Development of an IceCube Realtime Alert using Multiplet Signal for Optical Follow-Up

The IceCube Collaboration
(a complete list of authors can be found at the end of the proceedings)

E-mail: shimizu@hepburn.s.chiba-u.ac.jp

Multi-messenger observation of neutrino sources is a key for identifying the origin of astrophysical neutrinos, and it led to the identification of the blazar TXS 0506+056 as the first candidate in 2017. When the IceCube observatory detects neutrino signals, alerts are sent to the other telescopes to trigger follow-up observations. The newly proposed algorithm tries to find long time scale doublets and triplets, where two or three astrophysical neutrino event candidates are observed from the same direction within 30 days. This signature selects neutrino sources close to our galaxy, and makes it easier to do the follow-up observation by optical telescopes. Utilizing the 11 years archival data, the signal extraction efficiencies of time and spatial clustering algorithms were evaluated, and found that the new method showed significantly better performance for small number of detected events. Finally, the sensitivity of the new alert channel is presented in the phase space of neutrino source model parameters.

Corresponding authors: Nobuhiro Shimizu

1 Chiba University

* Presenter
1. Introduction

Since the first discovery of the high energy astrophysical neutrinos by the IceCube observatory in 2013 [1, 2], more detailed properties of the astrophysical neutrino fluxes have been clarified such as spectral index [3] and flavor ratio [4]. In 2017, an association of the realtime neutrino alert (IceCube-170922A), with the high-energy gamma-ray flare by Fermi-LAT was reported [5], and suggested the blazar TXS 0506+056 was likely an source of high energy neutrinos. However, it is known that the blazars cannot explain all the diffuse fluxes observed by the IceCube [6]. It is still unclear which astrophysical objects fill our universe with high energy neutrinos. Month-order transients, such as TDEs [7] are of particular interest as a possible association between an IceCube alert and TDE flare has been identified [8].

To identify the sources of high energy neutrinos, multi-messenger observation is a powerful strategy. In particular, a follow-up observation by optical telescopes can be a smoking gun to get a hint of the origin of neutrinos. The absence of the electromagnetic interaction of neutrinos allows them to propagate from the distant universe, but this feature makes the optical follow-up difficult. The rate density of supernovae is approximately $10^{-4} \text{ Mpc}^{-3} \cdot \text{yr}^{-1}$, and the number of the supernovae within redshift $z < 1$ results in $\sim 1 \text{ day}^{-1} \text{deg}^{-2}$. Since the angular resolution of neutrino events is $\sim 1^\circ$, a follow-up identification of optical counterpart using a long time window of more than a day becomes challenging due to contamination of irrelevant transients.

To successfully identify the neutrino source by multi-messenger observation, a multiplet signal is searched, where two or more neutrino events are observed in the same direction within a limited time window. By imposing the detection of multiple neutrinos, the distance to the neutrino source is biased to typically less than $z < 0.1$ [9]. The presented study aims to develop a new neutrino alert channel utilizing a month-time scale multiplet signal.

Figure 1 shows the expected number of multiplets in a year as a function of neutrino source model parameters when the timing window is 30 days [9]. Left is in neutrino source energy, $E_\nu$, and rate density, $\rho_\nu$ plane, while the vertical axis of the right panel is given as luminosity density, $E_\nu \rho_\nu$. The number already includes the efficiency of a neutrino event selection described in section 2.2. The colored bands correspond to the energy fluence of the diffuse neutrino flux of $E^2 \phi_{\nu_e} \in (10^{-9}, 10^{-7}) \text{ GeV} \cdot \text{s}^{-1} \cdot \text{cm}^{-2} \cdot \text{sr}^{-1}$. Given the measured value of $E^2 \phi_{\nu_e} = 1.44 \times 10^{-8}$ at 100 TeV [3], the bands are $\pm 1$ decade from the best fit. To calculate the number of multiplets, a neutrino flux from a given $E_\nu$ is bolometrically defined assuming isotopic luminosity distribution. The flavor ratio at the earth is set to be $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$. For every source, a probability to be observed as a multiplet is given by $p = 1 - e^{-\mu} - \mu e^{-\mu}$, where $\mu$ is an expected number of neutrino events in 30 days. This $p$ is convoluted when we integrate sources over distance. Here, the neutrino source evolution is assumed to be a star formation rate: $\propto (1 + z)^{3.4}$ for $z < 1$ and a constant for $z \geq 1$. The range of the integration over distance is from $(4\pi/3 \cdot \rho_\nu T_{30 \text{ day}})^{-1/3}$ (the distance that a single source is expected on average within this range) to a distance corresponding to $z = 4$. Further details are described in Ref. [9]. As can be seen in the figures, the expected number of the multiplet signal is small $< 1$ in most of the phase space, thus the algorithm to send a multiplet alert should be sensitive to such small number of detected events.
IceCube realtime neutrino alert using multiplet signal

2. Multiplet Selection in the IceCube

2.1 IceCube observatory

The IceCube observatory at the South Pole is a cubic-kilometer detector for astrophysical neutrinos. Cherenkov emissions from the charged particles generated by the interaction of high energy neutrinos are observed by 5160 digital optical modules (DOMs). The DOMs include 10" photomultiplier tubes in a high pressure-resistant vessel, and were deployed between 1450 m to 2450 m below the surface of ice by 86 strings. Detected photon signals are sent to a surface IceCube laboratory, and recorded when the trigger condition is satisfied. The patterns of detected photons by neutrino events are classified as either tracks or cascades. A track event is initiated by a charged current interaction of $\nu_\mu$, and the photon distributes along the line of the outgoing muon showing a good angular resolution of $\sim 1^\circ$ deg. A cascade event shows more isotropic hit distribution, and the angular resolution is $\sim 10^\circ$ deg.

2.2 Neutrino Event Selection

Major backgrounds to the astrophysical neutrinos are from atmospheric neutrinos for declination angle $\delta > -5^\circ$ while the other is from atmospheric muons for $\delta < -5^\circ$. Detected signals are first cleaned by rejecting isolated hits both in space and time, and then a track reconstruction is performed. Starting from a simplest geometrical line fit as a seed, several iterative fits are performed, and finally a sophisticated method SplineMPE fit determines the most probable direction of the neutrino event [10, 11]. Kinetic variables such as a geometrical center of gravity, total detected number of photons, likelihood of the reconstruction, etc, are used as inputs of the boosted decision tree to purify the astrophysical neutrino events (the event set is called GFU sample because of original purpose of a gamma-ray follow-up) [12]. This study uses only track type event from the northern direction $\delta > -5^\circ$ because the effective area to the neutrino signal is higher than that of southern sky. The total event rate of this selection is 3.3 mHz. The used dataset was collected since
the completion of the IceCube in 2011 until the end of 2022, corresponding 11.4 years of total live
time after excluding period of data acquisition failures.

2.3 Clustering of Multiplets

The conventional time-dependent all-sky realtime alert (hereafter referred to as all-sky) uses
the nested likelihood [13] to search for a population of neutrino events in time and position:

$$L_N(n_s, \gamma) = \prod_{i=1}^{N} \left[ \frac{n_s}{N} L_{\text{sig}}(\gamma) + \frac{N-n_s}{N} L_{\text{bg}} \right],$$

where $L_{\text{sig}}$ and $L_{\text{bg}}$ are signal and background likelihoods, $N$ is a total number of neutrinos used to
calculate the likelihood, $n_s$ is a number of signal source, and $\gamma$ is a spectral index of the energy of
neutrino signals ($dN/dE \propto E^{-\gamma}$). The signal and background likelihoods are defined as a product
of a spatial and an energy probability density functions. For a given maximum timing window $T_{\text{max}}$, a test statistics (TS) is constructed as a likelihood ratio against the null hypothesis:

$$\Lambda_{\text{all-sky}} = 2 \log \left( \frac{\max_{n_s, \gamma} L_N(n_s, \gamma)}{L_N(n_s = 0)} \cdot \frac{\Delta T}{T_{\text{max}}} \right),$$

where “max” means the nested likelihood is optimized with respect to $n_s$ and $\gamma$. Moreover, within $T_{\text{max}}$, a variable timing window $\Delta T$ and a direction $\vec{n}_\nu$ are fit in time and position space, and the combination having the largest $\Lambda_{\text{all-sky}}$ is selected. The factor $\Delta T/T_{\text{max}}$ compensates for the increase in background for larger windows by penalizing a short $\Delta T$.

Rather than using all the events in a variable timing window $\Delta T$, the newly proposed algorithm calculates a TS for every combination of two or three events. First, an incoming event is used as a seed, and multiplet candidate events are selected within $3^\circ$ of its vicinity in the record back to 30
days ($T_{\text{max}} = 30$ days). Then if the number of candidates satisfies $N \geq 2$ and $N \geq 3$, we calculate
doublet and triplet TSs, respectively, as:

$$\Lambda_{\text{doublet}} = 2 \log \left( \frac{L_{\text{sig}}^{(1)}(\gamma = 2.3)}{L_{\text{bg}}^{(1)}} \cdot \frac{L_{\text{sig}}^{(2)}(\gamma = 2.3)}{L_{\text{bg}}^{(2)}} \right),$$

$$\Lambda_{\text{triplet}} = 2 \log \left( \frac{L_{\text{sig}}^{(1)}(\gamma = 2.3)}{L_{\text{bg}}^{(1)}} \cdot \frac{L_{\text{sig}}^{(2)}(\gamma = 2.3)}{L_{\text{bg}}^{(2)}} \cdot \frac{L_{\text{sig}}^{(3)}(\gamma = 2.3)}{L_{\text{bg}}^{(3)}} \right),$$

where the upper numbers of the likelihoods represent indices of events. Similarly to the all-sky,
the most probable direction of neutrino source is determined by maximizing each TS whereas the
spectral index is not floated because $N = 2, 3$ is too small to determine $\gamma$. This $\gamma = 2.3$ was chosen
because of the best fit of the measured spectral index of the diffuse flux [3]. Finally, the combination
which has the largest TS is chosen.

3. Performance of the Multiplet Selection

To evaluate the distribution of the background of the TS, we generated a pseudo dataset by
scrambling the time of each neutrino event so that the statistics of the dataset were increased to
IceCube realtime neutrino alert using multiplet signal

Figure 2: Distribution of test statistics: left $\Lambda_{\text{doublet}}$ and right $\Lambda_{\text{triplet}}$. The black and colored histograms are distributions of background and signal, respectively. The number of detected signals are exactly two and three for doublet and triplet cases, respectively. The colors represent different spectral indices: red 2.0, blue 2.3, and green 2.7, respectively. The background histogram is scaled to be equivalent to one year while the scales of signal histograms are arbitrary. The smallest bin includes underflow. The vertical dashed lines represent the threshold of TS at FAR = 1 yr⁻¹.

more than an 8000-year equivalent value. The distribution of the astrophysical neutrino signal was evaluated by mixing simulated signal events on top of the pseudo dataset. Assuming the signal timing distribution is uniform, we mixed the signal events uniformly in $T_{\max}$ (= 30 days), and selected the maximum TS in $T_{\max}$. The direction of the signal source is uniformly distributed.

Figure 2 shows the distribution of $\Lambda_{\text{doublet}}$ and $\Lambda_{\text{triplet}}$ for both background and signal events.

A false alarm rate (FAR) at a given threshold of a TS is evaluated as an upper cumulative sum of the background histogram. If $N \geq 3$, we in general have both $\Lambda_{\text{doublet}}$ and $\Lambda_{\text{triplet}}$, so we unite them as TS-score by selecting more significant pair as $\text{TS-score} = \max \left\{ -\text{FAR}(\Lambda_{\text{doublet}}), -\text{FAR}(\Lambda_{\text{triplet}}) \right\}$, where the negative sign converts the TS-score such that larger value shows higher significance.

The performances of the algorithms are compared based on the efficiency of signals under a given value of the FAR. Here, we define one trial as an operation to add simulated $N_{\text{det}}$ signal events on the pseudo experimental data within $T_{\max}$. Here, the inserted events already pass the described event selection in section 2.2, hence $N_{\text{det}}$ is a number of detected signal events. After calculating the TS for all the events in $T_{\max}$, if the maximum TS exceeds the threshold of the TS, the signal is to be successfully detected. The signal efficiency is defined as a fraction of such successful trials out of the total number of trials simulated. As a benchmark, we set a threshold at FAR = 1 yr⁻¹.

Figure 3 shows the signal efficiency as a function of $N_{\text{det}}$ when the spectral index is $\gamma = 2.3$. The signal source direction is uniformly sampled. The red and blue curves are signal efficiencies when we use only $\Lambda_{\text{doublet}}$ and $\Lambda_{\text{triplet}}$, while the green is the case when we use the TS score. In spite of the multiplet selection, a nonzero efficiency at $N_{\text{det}} = 1$ is observed, but this occurs due to a combination of a background and a signal events. If we compare the performance to the all-sky, the new proposed method shows significantly higher efficiency when $N_{\text{det}}$ is small whereas it is smaller for large $N_{\text{det}}$. This trend is natural because the all-sky TS uses all events inside $\Delta T$ while the newly


Figure 3: Efficiency to detect multiplets as a function of the detected number for the all sky (black), doublet (red), triplet (blue), and TS-score (green) when the spectral index is -2.3. The threshold of the false alarm rate is 1.0 per year in each algorithm.

Figure 4: Angular resolution of the multiplets which have TS values higher than the threshold at $\text{FAR} = 1 \text{ yr}^{-1}$ as a function of the detected number of signal event for the all-sky (black) and this algorithm (red). The solid and dashed lines represent 50% and 90% containment radius, respectively.

proposed TSs selectively cherry-pick significant combinations out of many events.

When operated as an alert, a pointing accuracy of the neutrino source direction is important for follow-up observation. Figure 4 shows the angular resolution as a function of $N_{\text{det}}$ for multiplets which pass the threshold of TS at $\text{FAR} = 1 \text{ yr}^{-1}$. The solid and dashed lines represent 50% and 90% containment regions, respectively. Thanks to multiple events, the angular resolution of multiplets are much better than typical uncertainty of singlet events ($\sim 1^\circ$). Moreover, compared to the all-sky method, the proposed algorithm shows a stable angular resolution for small $N_{\text{det}}$. For $N_{\text{det}} \geq 2$, our method exhibits a 90% containment angular resolution of $0.3^\circ$. Even if we relax the threshold of the TS up to $\text{FAR} = 10 \text{ yr}^{-1}$ (the energies of passing events become smaller and the angular resolution becomes poorer), the degradation is only 20%.

4. Sensitivity

The new algorithm will be used not only to issue an alert, but also to search for multiplets in 11 years archival data of IceCube. A commonly used quantity to demonstrate the sensitivity is “$n$-$\sigma$ discovery potential”, which is defined as a signal strength such that with 50% of chance,
a signal trial exceeds $n$-$\sigma$ of a background TS distribution. The left of Fig. 5 shows $3\sigma$ and $5\sigma$ discovery potentials of a time integrated energy fluence for the all-sky and our proposed algorithms as a function of declination angle for $\gamma=2.0$. Our algorithm shows higher sensitivity particularly for looser condition of $3\sigma$. This is even better when we relax the requirement on the signal efficiency from the conventional $50\%$ to $5\%$ as shown in right panel. Because our method is more optimized for a small number of detection, it is more sensitive to a limited subset of significant events.

To constrain the allowable neutrino source model parameters, we use the maximum TS values of doublet and triplet in 11 years $\Lambda_{\text{max}} = \max \{ \Lambda_{\text{doublet}}, \Lambda_{\text{triplet}} \}$, where the maximum is selected among 11 years dataset. The distribution of $\Lambda_{\text{max}}$ is first evaluated by the pseudo experiments, and then the p-value is defined as an upper cumulative sum of the observed $\Lambda_{\text{max}}$. Since we have not yet unblinded the real experimental data, we assume our universe is characterized by the neutrino source model parameters $E_\nu, \rho_\nu$ (and $\gamma$), and then we emulate the maximum TS with existence neutrino source $\Lambda_{\text{max}}^{w/\text{sig}}$. By averaging p-values for many $\Lambda_{\text{max}}^{w/\text{sig}}$, we get an averaged confidence interval of neutrino source model parameters consistent with background-only hypothesis. Figure 6 shows contours of the neutrino source model parameters at $90\%$ confidence level when the spectral indices are 1.5 (red), 2.0 (orange), 2.3 (green), 2.5 (blue), and 2.7 (cyan): left in $E_\nu-\rho_\nu$ plane, and right in $E_\nu-E_\nu\rho_\nu$ plane. The archival 11 years dataset will be sensitive to neutrino source model parameters above the contours.

Figure 5: Discovery potential as a function of the declination angle for the all-sky (black), and this algorithm (red) when $\gamma$ is 2.0. Solid, and dashed lines represent $5\sigma$ and $3\sigma$, discovery potential, respectively. Left figure is a conventional definition where the signal efficiency is $50\%$, while right is at $5\%$.

5. Conclusion

A multi-messenger observation by optical telescopes can be a smoking gun to identify the origin of high energy astrophysical neutrinos. However, a large uncertainty of the neutrino reconstruction degrades the significance of observation due to many irrelevant transients in distant universe. It becomes further difficult to perform a search for long time scale transients. A multiplet signal automatically gives bias on the neutrino sources $z < 0.1$ and makes easy to identify optical counterparts. Moreover, multiplet provides much better angular uncertainty than singlets. The expected number of multiplet signals are small, $< 1$ yr$^{-1}$, so the algorithm to issue an alert needs to be sensitive to such small number of detected events. In comparison with the conventional all-sky clustering method, the newly proposed algorithm calculates doublet and triplet TSs for every
PoS(ICRC2023)1467

IceCube realtime neutrino alert using multiplet signal

Figure 6: Contours of excluded region of neutrino source model parameters at 90% confidence level when the spectral indices are 1.5 (red), 2.0 (orange), 2.3 (green), 2.5 (blue), and 2.7 (cyan): left in $E_\nu$-$\rho_\nu$ plane, right in $E_\nu$-$E_\nu \rho_\nu$ plane. The center bands correspond to the diffuse neutrino flux of $E^2\phi_\nu \in (10^{-9}, 10^{-7})$ GeV $\cdot$ s$^{-1}$ $\cdot$ cm$^{-2}$ $\cdot$ sr$^{-1}$.

combination of two and three events, and provides much higher efficiency for signal events under a given threshold of the TS. This algorithm will be used to search for multiplets in the 11 years archival data as well.

References

IceCube realtime neutrino alert using multiplet signal
Acknowledgements

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, Wisconsin Alumni Research Foundation, Center for High Throughput Computing (CHTC) at the University of Wisconsin–Madison, Open Science

...
IceCube realtime neutrino alert using multiplet signal

Grid (OSG), Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS), Frontera 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, and Astroparticle physics computational facility at Marquette University; Belgium – Funds for Scientific 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 Compute 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); Switzerland – Swiss National Science Foundation (SNSF); United Kingdom – Department of Physics, University of Oxford.