷,
A. Christov²¹, B. Christy¹⁶, K. Clark³⁸, F. Clevermann¹⁹, S. Coenders¹,
S. Cohen²⁴, D. F. Cowen^{38,37}, A. H. Cruz Silva⁴¹, M. Danninger³⁴,
J. Daughhetee⁵, J. C. Davis¹⁷, C. De Clercq¹³, S. De Ridder²², P. Desiati²⁷,
M. de With⁹, T. DeYoung³⁸, J. C. Díaz-Vélez²⁷, M. Dunkman³⁸, R. Eagan³⁸,
B. Eberhardt²⁸, J. Eisch²⁷, R. W. Ellsworth¹⁶, S. Euler¹, P. A. Evenson³¹,
O. Fadiran²⁷, A. R. Fazely⁶, A. Fedynitch¹⁰, J. Feintzeig²⁷, T. Feusels²²,
K. Filimonov⁷, C. Finley³⁴, T. Fischer-Wasels⁴⁰, S. Flis³⁴, A. Franckowiak¹¹,
R. Franke⁴¹, K. Frantzen¹⁹, T. Fuchs¹⁹, T. K. Gaisser³¹, J. Gallagher²⁶,
L. Gerhardt^{8,7}, L. Gladstone²⁷, T. Glüsenkamp⁴¹, A. Goldschmidt⁸,
G. Golup¹³, J. A. Goodman¹⁶, D. Góra⁴¹, D. Grant²⁰, A. Groß³⁰,
M. Gurtner⁴⁰, C. Ha^{8,7}, A. Haj Ismail²², P. Hallen¹, A. Hallgren³⁹,
F. Halzen²⁷, K. Hanson¹², D. Heereman⁴², P. Heimann¹, D. Heinen¹,
K. Helbing⁴⁰, R. Hellauer¹⁶, S. Hickford¹⁵, G. C. Hill², K. D. Hoffman¹⁶,
R. Hoffmann⁴⁰, A. Homeier¹¹, K. Hoshina²⁷, W. Huelsnitz^{16,2}, P. O. Hulth³⁴,
K. Hultqvist³⁴, S. Hussain³¹, A. Ishihara¹⁴, E. Jacobi⁴¹, J. Jacobsen²⁷,
K. Jagielski¹, G. S. Japaridze⁴, K. Jero²⁷, O. Jlelati²², B. Kaminsky⁴¹,
A. Kappes⁹, T. Karg⁴¹, A. Karle²⁷, J. L. Kelley²⁷, J. Kiryluk³⁵, F. Kislak⁴¹,
J. Kläs⁴⁰, S. R. Klein^{8,7}, J.-H. Köhne¹⁹, G. Kohnen²⁹, H. Kolanoski⁹,
L. Köpke²⁸, C. Kopper²⁷, S. Kopper⁴⁰, D. J. Koskinen³⁸, M. Kowalski¹¹,
M. Krasberg²⁷, K. Krings¹, G. Kroll²⁸, J. Kunnen¹³, N. Kurahashi²⁷,
T. Kuwabara³¹, M. Labare¹³, H. Landsman²⁷, M. J. Larson³⁶,
M. Lesiak-Bzdak³⁵, M. Leuermann¹, J. Leute³⁰, J. Lünemann²⁸, J. Madsen³³,
R. Maruyama²⁷, K. Mase¹⁴, H. S. Matis⁸, F. McNally²⁷, K. Meagher¹⁶,
M. Merck²⁷, P. Mészáros^{37,38}, T. Meures¹², S. Miarecki^{8,7}, E. Middell⁴¹,

*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 N. Milke¹⁹, J. Miller¹³, L. Mohrmann⁴¹, T. Montaruli^{21,3}, R. Morse²⁷,
 38 R. Nahnhauser⁴¹, U. Naumann⁴⁰, H. Niederhausen³⁵, S. C. Nowicki²⁰,
 39 D. R. Nygren⁸, A. Obertacke⁴⁰, S. Odrowski³⁰, A. Olivas¹⁶, M. Olivo¹⁰,
 40 A. O’Murchadha¹², L. Paul¹, J. A. Pepper³⁶, C. Pérez de los Heros³⁹,
 41 C. Pfendner¹⁷, D. Pieloth¹⁹, N. Pirk⁴¹, J. Posselt⁴⁰, P. B. Price⁷,
 42 G. T. Przybylski⁸, L. Rädcl¹, K. Rawlins³, P. Redl¹⁶, R. Reimann¹,
 43 E. Resconi³⁰, W. Rhode¹⁹, M. Ribordy²⁴, M. Richman¹⁶, B. Riedel²⁷,
 44 J. P. Rodrigues²⁷, C. Rott¹⁷, T. Ruhe¹⁹, B. Ruzybayev³¹, D. Ryckbosch²²,
 45 S. M. Saba¹⁰, T. Salameh³⁸, H.-G. Sander²⁸, M. Santander²⁷, S. Sarkar³²,
 46 K. Schatto²⁸, M. Scheel¹, F. Scheriau¹⁹, T. Schmidt¹⁶, M. Schmitz¹⁹,
 47 S. Schoenen¹, S. Schöneberg¹⁰, L. Schönherr¹, A. Schönwald⁴¹, A. Schukraft¹,
 48 L. Schulte¹¹, O. Schulz³⁰, D. Seckel³¹, S. H. Seo³⁴, Y. Sestayo³⁰,
 49 S. Seunarine³³, C. Sheremata²⁰, M. W. E. Smith³⁸, M. Soiron¹, D. Soldin⁴⁰,
 50 G. M. Spiczak³³, C. Spiering⁴¹, M. Stamatikos^{17,4}, T. Stanev³¹, A. Stasik¹¹,
 51 T. Stezelberger⁸, R. G. Stokstad⁸, A. Stöbl⁴¹, E. A. Strahler¹³, R. Ström³⁹,
 52 G. W. Sullivan¹⁶, H. Taavola³⁹, I. Taboada⁵, A. Tamburro³¹,
 53 S. Ter-Antonyan⁶, S. Tilav³¹, P. A. Toale³⁶, S. Toscano²⁷, M. Usner¹¹,
 54 D. van der Drift^{8,7}, N. van Eijndhoven¹³, A. Van Overloop²², J. van Santen²⁷,
 55 M. Vehringer¹, M. Voge¹¹, M. Vraeghe²², C. Walck³⁴, T. Waldenmaier⁹,
 56 M. Wallraff¹, R. Wasserman³⁸, Ch. Weaver²⁷, M. Wellons^{27,*}, C. Wendt²⁷,
 57 S. Westerhoff²⁷, N. Whitehorn²⁷, K. Wiebe²⁸, C. H. Wiebusch¹,
 58 D. R. Williams³⁶, H. Wissing¹⁶, M. Wolf³⁴, T. R. Wood²⁰, K. Woschnagg⁷,
 59 C. Xu³¹, D. L. Xu³⁶, X. W. Xu⁶, J. P. Yanez⁴¹, G. Yodh²³, S. Yoshida¹⁴,
 60 P. Zarzhitsky³⁶, J. Ziemann¹⁹, S. Zierke¹, A. Zilles¹, M. Zoll³⁴, B. Recht⁴²,
 61 C. Ré⁴²

62 ¹III. Physikalisches Institut, RWTH Aachen University, D-52056 Aachen, Germany

63 ²School of Chemistry & Physics, University of Adelaide, Adelaide SA, 5005 Australia

64 ³Dept. of Physics and Astronomy, University of Alaska Anchorage, 3211 Providence Dr.,
 65 Anchorage, AK 99508, USA

66 ⁴CTSPS, Clark-Atlanta University, Atlanta, GA 30314, USA

67 ⁵School of Physics and Center for Relativistic Astrophysics, Georgia Institute of
 68 Technology, Atlanta, GA 30332, USA

69 ⁶Dept. of Physics, Southern University, Baton Rouge, LA 70813, USA

70 ⁷Dept. of Physics, University of California, Berkeley, CA 94720, USA

71 ⁸Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

72 ⁹Institut für Physik, Humboldt-Universität zu Berlin, D-12489 Berlin, Germany

73 ¹⁰Fakultät für Physik & Astronomie, Ruhr-Universität Bochum, D-44780 Bochum,
 74 Germany

75 ¹¹Physikalisches Institut, Universität Bonn, Nussallee 12, D-53115 Bonn, Germany

76 ¹²Université Libre de Bruxelles, Science Faculty CP230, B-1050 Brussels, Belgium

77 ¹³Vrije Universiteit Brussel, Dienst ELEM, B-1050 Brussels, Belgium

78 ¹⁴Dept. of Physics, Chiba University, Chiba 263-8522, Japan

79 ¹⁵Dept. of Physics and Astronomy, University of Canterbury, Private Bag 4800,
 80 Christchurch, New Zealand

81 ¹⁶Dept. of Physics, University of Maryland, College Park, MD 20742, USA

82 ¹⁷Dept. of Physics and Center for Cosmology and Astro-Particle Physics, Ohio State
 83 University, Columbus, OH 43210, USA

84 ¹⁸Dept. of Astronomy, Ohio State University, Columbus, OH 43210, USA

85 ¹⁹Dept. of Physics, TU Dortmund University, D-44221 Dortmund, Germany

86 ²⁰Dept. of Physics, University of Alberta, Edmonton, Alberta, Canada T6G 2E1

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

116 **Abstract**

117 The IceCube detector is a high-energy neutrino telescope located at the geo-
118 graphic South Pole. Neutrinos cannot be directly observed and must be inferred
119 from their interactions with other particles. These interactions sometimes gener-
120 ate a muon, which in turn emits observable light. At the energies the IceCube
121 detector is sensitive to, the neutrino and generated muon have almost parallel
122 tracks, so the neutrino track can be extrapolated from a reconstruction of the
123 muon track. However, reconstructing the muon track from the observed light is
124 challenging due to noise, light scattering in the ice medium, and the possibility
125 of simultaneously having multiple muons inside the detector.

126 This manuscript describes our work on two problems: (1) the *track recon-*
127 *struction* problem, in which, given a set of observations, our goal is to recover
128 the track of a muon, and (2) the *coincident event* problem, which is to deter-
129 mine how many muons are active in the detector during a time window. Rather
130 than solving these problems by developing more complex physical models, our
131 approach is to augment the detector's current models with simple filters and
132 robust statistical techniques. Using the metric of median angular resolution,
133 a standard metric for track reconstruction, our solution improves the accuracy

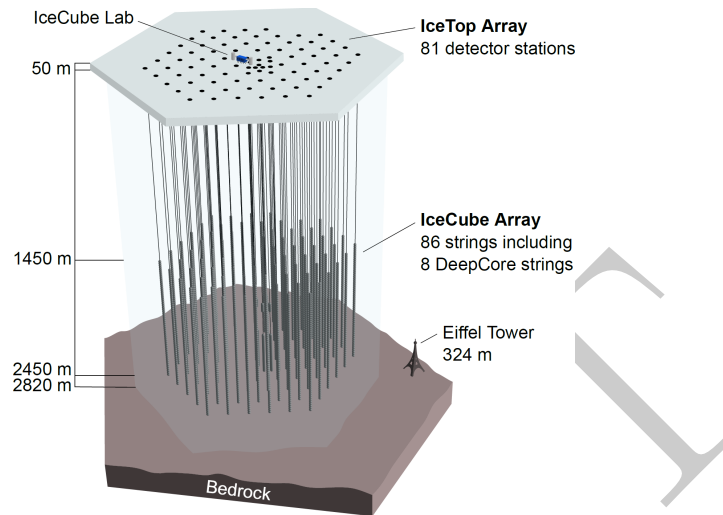


Figure 1: The IceCube neutrino detector in the Antarctic ice. A picture of the Eiffel Tower is shown for scale.

134 in the reconstruction direction by 13%. Our solution for the coincident-event
 135 problem accurately determines the number of muons 98% of the time, which is
 136 an improvement of 86% over the software previously used in IceCube.

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

139 **1. Introduction**

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

147 When a neutrino enters the telescope, it sometimes interacts with the ice and
 148 generates a muon. The neutrino track can be extrapolated from a reconstruction
 149 of the muon track. Muons are also generated by cosmic rays, and separation
 150 of the cosmic ray muons and neutrino muons is a necessary step for neutrino
 151 analysis. This separation is challenging, as the number of observed cosmic
 152 ray muons exceeds the number of observed neutrino muons by five orders of
 153 magnitude [3].

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

164 We examine two problems that arise in the IceCube detector’s separation of
165 cosmic rays muons from neutrino muons:

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

170 The IceCube Collaboration has spent considerable effort on both of these
171 problems over the last decade, as they are a critical step for data analysis.
172 They have developed sophisticated domain models that take into account the
173 interaction of near- and far-field effects of light, and have undertaken complex
174 mapping efforts to understand the effects of photon propagation in the ice [3,
175 4]. Our solutions do not further refine the detailed modeling of these physical
176 effects, but instead augment the models with off-the-shelf statistical techniques
177 combined with some simple data filtering to remove outliers.

178 *Related Work.* Track reconstruction and coincident event detection challenges
179 are ubiquitous in particle physics [5–7], both in particle accelerators and cosmic
180 particle detectors. While the work described in this manuscript builds on the
181 previous technique developed for the IceCube detector [3], our techniques are
182 general purpose, and potentially have applications in detectors beyond IceCube.

183 *Outline.* We begin by describing the background of the IceCube detector in
184 Section 2. In Section 3, we describe the reconstruction pipeline including the
185 prior IceCube software, then we discuss our work and its results. Section 4
186 describes our work on coincident event detection, and follows a parallel structure
187 to Section 3. We describe how in this application, a simple heuristic approach
188 is an improvement over the prior software. We concluded in Section 5.

189 2. Background

190 The IceCube detector is composed of 5,160 optical detectors, each composed
191 of a photomultiplier tube (PMT) and onboard digitizer[8]. The PMTs are spread
192 over 86 vertical strings arranged in a hexagonal shape, with a total instrumented
193 volume of approximately one cubic kilometer. The PMTs on a given string

194 are separated vertically by 17 m, and the string-to-string separation is roughly
195 125 m.

196 As the muon travels through the detector, it radiates light [9], which is
197 observed by the PMTs and broken down into discrete *hits* [10]. A collection of
198 hits is called an *event*, and if the number of hits in an event is sufficiently large,
199 the muon track reconstruction algorithm is triggered.

200 The track reconstruction algorithms used in the detector have several chal-
201 lenges that must be overcome. The underlying mechanics are stochastic and
202 incompletely modeled, the data are noisy and contains outliers, and the com-
203 putational abilities of the detector are limited.

204 *Modeling Difficulties.* The underlying physics of the system is nontrivial to
205 model. The muon’s light is scattered by the dust and air bubbles in the ice
206 medium. This scattering is both complex and stochastic, and the scattering
207 properties of the ice vary with depth [11].

208 *Noise.* The noise inherent in the data is another challenge. The PMTs are so
209 sensitive to light that they can record hits even in the absence of nearby muons.
210 These hits can arise from photons generated either by radioactive decay inside
211 the PMT [12].

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

216 3. Reconstruction Problem

217 By augmenting the reconstruction algorithm with some simple filters and
218 classical data analysis techniques, we show significant improvement in the re-
219 construction algorithm’s accuracy.

220 3.1. Prior IceCube Software

221 The muon track reconstruction process (outlined in Figure 2) starts when the
222 number of detected hits exceeds a preset threshold and initiates data collection.
223 After the initial data are collected, it then passes through a series of simple
224 filters to remove obvious outliers [13].

225 This is followed by a simple reconstruction algorithm, *linefit*, which simply
226 finds the track that minimizes the sum of the squares of the distances between
227 the track and the hits. More formally, assume there are N hits; denote the
228 position and time of the i th hit as \vec{x}_i and t_i , respectively. Let the reconstructed
229 muon track have a velocity of \vec{v} , and let the reconstructed track pass through
230 point \vec{x}_0 at time t_0 . Then linefit reconstruction solves the *least-squares* opti-
231 mization problem

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

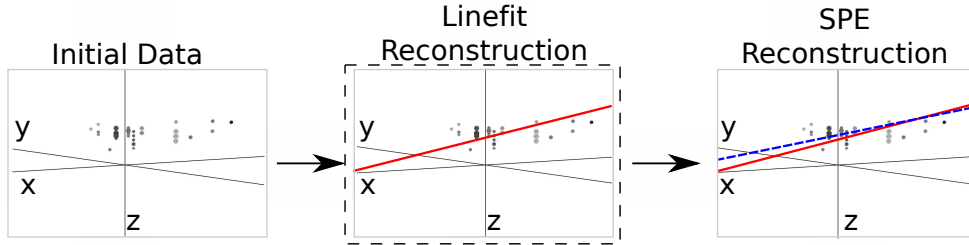


Figure 2: The reconstruction pipeline used to process data in the IceCube detector. After initial data are collected, it is then processed by some simple noise filters, which remove clear outliers. This cleaned data are processed by a simple reconstruction algorithm (solid line), which is used as the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

232 where

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

233 Linefit is primarily used to generate an initial track or *seed* for a more sophis-
 234 ticated reconstruction.

235 The reconstruction algorithm for the sophisticated reconstruction is *Single-*
 236 *Photo-Electron-Fit (SPE)* [3]. SPE uses the least-squares reconstruction, the
 237 event data, and a parameterized probability distribution function of scattering
 238 in ice [3] to reconstruct the muon track.

239 3.2. Algorithm Improvement

240 SPE is dependent on the seed. Given a seed that is inaccurate by 6° or
 241 more, SPE typically cannot recover, and produces a reconstruction with the
 242 same level of inaccuracy as the seed track. In addition, the likelihood space for
 243 SPE can contain multiple local maxima, so improving the accuracy of a seed
 244 that is already near the true solution improves the accuracy of SPE. Thus, we
 245 focused our work on improving the quality of the seed.

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

- 250 1. It ignores the scattering effects of the ice medium. Some of the photons can
 251 scatter for over a microsecond, which means that when they are recorded
 252 by a PMT, the muon will be over 300 m away.
- 253 2. While the noise reduction steps remove most of the outlier noise, the noise
 254 hits that survive can be far from the muon. Since these outliers are given
 255 quadratic weight, they exert a huge influence over the model.

256 The first pitfall occurs because the model is incomplete and does not ac-
 257 curately model the data, and the second demonstrates that the model is not

258 robust to noise. Our solution was twofold: improve the model and increase the
 259 noise robustness by replacing least squares with robust statistical techniques.

260 3.2.1. Improving the Model

261 The least-squares model does not model the scattering, and thus hits gen-
 262 erated by photons that scattered for a significant length of time are not useful
 263 predictors of the muon’s position. We found that a simple filter could iden-
 264 tify these scattered hits, and improve accuracy by of almost a factor of two by
 265 removing them from the dataset.

266 More formally, for each hit h_i , the algorithm looks at all neighboring hits
 267 within a neighborhood of r , and if there exists a neighboring hit h_j with a time
 268 stamp that is t earlier than h_i , then h_i is considered a scattered hit, and is not
 269 used in the simple reconstruction algorithm. Optimal values of r and t were
 270 found to be 156 m and 778 ns by tuning them on simulated muon data.

271 3.2.2. Adding Robustness to Noise

272 As described in equation 1, the least squares model gives outliers quadratic
 273 weight, whereas we would prefer that outliers had zero weight. There are robust
 274 models in classical statistics designed to marginalize outliers. We experimented
 275 replacing the least-squares model with a Huber fit [14], which improved the
 276 reconstruction accuracy.

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

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

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

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

279 Here, $\rho_i(t_0, \vec{x}, \vec{v})$ is defined in Equation 2 and μ is a constant calibrated to the
 280 data (on simulated muon events, the optimal value of μ is 153 m).

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

286 In addition to its attractive robustness properties, the Huber fit’s weight
 287 assignment also has the added benefit that it inherently labels points as outliers
 288 (those with $\rho \geq \mu$). Thus, once the Huber fit is computed, we can go one step
 289 farther and simply remove the labeled outliers from the dataset. A better fit is
 290 then obtained by computing the least-squares fit on the data with the outliers
 291 removed. The entire algorithm has a mean runtime that is approximately six
 292 times slower than Linefit’s mean runtime.

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 simple 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

293 3.3. Results

294 Our goal is to improve the accuracy of the reconstruction in order to better
 295 separate neutrinos from cosmic rays. Thus we present three measurements: (1)
 296 the accuracy change between linefit and the new algorithm, (2) the accuracy
 297 change when SPE is seeded with the new algorithm instead of linefit, and (3)
 298 the improvement in separation between neutrinos and cosmic rays.

299 To measure the improvement generated by our changes, we use the metric of
 300 *median angular resolution* θ_{med} , which is a standard metric used in the collab-
 301 oration. The angular resolution of a reconstruction is the arc-distance between
 302 the reconstruction and the true track. Our dataset is drawn from simulated
 303 neutrino data designed to be similar to that observed by the detector.

304 We can improve the median angular resolution of the simple reconstruction
 305 by 57.6%, as shown in Table 1. Seeding SPE with the improved simple recon-
 306 struction generates an improvement in the angular resolution of 12.9%. These
 307 improvements in the reconstruction algorithm result in 10% fewer atmospheric
 308 muons erroneously reconstructed as up-going, and 1% more muons correctly
 309 reconstructed as up-going.

310 4. Coincident Event Problem

311 In our second experiment, we look at the problem of determining when more
 312 than one muon has entered the detector. In the most common case, a single
 313 muon will pass through the detector and generate an event before exiting. These
 314 events are processed by the pipeline described in Figure 2. However, for roughly
 315 9% of the events collected by the data collection algorithm, more than one muon
 316 will be passing through the detector simultaneously, an occurrence known as a
 317 *coincident event*.

318 One of the primary sources of background noise in the scientific analyses
 319 of the IceCube Collaboration is coincident background muons that have been
 320 erroneously reconstructed as neutrinos. To see why this occurs, consider the
 321 coincident event shown in Figure 3. There are two clear groups of hits; how-
 322 ever, the reconstruction algorithm treats them as a single group, resulting in a
 323 erroneous reconstruction. In the ideal case, the reconstruction algorithm would
 324 identify coincident events and split them, as in Figure 4.

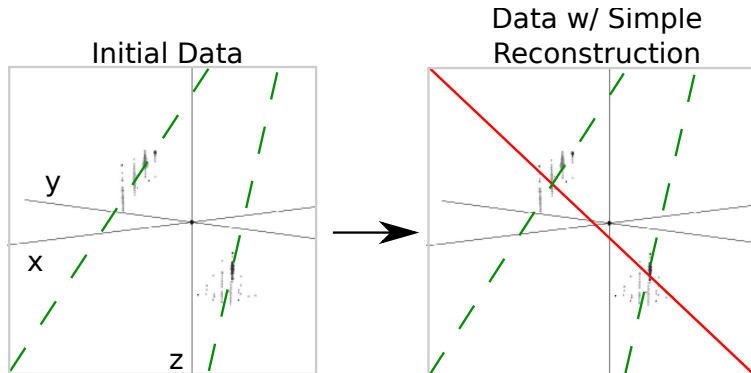


Figure 3: In this example, an event that is clearly composed of two muons (actual tracks shown as dashed lines) is treated as a single muon, and thus the reconstruction (solid line) is inaccurate.

325 The challenge in this example is determining the number of muons in an
 326 event. In our results, we find that a simple spatial clustering algorithm can
 327 solve this classification problem with less than 2% error.

328 4.1. Prior IceCube Software

329 Coincident events have been a concern in the IceCube analysis [15] for years,
 330 and some software has been developed to handle coincident events. As a baseline
 331 of comparison, we use the *TTrigger* software, which is described in [16].

332 4.2. Algorithm Improvement

333 Our solution to this problem is a proximal clustering algorithm. The intu-
 334 ition in proximal clustering is that points local in space and time are probably
 335 from the same muon. The proximal clustering algorithm iterates through each
 336 pair of hits (i, j) and builds an adjacency matrix \mathbf{A} as

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

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

342 4.2.1. Improving the Model

343 When implemented naively, proximal clustering succeeded for the majority
 344 of the events, but failed if there was a gap in the muon track, which can occur
 345 when the muon travels through dusty ice. If there is a significantly large gap,
 346 the algorithm erroneously separates the hits into two clusters.

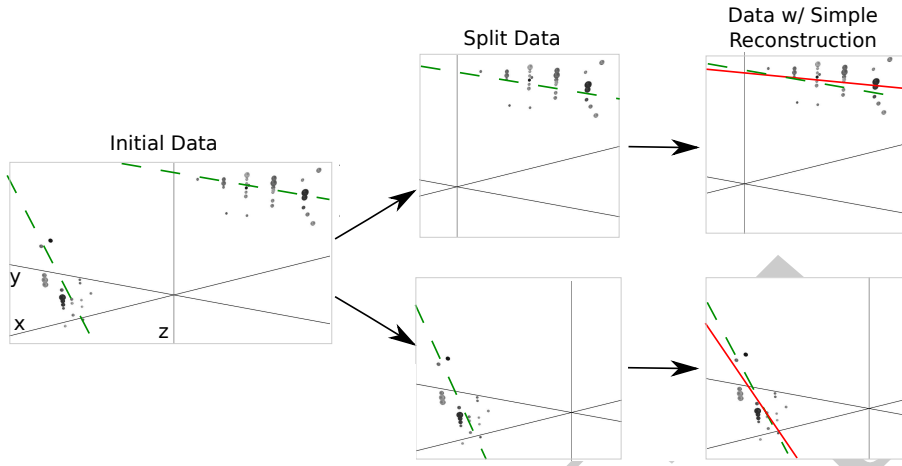


Figure 4: Ideally, the detector would split coincident events before computing the reconstruction. Splitting the event results in more accurate reconstructions (reconstructions shown as solid lines, true muon tracks shown as dashed lines). Note the difference in the reconstructions compared with Figure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

347 To get around this, an additional heuristic is added, *track connecting*. After
 348 the data segmentation is finished, track connecting determines if separate
 349 clusters should be combined. It computes the mean position and time of each
 350 cluster, and connects a hypothetical muon track T between each pair of sub-
 351 spaces.

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

355 4.2.2. Adding Robustness to Noise

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

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

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

368 4.3. Results

369 There were two competing goals for coincident event detection algorithms:
 370 the algorithm should be conservative enough that events containing single tracks

Table 2: Error Rates for Classification Algorithms

Algorithm	$E_{\text{Single}} \%$	$E_{\text{Multiple}} \%$	$E_{\text{tot}} \%$
Trivial	0.0	100.0	8.3
TTrigger	11.5	31.8	13.2
Proximal clustering	0.2	18.9	1.8

371 are not erroneously split, and aggressive enough that a useful fraction of coin-
372 cident events are split correctly. Erroneously discarding events containing neu-
373 trinos is worse than erroneously allowing additional noise into the data pool,
374 as noise can be eliminated by future filtering of the data pool. Our algorithm
375 is tuned to keep almost all of the single events correctly unsplit, while still
376 correctly splitting 80% of the coincident events.

377 4.3.1. Measurements

378 We modified the reconstruction pipeline shown in Figure 2, in between the
379 noise cleaning and the simple reconstruction, by adding a step for coincident
380 event detection, as shown in Figure 4. This step takes cleaned data and attempts
381 to classify the event as a single-track or multiple-track event.

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

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

387 For computing E_{tot} , we use $w_{\text{Single}} = 0.917$ and $w_{\text{Multiple}} = 0.083$, which is
388 the frequency in which single-muon and multiple-muon events appear in data
389 simulating the distribution of events that trigger the reconstruction algorithm.

390 We present our results for the coincident event problem by measuring how
391 well each algorithm performs at determining the number of subspaces in an
392 event.

393 There are two natural comparisons for our work: the prior software TTrigger,
394 as well as the trivial algorithm, which always classifies each event as a single-
395 track event. Clearly, the latter will always get the single-track events correct,
396 and always get the multiple-track events wrong. We provide a comparison of
397 these techniques in Table 2. As shown, our software classifies the number of
398 muons in the detector 86% better than TTrigger.

399 5. Conclusions

400 The challenges in the IceCube detector are complex. Despite this complexity,
401 we found that we can achieve significant improvement via classical data analysis
402 algorithms and simple models.

403 We looked at the problem of general reconstruction improvement, and found
404 that by applying a simple filter to the data and adding some robustness to the
405 fitting algorithm, we got superior reconstructions in the noisy environments of
406 the IceCube data. Our reconstruction software runs on-site, and is included in
407 all IceCube analysis.

408 We also looked at the problem of determining the number of muons in the de-
409 tector. We found that proximal clustering with some simple heuristics could cor-
410 rectly determine whether an event contained a single muon or multiple muons,
411 and this was an 86% improvement over the prior software.

412 References

- 413 [1] IceCube Collaboration, IceCube webpage, <http://icecube.wisc.edu/>.
- 414 [2] IceCube Collaboration, First year performance of the IceCube neutrino
415 telescope, *Astroparticle Physics* 26 (3) (2006) 155–173.
- 416 [3] IceCube Collaboration, Muon track reconstruction and data selection tech-
417 niques in AMANDA, *Nuclear Instruments and Methods in Physics Re-*
418 *search Section A* 524 (2004) 169–194.
- 419 [4] IceCube Collaboration, Measurement of South Pole ice transparency with
420 the IceCube LED calibration system IceCube Collaboration, *Nuclear In-*
421 *struments and Methods in Physics Research Section A*.
- 422 [5] ATLAS Collaboration, Tracking and vertexing with the ATLAS detector at
423 the LHC, *Nuclear Instruments and Methods in Physics Research Section A:*
424 *Accelerators, Spectrometers, Detectors and Associated Equipment* 650 (1)
425 (2011) 218–223.
- 426 [6] R. S. Chivukulaa, M. Goldena, E. H. Simmons, Multi-jet physics at hadron
427 colliders, *Nuclear Physics B* 363 (1) (1991) 83–96.
- 428 [7] S. Ellis, J. Huston, K. Hatakeyama, P. Loch, M. Tönnesmann, Jets in
429 hadron–hadron collisions, *Progress in Particle and Nuclear Physics* (60)
430 (2008) 484–551.
- 431 [8] IceCube Collaboration, Calibration and characterization of the IceCube
432 photomultiplier tube, *Nuclear Instruments and Methods in Physics Re-*
433 *search Section A* 618 (2010) 139–152.
- 434 [9] IceCube Collaboration, An improved method for measuring muon energy
435 using the truncated mean of dE/dx , *Nuclear Instruments and Methods in*
436 *Physics Research Section A*.
- 437 [10] IceCube Collaboration, The icecube data acquisition system: Signal cap-
438 ture, digitization, and timestamping, *Nuclear Instruments and Methods in*
439 *Physics Research Section A* 601 (3) (2009) 294–316.

- 440 [11] M. Wolf, E. Resconi, Verification of South Pole glacial ice simulations in Ice-
441 Cube and its relation to conventional and new, accelerated photon tracking
442 techniques, Master's thesis, Max-Planck-Institut für Kernphysik Heidelberg
443 (September 2010).
- 444 [12] IceCube Collaboration, IceCube sensitivity for low-energy neutrinos from
445 nearby supernovae, *Astronomy & Astrophysics* 535 (A109) (2011) 18.
- 446 [13] M. Ackermann, Searches for signals from cosmic point-like sources of high
447 energy neutrinos in 5 years of AMANDA-II data, Ph.D. thesis, Humboldt-
448 Universität zu Berlin (2006).
- 449 [14] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University
450 Press, 2009.
- 451 [15] IceCube Collaboration, Measurement of the atmospheric neutrino energy
452 spectrum from 100 GeV to 400 TeV with IceCube, *Physical Review D*
453 83 (1).
- 454 [16] D. Chirkin, Measurement of the atmospheric neutrino energy spectrum
455 with IceCube, Proceedings of the 31st ICRC.