ICECUBE MASTERCLASS
AN AUTHENTIC ASTROPHYSICS RESEARCH EXPERIENCE

Analysis activity

High-energy neutrinos from the cosmos

What we need

- One computer per every one or two students, connected to the internet.
- A copy of the *Science* paper for each student.

Resources provided

- General description of the activity, with links to tools, on the IceCube Masterclass website. (link)
- Example of an introductory talk on IceCube (link) and two talks on the analysis (link1, link2).
- Talk to guide the point source analysis (link).
- Randomized sky maps of the very high energy events detected in IceCube. (link)

Activity proposal

The goal is to allow students to replicate the search for a flux of astrophysical neutrinos published in *Science* (Nov 2013). Exercises and discussions grow in complexity while addressing the main ideas of the research.

The activity is designed in four parts, plus a bonus activity for afterward at home. We recommend covering at least three parts: i, ii and then iii or iv. Going through the four parts, with the proper guidance of tutors, should not take more than four hours.

Activity outline:

i) Introductory talk to explain the analysis goals and tools.
ii) Qualitative analysis to understand the signal we are looking for.
iii) Significance of the flux of very high energy events detected in IceCube.
iv) Point sources search.
v) At home: Read the paper and send comments/questions to masterclass@icecube.wisc.edu. We will open a two-week period after the masterclass to collect all questions, after which we will send out a final mailing with answers to all participants.

Activities in sections i,ii,iii designed by Sílvia Bravo and Claudio Kopper (WIPAC)

i) Introduction to the analysis

We suggest a 10- to 20-minute talk, where we introduce the analysis that the students will be working on. After questions, if any, from the students, tutors will outline the rest of the activity.

ii) Qualitative analysis: understanding the nature of our signal and background

- This will begin with a discussion of particles as detected by IceCube and a 5-minute introduction to the IceCube event display. We can use information on the masterclass website: "The Detection of Neutrinos in IceCube." If the event display has already been introduced in the introductory talks, students as a group could go through some of the events of the Particle ID Quiz.
- Tutors will start a discussion going through some of the ideas presented in the introductory talk. Students will talk about how signal and background are similar/different and identify the properties that allow us to distinguish between them.
- To help students guess some of the signal-vs.-background properties, students will discuss as a group the differences in the displays from a set of simulated signal and background events (Particle ID Quiz, background_signal).
- Tutors will guide students in designing an event selection for this analysis. We would like them to arrive at the idea of the veto implemented in the HESE analysis, and students can try to see how this veto works with the training Particle ID quiz, Selecting neutrinos in IceCube: the veto approach. Students should also arrive to the idea that we are expecting a few events.

iii) Significance of the flux of very high energy events detected in IceCube

- Tutors will present the display of the 28 very high energy events that passed the veto selection. Students will be given two minutes to guess which are the ones with the highest energy (using the Particle ID Quiz, hese_all).
- Tutors can spend a few minutes going through each of the 28 events, explaining why some of them are especially interesting (like tracks that point to sources, the highest energy ones because they are most probably signal, …) or just explaining singularities of the data (such as why we use time in Julian date, why declination and right ascension, …). Use Particle ID Quiz, hese.
- Tutors will provide the students with a table of properties of the 28 events (included in the talk) and will guide an initial discussion about which of the properties can be easily identified as "signal-like." At this point, we would like the students to focus on three properties: number of showers, declination and high-energy events.
- Tutors will guide students to think about how we can estimate the significance of this signal. We would like students to think about the relation between theory and experiments, the idea of proving that data is significantly different

from background—basically, counting how many events we observe for a given energy/declination and comparing these numbers to what we expect from a background-only hypothesis.

- Now it's time for students to play with the energy cut and see how the zenith distribution changes (link).
- Tutors will show a comparison of theoretical estimates for signal and background, explaining how the significance can be calculated. They will guide students to the idea of binned distributions, and to the importance of choosing the correct range and binning for each property.

iv) Quantitative analysis: search for point sources

Activity designed by Jacob Feintzeig (WIPAC)

Objective: Students should be able to:

1) Describe why we want to find point sources, and list examples of potential sources.
2) Describe or illustrate what a point source looks like in HESE data.
3) Perform a search for a point source using pre-generated sky maps and calculate a p-value.
4) Explain what the p-value means and how the analysis could be improved.

- Tutors will ask/remind to students why we are looking for a flux of very high energy neutrinos, i.e., that we want to find the sources of cosmic rays. What are some potential sources?
- Tutors will show two maps, one with positions of 28 random events and one with positions of events with a simulated point source. This will be followed by a discussion of how a point source appears in our data and the concept of angular resolution.
- Next will be a discussion about what a signal is, what background is, and how to distinguish between the two qualitatively.
- Tutors will describe a simple method of counting events inside a "signal region" and repeating the procedure on scrambled maps to estimate significance. We will choose as our region a circle around the galactic center with a 15-degree radius. Concepts of p-values and RA scrambling will be explained.
- Students will break up into groups of two (or four, depending on total class size). Each group will be provided with 10-20 randomized sky maps and will be asked to count events in the signal region.
- Groups will come up to the board to construct histograms of scrambled trials as a class.
- Students will be shown the real sky map and asked to count the number of events in the signal region and use a histogram to estimate p-value.
- The group will discuss: Is this significant?

- Tutors will lead a discussion on obvious flaws of the analysis: What if an event falls right outside the bin? If there's a track inside the bin, shouldn't that be more significant? Wouldn't it be better to include detailed information on angular resolution for each event individually? What if we don't know where the source is going to be? This leads to a basic description of what the likelihood is and does, guided by the paper. A sky map indicating likelihood (colors) will be shown.
- Tutors will explain how the p-values were calculated in the paper. ok
- Finally, tutors will guide a discussion about future searches for cosmic ray sources in IceCube, including the importance of statistics and correlations with results from other experiments.