

Improvement in Fast Particle 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. Bohm³⁴, 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,

*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

⁴Department of Physics, Sungkyunkwan University, Suwon 440-746, Korea

⁵NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

36 M. Merck²⁷, P. Mészáros^{37,38}, T. Meures¹², S. Miarecki^{8,7}, E. Middell⁴¹,
 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. Pfindner¹⁷, 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^{17,5}, T. Stanev³¹, A. Stasik⁴¹,
 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 ²¹Département de physique nucléaire et corpusculaire, Université de Genève, CH-1211
88 Genève, Switzerland
89 ²²Dept. of Physics and Astronomy, University of Gent, B-9000 Gent, Belgium
90 ²³Dept. of Physics and Astronomy, University of California, Irvine, CA 92697, USA
91 ²⁴Laboratory for High Energy Physics, École Polytechnique Fédérale, CH-1015 Lausanne,
92 Switzerland
93 ²⁵Dept. of Physics and Astronomy, University of Kansas, Lawrence, KS 66045, USA
94 ²⁶Dept. of Astronomy, University of Wisconsin, Madison, WI 53706, USA
95 ²⁷Dept. of Physics and Wisconsin IceCube Particle Astrophysics Center, University of
96 Wisconsin, Madison, WI 53706, USA
97 ²⁸Institute of Physics, University of Mainz, Staudinger Weg 7, D-55099 Mainz, Germany
98 ²⁹Université de Mons, 7000 Mons, Belgium
99 ³⁰T.U. Munich, D-85748 Garching, Germany
100 ³¹Bartol Research Institute and Department of Physics and Astronomy, University of
101 Delaware, Newark, DE 19716, USA
102 ³²Dept. of Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK
103 ³³Dept. of Physics, University of Wisconsin, River Falls, WI 54022, USA
104 ³⁴Oskar Klein Centre and Dept. of Physics, Stockholm University, SE-10691 Stockholm,
105 Sweden
106 ³⁵Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY
107 11794-3800, USA
108 ³⁶Dept. of Physics and Astronomy, University of Alabama, Tuscaloosa, AL 35487, USA
109 ³⁷Dept. of Astronomy and Astrophysics, Pennsylvania State University, University Park,
110 PA 16802, USA
111 ³⁸Dept. of Physics, Pennsylvania State University, University Park, PA 16802, USA
112 ³⁹Dept. of Physics and Astronomy, Uppsala University, Box 516, S-75120 Uppsala, Sweden
113 ⁴⁰Dept. of Physics, University of Wuppertal, D-42119 Wuppertal, Germany
114 ⁴¹DESY, D-15735 Zeuthen, Germany
115 ⁴²Dept. of Computer Sciences, University of Wisconsin, Madison, WI 53706, USA

116 **Abstract**

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

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

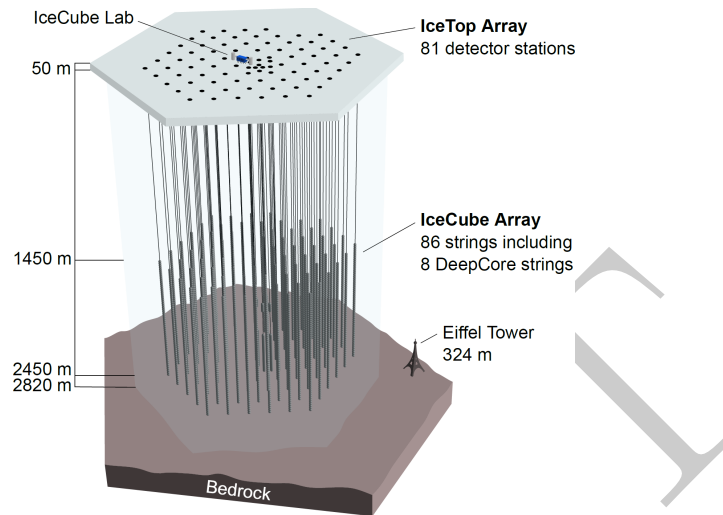


Figure 1: The IceCube neutrino detector in the Antarctic ice. A picture of the Eiffel Tower is shown for scale.

133 standard metric for track reconstruction, we improve the accuracy in the initial
 134 reconstruction direction by 13%. We also present improvements in measuring
 135 the number of muons in coincident events: we can accurately determine the
 136 number of muons 98% of the time, which is an improvement of 86% over the
 137 software previously used in IceCube.

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
157 muon was traveling downwards into the Earth or upwards out of the Earth.
158 Since neutrinos can penetrate through the Earth but cosmic ray muons cannot,
159 it 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. Since 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 onboard digitizer [9]. The PMTs are spread
196 over 86 vertical strings arranged in a hexagonal shape, with a total instrumented
197 volume of approximately one cubic kilometer. The PMTs on a given string
198 are separated vertically by 17 m, and the string-to-string separation is roughly
199 125 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) fired within 5 microseconds. When a trigger occurs, all data within
206 a 10 microsecond trigger window is saved, becoming an *event*. If the number of
207 hits in an event is sufficiently large, the muon track reconstruction algorithm is
208 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].

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

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

224 3. Reconstruction Improvement

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

228 3.1. Prior IceCube Software

229 The muon track reconstruction process (outlined in Figure 2) starts when the
230 number of detected hits exceeds a preset threshold and initiates data collection.

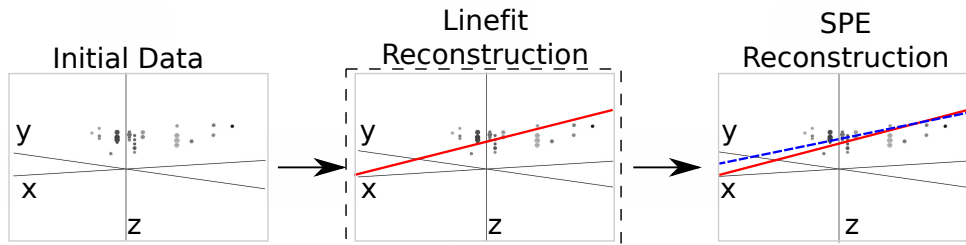


Figure 2: The reconstruction pipeline used to process data in the IceCube detector. After initial data are collected, it is then processed by some basic noise filters, which remove clear outliers. This cleaned data are processed by a basic reconstruction algorithm (solid line), which is used as the seed 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).

231 After the initial data are collected, the event then passes through a series of
 232 basic filters to remove obvious outliers [11].

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

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

240 where

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

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

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

250 3.2. Algorithm Improvement

251 SPE fit is dependent on the seed. Given a seed that is inaccurate by 6° or
 252 more, SPE fit typically cannot recover, and produces a reconstruction with the
 253 same level of inaccuracy as the seed track. In addition, the likelihood space for

254 SPE fit can contain multiple local maxima, so improving the accuracy of a seed
 255 that is already near the true solution improves the accuracy of SPE fit. Thus,
 256 the work here focuses on improving the quality of the seed.

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

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

267 The first pitfall occurs because the model is incomplete and does not accu-
 268 rately model the data, and the second demonstrates that the model is not robust
 269 to noise. The solution to this is twofold: improve the model and increase the
 270 noise robustness by replacing least squares with robust statistical techniques.

271 3.2.1. Improving the Model

272 The least-squares model does not model the scattering, and thus hits gen-
 273 erated by photons that scattered for a significant length of time are not useful
 274 predictors of the muon’s position. We found that a basic filter could identify
 275 these scattered hits, and improve accuracy by of almost a factor of two by
 276 removing them from the dataset.

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

283 3.2.2. Adding Robustness to Noise

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

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

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

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

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

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

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

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

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

306 3.3. Results

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

312 To measure the improvement generated by the changes, we use the metric of
 313 *median angular resolution* θ_{med} , which is a standard metric within the collab-
 314 oration. The angular resolution of a reconstruction is the arc-distance between
 315 the reconstruction and the true track. The dataset is drawn from simulated
 316 neutrino data designed to be similar to that observed by the detector.

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

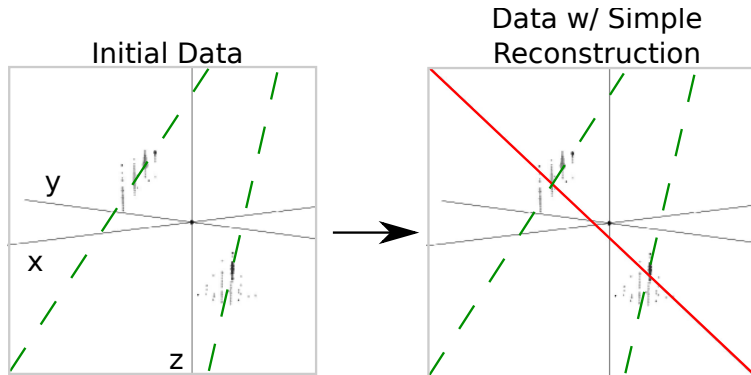


Figure 3: In this example, an event that is clearly composed of two muons (actual tracks shown as dashed lines) is treated as a single muon, and thus the reconstruction (solid line) is inaccurate.

323 4. Coincident Event Improvements

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

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

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

341 4.1. Prior IceCube Software

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

345 4.2. Algorithm Improvement

346 Here we present a proximal clustering algorithm. The intuition in proximal
 347 clustering is that points local in space and time are probably from the same

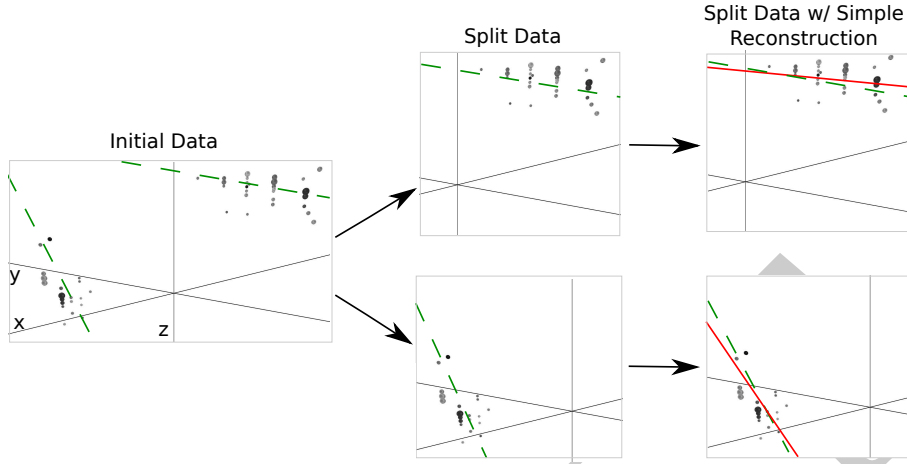


Figure 4: Ideally, the detector would split coincident events before computing the reconstruction. Splitting the event results in more accurate reconstructions (reconstructions shown as solid lines, true muon tracks shown as dashed lines). Note the difference in the reconstructions compared with Figure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

348 muon. The proximal clustering algorithm iterates through each pair of hits
 349 (i, j) and builds an adjacency matrix \mathbf{A} as

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

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

355 4.2.1. Improving the Model

356 When implemented naively, proximal clustering succeeded for the majority
 357 of the events, but failed if there was a gap in the muon track, which can occur
 358 when the muon travels through dusty ice layers with short scattering length. If
 359 there is a significantly large gap, the algorithm erroneously separates the hits
 360 into two clusters.

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

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

368 *4.2.2. Adding Robustness to Noise*

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

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

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

381 *4.3. Results*

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

388 *4.3.1. Measurements*

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

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

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

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

401 We present the results for the coincident event problem by measuring how
402 well each algorithm performs at determining the number of subspaces in an
403 event.

404 There are two natural comparisons for the work: the prior software TTrigger,
405 as well as the trivial algorithm, which always classifies each event as a single-
406 track event. Clearly, the latter will always get the single-track events correct,
407 and always get the multiple-track events wrong. We provide a comparison of
408 these techniques in Table 2. As shown, the new algorithm classifies the number
409 of muons in the detector 86% better than TTrigger.

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

410 5. Conclusions

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

417 We also looked at the problem of determining the number of muons in the
 418 detector. We found that proximal clustering with some basic heuristics could
 419 correctly determine whether an event contained a single muon or multiple muons
 420 with less than 2% error, yielding an 86% improvement over the prior software.

421 References

- 422 [1] IceCube Collaboration, IceCube webpage, <http://icecube.wisc.edu/>.
- 423 [2] IceCube Collaboration, First year performance of the IceCube neutrino
 424 telescope, *Astroparticle Physics* 26 (3) (2006) 155–173.
- 425 [3] IceCube Collaboration, Muon track reconstruction and data selection tech-
 426 niques in AMANDA, *Nuclear Instruments and Methods in Physics Re-*
 427 *search Section A* 524 (2004) 169–194.
- 428 [4] IceCube Collaboration, Measurement of South Pole ice transparency with
 429 the IceCube LED calibration system IceCube Collaboration, *Nuclear In-*
 430 *struments and Methods in Physics Research Section A* (2013) 73–89.
- 431 [5] The AMANDA Collaboration, Optical properties of deep glacial ice at the
 432 south pole, *Journal of Geophysical Research* 111 (D13) (2006) D13203.
- 433 [6] ATLAS Collaboration, Tracking and vertexing with the ATLAS detector at
 434 the LHC, *Nuclear Instruments and Methods in Physics Research Section A:*
 435 *Accelerators, Spectrometers, Detectors and Associated Equipment* 650 (1)
 436 (2011) 218–223.
- 437 [7] R. S. Chivukulaa, M. Goldenaa, E. H. Simmons, Multi-jet physics at hadron
 438 colliders, *Nuclear Physics B* 363 (1) (1991) 83–96.
- 439 [8] S. Ellis, J. Huston, K. Hatakeyama, P. Loch, M. Tönnesmann, Jets in
 440 hadron–hadron collisions, *Progress in Particle and Nuclear Physics* (60)
 441 (2008) 484–551.

- 442 [9] IceCube Collaboration, Calibration and characterization of the IceCube
443 photomultiplier tube, Nuclear Instruments and Methods in Physics Re-
444 search Section A 618 (2010) 139–152.
- 445 [10] IceCube Collaboration, The icecube data acquisition system: Signal cap-
446 ture, digitization, and timestamping, Nuclear Instruments and Methods in
447 Physics Research Section A 601 (3) (2009) 294–316.
- 448 [11] M. Ackermann, Searches for signals from cosmic point-like sources of high
449 energy neutrinos in 5 years of AMANDA-II data, Ph.D. thesis, Humboldt-
450 Universität zu Berlin (2006).
- 451 [12] V. Stenger, Track fitting for DUMAND-II Octagon Array, Tech. rep., Uni-
452 versity of Hawai'i at Manoa (1990).
- 453 [13] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University
454 Press, 2009.
- 455 [14] IceCube Collaboration, Measurement of the atmospheric neutrino energy
456 spectrum from 100 GeV to 400 TeV with IceCube, Physical Review D
457 83 (1).
- 458 [15] D. Chirkin, Measurement of the atmospheric neutrino energy spectrum
459 with IceCube, Proceedings of the 31st ICRC.