JOINT INSTITUTE FOR NUCLEAR RESEARCH

Dzelepov Laboratory of Nuclear Problems

FINAL REPORT ON THE SUMMER STUDENT PROGRAM

Improvement of supernova neutrino detection efficiency in NOνA experiment

---

Supervisor:
Andrey Sheshukov

Student:
Maria Petropavlova

Participation period:
June 25 – August 20

Dubna, 2017

Contents

Abstract

The NOνA experiment is capable of detecting a neutrino signal from future galactic core-collapse supernova. In this work we consider methods of improving the efficiency and range of such detection.

Chapter 1

Supernova Neutrinos

1.1 Core collapse

When a massive star has exhausted its nuclear fuel, it collapses to form a compact object such as a neutron star or a black hole. A prominent feature of the collapse is that 99% of the gravitational binding energy of the resulting remnant is converted to neutrinos with energies of a few tens of MeV over a timescale of a few tens of seconds. This highly efficient energy loss via neutrinos occurs because the neutrinos interact only via the weak interaction and can escape easily, whereas photons are trapped. A core-collapse supernova will produce an enormous burst of neutrinos of all flavors in the few-tens-of-MeV range.

The last Core-collapse supernova was obsereved 1987 (SN1987A) in the Large Magellanic Cloud (LMC), 50 kpc away from Earth. Worldwide capabilities for supernova neutrino detection have increased by orders of magnitude since 1987. The next observation of a nearby core-collapse supernova will provide a great deal of information for both physics and astrophysics. The rate of core-collapse supernovae is estimated to be a few per century in a galaxy such as the Milky Way, so the chance of observing one in the next few decades is not negligible. The most likely distance of the next core-collapse supernova from Earth is between 12 abd 15 kpc, according to the distribution of possible supernova progenitors in the Milky Way [1].

1.2 Neutrino interactions in the tens-of-MeV range

Neutrinos are detected via electromagnetically or strongly interacting products of weak charged-current (CC) and weak neutral-current (NC) interactions with electrons and nuclei. This section describes neutrino interactions in the tens-of-MeV range relevant for supernova detection, such as Inverse Beta Decay and Charged- and Neutral-Current Interactions with Nuclei [1].

Inverse Beta Decay

Relatively cheap detector materials such as water and hydrocarbon-based scintillator have many free protons. The most significant interaction in these materials is Inverse Beta Decay (IBD), which is the interaction between $\bar{\nu}_e$ and free protons,

$$\nu_e + p \rightarrow n + e^+. \tag{1.1}$$

The IBD kinematic threshold is $E_{\nu thr} = 1.8$ MeV. The positron's energy loss can typically be observed. In the supernova energy regime, to a good approximation $E_{e^+} = E_\nu - 1.3$ MeV. In scintillation detectors, the 0.511 MeV positron annihilation γs may also be observed. The neutron may be captured on free protons, with an approximately $200 - \mu s$ thermalization and capture time, producing a deuteron and a 2.2 MeV γ. The neutron may also be captured on another nucleus; in particular, the detector may be doped with some material with a high neutron capture cross section, such as gadolinium (Gd) [1].

A simple approximation which agrees with full result in [5] within few permille for $E_\nu < 300$ MeV is

$$\sigma(\bar{\nu}_e p) \approx 10^{-43}[cm^2] p_e E_e E_\nu^{-0,07056+0,02018 \ln E_\nu - 0,001953 ln^3 E_\nu}, \tag{1.2}$$

where all energies are expressed in MeV, $E_e = E_\nu - \Delta$, $\Delta = m_n - m_p \approx 1,293$ MeV [5].

Charged- and Neutral-Current Interactions with Nuclei

Neutrinos also interact with nucleons in nuclei via CC and NC interactions, although cross sections are typically somewhat smaller for bound than for free nucleons. CC interactions proceed via interaction of ν_e and $\bar{\nu}_e$ with neutrons and protons, respectively, in nuclei,

$$\nu_e + (N, Z) \rightarrow (N - 1, Z + 1) + e^-, \tag{1.3}$$

$$\bar{\nu}_e + (N, Z) \rightarrow (N + 1, Z - 1) + e^+. \tag{1.4}$$

The kinematic threshold is

$$E_{\nu thr} = \frac{M_f^2 + m_e^2 + 2M_f m_e - M_i^2}{2M_i} \sim M_f - M_i + m_e, \tag{1.5}$$

where M_f and M_i are the initial and final state nuclear masses and m_e is the electron mass. Note that at supernova energies, ν_μ and ν_τ are below the CC interaction threshold and thus are kinematically unable to produce their partner leptons. In nuclei, the $\bar{\nu}_e$ interaction is typically suppressed at a given energy with respect to the ν_e interaction due to Pauli blocking. Typically the energy loss of the charged lepton is observable. In the $\bar{\nu}_e$ case, the γs produced by annihilation of the positron may be observable. Ejecta (nucleons and γs)

produced by the final nucleus as it deexcites may also be observable and may help to tag the interaction. The neutrino interactions may furthermore leave radioactive nuclear products, for which decays correlated in space and time with the primary lepton could be observed.

NC interactions on nucleons in nuclei may also produce observable signals via ejected nucleons or deexcitation γs. Of particular note are the 15.1 MeV γ from ^{12}C deexcitation and the ejection of neutrons from lead.

In CC cases, produced leptons retain memory of the incoming neutrino energy, as the heavy recoil nucleus tends to take away little energy. In both CC and NC cases, specific interactions have different thresholds and energy-dependent cross sections, so even if there is no product to measure that remembers the neutrino energy on an event-by-event basis, there is neutrino spectral information to be had by measuring statistical distributions of products. Neutrino interactions on nuclei in the tens-of-MeV range are relatively poorly understood theoretically, in terms of both the interaction rate and the angular and energy distributions of the resulting interaction products. These interactions and their products tend to be sensitive to the details of the nuclear physics involved [1].

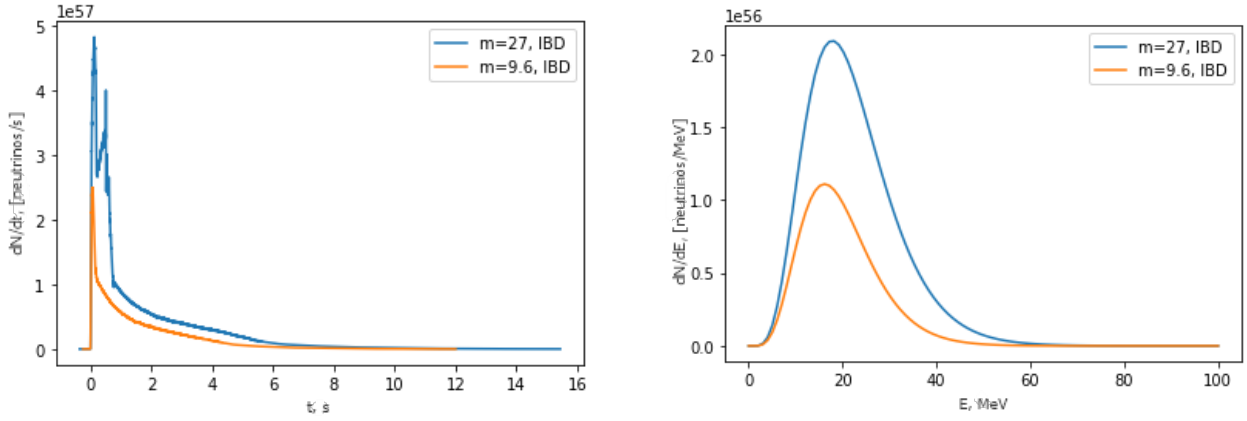
NC on carbon cross section independent of any nuclear model is given by

$$\sigma = 1.08 \times 10^{38} cm^2 \left(\frac{E_\nu - \omega^2}{M_N} \right) \tag{1.6}$$

where E_ν is the energy of the incident neutrino, ω is is the negative Q-value of the reaction (Q=-15.1 MeV), M_N the nucleon mass [6].

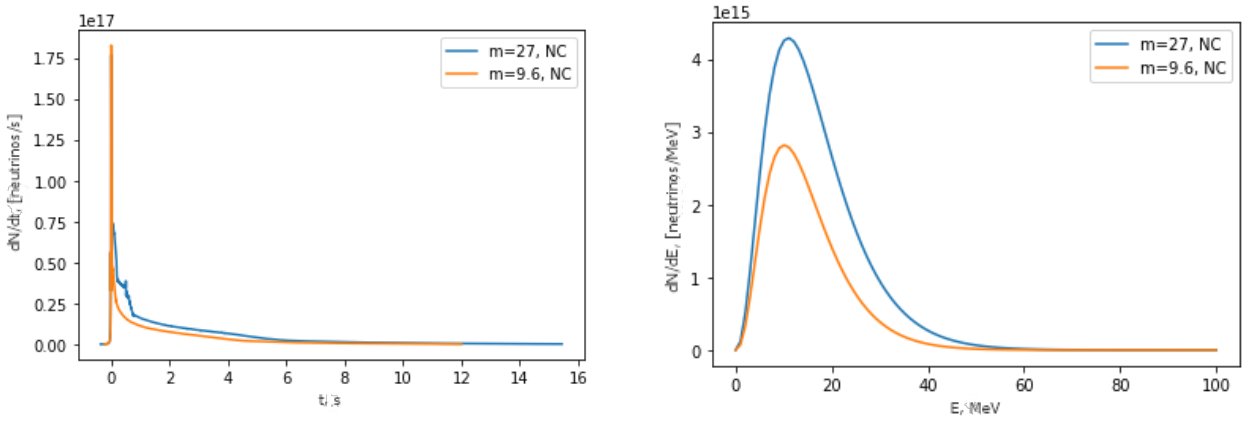
Expected fluxes from supernova for IBD and NC on carbon

We consider a flux on supernova surface, which could be introduced as $\frac{d^2N}{dEdt}$. Following pictures (1.1), (1.2) present expected differential fluxes $\frac{dN}{dt}$, $\frac{dN}{dE}$ for various masses of progenitor star and interactions.



a) b)

Figure 1.1: a) $\frac{dN}{dt}$ for IBD, b) $\frac{dN}{dE}$ for IBD



a) b)

Figure 1.2: a) $\frac{dN}{dt}$ for NC, b) $\frac{dN}{dE}$ for NC

Chapter 2

The NOνA experiment

2.1 Design

NOνA (NuMI Off-Axis ν_e Appearance) is a particle physics experiment designed to detect neutrinos in Fermilab's NuMI (Neutrinos at the Main Injector) beam.

The NuMI is a project which creates neutrino beam. Accelerator neutrinos are generated by a high-energy proton beam striking a nuclear target to produce pions and kaons, which in turn decay into neutrinos. A schematic view of a conventional neutrino beamline is shown in picture 2.1. First, energetic secondary pions and kaons are produced when high-energy proton beams interact with the nuclear target. Second, some of the charged pions and kaons within certain momentum range are focused by magnets (also known as horns), so that they are approximately traveling in parallel with the incident proton beam direction. The polarity of the magnet can be selected by changing the current in order to focus either the positive or negative charged particles. The charged pions and kaons then travel through a long decay pipe to provide enough time for them to decay. Finally, a thick absorber is placed at the end of decay pipe to absorb the muons (decay products of pions and kaons) and other remaining charged particles [4].

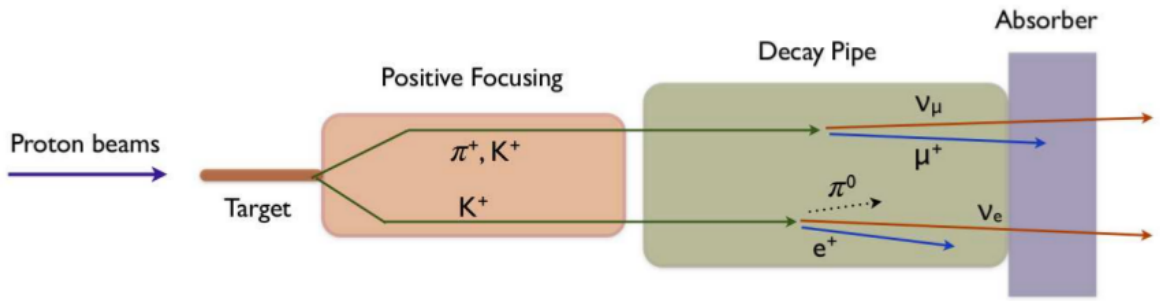


Figure 2.1: A schematic view of the conventional neutrino beamline

The neutrino beam is aimed downward at a 3.3° angle. Although the beam starts out 150 feet below ground at Fermilab, it will pass as much as six miles below the surface as it travels toward Ash River.

Figure 2.2: NuMI Beamline

The NOνA experiment uses two detectors: a 330 metric-ton near detector at Fermilab about 1 km from NuMI target and a much larger 14 metric-kiloton far detector in Minnesota just south of the U.S.-Canada border, 810 km from NuMI target. The NOνA Far Detector are made up of 344,000 cells of extruded, highly reflective plastic PVC filled with liquid scintillator. Each cell in the far detector measures 3.9 cm wide, 6.0 cm deep and 15.5 meters long. The Near Detector has 20,000 cells of the same.

When a neutrino strikes an atom in the liquid scintillator, it releases a burst of charged particles. As these particles come to rest in the detector, their energy is collected using wavelength-shifting fibers connected to photo-detectors. Using the pattern of light seen by the photo-detectors, we can determine what kind of neutrino caused the interaction and what its energy was.

A neutrino beam, much like a beam of light from a flashlight, gradually spreads apart as it travels. The width of the NuMI beam at Fermilab starts at about six feet and grows to several miles by the time it reaches the far detector in Minnesota. NOνA detectors are situated 14 mrad off-axis to the neutrino beam coming from Fermilab, optimizing the signal to background ratio. Due to the kinematics of neutrino production in pion decays, they are exposed to a relatively narrow band of neutrino energies centered at 2 GeV.

2.2 Goals in the NOνA experiment

Much of the NOνA's physics scope comes from the appearance channels, as the observed rates of ν_e and ν_e interactions provide information on the ordering of the neutrino masses (whether the $\nu 3$ state is heavier or lighter than other two), the amount of CP-violation present in the neutrino sector, th size of PMNS mixing angle θ_{13}, whether $\nu 3$ state has more ν_μ and ν_τ admixture (whether the θ_{23} is $>$ or $<$ than 45 degrees).

Through ν_{mu} and disappearance measurements, the NOνA experiment will provide improved precisions on the dominant atmospheric oscillation param-

eters θ_{23} and δm^2_{atm}, which is of utmost importance for the global fits and understanding of the full neutrino oscillation scheme.

Current work is focused on using NOνA to detect neutrinos from galactic supernovas. NOνA will form part of the Supernova Early Warning System.

Chapter 3

Reconstruction and classification

3.1 Reconstruction

Detectors consist of planes divided into cells with vertical or horizontal positions. Every cell transmits the light from a signal through wave-length shifting fibers on an avalanche photo-diode (APD) detector, which measures a signal amplitude. A signal in one scintillator cell is called hit. So we get timing measurement, amplitude and location of a hit.

An IBD positron almost simultaneously scintillates a few near cells. Based on this fact we reconstruct clusters with total amplitude of combined hits.

A cluster belongs to XY class if both vertical and horizontal cells were scintillated, to XX class or YY class if only vertical or only horizontal cells, respectively, were scintillated.

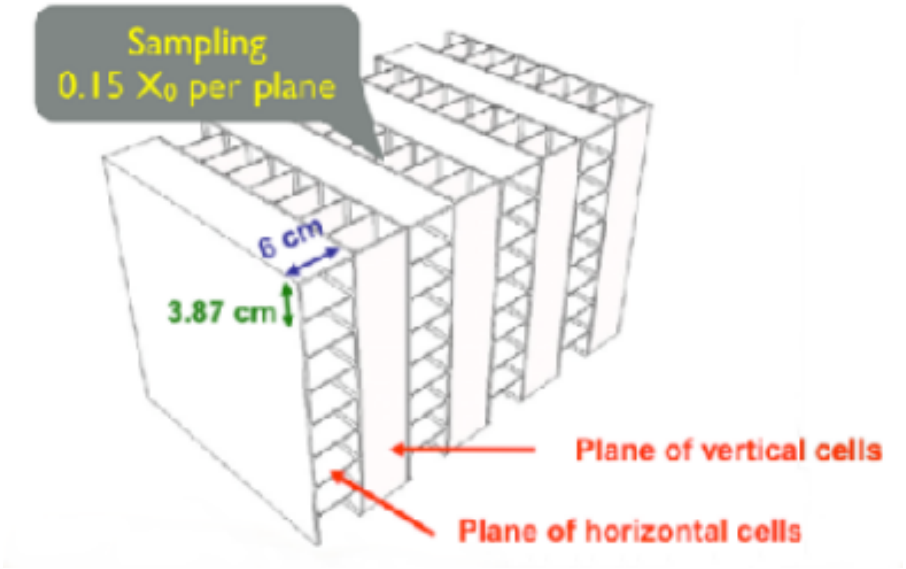


Figure 3.1: Illustration of planes in a detector

3.2 Classification

We need to divide clusters with signal from clusters with background. And as soon as we got reconstructed data: total amplitude, location, quantity of scintillated cells and timing measurement for each cluster, we could implement the procedure of classification.

We use support vector clustering(SVC) with radial basis function (RBF) kernel for our classifier. At this stage we have $recall = 0.94$ on testing data.

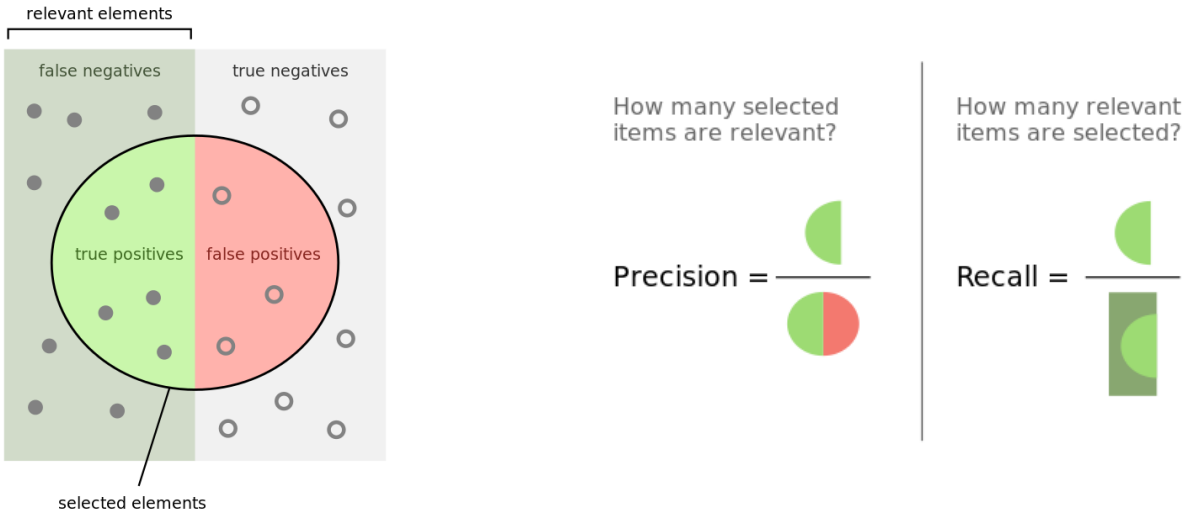


Figure 3.2: Precision and recall

At the output the classifier gives signal and background numbers of events for 1s and every 5μs.

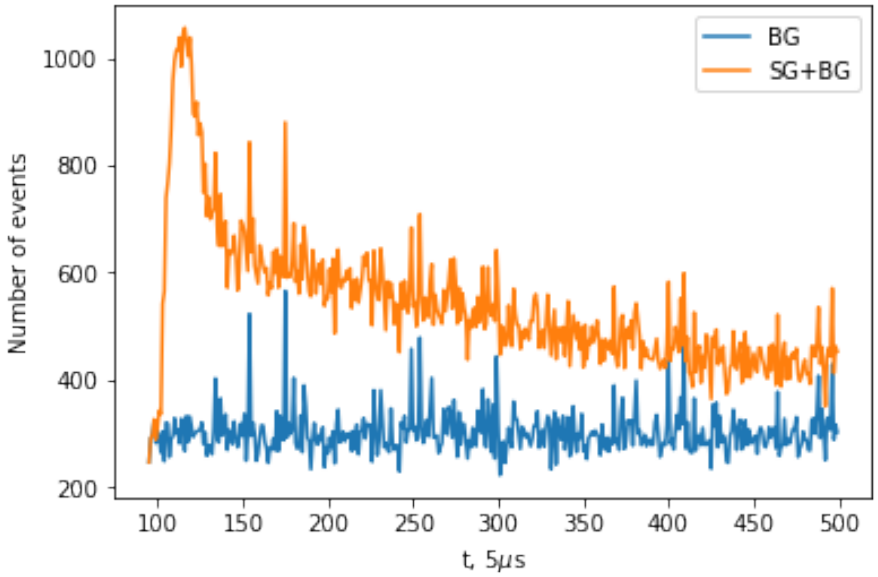


Figure 3.3: Number of events, classified as signal candidates vs time without the supernova (blue) and with supernova at 1kpc distance(orange)

Chapter 4

Trigger efficiency estimation

4.1 Simple model of finding a threshold

There are some average numbers of events for background(BG) and signal(SG). So we can use Poisson distribution with definite λ. In case we have only one detector and supernova signal notification have to responce once a week, we could define a probability of triggering:

$$\alpha = \frac{1}{7 * 24 * 60 * 60} \simeq 1,653 \cdot 10^{-6}. \tag{4.1}$$

Then a trigger response will depend from numbers of events on that detector. The distribution tail in turn have to be equal to α as the task requires.

$$\alpha = \int_x^{\infty} P(k)dk, \tag{4.2}$$

where x will be a threshold of triggering.

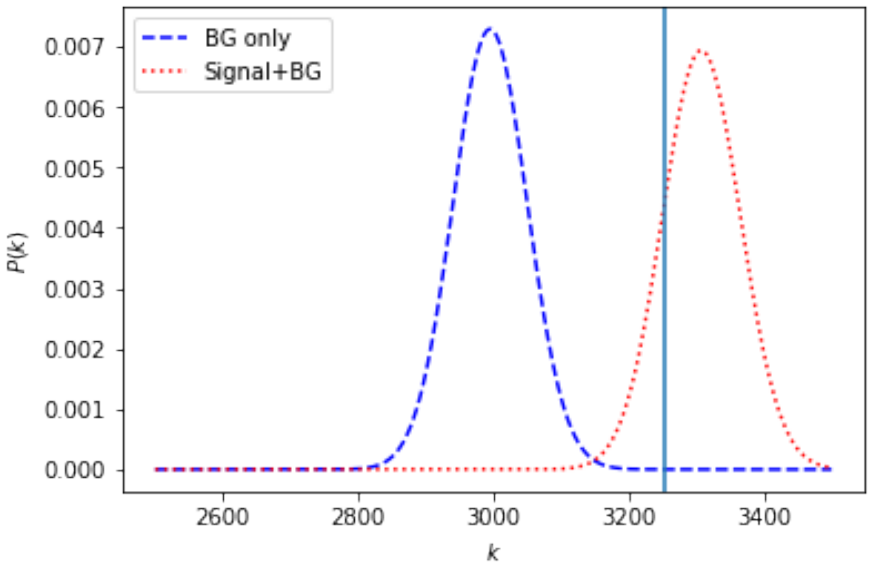


Figure 4.1: Threshold for one observable

4.2 Increasing of detection efficiency with adding new interaction channels

But in case we have 2 detectors, 2 channels of interactions (IBD, NC) on each detector and data from XY, YY, XX event classes the task becomes more complicated.

Now the NOvA experiment uses only two measurements: XY event class on IBD channel from Far and Near detectors. Let us consider the case where data from all event class are involved in analyzes and compare detection efficiency.

	Name	BG	SG
0	XY_FD	2951.400	312.2800
1	XY_ND	1.682	4.3932

a)

	Name	BG	SG
0	XY_FD	2951.400	312.2800
1	XX_FD	6370.400	178.0400
2	YY_FD	7817.700	169.3200
3	XY_ND	1.682	4.3932
4	XX_ND	11.572	1.0664
5	YY_ND	3.662	1.4784

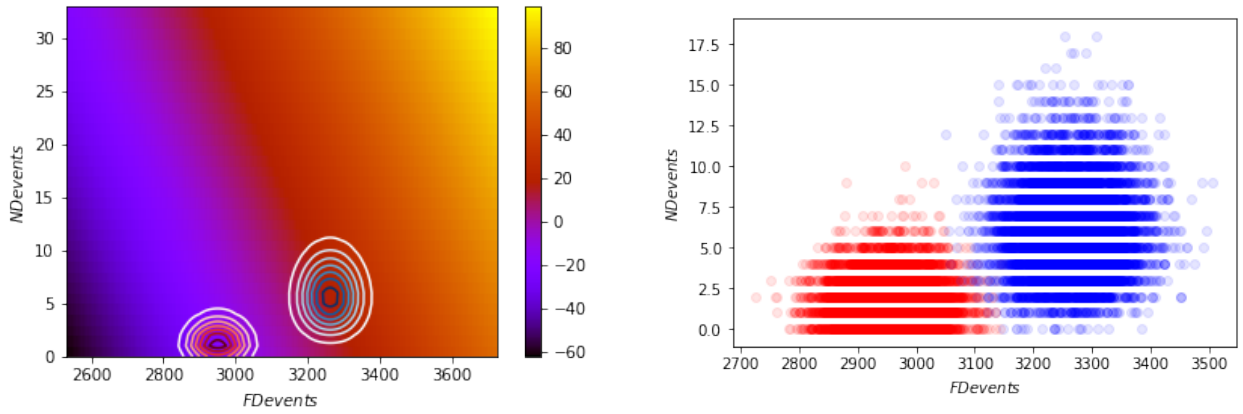
b)

Figure 4.2: Examples of data on IBD channel: a) only XY event class in FD, ND detectors, b) all event classes in FD, ND detectors

We use Poisson distribution to create samples for signal and background events. Our data are 2- and more dimensional therefore we take likelihood as follows:

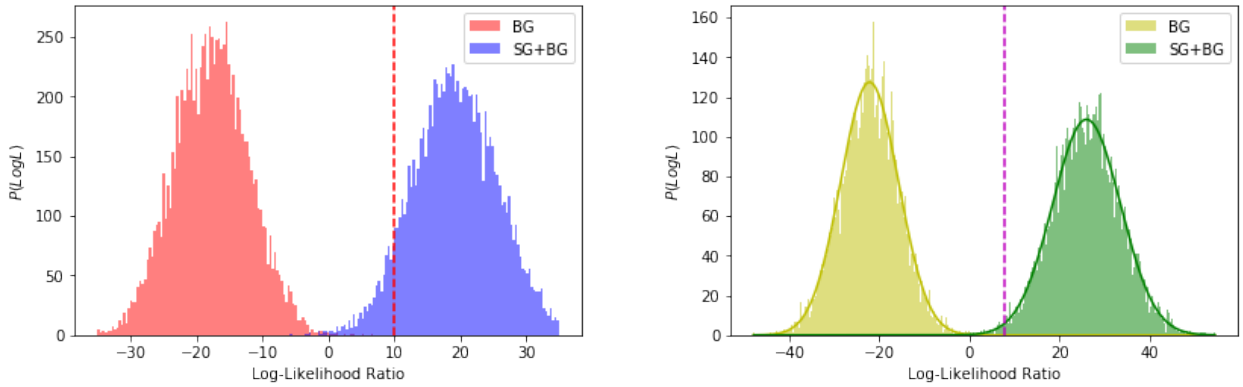
$$
\begin{aligned}
LogL &= \sum_{j}^{d} logL = \\
&= \sum_{j}^{d} \ln\left(\frac{pmf(\lambda = BG_j + SG_j, k = n)}{pmf(\lambda = BG_j, k = n)}\right) = \\
&= \sum_{j}^{d} n \cdot \ln\left(\frac{BG_j + SG_j}{BG_j}\right) - \sum_{j}^{d} SG_j
\end{aligned}
\tag{4.3}
$$

where n is a point from sample, d is quantity of measurments or dimensional of sample. So we get one dimensional log-likelihood ratio function ($Lsample(bg)$,$Lsample(sg)$) instead d-dimensional sample ($sample(bg)$, $sample(sg)$). As soon as $Lsample$ were found we could get its average and variance. Taking into account we used Poisson distribution with $N \to \infty$, $Lsample$ could be approximated with normal distribution. Substituting resulting probability density function into 4.2 we get threshold for the case with d measurements.



a) b)

Figure 4.3: a)Log-Likelihood ratio distribution for 2 observables, b)Poisson distribution for background and signal



a) b)

Figure 4.4: a) Threshold for 2 observables, b) Threshold for 6 observables

4.3 Comparison of the detection efficiencies for various SN distances

Our data were simulated for 10 kps. In point of observation radiation flux depends on square of distance r till radiation source. Therefore we could estimate detection efficiency converting data signal part in following way:

$$newSG = \frac{SG \cdot 10^2}{r^2}. \tag{4.4}$$

At the following picture 4.5 we took r in range from 0 to 17 kpc. It shows the distinction between detection efficiencies in cases of using one event class and three event classes.

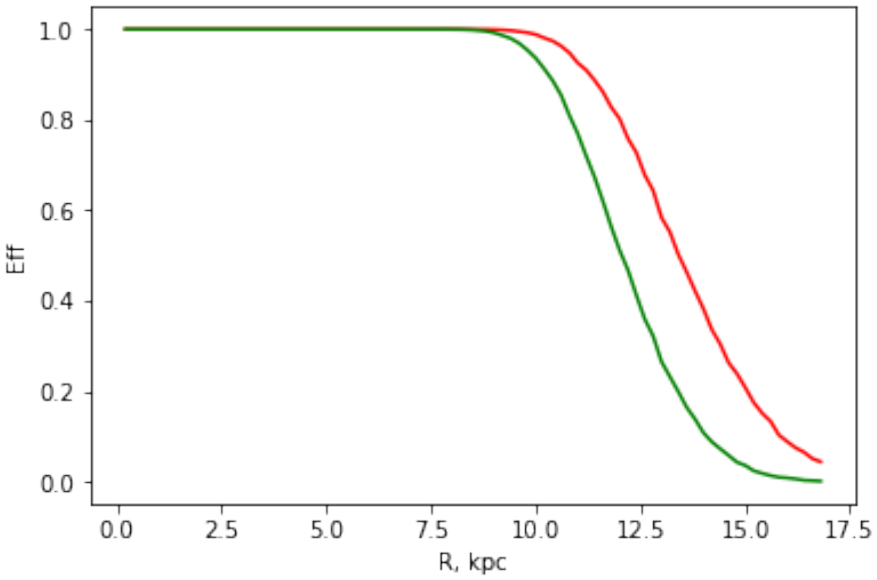


Figure 4.5: Dependence of the efficiency from R. Green - only XY event class, red - all event classes.

4.4 Accounting for the expected signal shape

Earlier, we had data with number of events for 1s. Now we will use data with number of events for every 5ms. So instead of an average number of events in 1s time interval we could consider 200 time intervals with duration of 5ms. Therefore one can use (4.3) as previously described:

$$LogL(\vec{N}) = \sum_i N_i \cdot \ln\left(\frac{S_i + B_i}{B_i}\right) - \sum_i S_i, \tag{4.5}$$

where S_i and B_i - average number of signal and background events in i-th 5 μs time bin, and N_i is a measured number of candidates.

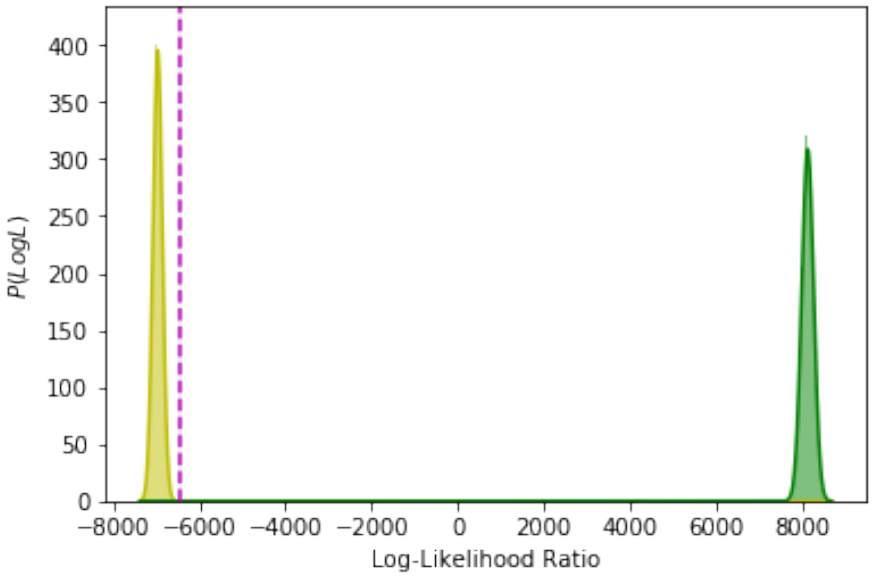


Figure 4.6: Threshold for 200 observables for 1 kpc

Conclusion

In current work we considered following methods of supernova detection efficiency improvement:

1. Taking into account additional event topologies (XX and YY candidates).
2. Accounting for the expected signal shape.

We evaluated the detection efficiency depending on a distance to a supernova explosions for various numbers of observables.

Also a new method of the selection of supernova interaction candidates selection was considered. This method, based on machine learning algorithms, could provide a significant increase in signal/background ratio.

Appendix

Poisson distribution

The probability of observing k events in an interval is given by the equation

$$P(k \text{ events in interval}) = e^{-\lambda}\frac{\lambda^k}{k!} \tag{4.6}$$

where λ is the average number of events per interval.

Likelihood function

In statistics, a likelihood function (often simply the likelihood) is a function of the parameters of a statistical model given data. In statistics, a distinction is made depending on the roles of outcomes vs. parameters. Probability is used before data are available to describe possible future outcomes given a fixed value for the parameter (or parameter vector). Likelihood is used after data are available to describe a function of a parameter (or parameter vector) for a given outcome.

Log-likelihood function

The natural logarithm of the likelihood function, called the log-likelihood, is more convenient to work with. Because the logarithm is a monotonically increasing function, the logarithm of a function achieves its maximum value at the same points as the function itself.

For example, some likelihood functions are for the parameters that explain a collection of statistically independent observations. In such a situation, the likelihood function factors into a product of individual likelihood functions. The logarithm of this product is a sum of individual logarithms, and the derivative of a sum of terms is often easier to compute than the derivative of a product. In addition, several common distributions have likelihood functions that contain products of factors involving exponentiation. The logarithm of such a function is a sum of products, again easier to differentiate than the original function.

$$L(\mathbf{x} \mid \theta) = \sum_{i=1}^{n} \ln f_X(x_i \mid \theta), \tag{4.7}$$

where events are independent.

Likelihood-ratio test

A statistical model is often a parametrized family of probability density functions or probability mass functions $f(x|\theta)$. A simple-vs.-simple hypothesis test has completely specified models under both the null and alternative hypotheses, which for convenience are written in terms of fixed values of a notional parameter θ:

$$H_0 : \theta = \theta_0,$$
$$H_1 : \theta = \theta_1.$$

Note that under either hypothesis, the distribution of the data is fully specified; there are no unknown parameters to estimate. The likelihood ratio test is based on the likelihood ratio, which is often denoted by Λ (the capital Greek letter lambda). The likelihood ratio is defined as follows:

$$\Lambda(x) = \frac{L(\theta_0 \mid x)}{L(\theta_1 \mid x)} = \frac{f(\bigcap_i x_i \mid \theta_0)}{f(\bigcap_i x_i \mid \theta_1)} \tag{4.8}$$

Acknowledgements

I would like to thank my supervisor Andrey Sheshukov for the patient guidance, encouragement and advice he has provided throughout my time as his student. I also must express my gratitude to Konstantin Treskov for valuable discussions and advises. Finally, I would like to thank JINR University Centre for providing the financial support during my summer practice.

Bibliography

[1] K. Scholberg - Supernova Neutrino Detection

[2] Supernova Neutrinos: Production, Oscilations and Detection //IL NUOVO CIMENTO

[3] https://www-nova.fnal.gov/how-nova-works.html

[4] Qian X. Neutrino Mass Hierarchy/ X. Qian, P. Vogel // arXiv:1505.0189[hep-ex]

[5] Strumia A. Precise quasielastic neutrino/nucleon cross-section/ Alessandro Strumia, Francesco Vissani

[6] Armbruster B. Measurement of the weak neutral current excitation ${}^{12}C(\nu_\mu, \nu_\mu^{'}){}^{12}C^*1(1^+, 1; 15.1 \text{ MeV})$ at $E_{\nu_\mu} = 29.8$ MeV / B. Armbruster , I. Blair et al. KARMEN Collaboration //Physics letters B 423(1998) 15-20