THE LIMIT OF TAU NEUTRINO DETECTION

GEANT4 SIMULATION

Introduction

The aim of this project is to simulate the interaction of high energy neutrinos in water (seawater) and evaluate the performance of a discrete detector in tau neutrino identification. High energy tau neutrino will produce a double-band signature in a water- or ice-based Cerenkov detector like IceCube. For tau neutrinos with lower energy, the two cascades will be very close and they cannot be separated spatially. However, there might be some PMTs that receive enough photons to see two response peaks, one from the hadronic shower produced by the charged current interaction and one from the tau decay. This study evaluated the performance of the detector for detecting tau neutrino (10 TeV to 60 TeV) based on event identification rate.

Simulation Set-up

GEANT4 is a toolkit for both full and fast Monte Carlo simulation of detectors in High Energy Physics. The documentation of GEANT4 can be found at:

http://geant4-userdoc.web.cern.ch/geant4-userdoc/UsersGuides/ForApplicationDeveloper/html/

Physics

The physics list is FTFP_BERT (valid within the energy range of 0 to 100 TeV) with G4EmStandard_option4 electromagnetic physics and G4OpticalPhysics. All physics processes are listed below. Detailed description of each physics process can be found in the Physics Reference manual.

GEANT4 Physics Reference Manual:

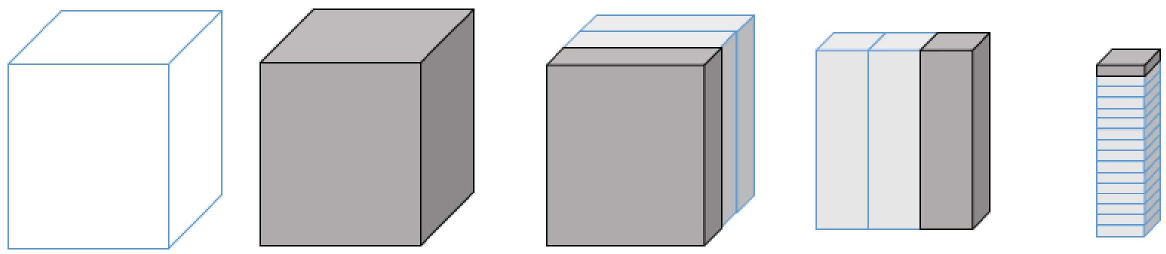
http://geant4-userdoc.web.cern.ch/geant4-userdoc/UsersGuides/PhysicsReferenceManual/html/

```
/process/list
     Transportation,              phot,             compt,              conv
               Rayl,               msc,             eIoni,             eBrem
          ePairProd,       CoulombScat,           annihil,            muIoni
            muBrems,        muPairProd,             hIoni,            hBrems
          hPairProd,           ionIoni,   nuclearStopping,     photonNuclear
    electronNuclear,    positronNuclear,      muonNuclear,             Decay
         hadElastic, neutronInelastic,          nCapture,  protonInelastic
        pi+Inelastic,      pi-Inelastic,    kaon+Inelastic,    kaon-Inelastic
    kaon0LInelastic,   kaon0SInelastic,  lambdaInelastic, anti-lambdaInelastic
   sigma-Inelastic, anti_sigma-Inelastic,  sigma+Inelastic, anti_sigma+Inelastic
        xi-Inelastic,    anti_xi-Inelastic,    xi0Inelastic,   anti_xi0Inelastic
   omega-Inelastic, anti_omega-Inelastic, anti_protonInelastic, anti_neutronInelastic
anti_deuteronInelastic, anti_tritonInelastic,  anti_He3Inelastic, anti_alphaInelastic
hFritiofCaptureAtRest, hBertiniCaptureAtRest, muMinusCaptureAtRest,       dInelastic
        tInelastic,       He3Inelastic,     alphaInelastic,       ionInelastic
            nKiller,       OpAbsorption,        OpRayleigh,            OpMieHG
         OpBoundary,             OpWLS,          Cerenkov,      Scintillation
```

All optical processes require the Material Properties Table to be filled will relevant data when construct the detector geometry. No default data will be provided by GEANT4. The only exception is the OpRayleigh process. Rayleigh scattering length will be automatically calculated for material named *Water* and materials with known isothermal compressibility provided that no Rayleigh scattering length is specified.

Detector construction

The detector is a cubic tank filled with water or seawater (two different sizes were used in my simulation: $300m \times 300m \times 300m$ and $150m \times 150m \times 150m$). The cubic tank is divided into several layers and each layer is divided into columns (represent strings) and then cells. There are three layers in the box, three columns in each layer, and twelve cells in each column. The box is located at the origin of the world coordinate.



World	Box	Layer	Column	Cell
(Word Volume)	G4PVPlacement	G4PVReplica	G4PVReplica	G4PVReplica

One sphere-shaped sensitive detector (G4VSensitiveDetector) of diameter 355mm is placed in each cell to simulate one PMT (or one DOM, digital optical module). G4SDParticleFilter is applied to those volumes to make the PMTs only sensitive to optical photons. All optical photons enter the sphere region will be killed (absorbed by the PMT).

The composition of the water and seawater used in the simulation are given in the table below. Water is defined as G4_WATER from the NIST data base. The optical properties of pure water and sea water will be discussed in later section.

Medium	Density (g/cm^3)	Elements	Fraction of Mass (%)
Water	1.0	O	88.81
		H	11.19
Seawater	1.04	O	86.04

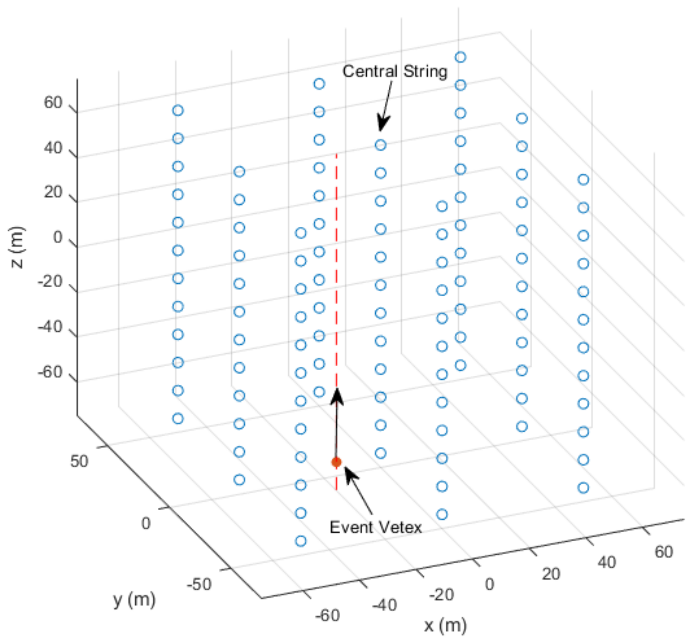
H	10.84
Cl	1.92
Na	1.07
Mg	0.13

Event Generation

GEANT4 cannot handle neutrino interaction with matters (this can be done by other tools like GENIE, NEUT, etc.). In CC interaction a charged lepton and a hadronic shower are produced at the event vertex. The charged lepton and the hadronic shower can be simulated separately in GEANT4. Since the neutrino is very high in energy, the momentum of secondary particles produced in the CC interaction is assumed to be in the same direction as the primary neutrino. The event vertex is located at (-20m, -10m, -62.5m) which is a point near the lower end of the central string (Figure).

The charged lepton will be injected into the detector at a point near the central string (Figure). 20 pions with the same energy and randomly assigned charge (-1, 0, or 1) are injected at the same location to fake the hadronic shower. The momentum direction of those pions follows exponential distribution with 1 milliradian mean.

To make each event reproduceable, the Event number will be used as the seed of the HepRandom engine. The default event number starts at zero and increases by one after each event. At the beginning of each event (before the primary event is generated), the random engine will be seed with the event number. The event number can be manually set using UI command (specified in PrimaryGeneratorMessenger.cc).



Data Taking

PMTs can be identified by their Layer ID, Column ID, and Cell ID. A Hit collection (PMTHit) will be created to temporarily store the hit information for each event. The Layer ID, Column ID, and Cell ID of the PMT, arrival time, energy, and location of the photon will be recorded once it hits a PMT. G4AnalysisManager will handle the output of the data. A Ntuple object is created at the beginning of each run. At the end of each event, data in the hit collection will be written to the Ntuple object by the analysis manager. After the run is finished, data in the Ntuple will be tabulated into a .csv file (The format of the output file can be altered by changing the included header files in Analysis.hh). The

Column	0	1	2	3	4	5	6	7	8
Data	Layer ID	Column ID	Cell ID	Time	Position x	Position y	Position z	Energy	Event ID

Monitoring the Process

StackingAction counts the number of Cerenkov photons produced in each event. The trajectory of every particle produced in the simulation will be temporarily saved in the memory and those trajectories can be visualized at the end of each run. For primary particles with high energy (>100GeV), many secondary particles are produced and the trajectory information will be very large. Since visualization is not necessary in many cases, the save trajectory option is disabled in TrackingAction.

Optical Properties of Water and Seawater

Cerenkov Effect

The Cerenkov photons are emitted when a charged particle travels faster than the local speed of light in a medium. G4Cerenkov calculates the number of photons produced each step. The refractive index (RINDEX) of the medium must be provided. For both pure water and seawater, the refractive index is set to 1.333.

Medium Boundary Effects

Assume no reflection at the boundary of the detector

Absorption

G4OpAbsorption: absorption of photons. The photon is simply killed after it travel a certain distance. The data of absorption length (ABSLENGTH) must be provided. The absorption length at different wavelength is given in the following tables.

Wavelength (nm)	Photon Energy (eV)	Absorption Length (m)
600	2.066	4.376
575	2.156	12.644
550	2.254	22.286
525	2.362	31.672
500	2.480	39.790
475	2.610	40.427
450	2.755	33.670
425	2.917	26.027
400	3.100	17.081
375	3.306	8.531
350	3.542	4.289
325	3.815	2.394
300	4.133	1.492

* “Optical Constants of Water in the 200-nm to 200-µm Wavelength Region” by George M. Hale and Marvin R. Querry

Wavelength (nm)	Photon Energy (eV)	Absorption Length (m)
465	2.666	55.0
405	3.061	35.0
365	3.396	30.5

* Mar-19 STRAW data by Matthew Man

Rayleigh Scattering

G4OpRayleigh: elastic scattering of photons by particles much smaller than the wavelength of the light. The Rayleigh scattering length is the average distance traveled by a photon before it is Rayleigh scattered in the medium. This distance is calculated automatically for the medium if the isothermal compressibility (ISOTHERMAL_COMPRESSIBILITY) and temperature of the medium are provided. Alternatively, the scattering length (RAYLEIGH) can be directly provided by the user.

$$\frac{d\sigma}{d\Omega} \sim 1 + \cos^2\theta$$

Mie Scattering

G4OpMieHG: elastic scattering of photon by particle that have a diameter similar to or larger than the wavelength of the light. The analytical express of Mie scattering is very complicated. G4OpMieHG utilize two Henyey-Greenstein functions to approximate the Mie solution.

$$\frac{d\sigma}{d\Omega} = r\frac{d\sigma}{d\Omega}\left(g_f, \theta_f\right) + (1-r)\frac{d\sigma}{d\Omega}(g_b, \theta_\mathrm{b})$$

$$\frac{d\sigma}{d\Omega} \sim \frac{1-g^2}{(1+g^2-2gcos\theta)^{3/2}}$$

The Mie scattering length (MIEHG) and three parameters g_f, g_b, and r (MIEHG_FORWARD, MIEHG_BACKWARD, and MIEHG_FORWARD_RATIO) must be provided.

To simulate the photon propagation, the distance a photon will travel before it getting absorbed is selected randomly from the probability distribution $\frac{1}{\lambda_{abs}}e^{-x/\lambda_{abs}}$. Similarly, the distance a photon traveled before scattering is selected from the probability distribution $\frac{1}{\lambda_{sct}}e^{-x/\lambda_{sct}}$. If the absorption distance is shorter, the photon will get absorbed after it traveled the selected distance. Otherwise, it gets scattered and the type of the scattering is selected based on the given fraction of molecular scattering η.

For pure water, Mie scattering is not significant; Therefore, only Rayleigh scattering is simulated. The temperature of the water is set to 10 degree Celsius, and the isothermal compressibility is set to $7.658 \times 10^{-23}\ m^3/MeV$. The generated scattering length is about 100m to 200m in the blue-violet range.

For seawater, both Mie and Rayleigh scattering must be taken into account. But in GEANT4, the fraction of molecular scattering cannot be controlled since G4OpRayleigh and G4MieHG are two isolate processes. GEANT4 cannot select the scattering type based on η. It is noticeable that for Mie scattering, G4OpMieHG uses two HG functions (called two term Henyey-Greenstein) to approximate the Mie Scattering phase function. Instead of approximate the Mie scattering phase function, one can approximate the total phase (The combination of Rayleigh and Mie scattering) function by setting the parameter differently. Only the G4MeiHG process needs to be enabled to simulate the combined scattering process. The only disadvantage for this method is that only a constant fraction of molecular scattering η can be used, since GEANT4 only accepts constant parameters for the TTHG. It is observed that the value of η does not vary a lot with the wavelength; therefore, this approximation should be accurate enough for our purpose.

$$\beta_{TTHG}(\alpha, g_1, g_2, cos\theta) = \alpha\beta_{HG}(g_1, cos\theta) + (1-\alpha)\beta_{HG}(g_2, cos\theta)$$

$$\beta_{HG}(g, cos\theta) \equiv \frac{1}{4\pi}\frac{1-g^2}{(1+g^2-2gcos\theta)^{3/2}}$$

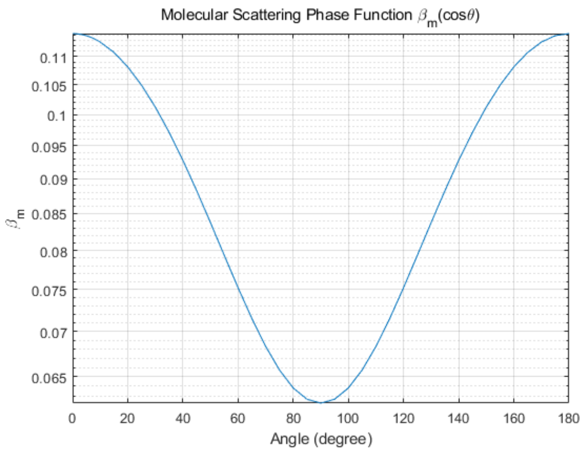
The corresponding input parameters for G4OpMieHG are $g_f = g_1$, $g_b = -g_2$, and $r = \alpha$.

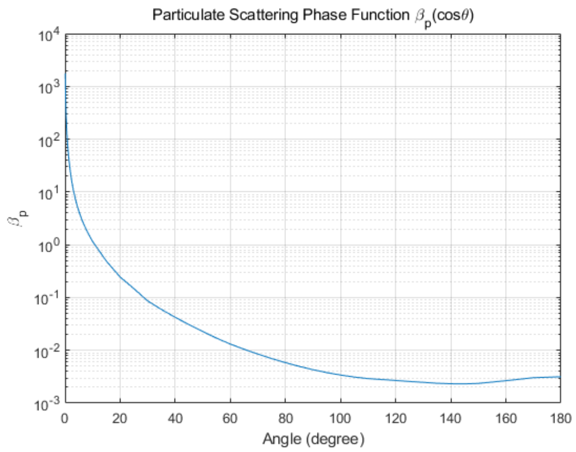
The scattering length of the seawater is found following the steps described in this paper: https://arxiv.org/pdf/astro-ph/0412126.pdf

The molecular scattering is described by the Einstein-Smoluchowski formula,

$$\beta_m(\cos\theta) = 0.06225(1 + 0.835\cos^2\theta)$$

And the particulate scattering phase function is obtained from the tabulate data by Mobley et al. The plots of the phase functions for the molecular and particulate scattering are shown below.

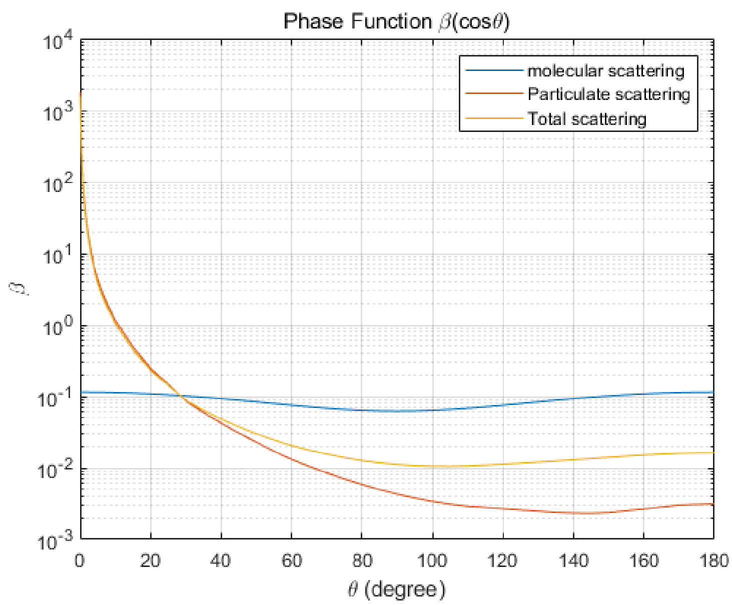




The total phase is a weighted sum of the molecular and the particulate scattering,

$$\beta(\cos\theta) = \eta\beta_m(\cos\theta) + (1-\eta)\beta_p(\cos\theta)$$

Where η is the fraction of molecular scattering. The plot of the total phase function for $\eta = 0.118\ (average)$ is shown below.



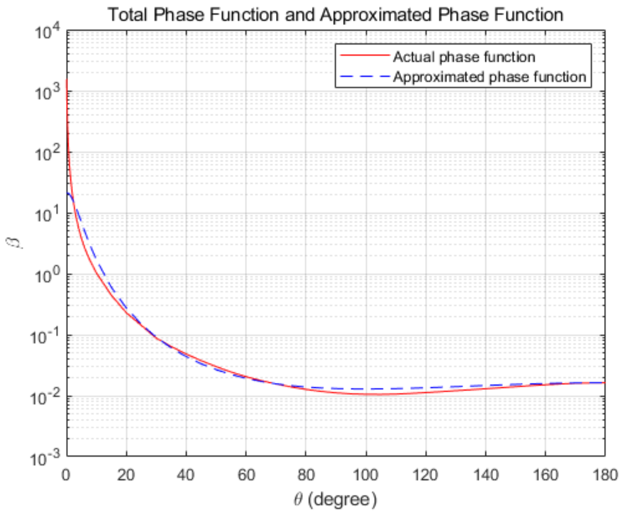
The best fit parameters are calculated using the following formula ("A Three-Parameter Analytic Phase Function for Multiple Scattering Calculations", by George W. Kattawar):

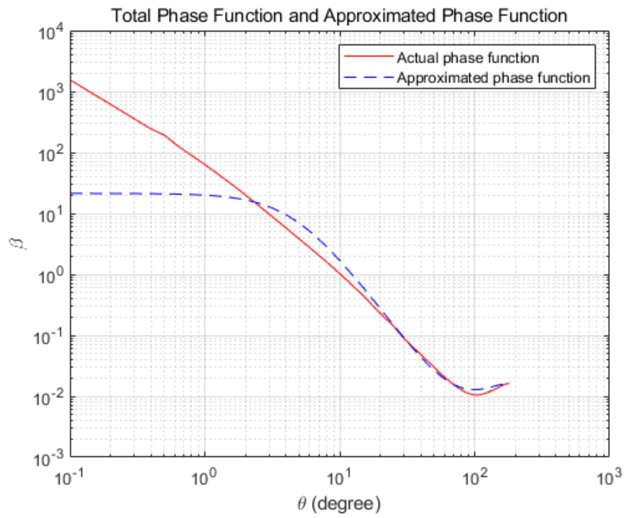
$$g_2 = \frac{t - hg - [(hg-t)^2 - 4(h-g^2)(tg-h^2)]^{1/2}}{2(h-g^2)}$$

$$g_1 = \frac{gg_2 - h}{g_2 - g}$$

$$\alpha = \frac{g - g_2}{g_1 - g_2}$$

Where g, h, and t are the projections of the phase function on a basis of Legendre polynomials (i.e. $g = \int_{-1}^{1} \beta(x) \cdot P_1(x)\, dx$, $h = \int_{-1}^{1} \beta(x) \cdot P_2(x)\, dx$, $t = \int_{-1}^{1} \beta(x) \cdot P_3(x)\, dx$). The calculated fitting parameters are: $g_2 = -0.1491$, $g_1 = 0.9204$, and $\alpha = 0.8831$ for $\eta = 0.1180$. The plots of the total phase function and the approximated phase function by TTHG are given below. The approximated phase function agrees with the actual phase function at large angle.



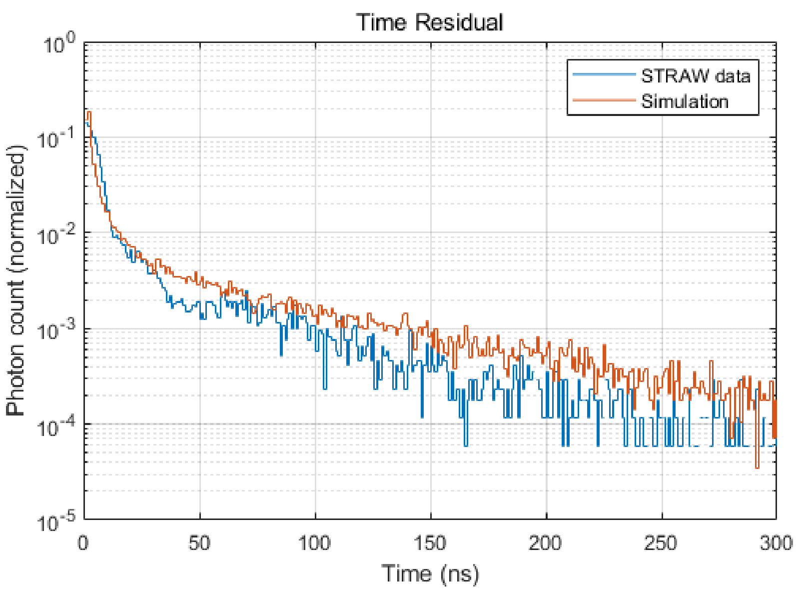


The scattering length data is given below:

Wavelength (nm)	Photon Energy (eV)	Scattering Length (m)
465	2.666	32.3
405	3.061	19.0
365	3.396	10.6

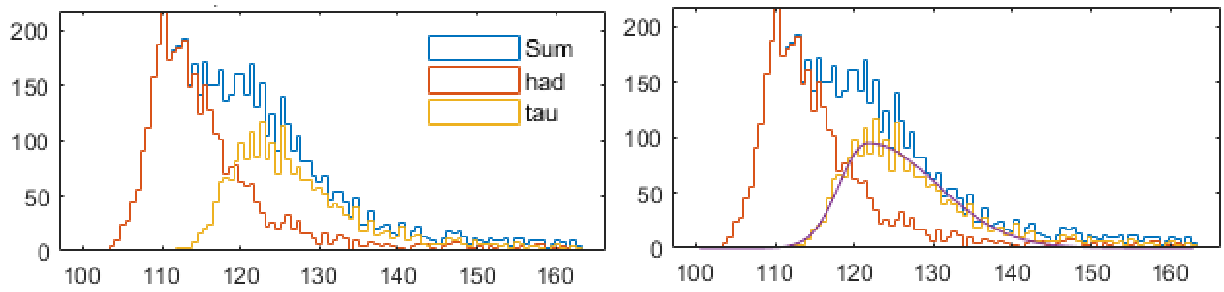
* Results from Mar-19 STRAW data by Matthew Man

The time residuals were compared with data. STRAW data used: SDOM1 _P2_violet _20V



Double-pulse Identification

The arrival time of each photon was binned and the arrival time distributions for each detector were plotted. Then an algorithm that can identify the double-peak pattern was developed. Observed that the shape of the pulse produced solely by hadronic shower or tau decay is a bifurcated Gaussian (Gaussian curve with different widths on the left and right side, parameterized by four variables: total width, amplitude, mean, width ratio between the two sides), the overall signal is fitted with the sum of two bifurcated gaussian.



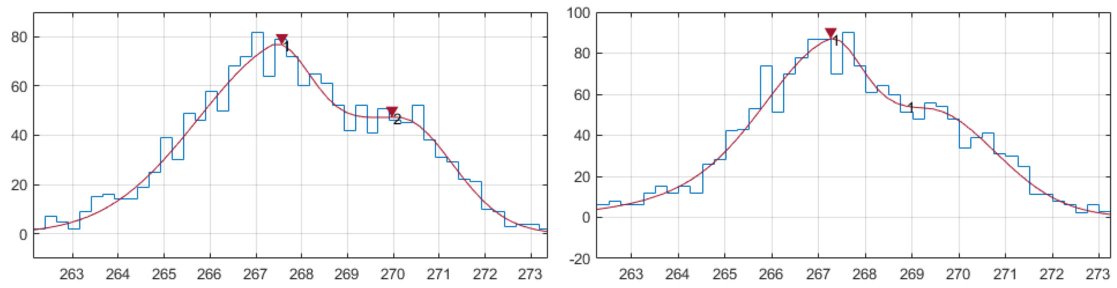
The time distributions of photons produced by the hadronic shower are shown in red. The time distributions of photons produced in the tau decay or electromagnetic cascade are shown in yellow. The combined distributions are shown in blue. (x: time in ns; y: photon count)

To select a suitable range of time, the time distribution curve is first fitted with a single bifurcated Gaussian. The location and width of the Gaussian curve is found. Then only the photon within 2.5 standard deviations of the bifurcated Gaussian is taken into account. The bin size is fixed at 0.25 ns. If the bin size is too small, 0.5 ns 0.75 ns or 1 ns bin will be used. The time distribution of those selected photons is then fitted with the sum of two bifurcated Gaussian. The fit function will find the location, width, height, and width ratio of the two bifurcated gaussian.

This curve fitting is required since it rejects the small pumps produced by scattered photons. There are several optional methods to identify the double-pulse signal. One of them is to identify the rising and falling edge of the signal. If there is a monotonic increase in a certain number of bins (during a certain time), this part can be identified as a rising edge.

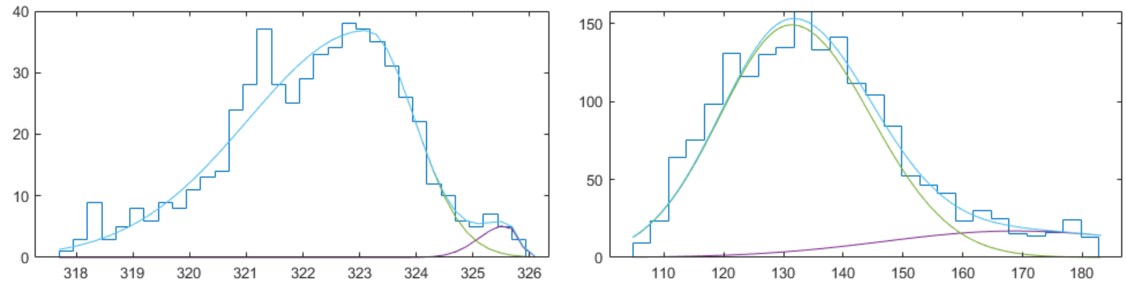
The following patterns are identified as double-peak:

1. Two explicit peaks can be found by the MATLAB peakfind function;
2. A bump can be found (by finding an inflection point where the first derivative is approximately zero).



If the fit results meet one of the conditions listed below, the pattern is identified as single-peak

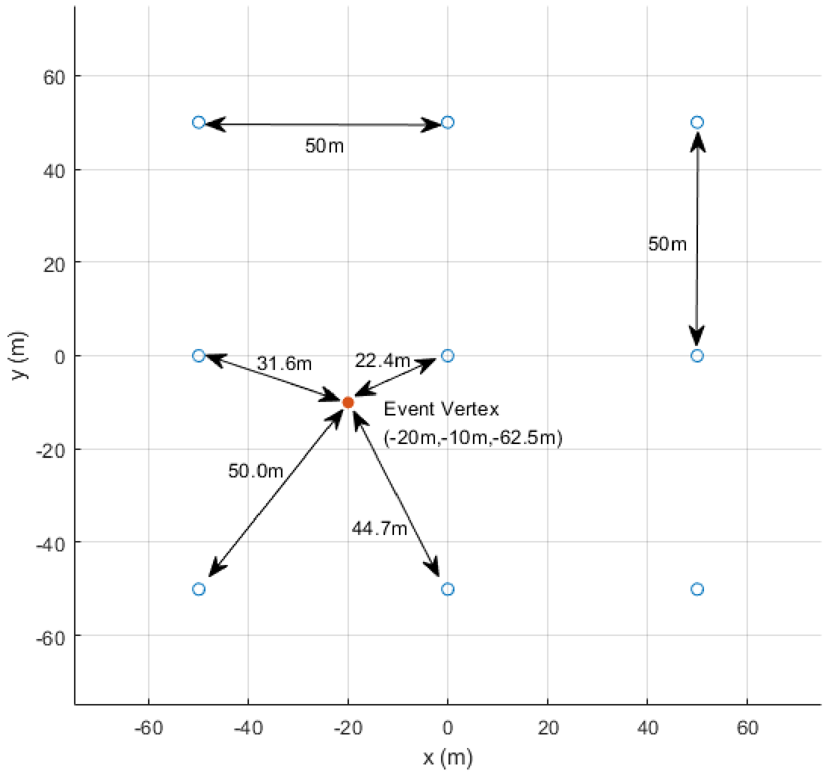
1. One of the bifurcated gaussian width < 3 or height < 5;
2. One of the bifurcated gaussian is too wide or narrow;
3. The PMT is located at 35~45 degree from momentum direction of the neutrino (to save time, since a large number of photons are received by the detectors at the Cherenkov angle, usually it is impossible to identify the double-peak pattern).



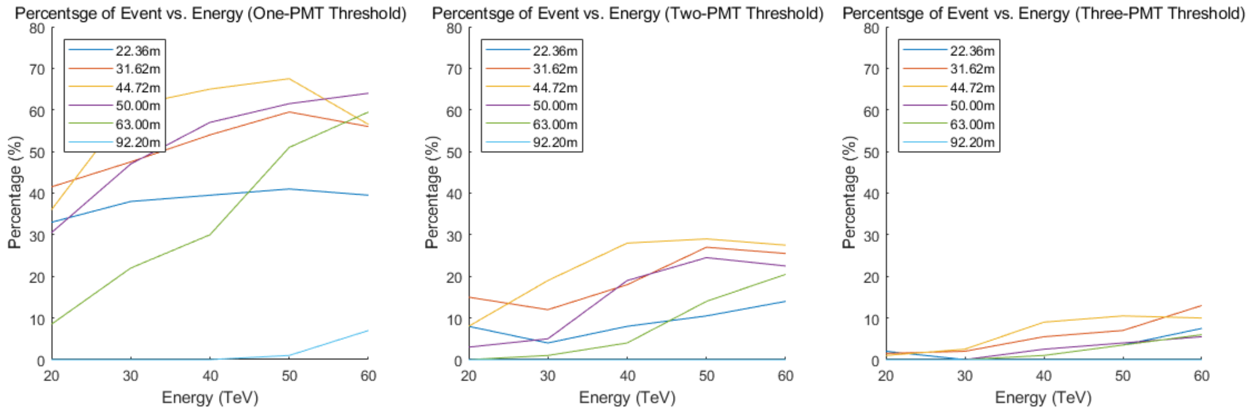
The double peak identification function was tested on 40TeV electron neutrino events and 40TeV tau neutrino events in pure water. Some parameters were adjusted to make the function pick up as many tau events as possible while only a few electron events will be identified as tau events.

Results

For the setup with string separation of 100m, only the PMTs on the string closest to event vertex can see the double pulse signal; Therefore, the desperation between strings are hafted and the following set up was used:



Total 200 event were simulated at each energy. The percentage of events can be identified versus the energy of the neutrino are plotted.



https://github.com/mx906/tau-seperation