Supernova triggering and signals combination for the NOvA detectors

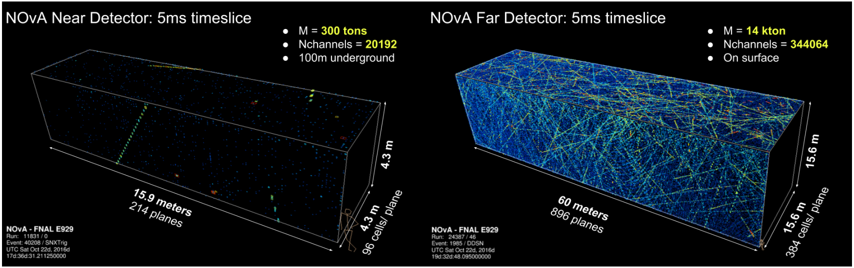
Andrey Sheshukov

Joint Institute for Nuclear Research, Dubna

15.06.2019

NOvA detectors

Main goal of the NOvA experiment: study neutrino oscillations in the muon (anti-)neutrino beam, with $\langle E_\nu \rangle = 2$ GeV.

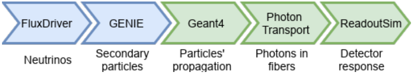


Two detectors, composed of extruded PVC cells, filled with liquid scintillator.

- Similar structure ⇒ almost the same reconstruction and data processing,
- Different size and overburden ⇒ very different BG conditions, statistics.

SN neutrino interactions simulation

We are developing **GenieSNova** package: feeding SN neutrino flux into GENIE.



Features:

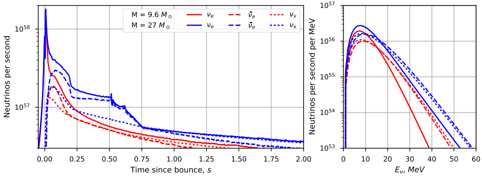
- Provides a flux driver for GENIE
- Reads GDML detector geometry
- Can read Garching and SND model files

Generates:

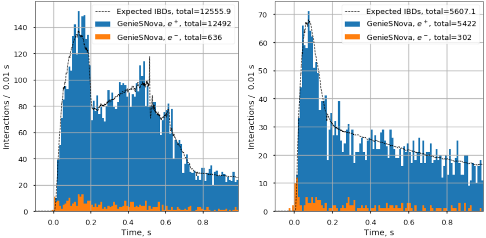
- T_ν: time ordered — important for pipeline frameworks (ART)
- E_ν: histogram or analytical formula
- Position: in a window near the detector

Interactions:

- Inverse beta decay
- Elastic scattering on electron



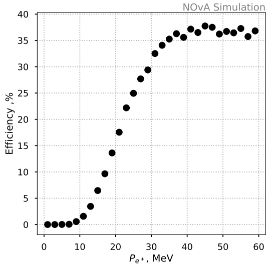
Neutrino signal from core-collapse supernova
(Garching arXiv:astro-ph/0604300)

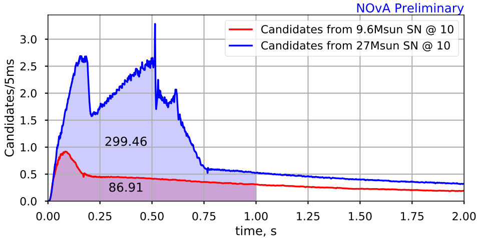


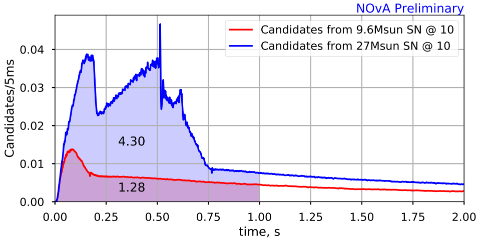
Generated secondaries (e^+, e^-) vs expected number

Detecting supernova neutrinos

Main detection channel: $\bar{\nu}_e + p \to e^+ + n$.







	N_{bg}	N_{sg}@10kpc
FarDet	2483.21	86.91
NearDet	0.52	1.28

Table: Expected number of candidates from the first second of SN, after the reconstruction and selection procedures

Supernova triggering system: overview

A dedicated triggering system was deployed on both NOvA detectors.
Main goal: process data in real time in search of SN neutrino signal.

- Split incoming data in 5ms slices
- Each slice is processed in one of 2200 parallel processes (169 for NearDet):
  - Background rejection
  - Reconstruction of neutrino interaction candidates
  - Send $\{N_{cands}, Timestamp\}$ to the trigger node.
- Trigger node:
  - Collects and sorts the data into a time series $\{n_i, t_i\}$
  - Analyzes the data to decide if there was significant SN-like signal.
  - If signal is found: tells the system to save the data for more thorough offline processing.

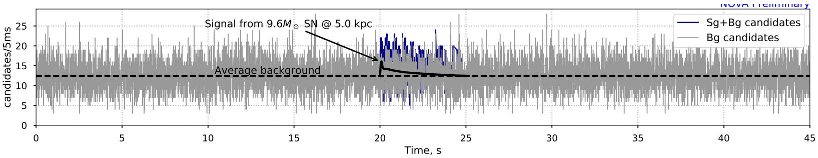
The triggering system operates since Nov 2017.

Each detector has a separate trigger pipeline, but the final trigger decision is shared: if FarDet or NearDet triggers, save data on both.

Trigger latency (time from SN start to the trigger decision) is **40-60s** for FarDet and **5.7s** for NearDet.

Analyzing time series

Input data from the candidates selection: time series: $n_i = b_i + s_i$



The triggering system needs to distinguish between the "Bg only" H_0 and "Bg+SN" H_1 hypotheses, within a sliding time window $\vec{n} = \{n_i\}$

- Define a test statistics: $X(\vec{n})$

Defining test statistics: log likelihood ratio

Best discrimination power obtained, if we use all available information: n_i, t_i
Use the information about the expected signal shape vs time.

$$P(n_i|H_0) = \frac{B^{n_i} \cdot e^{-B}}{n_i!}, \qquad P(n_i|H_1) = \frac{(B+S_i)^{n_i} \cdot e^{-(B+S_i)}}{n_i!}$$

Log likelihood ratio

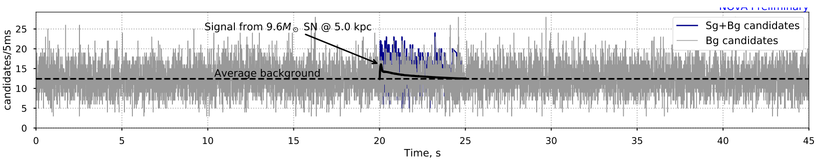
$$\ell(\vec{n}) \equiv \log \frac{P(\vec{n}|H_1)}{P(\vec{n}|H_0)} = \sum_i n_i \cdot A_i, \quad \text{where } A_i = \log\left(1 + \frac{S_i}{B}\right)$$

Very convenient for real-time computation:

- A linear combination of the data in considered time window.
- For trigger we want a sliding time window - this becomes a convolution of incoming data n_i with the kernel $\hat{A}$.
- Kernel $\hat{A}$ recalculated only when we remeasure the background level (every 1 minute for FarDet, 10 minutes for NearDet)

Analyzing time series

Input data from the candidates selection: time series: $n_i = b_i + s_i$



The triggering system needs to distinguish between the "Bg only" H_0 and "Bg+SN" H_1 hypotheses, within a sliding time window $\vec{n} = \{n_i\}$

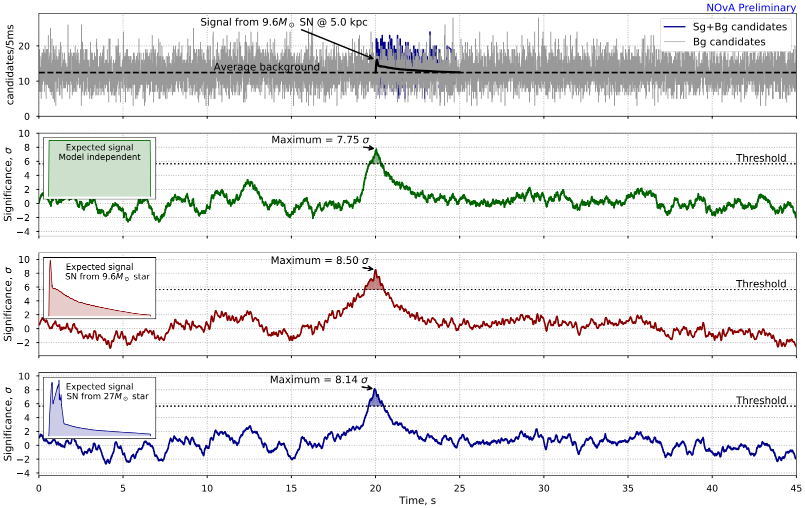
- Define a test statistics: $X(\vec{n}) = \sum_i n_i \cdot A_i$
- Significance is characterized by p-value: $p(X(\vec{n})) \equiv P(x > X(\vec{n})|H_0)$
- Significance in z-score ("sigmas"): $z(x) = \mathrm{erf}^{-1}((1 - p(x))/2)$.

Trigger fires if significance exceeds threshold:

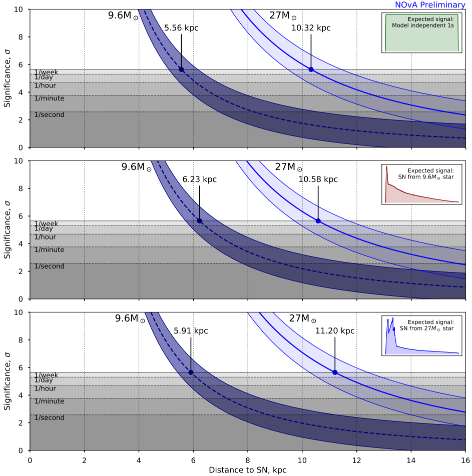
$\alpha = 1/week$ $\Longleftrightarrow$ $p_{thr} = 8.267 \cdot 10^{-9}/5ms$ $\Longleftrightarrow$ $z_{thr} = 5.645\sigma$

Note: z follows standard normal distribution $\mathcal{N}(\mu = 0, \sigma = 1)$ for H_0 — by construction.

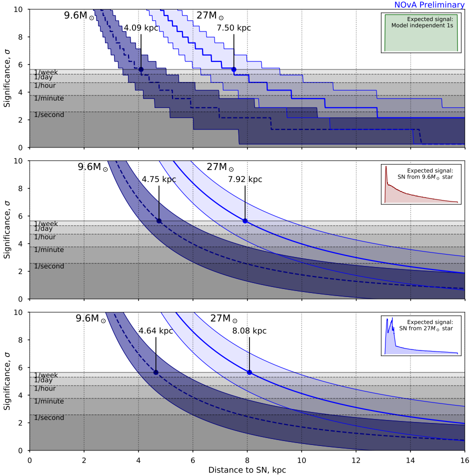
Example: triggering on Far Detector



Significance vs. distance



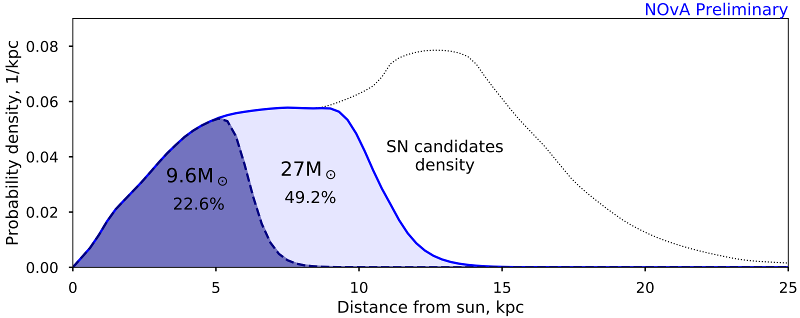
Far Detector.



Near Detector.

Results: Far detector reach in the galaxy

We use 9.6M model for our expected signal shape for the trigger setup. Using Far detector reach, we can calculate a fraction of SN candidates in the galaxy [arXiv:astro-ph/0604300] we can detect:

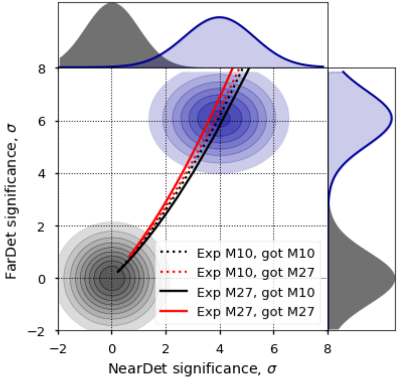


Joint significance

Currently the detectors perform hypothesis test separately.

If we can get the synchronized significance scores from both detectors, we can perform meta-analysis with these two values.

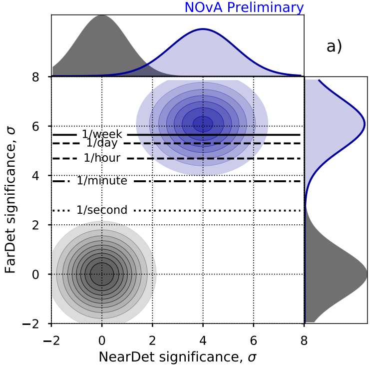
- Significance scores from detectors define a point on N-dimensional space.
- Define a combined "significance" function, which can then be used as test statistics: $S(z_{near}, z_{far})$.
- The ways to define this function are infinite, depending on what we want:
  - Maximize the efficiency for specific signal?
  - Keep as model independent as possible?
  - Minimize the data flow between detectors?



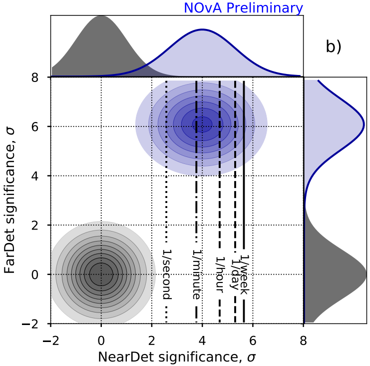
NOvA as a miniature coincidence network: easy to extend from 2 to N detectors.

Combination modes

Detectors triggering separately:



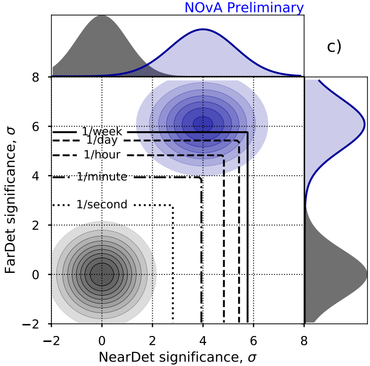
$S = z_{far}$

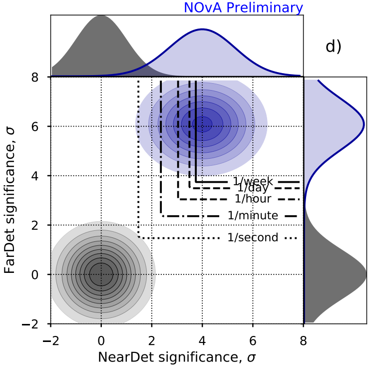


$S = z_{near}$

Combination modes

Trigger signals combined (each detector still makes its own decision, but we combine the results):

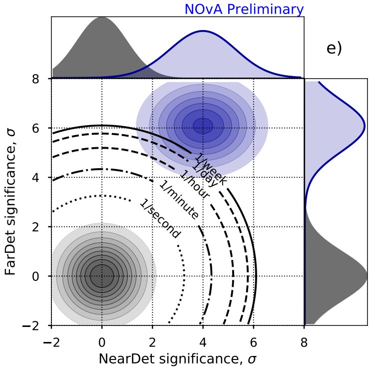




FD or ND: $S = \max(z_{far}, z_{near})$

FD and ND: $S = \min(z_{far}, z_{near})$

"Model independent" combination



$$S = \sum_k^N z_k^2$$

(Ficher's combination)

Background distribution:
χ^2 distribution with N degrees of freedom.

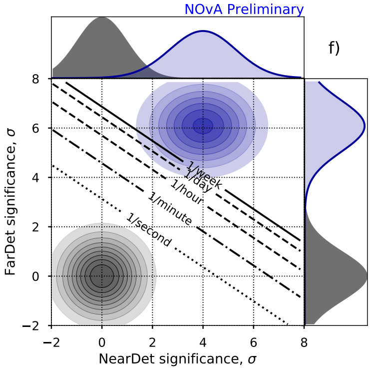
Advantages:

- A natural thing to do with p-values — multiply them.
- Independent of the SN model: just reject background.

Disadvantages:

- Non-sensitive detector can decrease total sensitivity of network (imagine NOvA NearDet vs SK).
- Nonlinear
- Background distribution depends on the number of detectors — change the thresholds when one of the detectors is down?

Linear combination



$$S = \sum_{k}^{N} w_k z_k$$

(Stouffer's z-score combination)

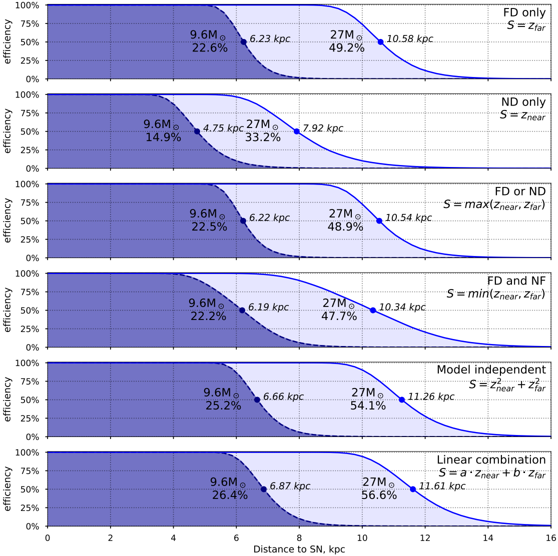
Basically, we project our significance to the vector $\vec{w}$.

Background distribution:
Standard normal distribution:
$\mathcal{N}(\mu = 0, \sigma = 1)$

Advantages:

- Weight for each detector is defined by average significance on "standard candle" (SN @ 10kpc).
- Background distribution is fixed for any number of detectors. So thresholds are fixed.
- Linear transformation: easy and efficient to calculate.
- We can account for possible correlations between detectors BG, by adjusting $\vec{w}$.

Combination modes efficiency comparison



Summary efficiency vs. time for all described combinations. Percents give the covered fraction of SN candidate in galaxy.

- "OR" and "AND" modes give worse sensitivity, than "FD only"
- "Linear combination" mode gives the best result.

SN coincidence server

We have developed a dedicated server application for receiving data from detectors and calculate combined significance.

Purpose:

- NOvA detector:
  2 clients sending data with 5ms interval
- Prototype for SNEWSv2:
  N clients sending data with various intervals and delays

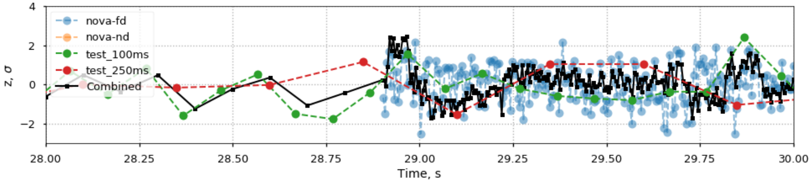
Requirements:

- Calculate combined significance in real time
- React fast! Use currently available data:
  - If the fast experiments provide high significance - trigger!
  - Don't have to wait for the very last experiment to send data
- We need to know which detectors are alive
- Streaming data (no connect-send-disconnect)
- Security: use SSL certificates to authorize clients

SN coincidence server

Features:

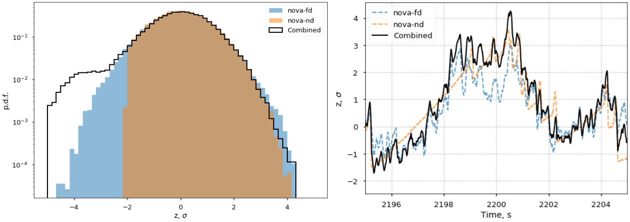
- A client (SN detector) connects to the server and authenticated by SSL certificate
- A client connects and starts streaming data to the server: sending arrays of data points z_i, t_i.
- Server handles each client in a separate thread, storing data points.
- When new data is received, the server updates the combined significance around this point.
    - This allows to react fast.
    - Also the result will be independent on the order in which the data arrives.
- Server adapts its time sampling to the clients currently connected.



SN coincidence server

Current status:

- First prototype works
- Tested by fake clients — stable
- Tested by real data from NOvA detectors — stable, BG distribution under control

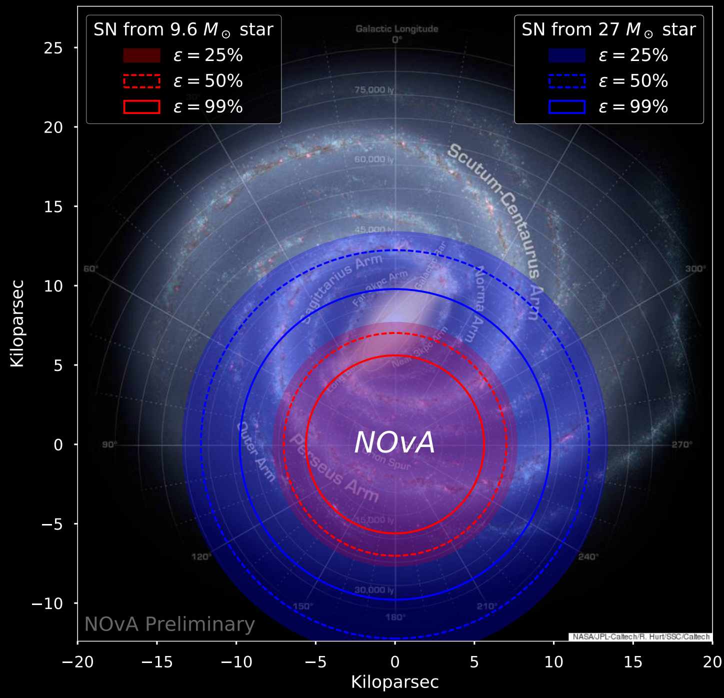


Future plans: this summer

- Tests with live data stream from detectors
- Implement monitoring
- Implement trigger logic

Summary

1. NOvA trigger system is operating since Nov 2017.
2. Using the log likelihood ratio as the test statistics allows to take into account the signal shape and enhance the sensitivity.
3. For coincidence network we propose the method: linear combination of significance scores. It allows a network to be extendable and stable for any number of experiments.
4. A working prototype of a coincidence server was developed using this method.



SN from 9.6 $M_\odot$ star
$\varepsilon = 25\%$
$\varepsilon = 50\%$
$\varepsilon = 99\%$
SN from 27 $M_\odot$ star
$\varepsilon = 25\%$
$\varepsilon = 50\%$
$\varepsilon = 99\%$
Galactic Longitude
0°
75,000 ly
60,000 ly
45,000 ly
30,000 ly
15,000 ly
60°
90°
120°
150°
180°
210°
240°
270°
300°
Scutum-Centaurus Arm
Sagittarius Arm
Norma Arm
Outer Arm
Perseus Arm
NOvA
NOvA Preliminary
NASA/JPL-Caltech/R. Hurt/SSC/Caltech
Kiloparsec
Kiloparsec
25
20
15
10
5
0
−5
−10
−20
−15
−10
−5
0
5
10
15
20