

Measurement of the mean number of muons with energies above 500 GeV in air showers detected with the IceCube Neutrino Observatory

R. Abbasi,¹⁶ M. Ackermann,⁶³ J. Adams,¹⁷ S. K. Agarwalla,^{39,*} J. A. Aguilar,¹⁰ M. Ahlers,²¹ J.M. Alameddine,²² N. M. Amin,⁴³ K. Andeen,⁴¹ C. Argüelles,¹³ Y. Ashida,⁵² S. Athanasiadou,⁶³ S. N. Axani,⁴³ R. Babu,²³ X. Bai,⁴⁹ J. Baines-Holmes,³⁹ A. Balagopal V.,³⁹ S. W. Barwick,²⁹ S. Bash,²⁶ V. Basu,⁵² R. Bay,⁶ J. J. Beatty,^{19,20} J. Becker Tjus,^{9,†} P. Behrens,¹ J. Beise,⁶¹ C. Bellenghi,²⁶ B. Benkel,⁶³ S. BenZvi,⁵¹ D. Berley,¹⁸ E. Bernardini,^{47,‡} D. Z. Besson,³⁵ E. Blaufuss,¹⁸ L. Bloom,⁵⁸ S. Blot,⁶³ I. Bodo,³⁹ F. Bontempo,³⁰ J. Y. Book Motzkin,¹³ C. Boscolo Meneguolo,^{47,‡} S. Böser,⁴⁰ O. Botner,⁶¹ J. Böttcher,¹ J. Braun,³⁹ B. Brinson,⁴ Z. Brisson-Tsavoussis,³² R. T. Burley,² D. Butterfield,³⁹ M. A. Campana,⁴⁸ K. Carloni,¹³ J. Carpio,^{33,34} S. Chattopadhyay,^{39,*} N. Chau,¹⁰ Z. Chen,⁵⁵ D. Chirkin,³⁹ S. Choi,⁵² B. A. Clark,¹⁸ A. Coleman,⁶¹ P. Coleman,¹ G. H. Collin,¹⁴ A. Connolly,^{19,20} J. M. Conrad,¹⁴ R. Corley,⁵² D. F. Cowen,^{59,60} C. De Clercq,¹¹ J. J. DeLaunay,⁵⁹ D. Delgado,¹³ T. Delmeulle,¹⁰ S. Deng,¹ P. Desiati,³⁹ K. D. de Vries,¹¹ G. de Wasseige,³⁶ T. DeYoung,²³ J. C. Díaz-Vélez,³⁹ S. DiKerby,²³ M. Dittmer,⁴² A. Domi,²⁵ L. Draper,⁵² L. Dueser,¹ D. Durnford,²⁴ K. Dutta,⁴⁰ M. A. DuVernois,³⁹ T. Ehrhardt,⁴⁰ L. Eidenschink,²⁶ A. Eimer,²⁵ P. Eller,²⁶ E. Ellinger,⁶² D. Elsässer,²² R. Engel,^{30,31} H. Erpenbeck,³⁹ W. Esmail,⁴² S. Eulig,¹³ J. Evans,¹⁸ P. A. Evenson,⁴³ K. L. Fan,¹⁸ K. Fang,³⁹ K. Farrag,¹⁵ A. R. Fazely,⁵ A. Fedynitch,⁵⁷ N. Feigl,⁸ C. Finley,⁵⁴ L. Fischer,⁶³ D. Fox,⁵⁹ A. Franckowiak,⁹ S. Fukami,⁶³ P. Fürst,¹ J. Gallagher,³⁸ E. Ganster,¹ A. Garcia,¹³ M. Garcia,⁴³ G. Garg,^{39,*} E. Genton,^{13,36} L. Gerhardt,⁷ A. Ghadimi,⁵⁸ C. Glaser,⁶¹ T. Glüsenkamp,⁶¹ J. G. Gonzalez,⁴³ S. Goswami,^{33,34} A. Granados,²³ D. Grant,¹² S. J. Gray,¹⁸ S. Griffin,³⁹ S. Griswold,⁵¹ K. M. Groth,²¹ D. Guevel,³⁹ C. Günther,¹ P. Gutjahr,²² C. Ha,⁵³ C. Haack,²⁵ A. Hallgren,⁶¹ L. Halve,¹ F. Halzen,³⁹ L. Hamacher,¹ M. Ha Minh,²⁶ M. Handt,¹ K. Hanson,³⁹ J. Hardin,¹⁴ A. A. Harnisch,²³ P. Hatch,³² A. Haungs,³⁰ J. Häußler,¹ K. Helbing,⁶² J. Hellrung,⁹ L. Hennig,²⁵ L. Heuermann,¹ R. Hewett,¹⁷ N. Heyer,⁶¹ S. Hickford,⁶² A. Hidvegi,⁵⁴ C. Hill,¹⁵ G. C. Hill,² R. Hmaid,¹⁵ K. D. Hoffman,¹⁸ D. Hooper,³⁹ S. Hori,³⁹ K. Hoshina,^{39,§} M. Hostert,¹³ W. Hou,³⁰ T. Huber,³⁰ K. Hultqvist,⁵⁴ K. Hymon,^{22,57} A. Ishihara,¹⁵ W. Iwakiri,¹⁵ M. Jacquart,²¹ S. Jain,³⁹ O. Janik,²⁵ M. Jeong,⁵² M. Jin,¹³ N. Kamp,¹³ D. Kang,³⁰ X. Kang,⁴⁸ A. Kappes,⁴² L. Kardum,²² T. Karg,⁶³ M. Karl,²⁶ A. Karle,³⁹ A. Katil,²⁴ M. Kauer,³⁹ J. L. Kelley,³⁹ M. Khanal,⁵² A. Khattee Zathul,³⁹ A. Kheirandish,^{33,34} H. Kimku,⁵³ J. Kiryluk,⁵⁵ C. Klein,²⁵ S. R. Klein,^{6,7} Y. Kobayashi,¹⁵ A. Kochocki,²³ R. Koirala,⁴³ H. Kolanoski,⁸ T. Kontrimas,²⁶ L. Köpke,⁴⁰ C. Kopper,²⁵ D. J. Koskinen,²¹ P. Koundal,⁴³ M. Kowalski,^{8,63} T. Kozynets,²¹ N. Krieger,⁹ J. Krishnamoorthi,^{39,*} T. Krishnan,¹³ K. Kruiswijk,³⁶ E. Krupczak,²³ A. Kumar,⁶³ E. Kun,⁹ N. Kurahashi,⁴⁸ N. Lad,⁶³ C. Lagunas Gualda,²⁶ L. Lallement Arnaud,¹⁰ M. Lamoureux,³⁶ M. J. Larson,¹⁸ F. Lauber,⁶² J. P. Lazar,³⁶ K. Leonard DeHolton,⁶⁰ A. Leszczyńska,⁴³ J. Liao,⁴ Y. T. Liu,⁶⁰ M. Liubarska,²⁴ C. Love,⁴⁸ L. Lu,³⁹ F. Lucarelli,²⁷ W. Luszczak,^{19,20} Y. Lyu,^{6,7} J. Madsen,³⁹ E. Magnus,¹¹ K. B. M. Mahn,²³ Y. Makino,³⁹ E. Manao,²⁶ S. Mancina,^{47,¶} A. Mand,³⁹ I. C. Maris,¹⁰ S. Marka,⁴⁵ Z. Marka,⁴⁵ L. Marten,¹ I. Martinez-Soler,¹³ R. Maruyama,⁴⁴ F. Mayhew,²³ F. McNally,³⁷ J. V. Mead,²¹ K. Meagher,³⁹ S. Mechbal,⁶³ A. Medina,²⁰ M. Meier,¹⁵ Y. Merckx,¹¹ L. Merten,⁹ J. Mitchell,⁵ L. Molchany,⁴⁹ T. Montaruli,²⁷ R. W. Moore,²⁴ Y. Morii,¹⁵ A. Mosbrugger,²⁵ M. Moulai,³⁹ D. Mousadi,⁶³ T. Mukherjee,³⁰ R. Naab,⁶³ M. Nakos,³⁹ U. Naumann,⁶² J. Necker,⁶³ L. Neste,⁵⁴ M. Neumann,⁴² H. Niederhausen,²³ M. U. Nisa,²³ K. Noda,¹⁵ A. Noell,¹ A. Novikov,⁴³ A. Obertacke Pollmann,¹⁵ V. O'Dell,³⁹ A. Olivas,¹⁸ R. Orsoe,²⁶ J. Osborn,³⁹ E. O'Sullivan,⁶¹ V. Palusova,⁴⁰ H. Pandya,⁴³ A. Parenti,¹⁰ N. Park,³² V. Parrish,²³ E. N. Paudel,⁵⁸ L. Paul,⁴⁹ C. Pérez de los Heros,⁶¹ T. Pernice,⁶³ J. Peterson,³⁹ M. Plum,⁴⁹ A. Pontén,⁶¹ V. Poojyam,⁵⁸ Y. Popovych,⁴⁰ M. Prado Rodriguez,³⁹ B. Pries,²³ R. Procter-Murphy,¹⁸ G. T. Przybylski,⁷ L. Pyras,⁵² C. Raab,³⁶ J. Rack-Helleis,⁴⁰ N. Rad,⁶³ M. Ravn,⁶¹ K. Rawlins,³ Z. Rechav,³⁹ A. Rehman,⁴³ I. Reistroffer,⁴⁹ E. Resconi,²⁶ S. Reusch,⁶³ C. D. Rho,⁵⁶ W. Rhode,²² B. Riedel,³⁹ A. Rifaie,⁶² E. J. Roberts,² S. Robertson,^{6,7} M. Rongen,²⁵ A. Rosted,¹⁵ C. Rott,⁵² T. Ruhe,²² L. Ruohan,²⁶ J. Saffer,³¹ D. Salazar-Gallegos,²³ P. Sampathkumar,³⁰ A. Sandrock,⁶² G. Sanger-Johnson,²³ M. Santander,⁵⁸ S. Sarkar,⁴⁶ J. Savelberg,¹ P. Schaile,²⁶ M. Schaufel,¹ H. Schieler,³⁰ S. Schindler,²⁵ L. Schlickmann,⁴⁰ B. Schlüter,⁴² F. Schlüter,¹⁰ N. Schmeisser,⁶² T. Schmidt,¹⁸ F. G. Schröder,^{30,43} L. Schumacher,²⁵ S. Schwirn,¹ S. Sclafani,¹⁸ D. Seckel,⁴³ L. Seen,³⁹ M. Seikh,³⁵ S. Seunarine,⁵⁰ P. A. Sevlé Myhr,³⁶ R. Shah,⁴⁸ S. Shefali,³¹ N. Shimizu,¹⁵ B. Skrzypek,⁶ R. Snihur,³⁹ J. Soedingrekso,²² A. Sogaard,²¹ D. Soldin,⁵² P. Soldin,¹ G. Sommani,⁹ C. Spannfellner,²⁶ G. M. Spiczak,⁵⁰ C. Spiering,⁶³ J. Stachurska,²⁸ M. Stamatikos,²⁰ T. Stanev,⁴³ T. Stezelberger,⁷ T. Stürwald,⁶² T. Stuttard,²¹ G. W. Sullivan,¹⁸ I. Taboada,⁴ S. Ter-Antonyan,⁵ A. Terliuk,²⁶ A. Thakuri,⁴⁹ M. Thiesmeyer,³⁹ W. G. Thompson,¹³ J. Thwaites,³⁹ S. Tilav,⁴³ K. Tollefson,²³ S. Toscano,¹⁰ D. Tosi,³⁹ A. Trettin,⁶³ A. K. Upadhyay,^{39,*} K. Upshaw,⁵ A. Vaidyanathan,⁴¹ N. Valtonen-Mattila,^{9,61} J. Valverde,⁴¹ J. Vandenbroucke,³⁹ T. Van Eeden,⁶³ N. van Eijndhoven,¹¹ J. van Santen,⁶³ J. Vara,⁴² F. Varsi,³¹ M. Venugopal,³⁰ M. Vereecken,³⁶ S. Vergara Carrasco,¹⁷ S. Verpoest,⁴³ D.

Veske,⁴⁵ A. Vijai,¹⁸ J. Villarreal,¹⁴ C. Walck,⁵⁴ A. Wang,⁴ E. Warrick,⁵⁸ C. Weaver,²³ P. Weigel,¹⁴ A. Weindl,³⁰ A. Y. Wen,¹³ C. Wendt,³⁹ J. Werthebach,²² M. Weyrauch,³⁰ N. Whitehorn,²³ C. H. Wiebusch,¹ D. R. Williams,⁵⁸ L. Witthaus,²² M. Wolf,²⁶ G. Wrede,²⁵ X. W. Xu,⁵ J. P. Yan,²⁴ Y. Yao,³⁹ E. Yildizci,³⁹ S. Yoshida,¹⁵ R. Young,³⁵ F. Yu,¹³ S. Yu,⁵² T. Yuan,³⁹ A. Zegarelli,⁹ S. Zhang,²³ Z. Zhang,⁵⁵ P. Zhelnin,¹³ and P. Zilberman³⁹

(IceCube Collaboration)

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

²*Department of Physics, University of Adelaide, Adelaide, 5005, Australia*

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

⁴*School of Physics and Center for Relativistic Astrophysics, Georgia Institute of Technology, Atlanta, GA 30332, USA*

⁵*Dept. of Physics, Southern University, Baton Rouge, LA 70813, USA*

⁶*Dept. of Physics, University of California, Berkeley, CA 94720, USA*

⁷*Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

⁸*Institut für Physik, Humboldt-Universität zu Berlin, D-12489 Berlin, Germany*

⁹*Fakultät für Physik & Astronomie, Ruhr-Universität Bochum, D-44780 Bochum, Germany*

¹⁰*Université Libre de Bruxelles, Science Faculty CP230, B-1050 Brussels, Belgium*

¹¹*Vrije Universiteit Brussel (VUB), Dienst ELEM, B-1050 Brussels, Belgium*

¹²*Dept. of Physics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada*

¹³*Department of Physics and Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge, MA 02138, USA*

¹⁴*Dept. of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

¹⁵*Dept. of Physics and The International Center for Hadron Astrophysics, Chiba University, Chiba 263-8522, Japan*

¹⁶*Department of Physics, Loyola University Chicago, Chicago, IL 60660, USA*

¹⁷*Dept. of Physics and Astronomy, University of Canterbury, Private Bag 4800, Christchurch, New Zealand*

¹⁸*Dept. of Physics, University of Maryland, College Park, MD 20742, USA*

¹⁹*Dept. of Astronomy, Ohio State University, Columbus, OH 43210, USA*

²⁰*Dept. of Physics and Center for Cosmology and Astro-Particle Physics, Ohio State University, Columbus, OH 43210, USA*

²¹*Niels Bohr Institute, University of Copenhagen, DK-2100 Copenhagen, Denmark*

²²*Dept. of Physics, TU Dortmund University, D-44221 Dortmund, Germany*

²³*Dept. of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA*

²⁴*Dept. of Physics, University of Alberta, Edmonton, Alberta, T6G 2E1, Canada*

²⁵*Erlangen Centre for Astroparticle Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg, D-91058 Erlangen, Germany*

²⁶*Physik-department, Technische Universität München, D-85748 Garching, Germany*

²⁷*Département de physique nucléaire et corpusculaire, Université de Genève, CH-1211 Genève, Switzerland*

²⁸*Dept. of Physics and Astronomy, University of Gent, B-9000 Gent, Belgium*

²⁹*Dept. of Physics and Astronomy, University of California, Irvine, CA 92697, USA*

³⁰*Karlsruhe Institute of Technology, Institute for Astroparticle Physics, D-76021 Karlsruhe, Germany*

³¹*Karlsruhe Institute of Technology, Institute of Experimental Particle Physics, D-76021 Karlsruhe, Germany*

³²*Dept. of Physics, Engineering Physics, and Astronomy, Queen's University, Kingston, ON K7L 3N6, Canada*

³³*Department of Physics & Astronomy, University of Nevada, Las Vegas, NV 89154, USA*

³⁴*Nevada Center for Astrophysics, University of Nevada, Las Vegas, NV 89154, USA*

³⁵*Dept. of Physics and Astronomy, University of Kansas, Lawrence, KS 66045, USA*

³⁶*Centre for Cosmology, Particle Physics and Phenomenology - CP3, Université catholique de Louvain, Louvain-la-Neuve, Belgium*

³⁷*Department of Physics, Mercer University, Macon, GA 31207-0001, USA*

³⁸*Dept. of Astronomy, University of Wisconsin—Madison, Madison, WI 53706, USA*

³⁹*Dept. of Physics and Wisconsin IceCube Particle Astrophysics Center, University of Wisconsin—Madison, Madison, WI 53706, USA*

⁴⁰*Institute of Physics, University of Mainz, Staudinger Weg 7, D-55099 Mainz, Germany*

⁴¹*Department of Physics, Marquette University, Milwaukee, WI 53201, USA*

⁴²*Institut für Kernphysik, Universität Münster, D-48149 Münster, Germany*

⁴³*Bartol Research Institute and Dept. of Physics and Astronomy, University of Delaware, Newark, DE 19716, USA*

⁴⁴*Dept. of Physics, Yale University, New Haven, CT 06520, USA*

⁴⁵*Columbia Astrophysics and Nevis Laboratories, Columbia University, New York, NY 10027, USA*

⁴⁶*Dept. of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom*

⁴⁷*Dipartimento di Fisica e Astronomia Galileo Galilei,
Università Degli Studi di Padova, I-35122 Padova PD, Italy*

⁴⁸*Dept. of Physics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA*

⁴⁹*Physics Department, South Dakota School of Mines and Technology, Rapid City, SD 57701, USA*

⁵⁰*Dept. of Physics, University of Wisconsin, River Falls, WI 54022, USA*

⁵¹*Dept. of Physics and Astronomy, University of Rochester, Rochester, NY 14627, USA*

⁵²*Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA*

⁵³*Dept. of Physics, Chung-Ang University, Seoul 06974, Republic of Korea*

⁵⁴*Oskar Klein Centre and Dept. of Physics, Stockholm University, SE-10691 Stockholm, Sweden*

⁵⁵*Dept. of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794-3800, USA*

⁵⁶*Dept. of Physics, Sungkyunkwan University, Suwon 16419, Republic of Korea*

⁵⁷*Institute of Physics, Academia Sinica, Taipei, 11529, Taiwan*

⁵⁸*Dept. of Physics and Astronomy, University of Alabama, Tuscaloosa, AL 35487, USA*

⁵⁹*Dept. of Astronomy and Astrophysics, Pennsylvania State University, University Park, PA 16802, USA*

⁶⁰*Dept. of Physics, Pennsylvania State University, University Park, PA 16802, USA*

⁶¹*Dept. of Physics and Astronomy, Uppsala University, Box 516, SE-75120 Uppsala, Sweden*

⁶²*Dept. of Physics, University of Wuppertal, D-42119 Wuppertal, Germany*

⁶³*Deutsches Elektronen-Synchrotron DESY, Platanenallee 6, D-15738 Zeuthen, Germany*

(Dated: October 20, 2025)

We present a measurement of the mean number of muons with energies larger than 500 GeV in near-vertical extensive air showers initiated by cosmic rays with primary energies between 2.5 PeV and 100 PeV. The measurement is based on events detected in coincidence between the surface and in-ice detectors of the IceCube Neutrino Observatory. Air showers are recorded on the surface by IceTop, while a bundle of high-energy muons (TeV muon) from the shower can subsequently produce a tracklike event in the IceCube in-ice array. Results are obtained assuming the hadronic interaction models Sibyll 2.1, QGSJet-II.04, and EPOS-LHC. The measured number of TeV muons is found to be in agreement with predictions from air-shower simulations. The results have also been compared to a measurement of low-energy muons by IceTop, indicating an inconsistency between the predictions for low- and high-energy muons in simulations based on the EPOS-LHC model.

I. INTRODUCTION

Cosmic rays are ionized nuclei coming from outer space with energies extending up to 10^{20} eV and beyond. A detailed knowledge of the properties of the cosmic-ray flux is important for understanding their origin. In addition to directly carrying information about the sources and propagation of cosmic rays, it is a major uncertainty in the calculation of the atmospheric neutrino flux, the main background for neutrino astronomy [1, 2]. The high energies of cosmic rays furthermore provide the opportunity to explore particle physics beyond the reach of human-made accelerators [3].

While the energy spectrum of cosmic rays has been measured with high precision over many orders of magnitude, the mass composition remains uncertain. This results from the fact that cosmic rays with energies above several 100 TeV are observed indirectly through the extensive air showers they produce when entering the Earth's atmosphere. The properties of the primary cos-

mic rays can be inferred from several observables accessible through ground-based experiments. One such observable is the number of muons in the air shower, which can be probed with arrays of particle detectors. Muons originate predominantly from the decay of charged pions and kaons produced as part of the hadronic cascade, and, together with an independent estimate of the primary energy, their number can be used to infer the mass of the cosmic-ray nucleus [4]. The interpretation of the air-shower observations in terms of properties of the primary particle relies, however, on the accurate simulation of the development of the air shower in the atmosphere. These Monte Carlo simulations make use of phenomenological models describing the high-energy hadronic interactions in the shower, which cannot be calculated in the context of perturbative quantum chromodynamics. The models are tuned to accelerator data, but have to extrapolate to a phase space not accessible to these experiments, leading to significant uncertainties [3, 5].

Several observatories have performed measurements of the muon content of air showers. A 2019 meta-analysis including data from eight observatories found a discrepancy between data and expectations from simulations with a significant dependence on the shower energy [6], with the Pierre Auger Observatory reporting evidence for a deficit of muons in simulations based on post-LHC hadronic interaction models at the highest energies [7, 8]. This observation is commonly referred to as *the Muon Puzzle*, its most plausible origin being the description of

* also at Institute of Physics, Sachivalaya Marg, Sainik School Post, Bhubaneswar 751005, India

† also at Department of Space, Earth and Environment, Chalmers University of Technology, 412 96 Gothenburg, Sweden

‡ also at INFN Padova, I-35131 Padova, Italy

§ also at Earthquake Research Institute, University of Tokyo, Bunkyo, Tokyo 113-0032, Japan

¶ now at INFN Padova, I-35131 Padova, Italy

hadronic interactions in the air showers [9]. The current experimental picture, however, is ambiguous; while some observatories report a discrepancy between measurements and simulations based on various hadronic interaction models, others do not observe any such disagreements [10]. The observatories all operate in different regions of the parameter space, such as observation altitude and muon energy threshold, and a consistent picture has yet to emerge. As a result, interpreting the muon measurements of different observatories in terms of the cosmic-ray mass composition leads to inconsistent results. Resolving this situation, i.e. the Muon Puzzle and the possibly related inconsistencies between observatories, is considered one of the most pressing problems in high-energy cosmic-ray physics and motivates performing a variety of muon measurements to probe air-shower development and hadronic interactions under different conditions [11]. Complementary information toward understanding the issues in the description of high-energy hadronic interactions will also be obtained by collider experiments, with those measuring in the forward region being particularly important for air-shower physics [9, 12].

The IceCube Neutrino Observatory [13] is a particle detector located at the geographic South Pole in Antarctica. Its combination of a surface detector array, IceTop, and the large-volume in-ice detector that is located between 1.5 km and 2.5 km below, allows it to perform unique air-shower measurements. A measurement of the density of mainly low-energy muons detected at the surface with IceTop, often referred to as GeV muons, has been previously published [14]. The thick ice sheet covering the IceCube in-ice detector absorbs all muons with energies below several hundred GeV, allowing for a measurement of purely the high-energy muon content in air showers. These muons, which we refer to as TeV muons, are the subject of this article. The TeV muons are of special interest as they are predominantly produced in the early stages of the shower development. Their number depends on the energy and mass of the primary. The mass dependence is stronger than that of low-energy muons observed with surface arrays, making them also a particularly interesting observable for mass composition studies [15] (see also Appendix A). In this work, we present a measurement of the mean number of muons with energies above 500 GeV in near-vertical air showers observed with both IceTop and the IceCube in-ice array in the primary energy range of 2.5 PeV to 100 PeV.

The article is structured as follows. In Section II, we introduce the IceCube Neutrino Observatory and its operation as an extensive-air-shower detector. Following this, the experimental data and simulations used in the analysis are described in Section III. Section IV describes in detail the analysis method, including neural network reconstructions of the primary cosmic-ray energy and TeV muon multiplicity, as well as the application of correction factors derived from simulations. The final results are presented in Section V. Two appendices are

included: Appendix A discusses the predictions for high-energy muons from simulations in detail; Appendix B includes various demonstrations of the robustness of the analysis results.

II. ICETOP AND ICECUBE IN-ICE

The IceCube Neutrino Observatory, illustrated in Fig. 1, is a kilometer-scale detector located at the geographic South Pole, constructed mainly for the detection of high-energy neutrinos and cosmic rays [13]. The IceCube in-ice array consists of 5160 digital optical modules (DOMs) deployed on 86 vertical strings at a depth between 1450 m and 2450 m with a horizontal spacing of 125 m, instrumenting a volume of one cubic kilometer of ice. Each DOM contains a photomultiplier tube sensitive to the Cherenkov photons produced by relativistic charged particles propagating through the ice. The surface component of the observatory, IceTop [16], is an array consisting of 81 stations of two ice-Cherenkov tanks approximately following the in-ice string locations, covering an area of about one square kilometer. Each tank contains two DOMs to record the Cherenkov light produced by air-shower particles penetrating the ice in the tank. The lids of the tanks were deployed flush with the snow surface at the site. Since then, however, several meters of snow have accumulated on top of the array.

The IceTop array is situated at an altitude of 2835 m above sea level, corresponding to an atmospheric depth of about 690 g/cm². IceTop signals are calibrated in terms of the typical signal produced by vertical muons penetrating the tanks, known as the vertical equivalent muon or VEM. The signals recorded in the tanks for an air shower whose impact point is located inside the array, are dominated by the electromagnetic shower component, except at large lateral distances, where a muonic signal component can be discerned. This feature has been used previously to measure the density of mainly low-energy ($\mathcal{O}(\text{GeV})$) muons at the surface [14]. For air showers whose shower axis also passes through the volume of the in-ice detector, a narrow bundle of high-energy muons can be observed in coincidence with the surface signals. Muons need an energy of at least several hundred GeV to travel from the surface to the top of the in-ice array. When requiring that the shower core is contained within the IceTop footprint, events with a coincident surface trigger and in-ice muon bundle are limited to zenith angles of $\theta \lesssim 38^\circ$. This class of events has been used before to perform a measurement of the cosmic-ray mass composition [17].

The IceTop signals are typically used to reconstruct various parameters, such as the shower core position and direction, with a maximum likelihood method which fits the time and charge distributions [16]. This procedure also reconstructs the shower size S_{125} , the signal at a distance of 125 m from the shower axis, which is a proxy for the primary cosmic-ray energy. A simple exponential

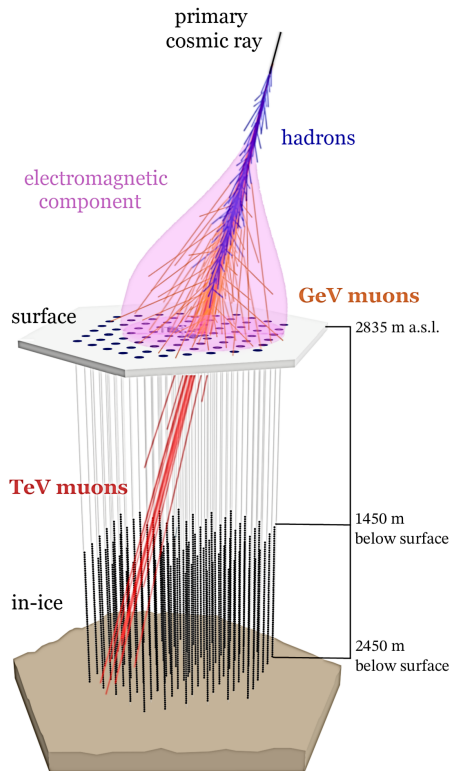


Figure 1. Schematic drawing of a cosmic-ray air shower observed in coincidence between IceTop, the surface component of the IceCube Neutrino Observatory, and the IceCube in-ice array.

attenuation function is used during the reconstruction to take the impact of the snow coverage on top of a tank on the observed signal into account [18]. Additional information about the primary particle and the air shower can be obtained from the in-ice detector, typically by studying the charge deposit of the high-energy muon bundle as it propagates through the ice. One method, which has been used in previous analyses, consists of determining the energy loss in equidistant segments along the line of propagation of the bundle. As the photons detected by a DOM can originate from different positions along the track, an unfolding is required to obtain the contributions from energy losses in all track segments to the observed light distribution [19]. These established reconstruction methods are applied to all air-shower events included in the analysis presented in this paper.

III. DATASETS

A. Experimental dataset

The analysis uses data collected between May 15, 2012 and May 2, 2013 with an effective live time of about 323 days. A total of 1 216 154 events with a reconstructed energy between 2.5 PeV and 100 PeV pass the selection

criteria described below.

The event selection is aimed at air showers which trigger IceTop and have a coincident bundle of high-energy muons in the in-ice detector. To obtain an event sample with a high quality of air-shower reconstructions, cuts established in previous IceCube cosmic-ray analyses are applied [17, 18]. Events must trigger at least five IceTop stations and at least one of the stations must have a signal greater than 6 VEM. The shower reconstruction is required to succeed, and the reconstructed shower core is required to be within a geometric boundary slightly smaller than the array. Furthermore, the station with the largest signal must not be on the edge of the array. For the in-ice part of the events, quality cuts are applied to the unfolded muon-bundle energy loss, requiring a successful reconstruction which has at least three reconstructed energy losses inside the detector volume that are nonzero (see also Section IV A). While this selection can include events up to about 38° in zenith, the analysis is further restricted to events with $\cos\theta > 0.95$ ($\theta \lesssim 18^\circ$). This simplifies the possible muon bundle geometries included in the analysis and ensures that all muon bundles have propagated through a similar amount of matter before reaching the detector, while still resulting in a systematics-dominated measurement.

As will be discussed in Section IV A, the resolution of the cosmic-ray energy reconstruction method that is used in this analysis, defined as the standard deviation of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$, is better than 0.1 above 1 PeV and about 0.05 above 10 PeV. The analysis only includes events above the threshold energy where the event selection reaches nearly full efficiency for all masses of cosmic-ray nuclei, i.e. 2.5 PeV for the level of snow coverage on the IceTop array in the included time interval [17], and includes events of energies up to 100 PeV. Around the threshold energy, a core resolution better than 12 m and an angular resolution better than 0.8° are obtained after quality cuts, improving to about 5 m and 0.4° at 30 PeV [20].

B. Simulated datasets

Air shower simulations are produced with CORSIKA v7.3700 [21], using an atmospheric model describing a typical South Pole atmosphere for the month of April.¹ This model closely approximates the yearly average South Pole atmosphere [22]. Primary energies are sampled according to an E^{-1} differential energy spectrum. Hadrons and (anti)muons are simulated down to energies of 0.05 GeV, while electrons/positrons and photons are included down to 0.01 GeV and 0.002 GeV respectively.

¹ This atmospheric model was included as a standard model in CORSIKA v7.4700 with identifier 33.

The model used for hadronic interactions below a laboratory energy of 80 GeV is FLUKA 2011.2c [23, 24]. Datasets with different high-energy hadronic interaction models are included in the analysis. The main dataset is based on Sibyll 2.1 [25] and includes four types of primary cosmic ray: p, He, O, and Fe. Two other datasets are based on QGSJet-II.04 [26] and EPOS-LHC [27] and include only proton and iron showers. Sibyll 2.1 is a *pre-LHC* model, while QGSJet-II.04 and EPOS-LHC are *post-LHC* models, taking into account high-energy data from the LHC.

The detector response to the shower particles is simulated using IceCube software including the entire hardware and data-acquisition chain, as in previous analyses [17, 18]. Each CORSIKA shower is resampled 100 times uniformly over an energy-dependent area larger than the IceTop detector area. The observation level in CORSIKA is set to 2837 m, which is several meters above the top of the IceTop tanks to allow for the inclusion of snow accumulated on top of the tanks in the simulation. The shower particles obtained from CORSIKA are propagated from the observation level through layers of air and snow before their interactions in the tanks are simulated, all using the GEANT4 [28] package. The simulated snow levels are those measured *in situ* in October 2012. Only muons with energies higher than 273 GeV are considered for simulation of the in-ice detector response, including their interactions with the ice and the resulting detection of Cherenkov photons. The same reconstructions and event selection applied to the experimental data are applied to the simulated events.

Additional CORSIKA-only datasets are produced to obtain predictions of the multiplicity of muons with $E_\mu > 500$ GeV with reduced statistical uncertainty, for comparison to the final analysis results. Predictions are derived for p and Fe primaries using the models Sibyll 2.1, QGSJet-II.04, and EPOS-LHC. They are described in more detail in Appendix A.

IV. ANALYSIS

The analysis presented in this paper determines the mean number of muons N_μ with an energy $E_\mu > 500$ GeV in air showers with $\cos\theta > 0.95$ for primary cosmic-ray energies E between 2.5 PeV and 100 PeV, where N_μ is defined at the surface.² The analysis starts from the initial air-shower and muon-bundle energy-loss reconstructions and the event selection discussed in Sections II and III. The reconstructed observables from IceTop and in-ice are used as inputs for a neural network reconstruction of the primary cosmic-ray energy and the high-energy

muon number. Correction factors are subsequently derived from simulations and used to correct for mass-dependent biases in the determination of the mean muon number $\langle N_\mu \rangle$ in bins of primary energy.

A. Neural network reconstruction

A neural network model is used to relate the signals generated by the bundle of high-energy muons in the in-ice IceCube detector to the number of muons in the air shower with an energy greater than 500 GeV at the surface. In addition, the shower size S_{125} reconstructed with IceTop provides the main sensitivity to the primary cosmic-ray energy. Various ways of combining the inputs in different neural networks were explored. For example, one could use only surface information for the reconstruction of the energy E and only in-ice information for the reconstruction of the muon number N_μ , or one could utilize them in a single model for combined reconstruction. Some of these approaches are discussed in more detail in Appendix B, demonstrating that the final results are invariant to such choices. We present in this section the approach that was used to obtain the nominal analysis results which are presented in Section V.

The muon bundle energy loss is reconstructed in segments of 20 m along the shower axis, which is obtained from the IceTop reconstruction and extended to the in-ice detector. Any segment not contained in the detector volume is removed. Events which end up with less than three segments that have a nonzero reconstructed energy loss in the detector are discarded. The first segment in the detector will correspond to a different slant depth traveled by the bundle for events with different zenith angles. To include this information, we create a fixed-length vector of reconstructed energy losses in such a way that each entry in the vector corresponds approximately to the same in-ice slant depth for all events, padding with zeros at the start or end of the vector based on the zenith angle. More specifically, a vertical event will have zeros at the end of the vector, while a more inclined event will have zeros at the start, in such a way that the first entry corresponds to a traveled distance of about 1450 m in the ice. This vector of energy losses is used as input to a recurrent neural network, more specifically a bidirectional gated recurrent unit layer [30], a common choice for sequential data. The output from the recurrent layer is fed into a fully connected (dense) layer combined with the shower size S_{125} and zenith angle θ reconstructed using IceTop. The neural network subsequently outputs predictions for both $\log_{10}(N_\mu)$ and $\log_{10}(E)$. The network is trained using a mean-squared error (MSE) loss function for each of the training targets, minimizing the combined loss function consisting of the sum of the two MSE losses. The implementation of the neural network is based on the KERAS [31] and TENSORFLOW [32] software libraries.

The neural network is trained on Sibyll 2.1 simula-

² Simulations of vertically down-going muons performed with PROPOSAL [29] show that muons with an energy of 500 GeV have a chance of about 75% to propagate down to at least 1450 m in the ice. This increases to 95% for muons with an energy of 1 TeV.

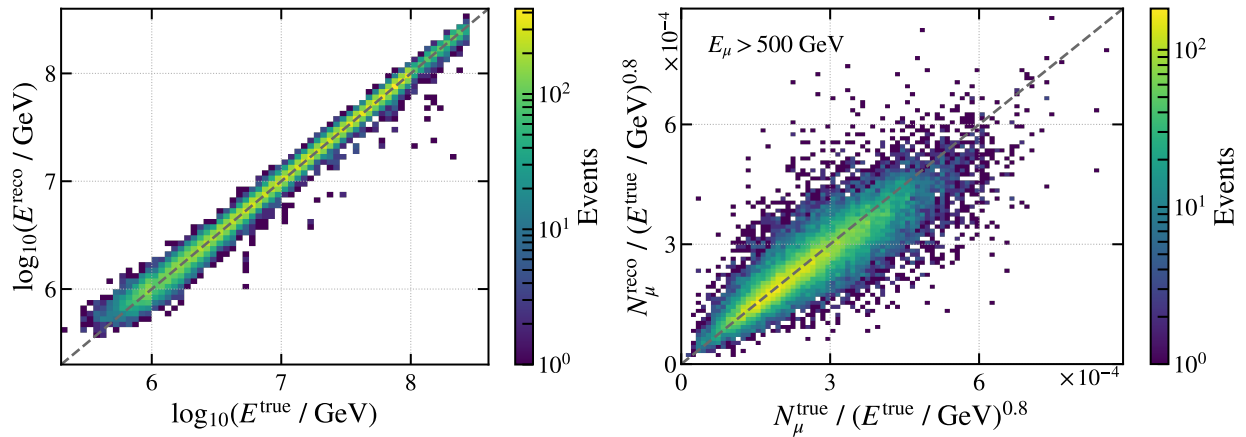


Figure 2. Correlation plots showing the relation between the true and neural-network reconstructed values for primary cosmic-ray energy (left) and high-energy muon multiplicity (right). The muon multiplicity has been divided by E^{β} to reduce the effect of its underlying energy dependence. The histograms include Sibyll 2.1 simulations of all four mass groups (p, He, O, and Fe).

tions with roughly equal amounts of the four primaries (p, He, O, and Fe), with true energies in the range $5.4 \leq \log_{10}(E / \text{GeV}) \leq 8.4$, corresponding to a range for N_{μ} of one muon up to about 2500. The quality of the reconstructions is examined by applying them to an independent test set. Fig. 2 shows the relations between the reconstructed and true values for E and N_{μ} , combined for all primaries. As the high-energy muon multiplicity approximately grows as E^{β} with $\beta \approx 0.8$, as discussed in Appendix A, the ratio $N_{\mu}/E^{0.8}$ has been plotted to prevent the correlation from simply reflecting the underlying energy dependence rather than the performance of the neural network. The bias and resolution of the reconstructions, defined as respectively the mean and the standard deviation of the logarithmic difference between reconstructed and true values, is shown in Fig. 3 for the different primary types separately. For most of the energy range included in the analysis, i.e. 6.4 to 8.0 in $\log_{10}(E/\text{GeV})$, the energy reconstruction has a resolution smaller than half the bin width of 0.1 in $\log_{10}(E)$, with some mass dependence in both the bias and the resolution. The muon number reconstruction, on the other hand, has a nearly mass-independent resolution, but has a clear mass-dependent bias.³ The number of muons with $E_{\mu} > 500 \text{ GeV}$ in simulations inside the primary energy range used in the analysis is between 5 and about 670 for proton showers and between 30 and about 1330 for iron showers.

B. Monte Carlo correction

The analysis determines the average high-energy muon number $\langle N_{\mu} \rangle$ in bins of primary cosmic-ray energy E based on the event-by-event reconstructions described above. The accuracy with which it can be obtained is tested based on simulations. In Fig. 4, the average of the reconstructed muon number values is shown in bins of reconstructed energy for different primaries, compared to the true average muon number in bins of true simulated energy. These values were determined from the subset of the Sibyll 2.1 dataset which was not used in the training of the neural network. While the energy dependence is captured rather well for all masses, systematic offsets can be seen between the reconstructed and true values. The bottom panel of the figure shows the ratio between the reconstructed and true values, indicating biases up to about $\pm 15\%$ with a clear mass dependence. The systematic overestimation for proton and underestimation for iron arises from the tendency of neural networks trained with MSE loss to predict values closer to the average behavior of the training data. The ratios are fit with quadratic functions for each mass separately, defining the correction factors that will be applied to the results. We note that the offsets shown in these plots result from both the imperfections of the muon multiplicity reconstruction, which has the largest impact, as well as from the energy reconstruction as a result of bin migration, which is a smaller effect.

³ Training two separate neural networks, one for E reconstruction using only IceTop information, and one for N_{μ} reconstruction using only in-ice information, changes this behavior. In that case, a larger mass dependence in the E reconstruction bias is obtained, while the N_{μ} network produces a reduced mass-dependent bias. The combined neural network presented in the main body of this work allows, however, for the more precise correction method for the resulting biases discussed in Section IV B. See Appendix B for a more detailed discussion.

The correction factors derived from simulations can be used to unbiased the results that will be obtained from experimental data. A complication is that the correction factors are clearly mass dependent, while the cosmic-ray composition at high energies is not precisely known. Rather than assuming a specific composition model to derive corrections, a model-independent approach is applied. It exploits the fact that the muon number itself is a measure of composition, and that the values of the

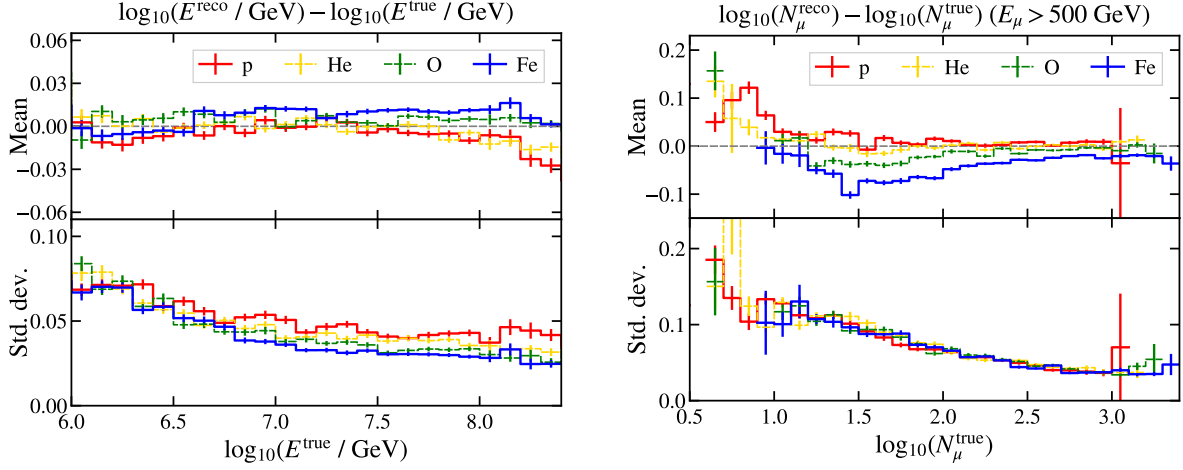


Figure 3. Bias and resolution of the neural network reconstructions, defined as the mean and standard deviation of the difference between the logarithms of the reconstructed and true values, shown separately for different primary cosmic-ray masses. Primary energy reconstruction is shown on the left, muon multiplicity ($E_\mu > 500$ GeV) reconstruction on the right. The analysis includes events between 6.4 and 8.0 in $\log_{10}(E/\text{GeV})$, corresponding to $0.7 \lesssim \log_{10}(N_\mu) \lesssim 2.8$ for proton showers and $1.5 \lesssim \log_{10}(N_\mu) \lesssim 3.1$ for iron showers.

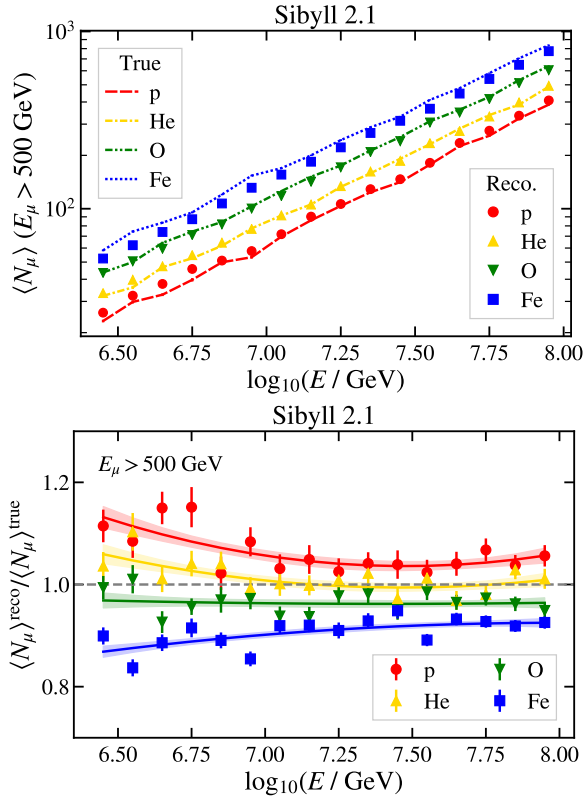


Figure 4. Top: comparison between the average reconstructed TeV muon number in bins of reconstructed primary energy and the true muon number in bins of true energy in Sibyll 2.1 simulation for four different primaries. Bottom: ratio of the reconstructed and true values from the top plot, fitted with quadratic functions defining correction factors based on Sibyll 2.1.

correction factors at a specific energy depend, to a good approximation, linearly on $\ln(A)$, with A being the nuclear mass number. This works as follows. The Heitler-Matthews model [33] predicts that the number of muons in an air shower depends on A and the cosmic-ray energy E as

$$N_\mu(E, A) = A^{1-\beta} \left(\frac{E}{\xi} \right)^\beta, \quad (1)$$

with an exponent $\beta < 1$ and ξ a constant.⁴ This relation inspires the definition of the z value, a common way of representing muon data by scaling it according to expectations from simulations [6],

$$z = \frac{\ln\langle N_\mu \rangle - \ln\langle N_\mu \rangle_{\text{p}}}{\ln\langle N_\mu \rangle_{\text{Fe}} - \ln\langle N_\mu \rangle_{\text{p}}}, \quad (2)$$

where $\langle N_\mu \rangle_{\text{p}}$ and $\langle N_\mu \rangle_{\text{Fe}}$ are the predictions from proton and iron simulations, respectively. The relation of Eq. (1) can thus be used to estimate $\ln(A)$ based on the $\langle N_\mu \rangle$ derived from the data as $z \approx \ln(A)/\ln(56)$. With the estimate of $\ln(A)$ obtained from the initial $\langle N_\mu \rangle$ estimate, a correction factor can be obtained by linearly interpolating the correction factors derived for proton and iron, C_{p} and C_{Fe} , in $\ln(A)$,

$$C_{\ln(A)}(E) = C_{\text{p}}(E) + \frac{C_{\text{Fe}}(E) - C_{\text{p}}(E)}{\ln(56)} \ln(A). \quad (3)$$

The interpolated correction factor, $C_{\ln A}$, can then be applied to the initial estimate of $\langle N_\mu \rangle$ to obtain a corrected

⁴ In Appendix A, we discuss the validity of this relation for the high-energy ($E_\mu > 500$ GeV) muons in the shower.

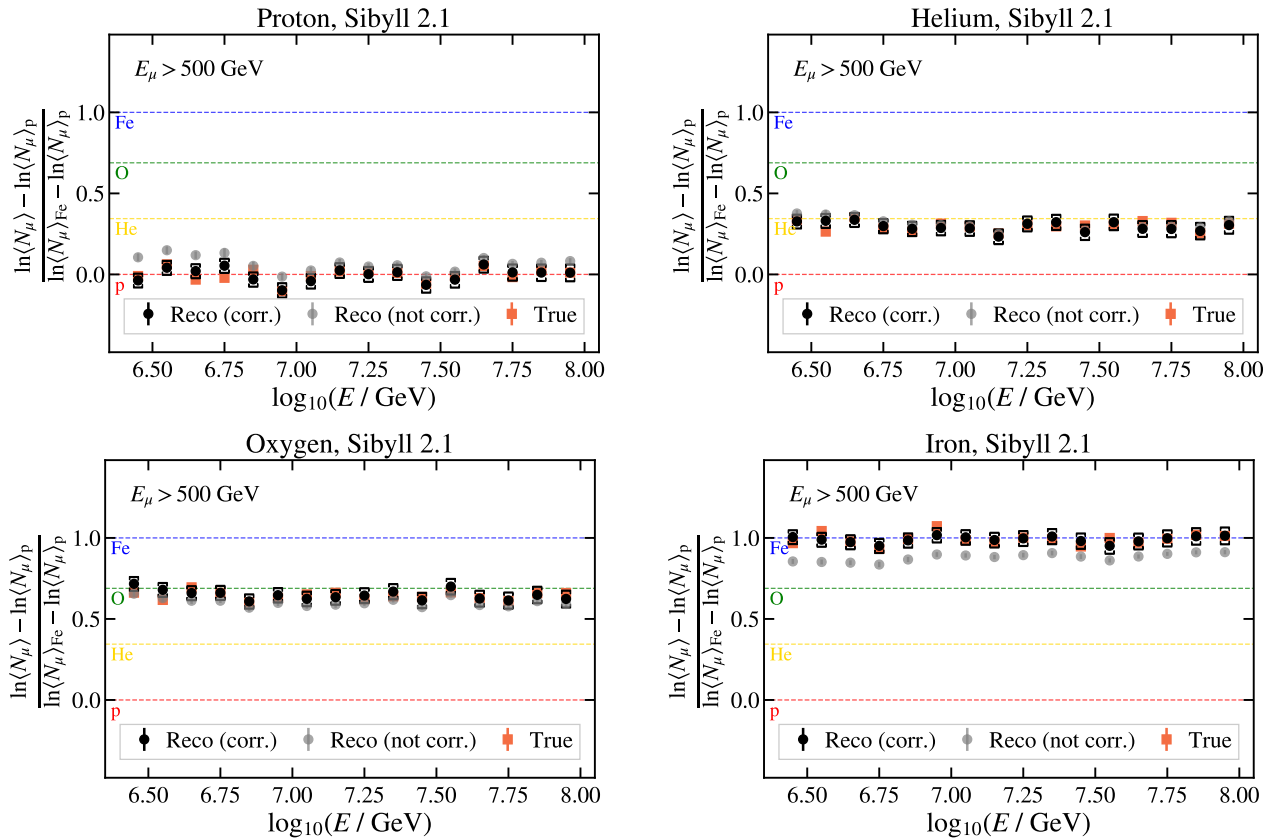


Figure 5. Comparison between the initial reconstruction of the average muon number (in bins of reconstructed energy), the corrected results, and the true muon number (in bins of true energy) obtained from air showers simulated with Sibyll 2.1. The values have been scaled according to Eq. (2), so that the true value for proton is at zero and the true value for iron is at 1. The brackets show the uncertainty assigned to the final result to account for the typical offsets remaining after the correction. This comparison is shown for four pure composition scenarios.

estimate. This process is iterated, using the newly corrected $\langle N_\mu \rangle$ values to estimate the mass composition and in turn derive a new correction factor. After several iterations, the process converges to what is then the final result for the average muon number. The approximate dependence of the correction factors on $\ln(A)$ has been tested by deriving correction factors for He and O by interpolation (i.e. Eq. (3) for $A = 4$ and 16) and comparing them to the ones derived directly from simulation (which are shown in Fig. 4).

The method has been verified using simulations. An example is shown in Fig. 5, where results were derived from simulations of a single primary mass which have undergone the whole analysis process. The corrected results can be seen to closely align with the Monte Carlo truth. This was repeated for a variety of possible mass compositions, and the method was found to reproduce the true $\langle N_\mu \rangle$ well, independently of the injected composition. The typical difference remaining between the corrected and true values after the correction are taken into account in the final result: the muon number is corrected for an average offset of 0.2%, and a spread of 4% is included as a systematic uncertainty, shown as the brackets in Fig. 5.

ets in Fig. 5.

The correction factors depend on the hadronic interaction model used in the simulations they are derived from. Correction factors are obtained for QGSJet-II.04 and EPOS-LHC simulations in addition to those for Sibyll 2.1; they are shown in Fig. 6. The correction factors effectively remove the dependence of the result on the hadronic model that the initial neural network reconstructions were trained on, and replaces it by a dependence on the model that the correction factors are derived from. They therefore allow us to interpret the experimental data under the assumption of different hadronic interaction models. This is shown in more detail in Appendix B, where results are obtained using a neural network trained on EPOS-LHC, leading to results consistent with the nominal results presented in the following section.

V. RESULTS AND DISCUSSION

Figure 7 shows the mean number of muons $\langle N_\mu \rangle$ with energies greater than 500 GeV in air showers with ener-

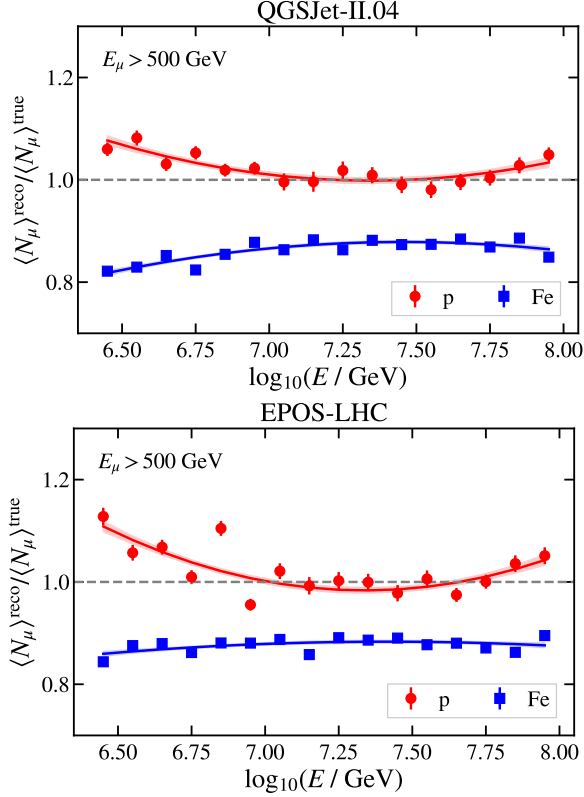


Figure 6. Correction factors derived from simulations based on QGSJet-II.04 and EPOS-LHC, equivalent to the Sibyll 2.1 correction factors shown in Fig. 4.

gies between 2.5 PeV and 100 PeV. These results were obtained in a model-dependent way by using the correction factors derived from simulations based on Sibyll 2.1, QGSJet-II.04, and EPOS-LHC (shown in Figs. 4 and 6). Also shown are the predictions of the muon number in proton and iron showers derived from CORSIKA simulations using the respective hadronic interaction models, as described in Appendix A. The shaded area around the points represents the total systematic uncertainties, while statistical uncertainties are too small to be visible in the figures. The central values are bracketed by the proton and iron predictions for all models, and qualitatively indicate that the mass composition becomes heavier with increasing primary energy.

As shown in Fig. 8, the muon numbers obtained using QGSJet-II.04 and EPOS-LHC are approximately 5% higher than the result obtained with Sibyll 2.1. As this spread is significantly smaller than the total systematic uncertainty (see below), we derive in addition a model-averaged muon multiplicity, shown in Fig. 9. This result includes an additional contribution to the total systematic uncertainty, representing the spread of the model-specific $\langle N_\mu \rangle$ results around the average. The average result is plotted together with the proton and iron predictions based on the different hadronic models. The $\langle N_\mu \rangle$ predictions for QGSJet-II.04 and EPOS-LHC are, respec-

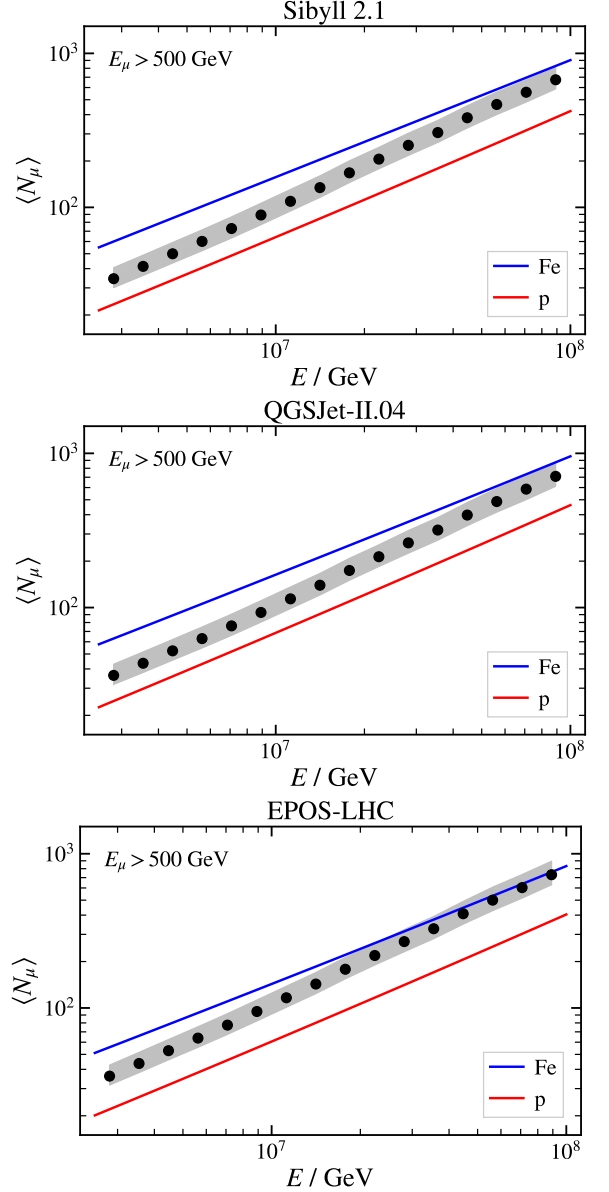


Figure 7. Average number of muons with energy greater than 500 GeV in near-vertical air showers as a function of the primary cosmic-ray energy obtained using the hadronic interaction models Sibyll 2.1, QGSJet-II.04, and EPOS-LHC. The shaded region indicates the systematic uncertainty, statistical uncertainties are not visible. The muon number expected from proton and iron simulations performed with the corresponding hadronic models are shown for comparison. These data are made available in a public data release [34].

tively, about 5% higher and lower than for Sibyll 2.1; a more detailed comparison between the predictions is shown in Fig. 13. The numerical values for the individual and model-averaged $\langle N_\mu \rangle$ results are made publicly available in electronic format in Ref. [34].

The total systematic uncertainty on the muon number measurement originates from different sources, shown in Fig. 10. Detector uncertainties originate from both

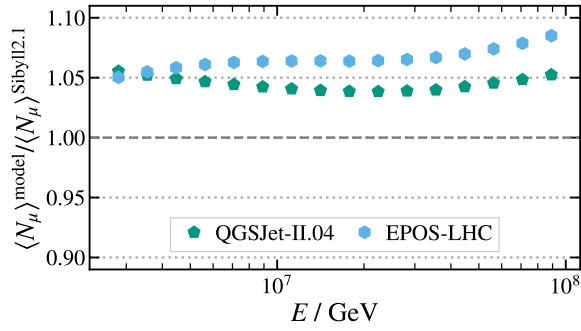


Figure 8. Ratio of the muon multiplicity ($E_\mu > 500$ GeV) results obtained with QGSJet-II.04 and EPOS-LHC to those obtained with Sibyll 2.1, as shown in Fig. 7.

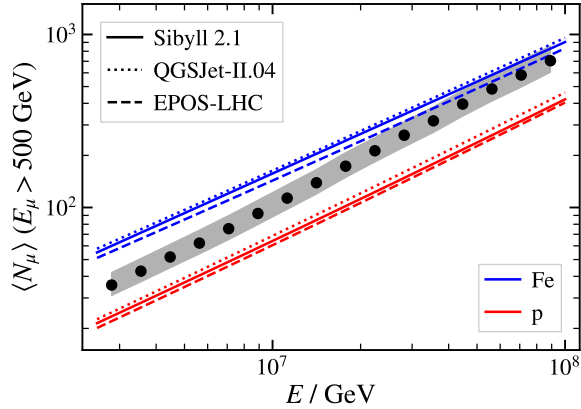


Figure 9. Average number of muons with energy greater than 500 GeV in near-vertical air showers as a function of primary cosmic-ray energy, averaged over the individual results obtained with Sibyll 2.1, QGSJet-II.04, and EPOS-LHC. The systematic uncertainty band includes the deviation of the individual results from the average added in quadrature with the other systematic uncertainties (see text for details). The corresponding simulated muon numbers for proton and iron are shown for comparison. These data are made available in a public data release [34].

IceTop and the in-ice array, and are essentially the same as those described in previous works [17]. The largest uncertainty is related to the modeling of the properties of the deep ice used in simulations (scattering and absorption of both the bulk ice and the refrozen ice in the drill holes around the DOMs). In Ref. [17], the uncertainties related to the ice model were combined with a 3% uncertainty on the DOM efficiency to determine the total in-ice light-yield uncertainty for down-going muon bundles. We opt here for a more conservative DOM efficiency uncertainty of 10%, leading to a total light-yield uncertainty of (+13.4%, -15.7%). The impact on the result has been determined by repeating the energy-loss reconstruction with a modified light-collection efficiency parameter. For IceTop, an uncertainty of ± 0.2 m is assumed for the effective attenuation length ($\lambda = 2.25$ m) of air shower particles in the snow [16, 18]. This impacts

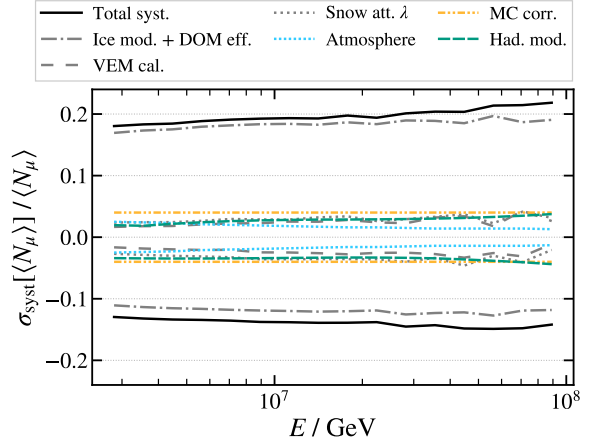


Figure 10. Relative size of the systematic uncertainties of the high-energy muon multiplicity measurement. For the model-average result of Fig. 9, all shown uncertainties are included and added in quadrature to obtain the total systematic uncertainty. For the individual model results shown in Fig. 7, all but the hadronic model uncertainty ('Had. mod.') are included.

both the shower size reconstruction, as well as the core and direction reconstruction, which in turn impact the track used in the muon-bundle energy-loss reconstruction. The impact of this uncertainty is evaluated by redoing the whole analysis chain after running an air-shower reconstruction with a different value of λ . Furthermore, a $\pm 3\%$ uncertainty on the calibration of the VEM charge unit for IceTop was included, based on simulations of the calibration process with different atmospheres, hadronic interaction models, and other systematic variations [35]. The impact of energy-bin migration on the result is included in the definition of the correction factors discussed in Section IV B. An uncertainty of 4% is assigned to cover any offsets that may remain after the correction procedure. Finally, the difference between the simulated atmosphere (Section III) and the actual atmosphere measured by the AIRS satellite [36], averaged over the data-taking period, is evaluated. Using the muon-production parametrization from Ref. [37] and the uncertainty on the atmospheric temperature measurements, we conservatively estimate that $\langle N_\mu \rangle$ might deviate by up to 2.5% between the simulated and real atmosphere. All uncertainties are added in quadrature and shown as bands in the main results of Fig. 7. For the model-average result shown in Fig. 9, the maximal deviation of the individual results was included as an additional systematic uncertainty, shown as the green lines in Fig. 10. Note that this hadronic interaction model uncertainty only reflects the spread between the models included in this analysis and that other models may produce larger differences.

Figure 11 shows the z -value representation of the results as defined in Eq. (2), i.e. the logarithms of the measured high-energy muon number derived using different hadronic interaction models scaled against the ex-

pectations from proton and iron simulations using the corresponding model. The result based on EPOS-LHC indicates a slightly heavier mass composition than the results based on Sibyll 2.1 and QGSJet-II.04, which are very similar in the z -values. Note that, in this representation, the difference between the values obtained for the various hadronic interaction models results from the model dependence of both the determination of the $\langle N_\mu \rangle$ results and of the predictions, which are of similar size (see Figs. 8 and 13). The experimental results are compared to predictions based on three composition models, commonly known as H3a [38], GST [39], and GSF [40].⁵ The observed muon number and its evolution with energy generally agrees with the expectations from these models within uncertainties.⁶

It is of interest to compare these results to a muon analysis performed with IceTop alone, mentioned earlier in Section I, and described in detail in Ref. [14]. In that analysis, the density of muons at the surface ρ_μ was determined at lateral distances of 600 m and 800 m from the shower axis. These are mainly low-energy muons, with a threshold of several 100 MeV, commonly referred to as GeV muons. The GeV muon density analysis covers the primary energy range used in the TeV muon analysis presented in this work, and uses the same zenith range ($\cos\theta > 0.95$) and hadronic interaction models. The results of the GeV muon density analysis and the TeV muon multiplicity analysis are shown together in terms of z values in Fig. 12. If the simulations consistently describe the experimental data, the two results should be consistent, as they are measures of the same primary cosmic-ray flux arriving at Earth. This is the case for the Sibyll 2.1 results, where we observe excellent agreement over the entire energy range. The increased production of low-energy muons in the post-LHC models QGSJet-II.04 and EPOS-LHC results in the GeV muon measurement being closer to expectations for lighter primaries. The tension with the TeV muon result is strongest for EPOS-LHC, where the z -values obtained from the GeV and TeV muon measurements are outside each other's uncertainty bands over nearly the entire energy range. We note that there are also preliminary indications for inconsistencies related to the slope of the lateral charge distribution observed in IceTop, most prominently for Sibyll 2.1 simulations, as reported in Ref. [41].

Several tests have been performed to ensure the robustness of the TeV muon result. First, it was checked that the nominal result is consistent with results obtained

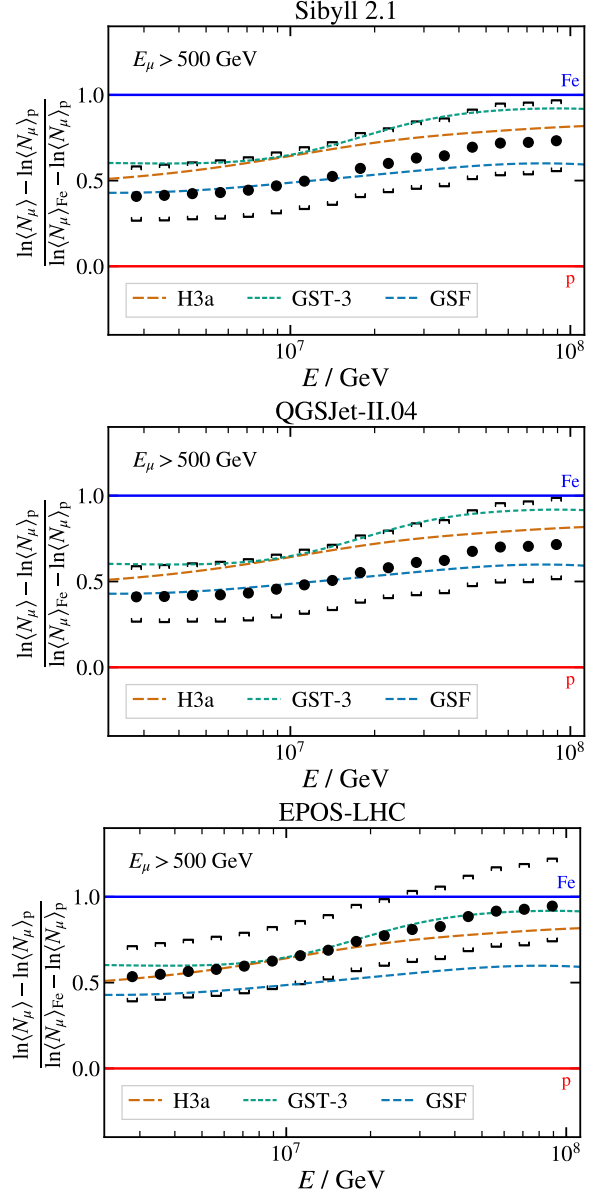


Figure 11. Measured muon multiplicities of Fig. 7 scaled with respect to the expectations from proton and iron simulations (z values) using different hadronic interaction models, as in Eq. (2). Systematic uncertainties are indicated by the brackets. The dashed lines display the expected muon multiplicities according to the cosmic-ray flux models H3a, GST, and GSF.

when the energy and muon multiplicity reconstructions are performed separately based on only IceTop and in-ice information respectively. Second, it was confirmed that a neural network trained on simulations based on a different hadronic model produces consistent results after deriving and applying the correction factors. Third, the analysis was performed replacing the neural-network estimated primary energy by a simple energy estimate based on the shower size S_{125} , derived for an analysis of the energy spectrum, reported in Ref. [18]. This relation was also used in the GeV muon density analysis. The

⁵ The predictions are calculated as $\sum_i f_i \langle N_\mu \rangle_i$, where the sum runs over the different mass groups i included in the model and f_i is the corresponding fraction of the total flux. The expected muon multiplicities $\langle N_\mu \rangle_i$ are the values derived from simulations for $i = (p, Fe)$, and are linearly interpolated in $\ln(A)$ for intermediate mass groups.

⁶ Note that while H3a is independent of IceCube results, the GST and GSF models include IceCube composition results derived from high-energy muons in their fit.

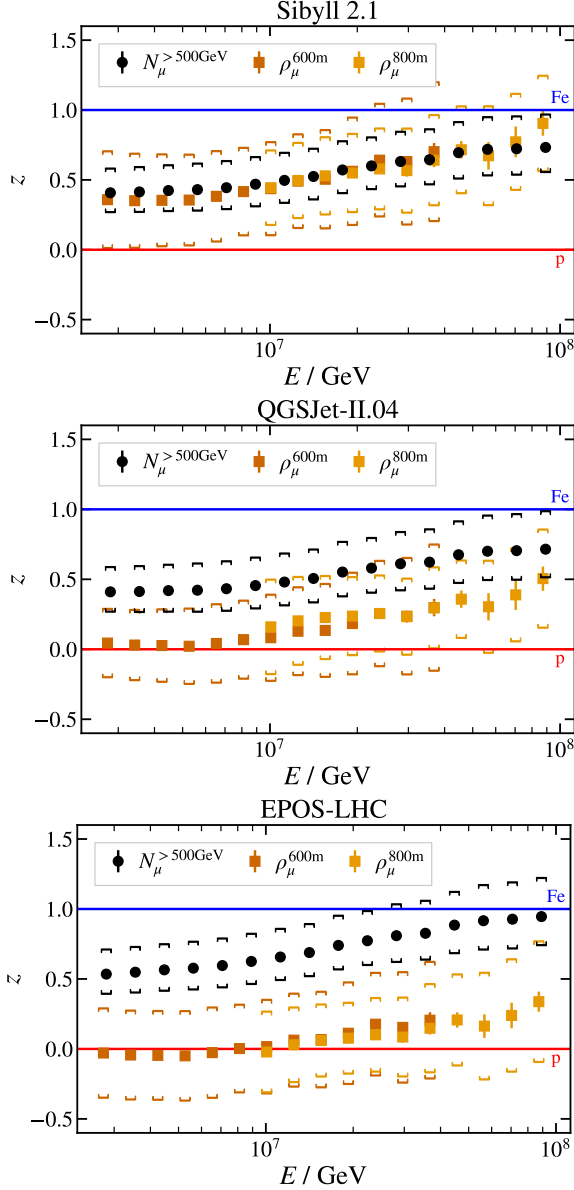


Figure 12. Comparison between the TeV muon data obtained in this work (N_μ , black), and the density of GeV muons measured with IceTop at different lateral distances (ρ_μ , shades of orange), represented as z -values (Eq. (2)). See the text for details.

results obtained with this approach are consistent with the nominal result of Fig. 7. More details about these consistency tests are given in Appendix B.

VI. CONCLUSION

We have presented a measurement of the average number of muons with energies greater than 500 GeV in near-vertical air showers observed in both IceTop and the IceCube in-ice array for cosmic-ray energies between 2.5 PeV and 100 PeV. The results were derived using the hadronic

interaction models Sibyll 2.1, QGSJet-II.04, and EPOS-LHC. They were compared with predictions from simulations, based on the respective hadronic models, and were found to agree within uncertainties with expectations from realistic primary cosmic-ray flux models.

The high-energy muon measurement was in addition compared to an earlier measurement of the density of low-energy muons at the surface performed with IceTop alone, by scaling both to expectations of proton and iron simulations. Represented in this fashion, both results should be in agreement and should point toward the same underlying mass composition. This is the case for the Sibyll 2.1 results. Interpreting the low-energy muon density data with the post-LHC models QGSJet-II.04 and EPOS-LHC yields, however, lighter mass compositions than the TeV muon measurement. While both measurements are still consistent within systematic uncertainties for QGSJet-II.04, this is not the case for EPOS-LHC. This tension indicates that the data are not described consistently by the simulations using this model, and confirms again the challenges in correctly describing the development of extensive air showers, as discussed in Section I. In addition, we note that preliminary work also indicates a possible inconsistency between the muon measurements and the slope of the lateral charge distribution in IceTop, in particular for Sibyll 2.1 [41].

The measurement of high-energy muons has considerable potential for future studies of air-shower development and tests of hadronic interaction models, including newer iterations of the models included in this work. The measurement of the average muon multiplicity can be extended toward higher primary energies, closing the gap to measurements at ultrahigh energy observatories, and a larger zenith range, probing different muon energies and atmospheric depths. It is of interest to study the distribution of high-energy muons in more detail, e.g. by considering in addition to the average number also its fluctuations [8, 42]. More precise combined measurements of the low- and high-energy muon content of air showers will be important toward resolving the Muon Puzzle, as they probe the muon energy spectrum, which differs between hadronic interaction models [43]. Of particular interest are measurements of the correlation between the low- and high-energy muons on an event-by-event basis; while this work already presented an event-by-event reconstruction of the high-energy muon number, a technique to get an improved low-energy muon estimator based on separately fitting the electromagnetic and muonic lateral distribution functions at the surface is under development [44].

The plans for IceCube-Gen2 and its surface array will increase the opening angle for possible coincident measurements between the surface and deep detector [45]. The planned surface radio antennas would offer sensitivity to the depth of shower maximum X_{max} in addition to the muon measurements, enabling more stringent tests of the consistency of hadronic models, an important condition for the unambiguous interpretation of indirect cosmic-ray measurements [11].

ACKNOWLEDGMENTS

The IceCube collaboration acknowledges the significant contributions to this manuscript from Stef Verpoest. The authors gratefully acknowledge the support from the following agencies and institutions: USA – U.S. National Science Foundation-Office of Polar Programs, U.S. National Science Foundation-Physics Division, U.S. National Science Foundation-EPSCoR, U.S. National Science Foundation-Office of Advanced Cyberinfrastructure, Wisconsin Alumni Research Foundation, Center for High Throughput Computing (CHTC) at the University of Wisconsin–Madison, Open Science Grid (OSG), Partnership to Advance Throughput Computing (PATH), Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS), Frontera and Ranch computing project at the Texas Advanced Computing Center, U.S. Department of Energy-National Energy Research Scientific Computing Center, Particle astrophysics research computing center at the University of Maryland, Institute for Cyber-Enabled Research at Michigan State University, Astroparticle physics computational facility at Marquette University, NVIDIA Corporation, and Google Cloud Platform; Belgium – Funds for Scien-

tific Research (FRS-FNRS and FWO), FWO Odysseus and Big Science programmes, and Belgian Federal Science Policy Office (Belspo); Germany – Bundesministerium für Bildung und Forschung (BMBF), Deutsche Forschungsgemeinschaft (DFG), Helmholtz Alliance for Astroparticle Physics (HAP), Initiative and Networking Fund of the Helmholtz Association, Deutsches Elektronen Synchrotron (DESY), and High Performance Computing cluster of the RWTH Aachen; Sweden – Swedish Research Council, Swedish Polar Research Secretariat, Swedish National Infrastructure for Computing (SNIC), and Knut and Alice Wallenberg Foundation; European Union – EGI Advanced Computing for research; Australia – Australian Research Council; Canada – Natural Sciences and Engineering Research Council of Canada, Calcul Québec, Compute Ontario, Canada Foundation for Innovation, WestGrid, and Digital Research Alliance of Canada; Denmark – Villum Fonden, Carlsberg Foundation, and European Commission; New Zealand – Marsden Fund; Japan – Japan Society for Promotion of Science (JSPS) and Institute for Global Prominent Research (IGPR) of Chiba University; Korea – National Research Foundation of Korea (NRF), Chung-Ang University Research Grant; Switzerland – Swiss National Science Foundation (SNSF).

-
- [1] J. Evans, D. G. Gamez, S. D. Porzio, S. Söldner-Rembold, and S. Wren, *Phys. Rev. D* **95**, 023012 (2017), arXiv:1612.03219 [astro-ph.HE].
 - [2] J. P. Yañez and A. Fedynitch, *Phys. Rev. D* **107**, 123037 (2023), arXiv:2303.00022 [hep-ph].
 - [3] R. Engel, D. Heck, and T. Pierog, *Annu. Rev. Nucl. Part. Sci.* **61**, 467 (2011).
 - [4] K.-H. Kampert and M. Unger, *Astropart. Phys.* **35**, 660 (2012), arXiv:1201.0018 [astro-ph.HE].
 - [5] T. Pierog, *EPJ Web Conf.* **208**, 02002 (2019).
 - [6] H. P. Dembinski *et al.* (EAS-MSU, IceCube, KASCADE-Grande, NEVOD-DECOR, Pierre Auger, SUGAR, Telescope Array, Yakutsk EAS Array), *EPJ Web Conf.* **210**, 02004 (2019), arXiv:1902.08124 [astro-ph.HE].
 - [7] A. Aab *et al.* (Pierre Auger), *Phys. Rev. D* **91**, 032003 (2015), [Erratum: *Phys. Rev. D* **91**, 059901 (2015)], arXiv:1408.1421 [astro-ph.HE].
 - [8] A. Aab *et al.* (Pierre Auger), *Phys. Rev. Lett.* **126**, 152002 (2021), arXiv:2102.07797 [hep-ex].
 - [9] J. Albrecht *et al.*, *Astrophys. Space Sci.* **367**, 27 (2022), arXiv:2105.06148 [astro-ph.HE].
 - [10] J. C. Arteaga Velazquez, *PoS ICRC2023*, 466 (2023).
 - [11] A. Coleman *et al.*, *Astropart. Phys.* **149**, 102819 (2023), arXiv:2205.05845 [astro-ph.HE].
 - [12] D. Soldin, *PoS ICRC2023*, 327 (2023), arXiv:2308.09079 [hep-ex].
 - [13] M. G. Aartsen *et al.* (IceCube), *JINST* **12** (03), P03012, arXiv:1612.05093 [astro-ph.IM].
 - [14] R. Abbasi *et al.* (IceCube), *Phys. Rev. D* **106**, 032010 (2022), arXiv:2201.12635 [hep-ex].
 - [15] B. Flaggs, A. Coleman, and F. G. Schröder, *Phys. Rev. D* **109**, 042002 (2024), arXiv:2306.13246 [hep-ph].
 - [16] R. Abbasi *et al.* (IceCube), *Nucl. Instrum. Meth. A* **700**, 188 (2013), arXiv:1207.6326 [astro-ph.IM].
 - [17] M. G. Aartsen *et al.* (IceCube), *Phys. Rev. D* **100**, 082002 (2019), arXiv:1906.04317 [astro-ph.HE].
 - [18] M. G. Aartsen *et al.* (IceCube), *Phys. Rev. D* **88**, 042004 (2013), arXiv:1307.3795 [astro-ph.HE].
 - [19] M. G. Aartsen *et al.* (IceCube), *JINST* **9**, P03009, arXiv:1311.4767 [physics.ins-det].
 - [20] S. Verpoest, *Studies of TeV muons in cosmic-ray air showers detected with IceTop and IceCube*, Ph.D. thesis, Ghent University, Ghent, Belgium (2022).
 - [21] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw, *FZKA-6019* (1998).
 - [22] S. De Ridder, *Sensitivity of IceCube cosmic ray measurements to the hadronic interaction models*, Ph.D. thesis, Ghent University, Ghent, Belgium (2019).
 - [23] A. Ferrari, P. R. Sala, A. Fasso, and J. Ranft, *CERN-2005-010, SLAC-R-773, INFN-TC-05-11, CERN-2005-10 10.2172/877507* (2005).
 - [24] T. T. Böhlen, F. Cerutti, M. P. W. Chin, A. Fassò, A. Ferrari, P. G. Ortega, A. Mairani, P. R. Sala, G. Smirnov, and V. Vlachoudis, *Nucl. Data Sheets* **120**, 211 (2014).
 - [25] E.-J. Ahn, R. Engel, T. K. Gaisser, P. Lipari, and T. Stanev, *Phys. Rev. D* **80**, 094003 (2009), arXiv:0906.4113 [hep-ph].
 - [26] S. Ostapchenko, *Phys. Rev. D* **83**, 014018 (2011), arXiv:1010.1869 [hep-ph].
 - [27] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, *Phys. Rev. C* **92**, 034906 (2015), arXiv:1306.0121 [hep-ph].
 - [28] S. Agostinelli *et al.* (GEANT4), *Nucl. Instrum. Meth. A*

- 506**, 250 (2003).
- [29] J. H. Koehne, K. Frantzen, M. Schmitz, T. Fuchs, W. Rhode, D. Chirkin, and J. Becker Tjus, *Comput. Phys. Commun.* **184**, 2070 (2013).
 - [30] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio (2014) arXiv:1412.3555 [cs.NE].
 - [31] F. Chollet *et al.*, Keras, <https://keras.io> (2015).
 - [32] M. Abadi *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems (2015), software available from tensorflow.org.
 - [33] J. Matthews, *Astropart. Phys.* **22**, 387 (2005).
 - [34] Replication Data for: Measurement of the mean number of muons with energies above 500 GeV in air showers detected with the IceCube Neutrino Observatory (2025).
 - [35] A. Van Overloop (IceCube), in *32nd International Cosmic Ray Conference*, Vol. 1 (Beijing, China, 2011) pp. 97–101.
 - [36] AIRS Science Team/Joao Teixeira (2013), AIRS/Aqua L3 Daily Standard Physical Retrieval (AIRS+AMSU) 1 degree x 1 degree V006, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC).
 - [37] T. K. Gaisser and S. Verpoest, *Astropart. Phys.* **133**, 102630 (2021), arXiv:2106.12247 [astro-ph.HE].
 - [38] T. K. Gaisser, *Astropart. Phys.* **35**, 801 (2012), arXiv:1111.6675 [astro-ph.HE].
 - [39] T. K. Gaisser, T. Stanev, and S. Tilav, *Front. Phys. (Beijing)* **8**, 748 (2013), arXiv:1303.3565 [astro-ph.HE].
 - [40] H. P. Dembinski, R. Engel, A. Fedynitch, T. Gaisser, F. Riehn, and T. Stanev, *PoS ICRC2017*, 533 (2018), arXiv:1711.11432 [astro-ph.HE].
 - [41] S. Verpoest *et al.* (IceCube), *PoS ICRC2021*, 357 (2021), arXiv:2107.09387 [astro-ph.HE].
 - [42] L. Cazon, R. Conceição, and F. Riehn, *Phys. Lett. B* **784**, 68 (2018), arXiv:1803.05699 [hep-ph].
 - [43] F. Riehn, R. Engel, A. Fedynitch, T. K. Gaisser, and T. Stanev, *Phys. Rev. D* **102**, 063002 (2020), arXiv:1912.03300 [hep-ph].
 - [44] R. Abbasi *et al.* (IceCube), *PoS ICRC2023*, 357 (2023), arXiv:2309.00741 [astro-ph.HE].
 - [45] M. G. Aartsen *et al.* (IceCube-Gen2), *J. Phys. G* **48**, 060501 (2021), arXiv:2008.04323 [astro-ph.HE].

Appendix A: Monte Carlo predictions of TeV muon multiplicity

To obtain predictions of the number of high-energy muons with negligible statistical uncertainty from simulations for comparison with the final analysis results, dedicated high-statistics CORSIKA-only simulations were produced, tracking particles down to 500 GeV only. Figure 13 shows the predictions for $\langle N_\mu \rangle$ ($E_\mu > 500$ GeV) at an observation level of 2837 m obtained from simulations of near-vertical proton and iron showers using different hadronic interaction models. QGSJet-II.04 predicts the highest number of muons, EPOS-LHC the lowest.

The Heitler-Matthews model of air-shower development, combined with the superposition assumption for showers initiated by nuclei of mass number A , predicts that the total number of muons in the shower depends on the cosmic-ray energy E and mass A as

$$N_\mu(E, A) = A^{1-\beta} \left(\frac{E}{\xi} \right)^\beta \quad (\text{A1})$$

where $\beta \approx 0.9$ is related to the multiplicity of the hadronic interactions and ξ is a constant called the critical energy [33]. From Fig. 13, it can be seen that the high-energy muon number has a similar dependence on energy. Fitting the slope, the value of β is found to be about 0.82 for proton and 0.76 for iron. According to Eq. (A1), the exponent also determines the difference between N_μ for proton and heavier nuclei. From the difference between the proton and iron predictions at fixed E , β is derived to be about 0.77 at the lower primary energies increasing to about 0.82 at the higher end.⁷ The fact that the β values derived from the energy dependence and the mass dependence are numerically close implies that the superposition model also holds well for the description of the high-energy muons. It is interesting to note that this different behavior compared to the total muon number makes the interpretation of a high-energy muon measurement less sensitive to possible offsets in the energy scale of an experiment [6]. Whereas a 10% shift in energy scale would result in a difference of about 0.2 in z (Eq. (2)) for a surface muon measurement, this would cause a shift of only about 0.08 in z for the high-energy muons.

While the relation between $\langle N_\mu \rangle$ and E is close to a straight line on the logarithmic plot, Fig. 13, the residuals after a linear fit show small deviations with a quadratic behavior on the level of 2%. For the accuracy of the calculations of e.g. the final z -values of Fig. 11, a quadratic term was added to the fit describing $\langle N_\mu \rangle(E)$.

⁷ This smaller value of β compared to the ~ 0.9 corresponds to the proton-iron difference for high-energy muons being larger than for the total number of muons at the surface. The high-energy muons are thus more sensitive to the mass composition.

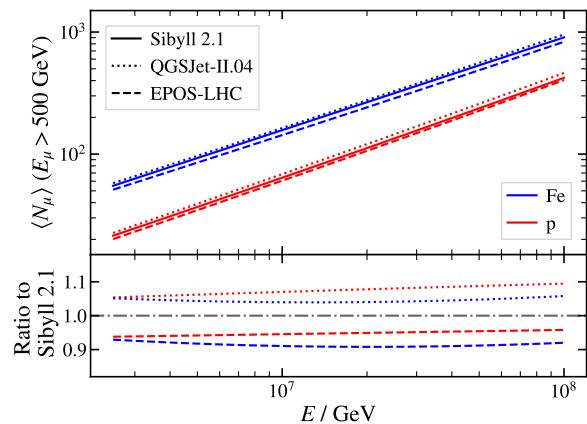


Figure 13. Average multiplicity of muons with an energy above 500 GeV in near-vertical air showers ($\cos \theta > 0.95$) as a function of cosmic-ray energy. The lines are quadratic fits to predictions obtained from CORSIKA simulations for proton and iron primaries based on different hadronic interaction models.

Appendix B: Neural network consistency checks

In this section of the Appendix, various checks are presented showing the robustness of the analysis.

1. Separate E and N_μ networks

The neural network model in this analysis uses three inputs to reconstruct the primary cosmic-ray energy E and the high-energy muon number N_μ ($E_\mu > 500$ GeV): the shower size S_{125} and zenith angle θ from IceTop, and a vector of energy losses from the IceCube in-ice array. To ensure that consistent results are obtained when the primary energy and the muon number reconstructions are performed completely independently, two different neural networks were trained. One uses S_{125} and θ to reconstruct E , the other uses the energy losses to reconstruct N_μ . The reconstruction quality for these networks is shown in Fig. 14. Compared to the performance of the combined neural network, shown in Fig. 3, the E reconstruction has an increased mass-dependent bias, indicating that the combined neural network uses the composition information present in the muon-bundle signal to improve the E reconstruction. The N_μ neural network, which does not use any surface information, shows, on the other hand, a reduced mass dependence in the bias.

As before, correction factors can be determined for each primary type by comparing the reconstructed and true average muon number $\langle N_\mu \rangle$. The resulting biases are shown in Fig. 15. They are found to be smaller than those in Fig. 4, however, they do not have a monotone relation with the mass of the primary. Therefore, an iterative correction where one interpolates between the proton and iron correction factors as in Section IV B cannot be applied; this is the main reason for using the combined

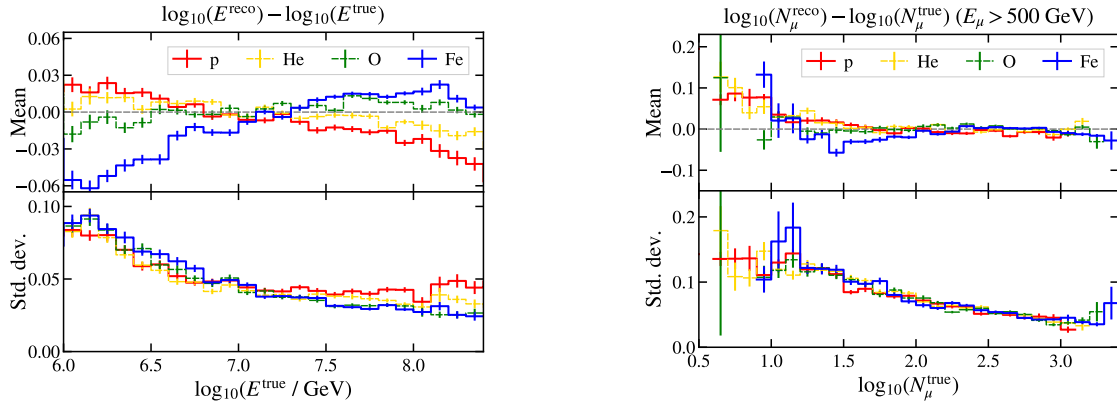


Figure 14. Bias and resolution of the energy reconstruction based only on IceTop information (left) and the TeV muon number reconstruction based only on in-ice information (right).

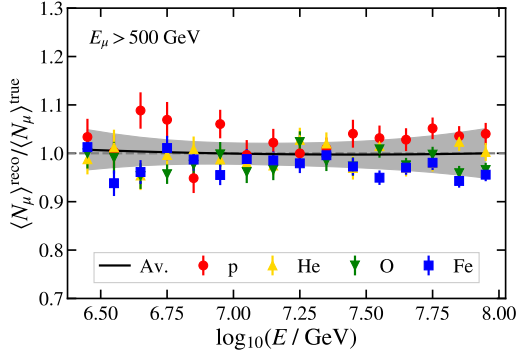


Figure 15. Ratio of reconstructed and true average muon number obtained from the reconstructions with separate neural networks for E and N_μ . The black line defines the average of fits to proton and iron, the shaded area shows the uncertainty assigned to cover the unknown mass composition.

neural network in the main body of this work. To derive results based on the separate reconstruction, a single correction factor is derived as the average between the proton and iron correction factors, and the spread between the two is added as an extra systematic uncertainty on the result obtained with this correction factor.

The result obtained with the separate neural networks and the average p-Fe correction factor is shown in Fig. 16 compared to the nominal result presented in Fig. 7. The two are in agreement for the entire energy range.

We note that, while no iterative correction as presented in Section IV B can be applied here, the smaller biases observed in Fig. 15 as compared to Fig. 4 would lead to uncertainties of a similar size on the final result.

2. Training on different hadronic models

The nominal results presented in the main body of this paper are obtained using a neural network trained

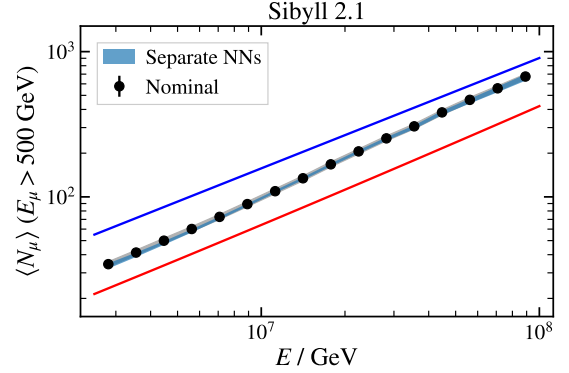


Figure 16. Comparison between the nominal result of the analysis and the result obtained using separate neural networks with only IceTop input for the energy reconstruction and only IceCube in-ice array input for the muon multiplicity reconstruction.

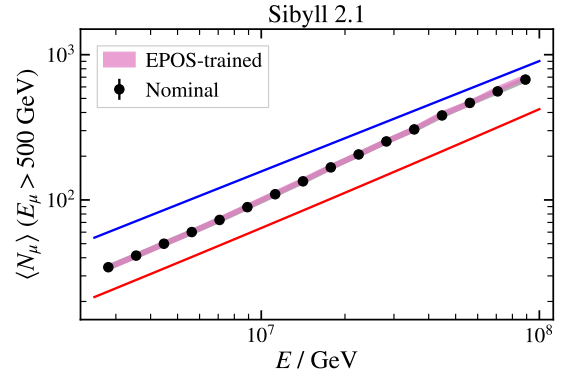


Figure 17. Comparison between the nominal result of the analysis and the result obtained using a neural network that was trained on simulations based on EPOS-LHC. In both cases, correction factors were derived from Sibyll 2.1 simulations.

on Sibyll 2.1 after which correction factors derived from Sibyll 2.1, QGSJet-II.04, and EPOS-LHC are applied. We show here that the results depend only on the model that the correction factors are derived from and are robust against changes in the model that the neural network is trained on. To this end, a neural network was trained on EPOS-LHC simulations. A correction factor was then derived by applying this network to Sibyll 2.1 simulations. The result obtained from experimental data with this combination of neural network and correction factors is shown in Fig. 17. The result is consistent with the nominal result obtained from training a network on Sibyll 2.1 and then deriving a correction factor from Sibyll 2.1. Similarly, training the network on EPOS-LHC and deriving correction factors based on EPOS-LHC gives consistent results as using the Sibyll 2.1-trained network and the EPOS-LHC-derived correction factors from Section IV.

3. Simple shower-size energy estimator

Previous IceTop analyses typically estimate the primary cosmic-ray energy through a conversion function based on the shower size S_{125} defined for a specific range of zenith angles. The IceTop analysis of the surface muon density, of which some results were shown in Section V, uses for its energy estimate a conversion function derived for the energy-spectrum analysis performed with IceTop-73, before the completion of the full IceTop detector, as described in Ref. [18]. The energy estimate in the high-energy muon analysis of this work comes from a neural network which also gets input from the in-ice detector. Earlier in this Appendix it was already confirmed that obtaining an energy estimate from a separate neural network based only on S_{125} and θ from IceTop does not produce a significantly different result. Here, we test the effect of using the simple S_{125} conversion function used in the GeV muon density analysis, while leaving the N_μ reconstruction unchanged. In Fig. 18, it is observed that the resulting muon measurement is again consistent with the nominal result.

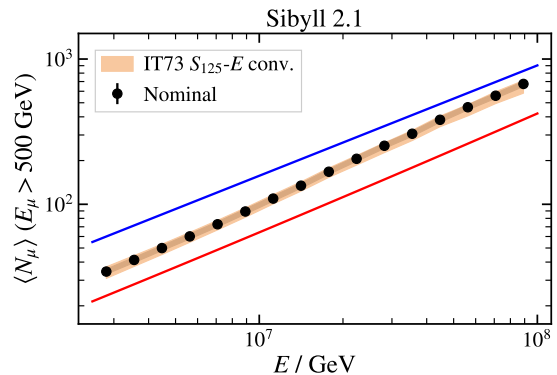


Figure 18. Comparison between the nominal result of the analysis and the result obtained when the energy is reconstructed with a simple conversion from S_{125} derived for the IceTop-73 energy-spectrum analysis (see text for details).