

First Double Cascade Tau Neutrino Candidates in IceCube and a New Measurement of the Flavor Composition

The IceCube Collaboration*

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

E-mail: juliana.stachurska@icecube.wisc.edu

The IceCube Neutrino Observatory at the South Pole, which detects Cherenkov light from charged particles produced in neutrino interactions, firmly established the existence of an astrophysical high-energy neutrino component. The expected neutrino flavor composition on Earth is $\nu_e : \nu_\mu : \nu_\tau$ of about 1:1:1 for neutrinos produced in astrophysical sources through pion decay. A measurement of the flavor composition on Earth can provide important constraints on sources and production mechanisms within the standard model, and can also constrain various beyond-standard-model processes. Here the measurement of the flavor composition performed on IceCube's High-Energy Starting Events sample with a livetime of about 7.5 years is presented. IceCube is directly sensitive to each neutrino flavor via the single cascade, track and double cascade event topologies. In IceCube, ν_τ -CC interactions above ~ 100 TeV can produce resolvable double cascades, breaking the degeneracy between ν_e and ν_τ present at lower energies. IceCube's first two identified double cascades are presented and the properties of the two ν_τ candidates are discussed.

Corresponding author: J.Stachurska[†]

DESY, Platanenallee 6, D-15738 Zeuthen, Germany

*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 Neutrino Observatory [1] reported the discovery of a diffuse astrophysical neutrino flux in 2012 using the High-Energy Starting Events (HESE) selection [2]. The HESE selection is an all-flavor event selection, with events previously being classified by eye as cascade-like or track-like. In the newest iteration of HESE with 7.5 years of data, older IceCube data have been reprocessed and re-analyzed following a new detector calibration, and new data have been added. Further improvements have been made to the atmospheric background modeling and to the likelihood used for the analyses. Details as well as results of the diffuse flux measurement can be found in [3]. We have further incorporated an algorithmic topology identification [4] into the HESE processing, and performed a new flavor composition measurement. We classify two events as double cascades for the first time, yielding the first candidates for astrophysical tau neutrinos. The results were first presented in [5]. After summarizing the results, we here present the details of follow-up work, aimed at assessing the probability of each of the two candidates to stem from a ν_τ charged current (CC) interaction in greater detail.

2. The Flavor Composition in the High-Energy Starting Events

There are three event topologies in IceCube to be distinguished and matched to the three neutrino flavors for a flavor composition measurement. Light depositions along an infinite track traversing the detector are called tracks and stem from ν_μ -CC interactions and atmospheric muons, as well as ν_τ -CC interactions where the τ lepton decays to a muon. Single cascades consist of one energy deposition and are produced by ν_e -CC and all flavor neutral current (NC) interactions. Double cascades are two energy depositions connected in space and time by a (dim) track. They are produced by ν_τ -CC interactions where the first cascade comes from the neutrino interaction and τ production, and the second from the τ decaying electromagnetically or hadronically. However, the short tau decay length of $\langle L_\tau \rangle \sim 50 \text{ m} \cdot E_\tau / \text{PeV}$, where E_τ is the τ energy, makes the distinction between single and double cascades very challenging in IceCube, where the mean horizontal distance between sensors is 125 m. In IceCube, the double cascade topology opens up at deposited energies $\mathcal{O}(100 \text{ TeV})$, breaking the degeneracy between ν_e and ν_τ flavors present at lower energies. All events are reconstructed using maximum likelihood fits for different hypotheses: single cascade [6], track [7] and double cascade [6, 8]. For the fits, the timing and spatial information of all light collected in an event is used. For a double cascade reconstruction to be deemed of sufficient quality, the fit must converge, a minimum of 1 TeV reconstructed energy is required for each of the cascades and the opening angle between the reconstructions using a double cascade and a track hypothesis cannot exceed 30° . Additionally, neither cascade can be more than 50 m outside of the instrumented volume. This soft containment requirement becomes increasingly constraining with increasing tau decay length, leading to a drop of successfully identifiable double cascades above lengths of 300 – 400 m. Observables used for classification are the distance between the reconstructed cascades (called *double cascade length* hereafter), the asymmetry between the two cascades' energies (called *energy asymmetry* hereafter), and the fraction of the total energy deposited close to the cascade vertices (called *energy confinement* hereafter). The energy asymmetry is defined as $A_E = (E_1 - E_2)/(E_1 + E_2)$, where $E_{1,2}$ are the reconstructed energies of the two cas-

Observable	Min	Max
Quality	good	
Length	10 m	–
Energy Asymmetry	-0.98	0.30
Energy Confinement	0.99	1.0

Table 1: Double cascade cuts. Note that while there is no upper limit on the length, the requirement that both cascades need to be contained within 50 m of the detector boundaries leads to an effective cutoff above $\sim 300 - 400$ m.

acades, and can take values $-1 \leq A_E \leq 1$. Events passing the quality, length, energy asymmetry and energy confinement cuts given in Table 1 are classified as double cascades. While the length resolution is ~ 2 m, the minimum required length for double cascades is 10 m. The larger the minimum required length, the smaller is the contamination by true single cascades, but also the signal expectation drops rapidly above lengths of $\sim 20 - 30$ m. Events failing any of the double cascade cuts are classified as indicated in Table 2.

Failed double cascade cut	resulting classification
Quality	depending on fit likelihoods
Length	single cascade
Energy Asymmetry	single cascade
Energy Confinement	track

Table 2: Track and single cascade classification. The cuts are given in order of precedence, such that the first failed cut determines the resulting classification. For events failing the double cascade fit quality cuts, the likelihoods of the track and single cascade fits are compared, and the topology with the corresponding fit with the higher likelihood is chosen.

Above 60 TeV reconstructed deposited energy, the HESE sample contains 60 events which we classify into 42 single cascades, 16 tracks, and 2 double cascades. A multi-component maximum-

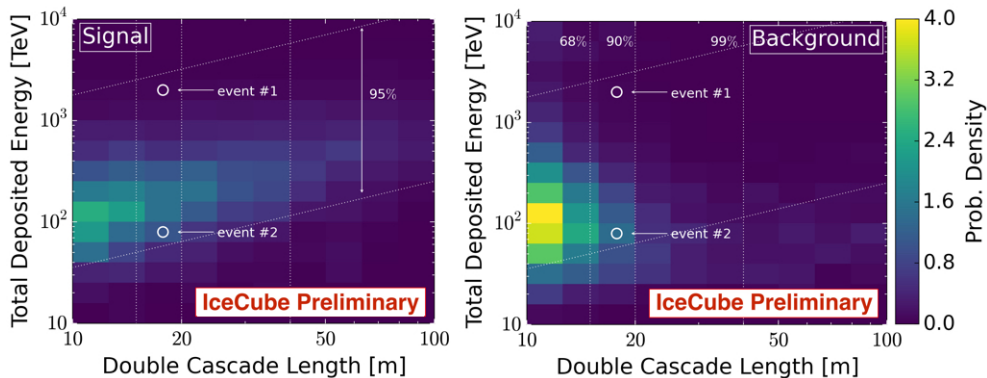


Figure 1: Total deposited energy against reconstructed double cascade length for the double cascade sample. Signal (ν_τ -induced double cascade events) histogram (left). Background (all remaining events) histogram (right). The two tau-neutrino candidate events are overlaid as white circles.

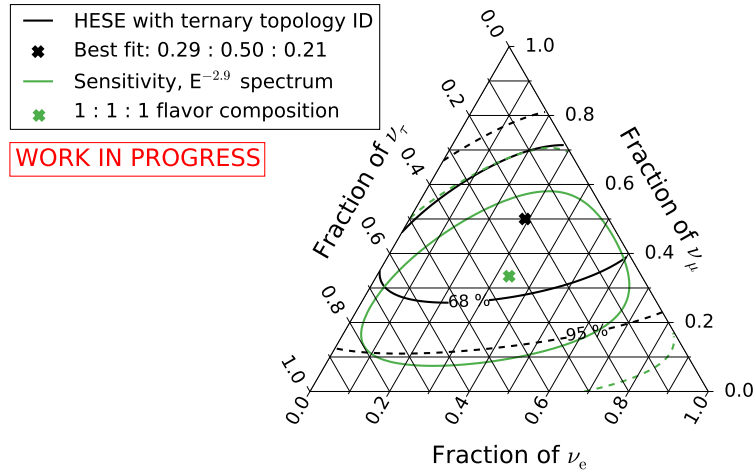
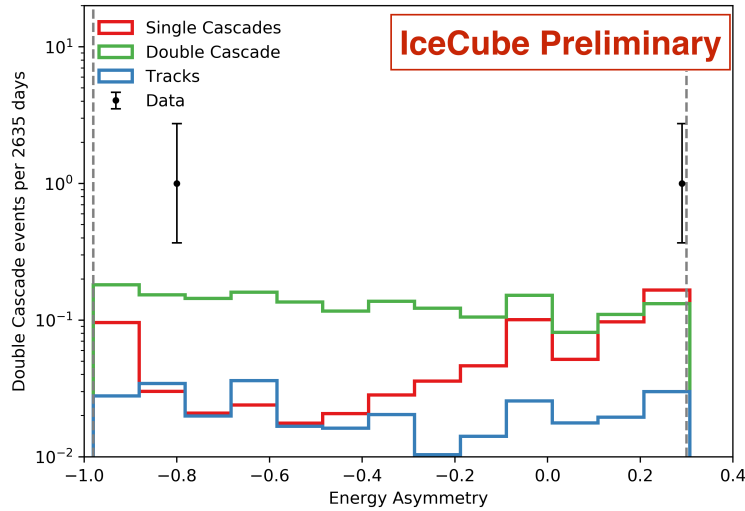


Figure 2: Measured flavor composition of IceCube HESE events with ternary topology ID and sensitivity at the best fit spectrum. Contours obtained using Wilks’ theorem [9].

likelihood fit is performed on the three topology samples using two-dimensional PDFs obtained from Monte Carlo simulations, as shown in Fig. 1 for the double cascade topology sample. The likelihood used is the SAY-likelihood [10], developed for the 7.5 year HESE update, which takes limited Monte Carlo statistics into account. For single cascades and tracks, the observables *total deposited energy* and *cosine θ* are used, where θ is the zenith angle. For double cascades, total deposited energy and double cascade length are used. For ν_τ induced double cascades, we expect a correlation between the energy of the event and the tau decay length. Events stemming from other flavors typically cluster at the thresholds, both in total deposited energy due to the falling spectrum and in double cascade length due to the very small mean reconstructed double cascade length for true single cascades. The total likelihood for the three topologies single cascade, track, and double cascade is $\mathcal{L} = \mathcal{L}^{\text{SC}} \mathcal{L}^{\text{T}} \mathcal{L}^{\text{DC}}$, where $\mathcal{L}^{\text{SC,T,DC}}$ is the HESE binned-likelihood over the single cascade, track and double cascade events. While in [3], the total likelihood is maximized assuming an equipartition of the flavors, here we fit the three flavors’ individual contributions f_α to the overall astrophysical normalization. f_α is the fraction of ν_α observed on Earth, with the constraint $f_e + f_\mu + f_\tau = 1$. The result of the flavor composition measurement and the sensitivity are shown in Figure 2, the best-fit point is a composition of $\nu_e : \nu_\mu : \nu_\tau = 0.29 : 0.50 : 0.21$. Our fit flavor composition is consistent with previously published results by IceCube [11, 12, 4], as well as with the expectation of $\sim 1 : 1 : 1$ flavor composition on Earth coming from a pion decay production mechanism. It is also consistent with a zero astrophysical ν_τ component. The double cascade events’ observables are shown in Table 4. “Big Bird” (Event#1) has a large positive energy asymmetry, very close to the cut value of 0.3 where the background contribution from single cascades is significant. Figure 3 shows the energy asymmetry for the best fit spectrum. It can be seen that “Big Bird” is in a background dominated region while in the case of “Double Double” (Event#2), the observables are in a signal-dominated region. To firmly conclude how compatible each of the double cascades is with a background hypothesis, i.e. with not being due to a ν_τ -CC interaction, an a posteriori analysis was performed.

Table 3: Observables of the two Double Cascades

	Event#1 (Big Bird)	Event#2 (Double Double)
Energy of 1st cascade	1.2 PeV	9 TeV
Energy of 2nd cascade	0.6 PeV	80 TeV
Energy Asymmetry	0.29	-0.80
Length	16 m	17 m

Table 4: Observables of the two Double Cascades. “Big Bird” has a large positive energy asymmetry, very close to the cut at 0.3.**Figure 3:** Distribution of the reconstructed energy asymmetry in the double cascade sample split by true topology for the best-fit astrophysical and atmospheric spectra based on simulation. The two double cascades (“Double Double” at -0.80, “Big Bird” at 0.29) are shown.

3. Detailed study of candidate events

The MC statistics available at the time of the unblinding were absolutely insufficient to incorporate the full information from the energy asymmetry observable. Further, as can be seen from Fig. 1, “Big Bird” lies in a region with very limited MC statistics, and “Double Double” lies in a region with both signal and background according to the MC templates. In order to incorporate the energy asymmetry observable, and to properly evaluate the signal and background contributions for the tau candidate events observed, a dedicated resimulation of each of the events was performed in a restricted parameter space, shown in Table 5. The mapping between true and reconstructed quantities is not straightforward. The interaction vertex in the resimulation was restricted to a cylinder with radius 50 m and height 50 m, the direction of the incoming neutrino spans 70° in zenith and 220° in azimuth, centered on the reconstructed interaction vertex and direction of the events. For the zenith and azimuth angles, the resolution depends on the event topology. While the azimuth was chosen to cover a wide range to account for possible contributions from azimuthal regions affected by the ice anisotropy and due to the limited azimuthal resolution for single cascades, the

Variable	“Big Bird”	“Double Double”
Primary Energy	>65 TeV	>1.5 PeV
Visible Energy	60 - 300 TeV	1 - 3 PeV
Vertex, $r - r_{\text{evt}}$		50 m
Vertex, $z - z_{\text{evt}}$		25 m
Azimuth $\phi - \phi_{\text{evt}}$		$\sim 110^\circ$
Zenith $\theta - \theta_{\text{evt}}$		$\sim 35^\circ$

Table 5: Restricted parameter space for resimulated events. The upper value of the primary energy depends on the interaction type, reflecting the spread of visible energy losses typical for that interaction. $r - r_{\text{evt}}$ is the two-dimensional distance in the x, y -plane. For “Big Bird”-like events, the zenith and azimuth regions were further restricted for ν_μ -CC interactions to maximize the number of events with reconstructed values close to the observed ones, as the direction resolution is very good for tracks.

zenith region was restricted slightly more as the zenith resolution is better. With an angular resolution of $\sim 10^\circ$ for cascades, this seemed sufficient. However, the simulation did not include very down-going events, which could contribute up to $\sim 3\%$ to the signal and $\sim 6\%$ to the single cascade background classified as double cascades. In case of the primary energy, the mapping depends on the neutrino spectrum and the interaction type, and is only well-correlated to the reconstructed deposited energy for ν_e -CC interactions, as only in this case the neutrino deposits its entire energy in the form of visible energy in the detector. All other interactions have some non-visible energy losses – final state neutrinos, intrinsically darker hadronic cascades, muons leaving the detector – such that it is not a priori known what primary energy range will significantly contribute to the region around the reconstructed values. The primary neutrino energy was restricted such as to cover all energies that can contribute to the observed reconstructed energies in reconstructed space, which had to be determined by trial and error. True quantities for the energy asymmetry and double cascade length are only defined for ν_τ -CC interactions. Those variables were left unconstrained during resimulation.

3.1 The expected fraction of ν_τ -CC events

The posterior probability for each event to have originated from a ν_τ CC interaction is:

$$P(\nu_\tau | \vec{\eta}_{\text{evt}}) = \frac{P_{\nu_\tau}(\vec{\eta}_{\text{evt}}) N_{\nu_\tau}^{\text{DC}}}{P_{\nu_\tau}(\vec{\eta}_{\text{evt}}) N_{\nu_\tau}^{\text{DC}} + P_{\nu_\tau^c}(\vec{\eta}_{\text{evt}}) N_{\nu_\tau^c}^{\text{DC}}}, \quad (3.1)$$

where N_{ν_τ} and $N_{\nu_\tau^c}$ are the total expected number of events in the double cascade bin stemming from ν_τ -CC and non- ν_τ -CC interactions and taken from the total simulation. P_{ν_τ} and $P_{\nu_\tau^c}$ are the PDFs for the ν_τ -CC and non- ν_τ -CC components in the parameter space vector of each event, $\vec{\eta}_{\text{evt}}$. We can approximate this expression as the ratio of the differential rates at $\vec{\eta}_{\text{evt}}$,

$$P(\nu_\tau | \vec{\eta}_{\text{evt}}) \approx \frac{dN_{\nu_\tau}^{\text{DC}}/d\vec{\eta}_{\text{evt}}}{dN_{\nu_\tau}^{\text{DC}}/d\vec{\eta}_{\text{evt}} + dN_{\nu_\tau^c}^{\text{DC}}/d\vec{\eta}_{\text{evt}}} \equiv \tau. \quad (3.2)$$

In the last step we have defined the tauness τ , the fraction of events close to the reconstructed observables $\vec{\eta}_{\text{evt}}$ of the data events, which are expected to be of the ν_τ -CC interaction type. Note that

the tauness is always evaluated under certain assumptions for the physics parameters $\vec{\theta}$, e.g. for $\vec{\theta} = \hat{\theta}$ maximizing $\mathcal{L} = \mathcal{L}^C \mathcal{L}^T \mathcal{L}^{\text{DC}}$. We adopt this assumption in the following, making the prior probability for a ν_τ interaction to be the ν_τ content in the double cascade topology subsample according to the best-fit with a 1:1:1 composition [3]. Substituting $dN_{\nu_\tau}^{\text{DC}}/d\vec{\eta}_{\text{evt}} = N_{\nu_\tau} P_{\nu_\tau}(\vec{\eta}_{\text{evt}})$, where $N_{\nu_\tau} P_{\nu_\tau}(\vec{\eta}_{\text{evt}})$ is the differential expected number of ν_τ -CC events at the point $\vec{\eta}_{\text{evt}}$, and similarly for non- ν_τ -CC interactions (ν_τ), we can write

$$\tau = \frac{N_{\nu_\tau} P_{\nu_\tau}(\vec{\eta}_{\text{evt}})}{N_{\nu_\tau} P_{\nu_\tau}(\vec{\eta}_{\text{evt}}) + N_{\nu_\tau} P_{\nu_\tau}(\vec{\eta}_{\text{evt}})}. \quad (3.3)$$

The differential expected number of events at the point $\vec{\eta}_{\text{evt}}$, $N_{\nu_\tau} P_{\nu_\tau}(\vec{\eta}_{\text{evt}})$ and $N_{\nu_\tau} P_{\nu_\tau}(\vec{\eta}_{\text{evt}})$ can be computed from the resimulation sets using a multidimensional KDE with a gaussian kernel and the Rodeo algorithm [13]. The dimensionality necessarily needs to be extended to more than the 2 dimensions used in the original double cascade classification before, as the resimulation was carried out in the restricted parameter space. The additional dimensions are 3 dimensions for the vertex x, y, z and 2 dimensions for the direction θ, ϕ . With the total deposited energy we have 6 resulting dimensions. We further define a region of interest in the parameters not restricted during resimulation, double cascade length and energy asymmetry.

3.2 Results from the a-posteriori evaluation

The dedicated resimulation for the a posteriori evaluation yields a total of $\sim 13 \cdot 10^6$ ‘‘Double-Double’’-like events passing the HESE selection and $\sim 1 \cdot 10^6$ ‘‘Big-Bird’’-like events passing the HESE selection. We compute the tauness for each of the events to stem from a ν_τ -CC interaction for the best-fit spectrum and 1 : 1 : 1 flavor composition as given in [3]. The tauness for ‘‘Big Bird’’ is $\tau_{\text{best fit}}^{\text{BB}} \approx 75\%$, the tauness for ‘‘Double Double’’ is $\tau_{\text{best fit}}^{\text{DD}} \gtrsim 97\%$. That means, that we expect ~ 75 ($\gtrsim 97$) out of 100 observed events with Big Bird (Double Double) observables to be stemming from ν_τ -CC interactions. For Double Double, the statistics of the generated MC are not sufficient to evaluate the tauness to a higher precision. Note that the uncertainties in the fit spectrum and flavor composition have not yet been taken into account. For both events, the major background component misclassified as double cascades are true single cascades, predominantly from ν_e interactions. In the evaluated region of interest, the single cascade background is strongly suppressed, but still dominant for ‘‘Big Bird’’ for which single cascades are making up $\gtrsim 4/5$ of the background. Note also, that while for ‘‘Big Bird’’ the single cascade hypothesis fit yields a satisfying result, ‘‘Double Double’’ cannot be described well with a single cascade hypothesis fit.

4. Summary and Discussion

The latest update to the HESE sample and analyses with 7.5 years of livetime includes a dedicated algorithm to determine the topology of the observed events and can distinguish between single cascades, double cascades, and tracks. Double cascades are created by ν_τ -CC interactions, but due to the fast decay of the produced tau lepton, they are very difficult to identify. We have for the first time classified two of the IceCube data events as double cascades, thus finding promising candidates for astrophysical tau neutrinos. An a-posteriori analysis was performed to better assess the probability of each of the candidates to be stemming from a ν_τ -CC interaction, based on newly

produced MC events covering a similar parameter space as the observed data. For the single power-law spectrum fit to the HESE sample, we find that “Double Double” has a tauness of $\gtrsim 0.97$ and “Big Bird” has a tauness of ≈ 0.75 . This means that in a very long-running experiment we would expect at most one event for every 30 events observed with reconstructed values similar to “Double Double” to be due to a non- ν_τ -CC interaction, while a background scenario would account for one out of every four events seen with reconstructed values similar to “Big Bird”. Even with the resimulated MC events, the MC statistics are still the limiting factor for Double Double. In the meantime, two complementary tau searches were performed. Looking for two resolved pulses on single sensors, they use different information than the double cascade method discussed here. Both analyses [14, 15] also find “Double Double”.

References

- [1] **IceCube** Collaboration, M. G. Aartsen et al., *JINST* **12** (2017) P03012.
- [2] **IceCube** Collaboration, M. G. Aartsen et al., *Science* **342** (2013).
- [3] **IceCube** Collaboration, *PoS (ICRC2019) 1004* (these proceedings).
- [4] **IceCube** Collaboration, *PoS (ICRC2017) 974* (2018).
- [5] **IceCube** Collaboration, J. Stachurska, <https://doi.org/10.5281/zenodo.1301122>, (2018).
- [6] **IceCube** Collaboration, M. G. Aartsen et al., *JINST* **9** (2014) P03009.
- [7] **AMANDA** Collaboration, J. Ahrens et al., *Nucl. Instrum. Meth.* **A524** (2004) 169–194.
- [8] Patrick Hallen, *On the Measurement of High-Energy Tau Neutrinos with IceCube*, Master’s thesis, RWTH Aachen, 2013.
- [9] S. S. Wilks, *Annals Math. Statist.* **9** (1938) 60–62.
- [10] C. A. Argüelles, A. Schneider and T. Yuan, *Journal of High Energy Physics* **2019** (2019) 30.
- [11] **IceCube** Collaboration, M. G. Aartsen et al., *Astrophys. J.* **809** (2015) 98.
- [12] **IceCube** Collaboration, M. G. Aartsen et al., *Phys. Rev.* **D99** (2019) 032004.
- [13] H. Liu, J. Lafferty, and L. Wasserman in *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics*, vol. 2, pp. 283–290, (2007).
- [14] **IceCube** Collaboration, *PoS (ICRC2019) 960* (these proceedings).
- [15] **IceCube** Collaboration, *PoS (ICRC2019) 1036* (these proceedings).