

Search for Astrophysical Tau Neutrinos with an Improved Double Pulse Method

The IceCube Collaboration*

http://icecube.wisc.edu/collaboration/authors/icrc19_icecube

E-mail: mmeier@icecube.wisc.edu, jsoedingrekso@icecube.wisc.edu

Tau neutrino identification with the IceCube experiment would open new windows to neutrino physics as well as enable novel searches for cosmic neutrino sources. This work aims at a identification of tau neutrinos with astrophysical origin at energies above $\sim 100\text{TeV}$. For identification, we search for a double pulse structure in the signal of one of IceCubes Digital Optical Modules originating from a tau neutrino interaction and a subsequent tau decay within the detector. In this work, we present constraints on the tau neutrino flux based on an event sample with a livetime of about 7.5 years of IceCube data.

Corresponding authors: Maximilian Meier¹, Jan Soedingrekso^{†1}

¹ *TU Dortmund University*

*36th International Cosmic Ray Conference -ICRC2019-
July 24th - August 1st, 2019
Madison, WI, U.S.A.*

*For collaboration list, see PoS(ICRC2019) 1177.

[†]Speaker.

1. Introduction

The IceCube collaboration first observed an astrophysical flux of high-energy neutrinos in 2013 [1]. Although the flavor ratio of neutrinos produced by pion decays would be $\nu_e : \nu_\mu : \nu_\tau = 1 : 2 : 0$ at a source, due to neutrino oscillations over astrophysical baselines the flavor ratio at Earth is modified to $\nu_e : \nu_\mu : \nu_\tau \simeq 1 : 1 : 1$. Tau neutrino production in atmospheric air showers is negligible [2], which makes the detection of tau neutrinos another important verification of the astrophysical origin of the observed high-energy neutrinos. The tau neutrino component of the astrophysical flux has yet to be identified. Previous searches have only produced upper limits on the astrophysical tau neutrino flux [3, 4] but, recently, the first tau neutrino candidates were presented [5].

IceCube is a neutrino detector deployed deep in the ice at the geographic south pole instrumenting a volume of 1 km^3 [6]. IceCube consists of digital optical modules (DOMs) that observe Cherenkov light produced by secondary particles from neutrino interactions. This analysis aims to identify tau neutrinos via the so called double pulse signature [7], where a single DOM records two distinct light depositions: the first from the hadronic cascade induced by a charged current (CC) ν_τ interaction and the second one from the subsequent decay of the tau lepton (into an electron or into hadrons). This signature has two different backgrounds: single cascades produced by ν_e CC and all flavor neutral current (NC) interactions and tracks created by atmospheric muons and ν_μ CC interactions (and ν_τ CC interactions followed by a tau decaying into a muon). Section 2 introduces an event selection removing single cascade events by identifying double pulse waveforms and track events by their topological differences to cascades using machine learning.

Based on the ν_τ dominated event sample the astrophysical ν_τ flux normalization is measured in a binned Poisson likelihood fit, described in section 3. In this analysis, 7.5 years of IceCube data¹ recorded between May 2011 and December 2018 (with an effective livetime of 2666.8 days) are analyzed. Two double pulse events found by this analysis are characterized in section 4.

2. Event Selection

The event selection presented here improves the expected tau neutrino event rate of the three year double pulse search [3] by a factor of 2.5. The event selection is structured in the following way. First, IceCube waveforms are analyzed to select events containing at least one double pulse waveform. Double pulses are identified by training a Random Forest (RF) [8] with Monte Carlo simulations for signal (double pulse waveforms from ν_τ CC interactions) and cascade background (single cascade waveforms from ν_e CC and all-flavor ν NC interactions). For the RF the observables from [3] are used to characterize the waveforms by detecting rising and trailing edges based on the waveform derivative. Additional observables are used, which are described in [9]. These observables are summary statistics (e.g. the mean) of the waveform, the number of local maxima, the difference between unsmoothed and smoothed waveforms, and the compatibility of the waveform with a fit to an exponential function starting from the first local maximum. The RF assigns a score value to each waveform called “Double Pulse score”, where a value of 1 represents a signal-like waveform. The performance of the RF is illustrated in Fig. 1. It shows the remaining ν_τ CC

¹ All of the analyzed data was recorded while the detector has been in full operation with 86 strings.

events and cascade-like events containing at least one double pulse waveform as a function of the Double Pulse score cut, which is the threshold at which a waveform is accepted as a double pulse waveform. Here, a threshold of 0.2 is chosen, reducing the background from single cascades by more than three orders of magnitude to a subdominant level.

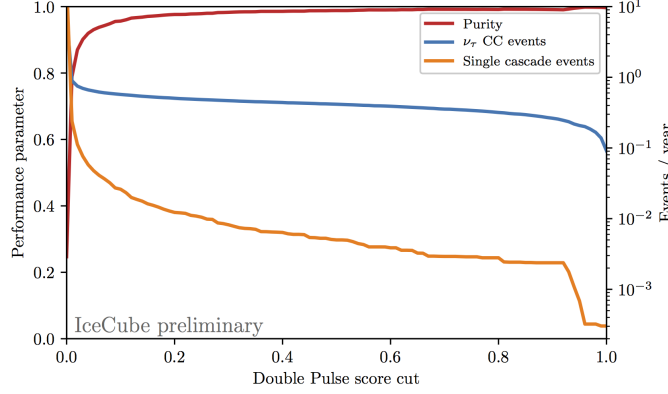


Figure 1: Influence of the Double Pulse score cut on expected events per year and the purity achieved for ν_τ CC with respect to all cascade-like events.

After the process of double pulse identification, the track-like background events are addressed. The sought after signal of ν_τ CC events is still dominated by track-like events originating from atmospheric muons and ν_μ CC interactions. The reason for that is that a muon can produce double pulses by a combination of Cherenkov light emission and high energy stochastic energy losses close to a DOM. In case of a ν_μ CC interaction the hadronic cascade can also be responsible for one of the pulses.

Atmospheric muons from cosmic ray air showers are primarily described with CORSIKA simulations [10]. For this analysis the most difficult background component is single muons depositing most of their energy in a single energy loss near the edge of the detector. To simulate this background more efficiently we rely on a parametrization of the single muon yield from CORSIKA, which is introduced in [11]. The separation is, again, carried out by a RF using observables focusing on the event topology and the location of the event in the detector.

The left panel of Fig. 2 shows the distribution of the “Cascade score” for different simulated components. The signal ν_τ CC interactions are split up into a contained and an uncontained fraction, as only the contained events are used as signal events in the training. Besides that no explicit containment cut or veto is used. The right panel of Fig. 2 compares the sum of all simulated components with 7.5 years of data, which shows good agreement. The error bands on the expectation are a combination of Poisson uncertainties and uncertainties due to limited MC simulations [12]². The bottom panel shows the ratio between data and simulations. The “p-value*” is the p-value to observe the value found in data with respect to the expected value.

The threshold on the Cascade score was chosen by optimizing the model rejection factor [14] keeping all events with a score above 0.62 for further analysis. For an astrophysical per-flavor flux of $1.01 \times 10^{-18} \text{ GeV}^{-1} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} (E_\nu/100 \text{ TeV})^{-2.19}$ [15], the conventional atmospheric

²This way of comparing histograms of data and simulation is suggested in [13].

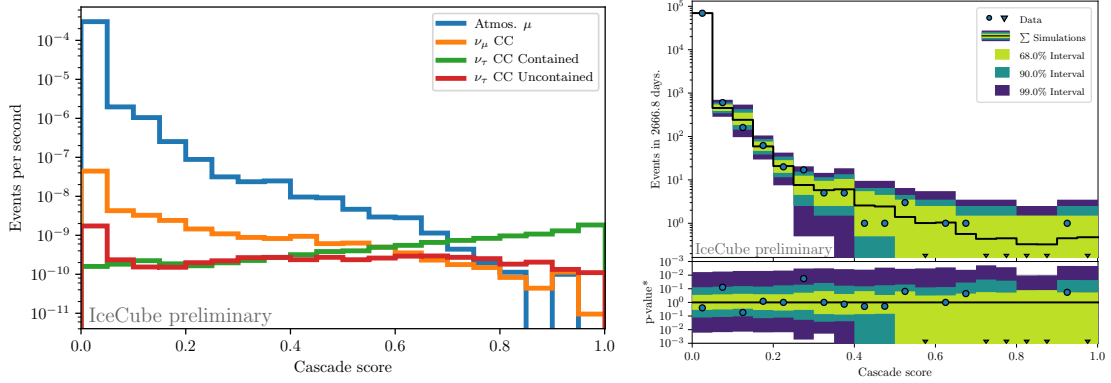


Figure 2: Cascade score distribution for different simulated components (left) and a comparison between the sum of simulations and 7.5 years of IceCube data with an effective livetime of 2666.8 days (right).

neutrino flux from [16] and the prompt neutrino flux from [17] the expected number of signal and background events is 2.10 and 0.98 respectively in the final sample for the analyzed time period of 7.5 years of data.

3. Analysis Method

The goal of this analysis is to constrain the tau neutrino flux normalization. This is done with a binned Poisson likelihood maximization with the bin-wise expectation μ_i defined as:

$$\mu_i(\lambda) = \mu_{B,i} + \lambda \mu_{S,i}, \quad (3.1)$$

with the background expectation μ_B , the signal expectation μ_S and λ as the only free parameter scaling the tau neutrino flux normalization. The astrophysical spectral index is kept fixed during the analysis as the number of overall expected events is small (~ 3) and thus not sufficient to measure multiple parameters. Instead, three spectral indices are tested ($E^{-2.19}$ [15], $E^{-2.50}$ [18] and $E^{-2.89}$ [19]). The final sample also contains a considerable fraction of background events, so favoring observables with strong separation power between signal and background is beneficial.

The observables used are the Double Pulse score and the Cascade score, that were already introduced earlier. One event can have multiple double pulse waveforms, therefore only the highest value per event for the Double Pulse score is used. Their 2D-distributions are presented in Fig. 3. The binning is chosen to have two irregular bins in Double Pulse score and seven equidistant bins in Cascade score, where the total number of bins is mainly restricted by limited simulations. The confidence interval for the likelihood fit is constructed via hypothesis test inversion [20]

$$C(x) = \{\lambda_0 \in \Lambda : -\Delta \ln(\mathcal{L}(\lambda_0|x)) \geq k_\alpha\} \quad (3.2)$$

where x denotes the observed data sample, Λ the allowed parameter space for λ , $-\Delta \ln(\mathcal{L}(\lambda_0|x))$ the value of the likelihood ratio test statistic for the null hypothesis $\lambda = \lambda_0$ given the observation x and k_α the critical value for the likelihood ratio test at confidence level (CL) $1-\alpha$.

Based on the same observable distributions observed event candidates are assigned a p-value to describe their compatibility with originating from the background distribution. This is constructed from the likelihood

$$\mathcal{L}(\lambda|i) = \frac{P_{B,i} + \lambda P_{S,i}}{1 + \lambda}, \quad (3.3)$$

with the probability distributions P_B and P_S describing the simulated events in the described binned 2D observable space. The parameter λ is bound to the interval $[0, 1]$ and is used to fit the “signalness” in the corresponding observable bin i . The background distribution P_B is confirmed to be consistent with data in a Cascade score region close to the signal region.

4. Results

IceCube data recorded between May 2011 and December 2018 was analyzed and two tau neutrino candidates were observed. Event#1, observed in the 2014 season³, is presented in the left panel of Fig. 4⁴. The right panel shows the three observed double pulses, located on three adjacent DOMs, with double pulse scores between 0.45 and 1.0. The Cascade score for this event is found to be 0.92. The resulting p-value is $p \simeq 0.035$, almost independent of the assumed astrophysical spectrum. This event is also observed as a double cascade tau neutrino candidate by [21] and as a double pulse tau neutrino candidate by [22].

Event#2 was recorded in the 2015 season and is depicted in the left panel of Fig. 5. The right panel shows the observed double pulse waveform with a double pulse score of 0.565. The Cascade score for this event is 0.675. The resulting p-value is $p = 1.0$ for all tested astrophysical spectra. The cascade score for both events is indicated in Fig. 2 (right) and Fig. 3.

The resulting data sample was used to measure the tau neutrino flux normalization with a binned Poisson likelihood fit as described in the previous section for three different assumptions on the shape of the astrophysical spectrum: $E^{-2.19}$, $E^{-2.50}$ and $E^{-2.89}$. Fig. 6 shows the likelihood scan when assuming an astrophysical spectral index of $E^{-2.19}$ as a black line. The color scale in the background indicates the TS distributions obtained by pseudo-experiments for different injected

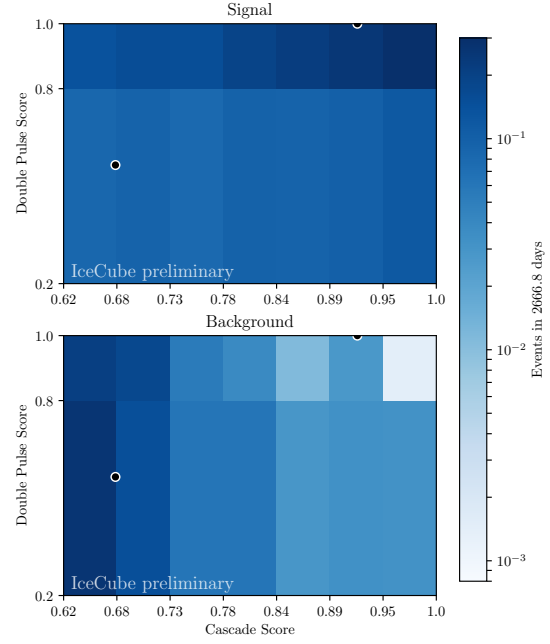


Figure 3: Binned 2D-distribution of the Double Pulse score and the Cascade score. The top panel shows the signal component and the bottom panel the sum of all background components. The depicted distributions assume an astrophysical $E^{-2.19}$ spectrum. The two observed events are indicated by black dots. Event#1 is located in the signal dominated region at (0.92, 1.0) and Event#2 in the background dominated region at (0.675, 0.565).

³A south pole season starts in spring and ends the spring of the following year.

⁴IceCube event views represent observed charge as a sphere around the DOM observing it. The size of the sphere is proportional to the deposited charge and the arrival time is encoded with different colors from early (red) to late (blue).

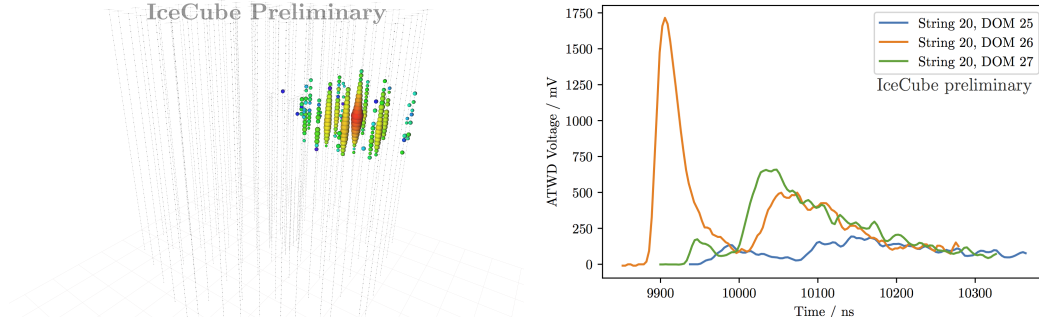


Figure 4: Left: Event view of the double pulse event recorded in the 2014 season. Right: Double pulse waveforms recorded for this event. The corresponding double pulse scores are 0.45 (blue), 1.0 (orange) and 0.81 (green). This event was also observed by [22]. The orange and green waveforms are seen by that analysis as well. The blue waveform is only classified as a double pulse waveform in this analysis.

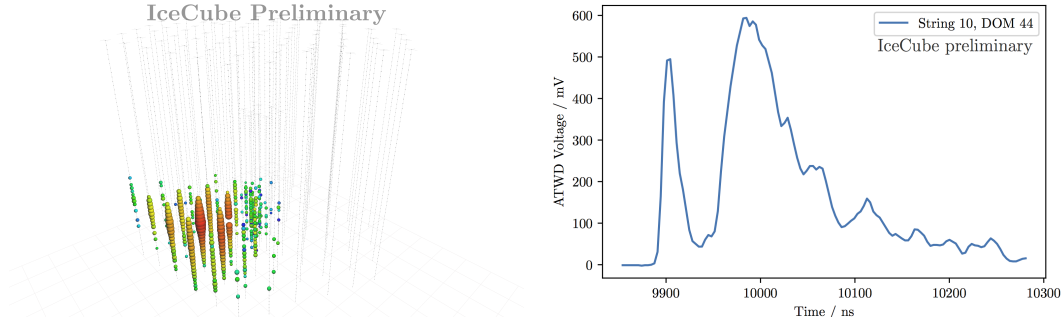


Figure 5: Left: Event view of the double pulse event recorded in the 2015 season. Right: Double pulse waveform recorded for this event with a double pulse score of 0.565.

values for the mean signal expectation μ_s . The red line depicts the critical values to reject a certain value of μ_s at 68 % CL. Then the confidence interval is constructed by finding the set of μ_s values that can not be rejected at a certain CL.

The measured tau normalizations are presented in Tab. 1. Independent of the spectral index the fit prefers a non-zero normalization. They are smaller than the baseline normalizations due to one observed event being a small, yet insignificant underfluctuation relative to the expectation. Within the 1σ uncertainty they agree with a flavor composition of $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$. In each case, the observation is incompatible with a tau normalization of 0 at the $\sim 1.9\sigma$ level. As already mentioned, systematic uncertainties are not included in this measurement. Their influence is small compared to the statistical uncertainty due the small data sample, and they will be discussed in more detail in a forthcoming publication.

Besides the model dependent constraints, based on the findings a differential upper limit was constructed to obtain a less model dependent constraint. The differential upper limit is constructed the same way as described in [26] and [3]. The same likelihood formalism as described in section 3 is used, but for each tested energy E_c a signal flux proportional to E^{-1} centered around E_c with a

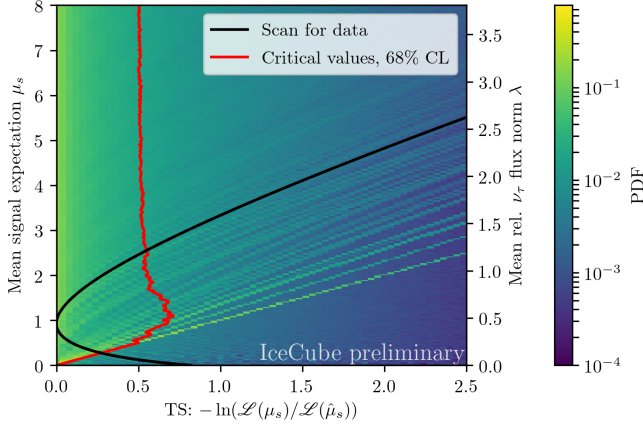


Figure 6: Likelihood scan for the observed data. The red line shows critical values at 68 % CL obtained from the underlying TS distributions for each value of μ_s .

γ	Φ_{0,ν_τ} at 100 TeV / ($10^{-18} \text{ GeV}^{-1} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$)
2.19	$0.45^{+0.79}_{-0.31}$
2.50	$0.83^{+1.46}_{-0.59}$
2.89	$1.62^{+2.78}_{-1.11}$

Table 1: Measurement of the astrophysical tau neutrino flux normalization for different astrophysical spectra.

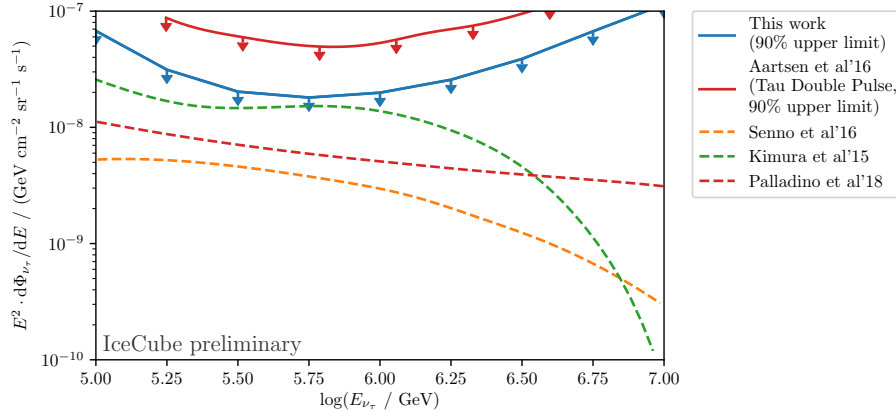


Figure 7: Differential upper limit on the tau neutrino flux between 100 TeV and 10 PeV. The red solid line shows the differential upper limit observed by a previous search for astrophysical tau neutrinos conducted on three years of IceCube data [3]. As a comparison, models for per-flavor astrophysical fluxes from choked jets and low luminosity GRBs [23] (orange), low luminosity AGNs [24] (green) and from a multi-component model [25] (red) are shown.

width of one decade is injected and the resulting 90 % upper limits are calculated. The differential upper limit is presented in Fig. 7 and compared to the differential limit of the previous 3 year double pulse analysis [3] and to three models predicting the astrophysical neutrino flux from choked jets and low luminosity gamma-ray bursts [23], low luminosity AGNs [24] and a multicomponent model [25].

5. Summary

We have performed a new measurement of the astrophysical tau neutrino flux based on 7.5 years

of IceCube data. An improved method using machine learning to identify tau neutrinos via their double pulse signature is presented. Two tau neutrino double pulse candidates were found. One candidate, also observed in the IceCube double cascade analysis [21], is found to be signal-like with a p-value of $p \simeq 0.035$ based on the observables used in this analysis. A model dependent measurement of the astrophysical tau neutrino flux normalization is presented in Tab. 1, finding a non-zero normalization with a statistical significance on the 1.9σ level. The observed values are compatible with a flavor ratio of $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$ and are not in conflict with previously reported upper limits from IceCube. Albeit two tau neutrino candidates were detected the differential upper limit on the tau neutrino flux, previously reported in [3], was also improved.

References

- [1] **IceCube** Collaboration, M. G. Aartsen et al., *Science* **342** (2013).
- [2] R. Enberg, M. H. Reno, and I. Sarcevic, *Phys. Rev. D* **78** (Aug, 2008) 043005.
- [3] **IceCube** Collaboration, M. G. Aartsen et al., *Phys. Rev.* **D93** (2016) 022001.
- [4] **IceCube** Collaboration, [PoS \(ICRC2017\) 974](#) (2018).
- [5] **IceCube** Collaboration, M. G. Aartsen et al., *EPJ Web Conf.* **207** (2019) 02005.
- [6] **IceCube** Collaboration, M. G. Aartsen et al., *JINST* **12** (2017) P03012.
- [7] D. Cowen, *Journal of Physics: Conference Series* **60** (3, 2007) 227–230.
- [8] L. Breiman, *Machine Learning* **45** (2001) 5–32.
- [9] **IceCube** Collaboration, [PoS \(ICRC2017\) 1009](#) (2018).
- [10] D. Heck et al., *CORSIKA: a Monte Carlo code to simulate extensive air showers*. Feb., 1998.
- [11] **IceCube** Collaboration, M. G. Aartsen et al., *Phys. Rev.* **D91** (2015) 022001.
- [12] C. A. Argüelles, A. Schneider, and T. Yuan, [arXiv:1901.04645](#).
- [13] R. Aggarwal and A. Caldwell, *Eur. Phys. J. Plus* **127** (2012) 24.
- [14] G. C. Hill and K. Rawlins, *Astropart. Phys.* **19** (2003) 393–402.
- [15] **IceCube** Collaboration, [PoS \(ICRC2017\) 1005](#) (2018).
- [16] M. Honda et al., *Phys. Rev. D* **75** (Feb, 2007) 043006.
- [17] A. Bhattacharya et al., *Journal of High Energy Physics* **2015** (Jun, 2015) 110.
- [18] **IceCube** Collaboration, M. G. Aartsen et al., *Astrophys. J.* **809** (2015) 98.
- [19] **IceCube** Collaboration, [PoS \(ICRC2019\) 1004](#) (these proceedings).
- [20] G. Casella and R. L. Berger, *Statistical Inference*. 2nd. Duxbury, 2002.
- [21] **IceCube** Collaboration, [PoS \(ICRC2019\) 1015](#) (these proceedings).
- [22] **IceCube** Collaboration, [PoS \(ICRC2019\) 1036](#) (these proceedings).
- [23] N. Senno, K. Murase, and P. Meszaros, *Phys. Rev.* **D93** (2016) 083003.
- [24] S. S. Kimura, K. Murase, and K. Toma, *Astrophys. J.* **806** (2015) 159.
- [25] A. Palladino and W. Winter, *Astron. Astrophys.* **615** (2018) A168.
- [26] **IceCube** Collaboration, M. G. Aartsen et al., *Phys. Rev.* **D98** (2018) 062003.