Background Rejection in the ARA Experiment

Carl Pfendner

Ohio State University

Supported by
NSF CAREER Award 1255557,
NSF ARA Grant 1404266,
BigData Grant 1250720

Contents

Part 1: Background rejection in the ARA Testbed station

- Event analysis techniques
- GRB timing rejection
- Optimization of cuts

Part 2: New algorithm for background rejection in stations with regular geometry

- Regular geometry advantages
- Efficiencies

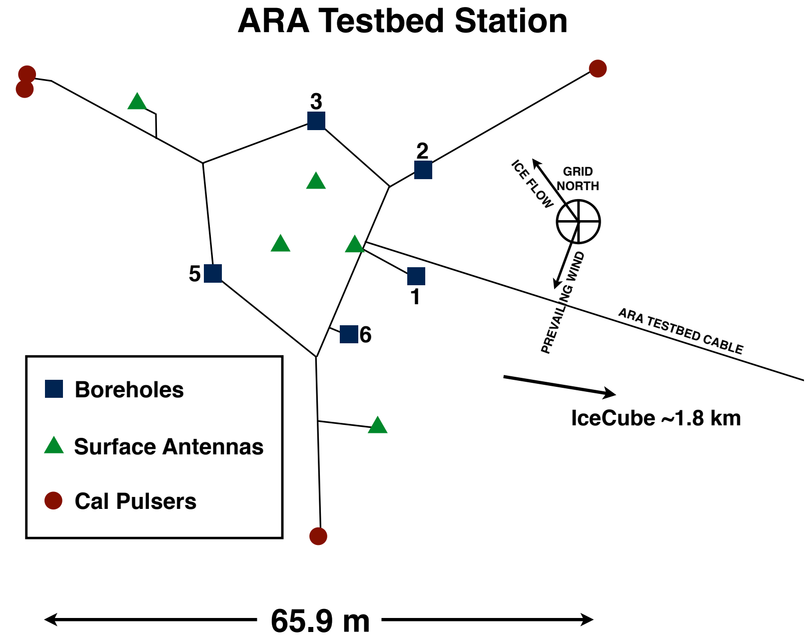
BACKGROUND REJECTION IN THE ARA TESTBED GRB NEUTRINO SEARCH

Background Rejection

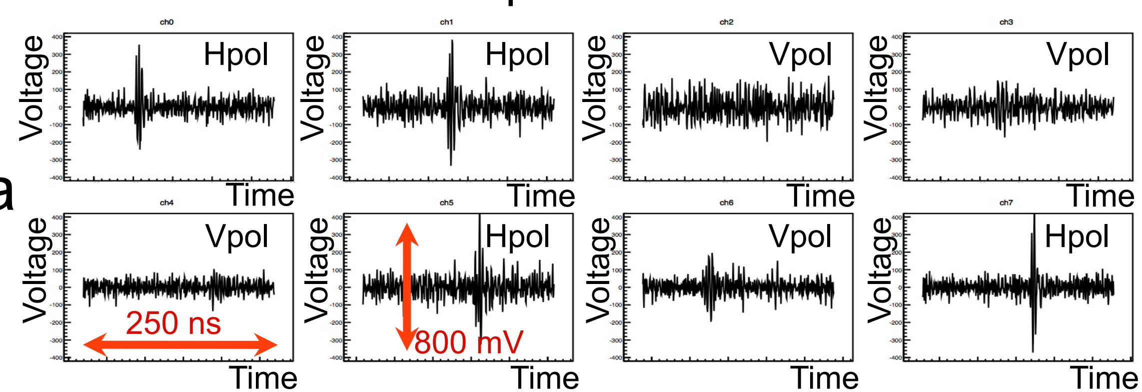
- 2 basic types of noise
  - CW
  - Thermal
    - Characterized by (semi-)random fluctuations from surrounding environment
- ARA trigger – based on tunnel diode output
  - Acts as a few-ns power integrator
  - Trigger rides a threshold determined by the thermal noise level
  - 100's of millions of events – almost all thermal noise
- How to reject these signals efficiently?
  - For analysis cuts
  - For filtering before transmission to the North

Testbed Station

- Total 16 antennas, 8 borehole antennas at 150 MHz to 850 MHz
- Maximum depth of antennas ~ 30 m
- 3 sets (Vpol + Hpol) of calibration pulsers
- Deployed 2010-2011
- Ran for 2 years (2011 – 2012)
  - Not intended for long-term operation



Calibration pulser event waveform from 8 deep antennas in Testbed



- First ARA neutrino searches carried out with Testbed station data

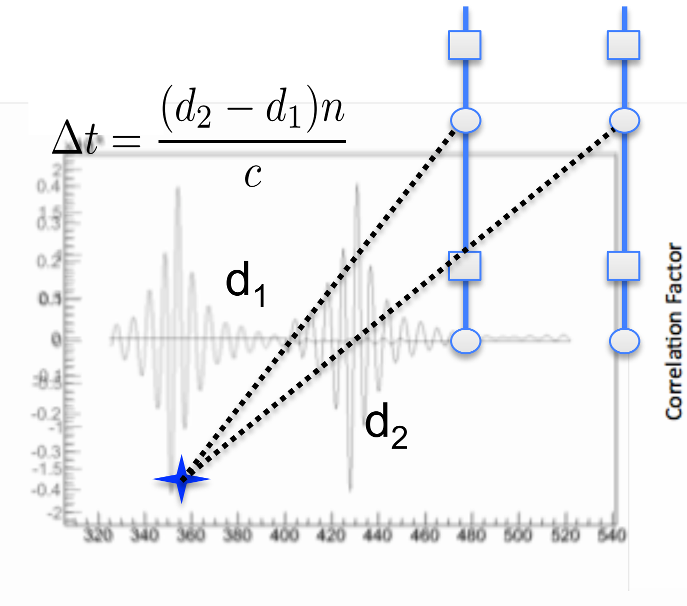
Diffuse: arxiv:1404.5285

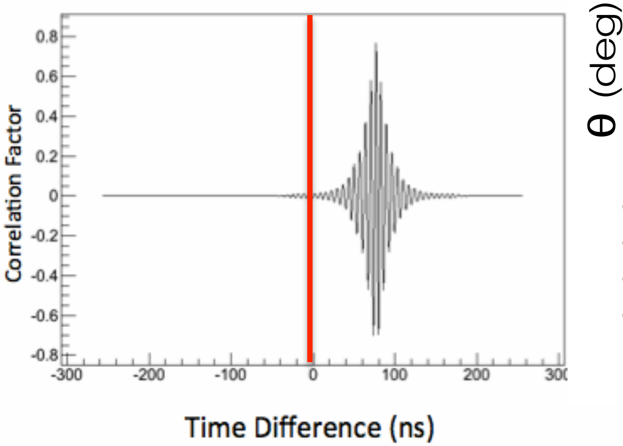
GRB: arxiv:1507.00100

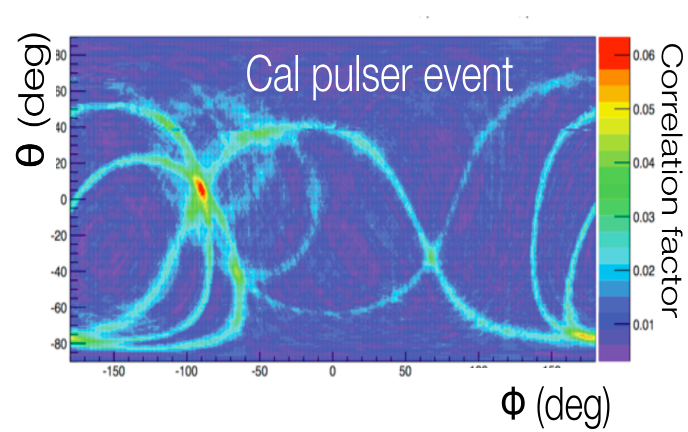
Testbed Analysis

Adapted interferometric technique from diffuse search for GRB search

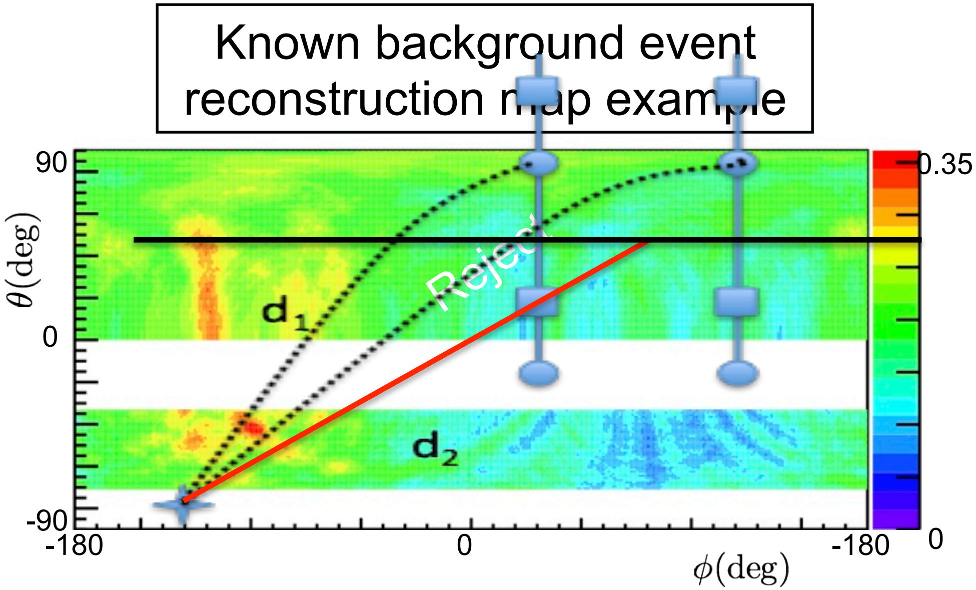
1. Impulsive waveform – ~1-10 ns time scale
2. Correlation factor - Convolution of the two waveforms including a timing offset
3. Calculate timing delays for all angles of approach
4. Sample correlation plot at these delays
5. Create a map for all pairs of antennas and the correlation

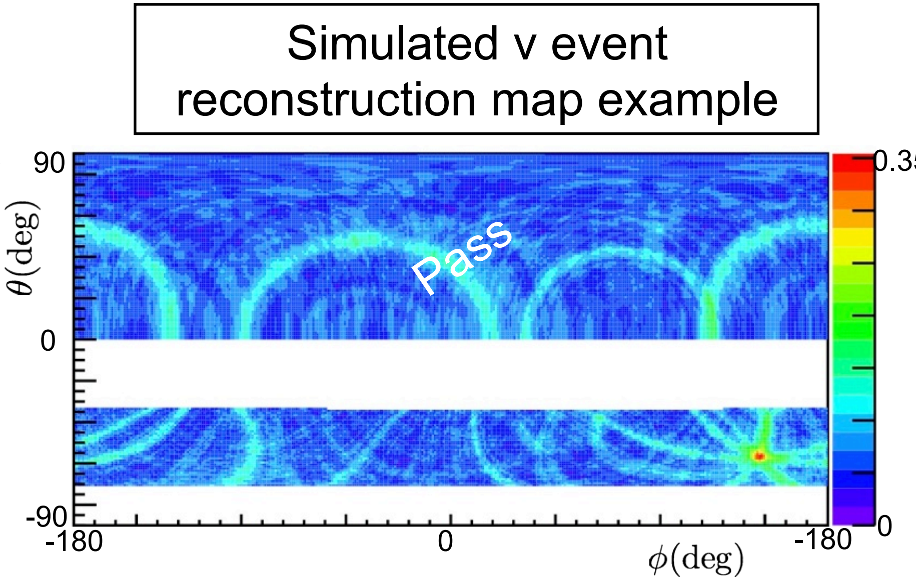






Reconstruction Quality Cut



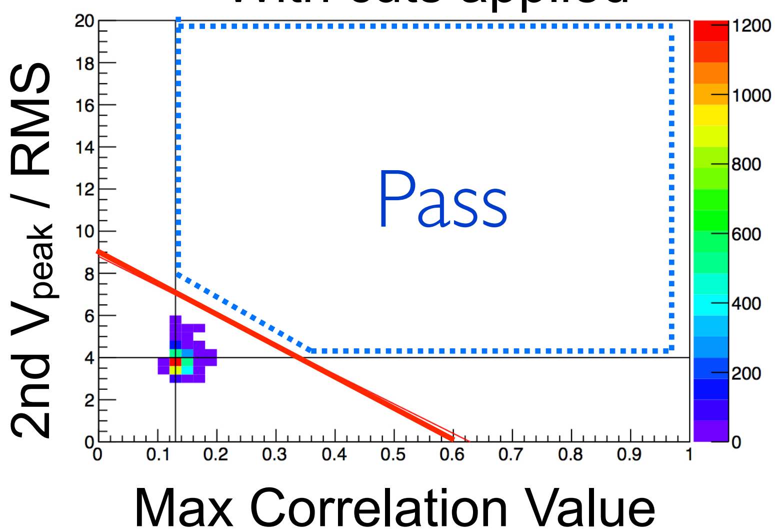


- Rejected thermal noise by requiring strong reconstruction map peak that is unique
- Reconstruction based on timing from ray-tracing
  - Use 30 m and 3 km maps in Hpol and Vpol
- Requires at least one reconstruction map to be of good quality
  - 1 deg^2 < Area of 85% contour surrounding the peak < 70 deg^2
  - Total 85% contour peak area < 16.2 x Area of 85% contour surrounding the peak
- Depending on the polarizations which pass the cut, the event is separated into Vpol and/or Hpol channels
- Rejects ~95% of noise-dominated events after initial quality cuts

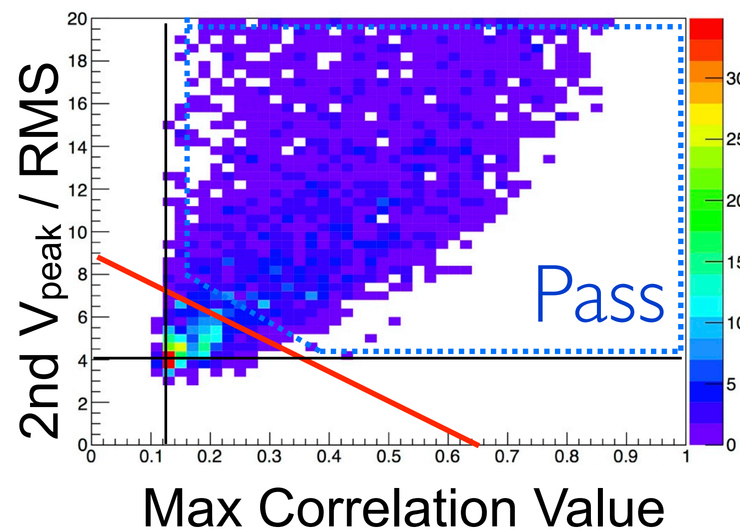
Peak/Correlation Cut

- Expect a correlation between signal strength from waveform and correlation value from reconstruction map for an impulsive event
- After removing known background events with other cuts, use this relation to get background estimation
  - Other cuts made: most reject specific anthropogenic signals

Testbed 10% data set
With cuts applied

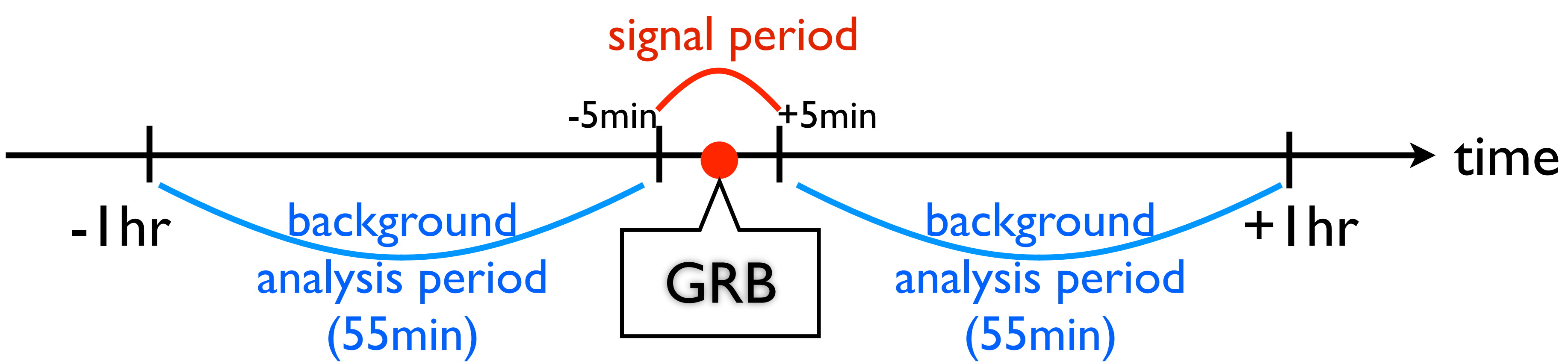


Simulated 10^{18} eV ν set
with cuts applied

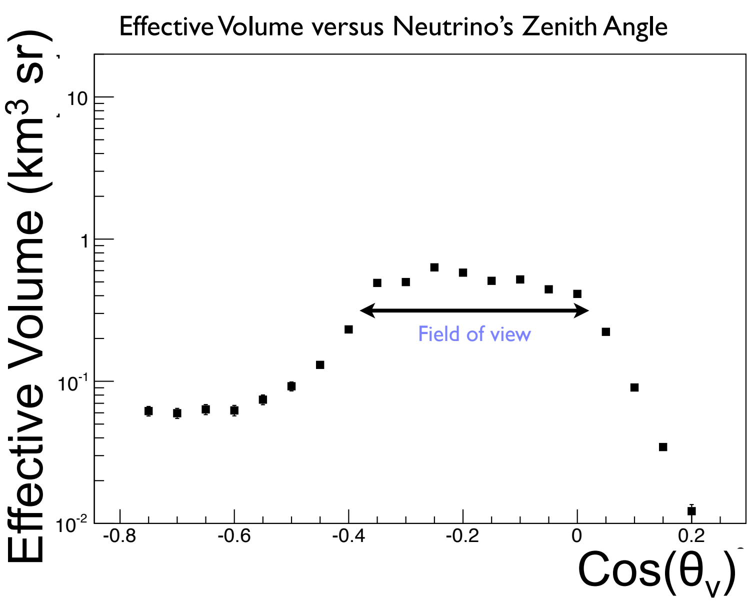


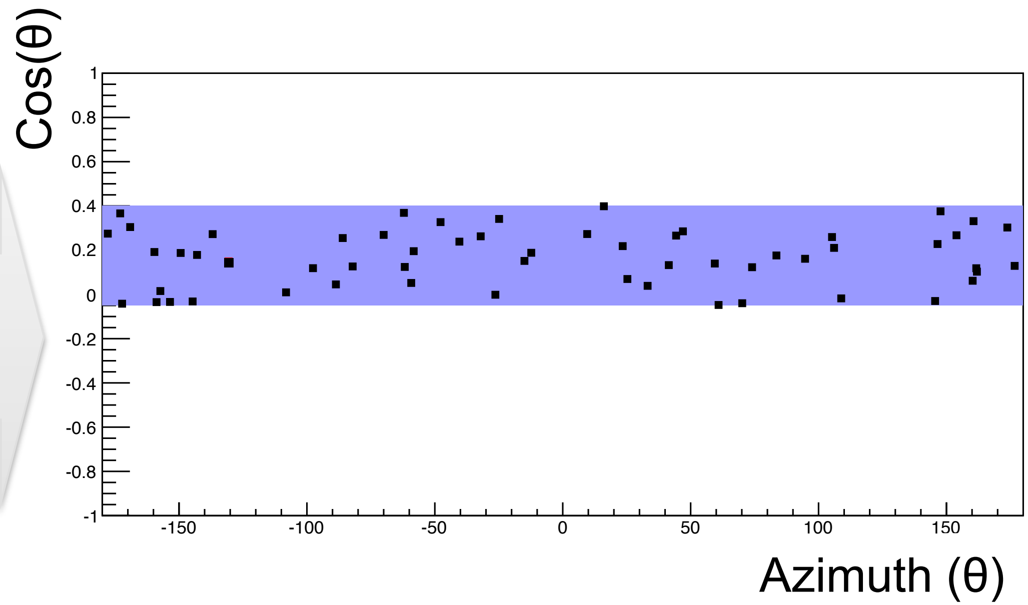
Testbed GRB analysis

- Adapt the above techniques from the Testbed diffuse neutrino search (arxiv:1404.5285) to search for events coincident with known Gamma Ray Bursts
  - Stricter requirements in time → relaxation of cut values
- 2 unblinding stages
  - Tune cuts on 10% of data in the background estimation window
  - 1: Check remaining 90% in background estimation window
  - 2: Signal search – 100% of data +/- 5 minutes around GRB event
  - Timing technique adapted from ANITA (arxiv: 1102.3206)

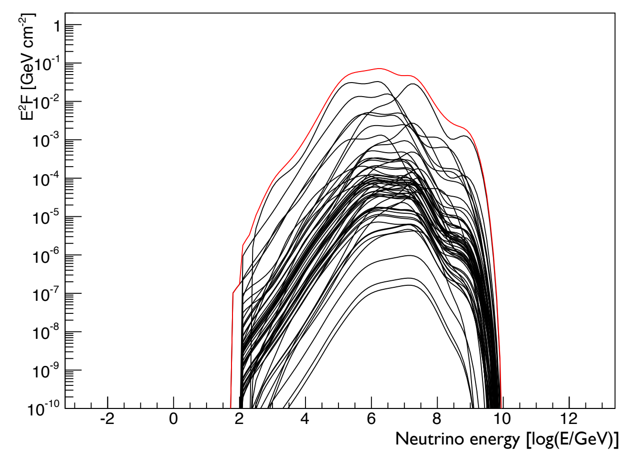


GRB Selection

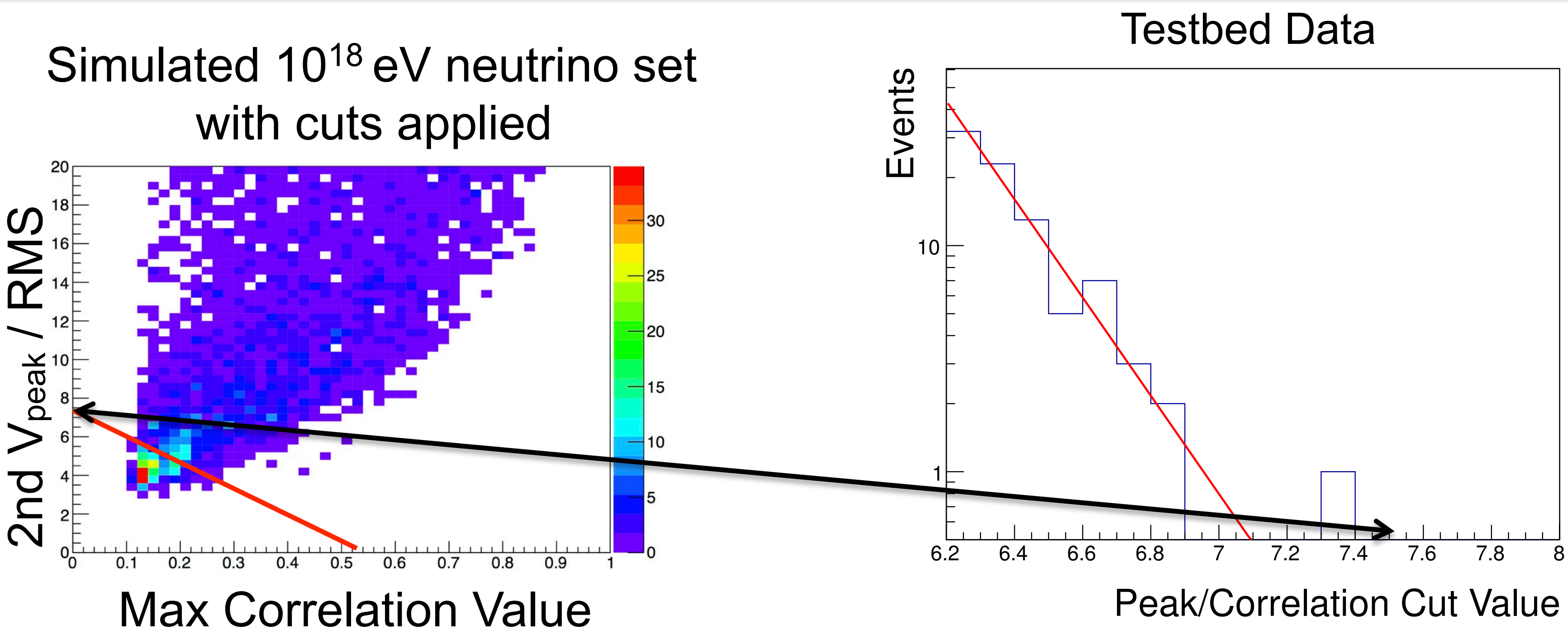




- Selected 57 GRBs based on livetime and geometric acceptance
- Get fluences for each GRB from NeuCosmA simulation and then total
- Tune cuts based on modeled neutrino fluence



Optimization



- Optimize the cut parameters:
  - Fit the background distribution with an exponential
  - Integrate extrapolation to get expected background
  - S_{upper} is the 90% confidence limit on the signal for an expected background
  - $N_{passed,sim}$ is the weighted number of passed simulated neutrinos from an expected flux
  - Maximize R to optimize for best limit

$$R = \frac{N_{\text{passed,sim}}}{S_{\text{upper}}}$$

Optimized Cut Values

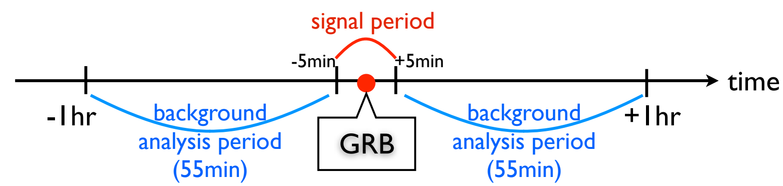
- All optimized cut parameters relaxed for GRB neutrino search when compared with diffuse neutrino search
- Factor of 2.4 improvement in efficiency against a simulated GRB flux
- Another cut for rejecting CW was removed

Cut	Reconstruction Quality Cut		Peak/Correlation Cut
Parameter	A_{peak}	A_{peak}/A_{total}	Peak/Correlation Cut Value
Diffuse Neutrino Search	50 deg^2	1.5	8.8
GRB Neutrino Search	70 deg^2	16.2	7.5

Preliminary Results

Stage 1 (90% background period unblinding):

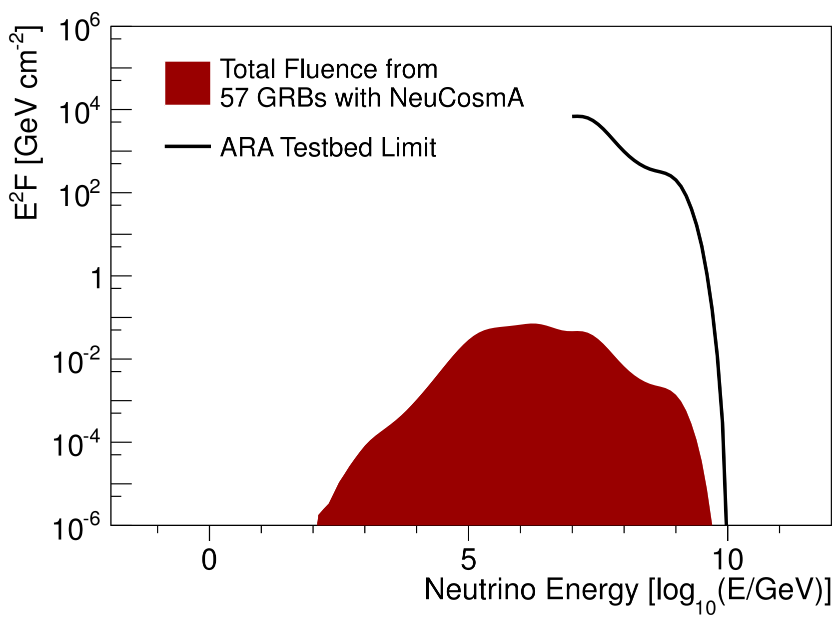
- Expected background events: 1.2
- 2 events survived

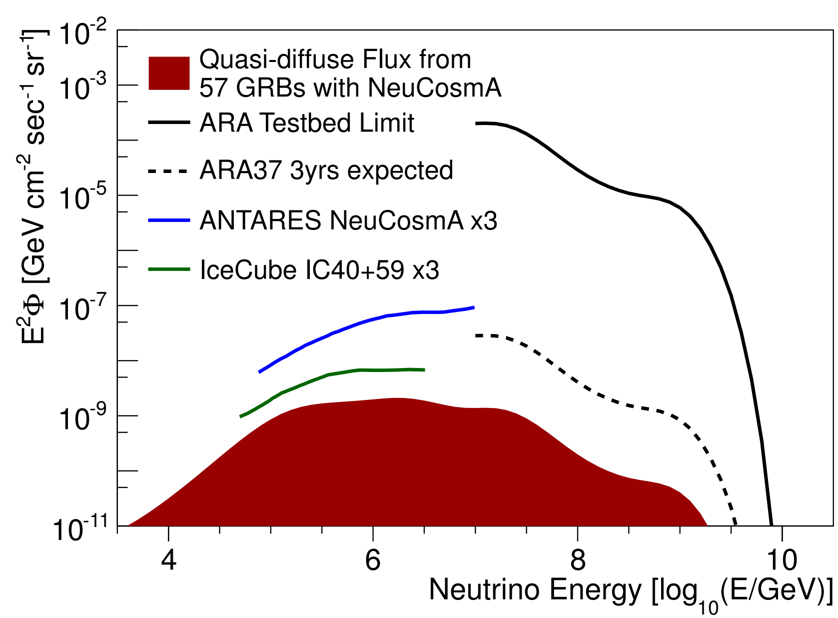


Stage 2 (signal period unblinding):

- Expected background: 0.12, Expected neutrinos: 1.7e-5
- 0 events survived

First quasi-diffuse flux limit above 10^{16} eV (arxiv:1507.00100)



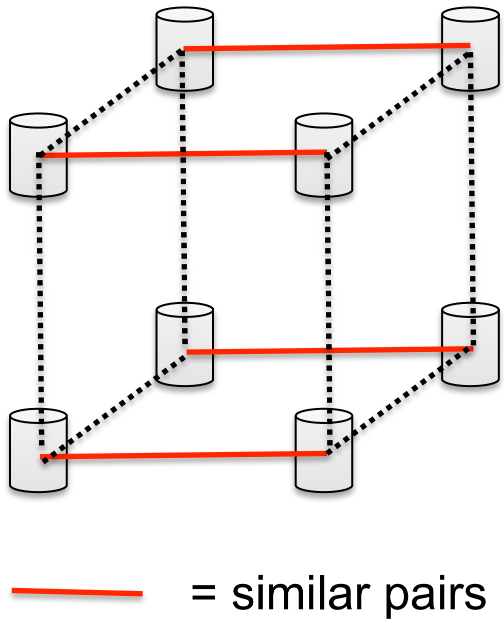


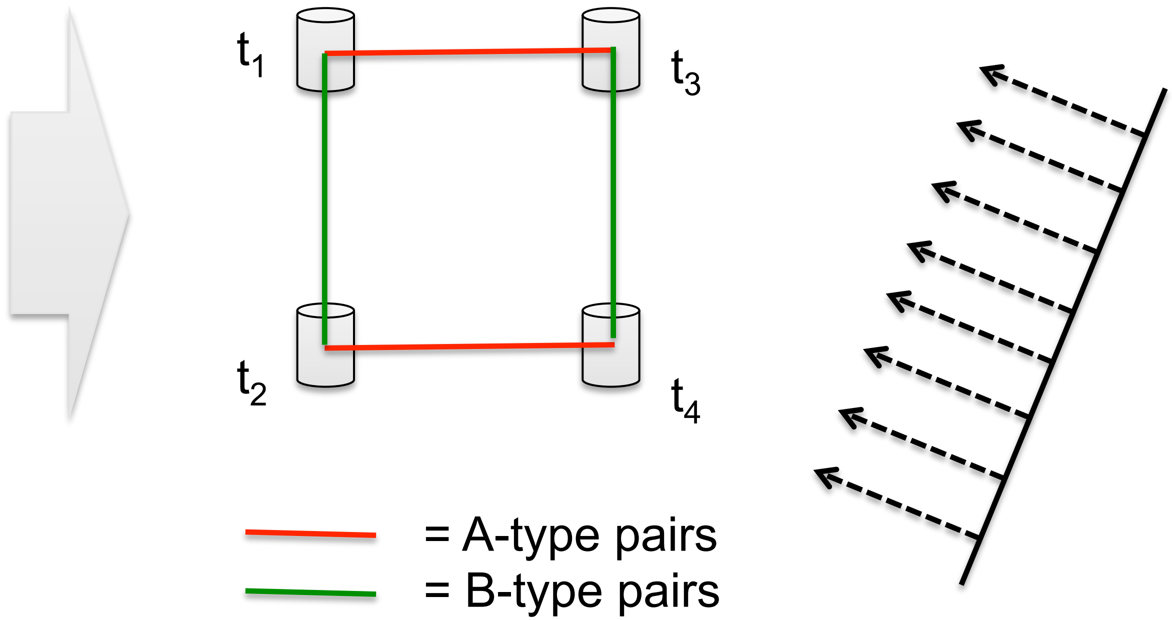
BACKGROUND REJECTION FOR A REGULAR ARRAY

Filter-level Algorithm

- 100's of millions of events – too many to efficiently use complex reconstruction methods
  - Need < 0.1% thermal acceptance to be efficient
- Can we create an adaptable, efficient filter-level algorithm
- Goals:
  - Computationally simple
  - Easily differentiates between signal and noise
  - Decrease volume of data to then use more computationally intensive techniques (ray-tracing, etc)
  - Single understandable output
  - Easily optimizable
- Ultimate goal is a deep station analysis of current data
  - Perhaps use algorithm as a trigger or filter to the North?

Planar Signal Wavefront





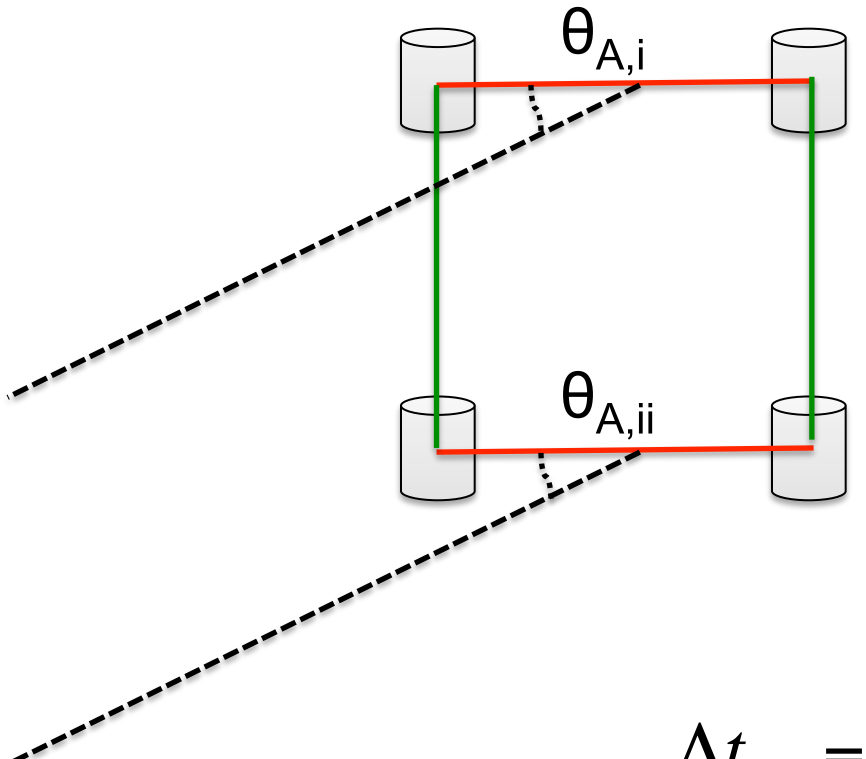
$$\Delta t_{A,i} = t_3 - t_1$$

$$\Delta t_{A,ii} = t_4 - t_2$$

$$\Delta t_{A,i} \approx \Delta t_{A,ii}$$

- Divide array into faces
- Difficult to directly compare timing from different sets of pair-types – what to do?

Angle of Incidence



- Use the angle from the baseline
- Comparable between different pair types

$$\theta_{A,i} \approx \theta_{A,ii} \qquad \cos\left(\theta_{A,i}\right) \approx \cos\left(\theta_{A,ii}\right)$$

$$\Delta t_{A,i} = \frac{n}{c}\cos\left(\theta_{A,i}\right)\Delta d_{A,i}$$

$$\cos\left(\theta_{A,i}\right) = \frac{c\Delta t_{A,i}}{n\Delta d_{A,i}}$$

Angular Variation - RMS

- Similar time differences → small variation
  - Find the “RMS” around their average

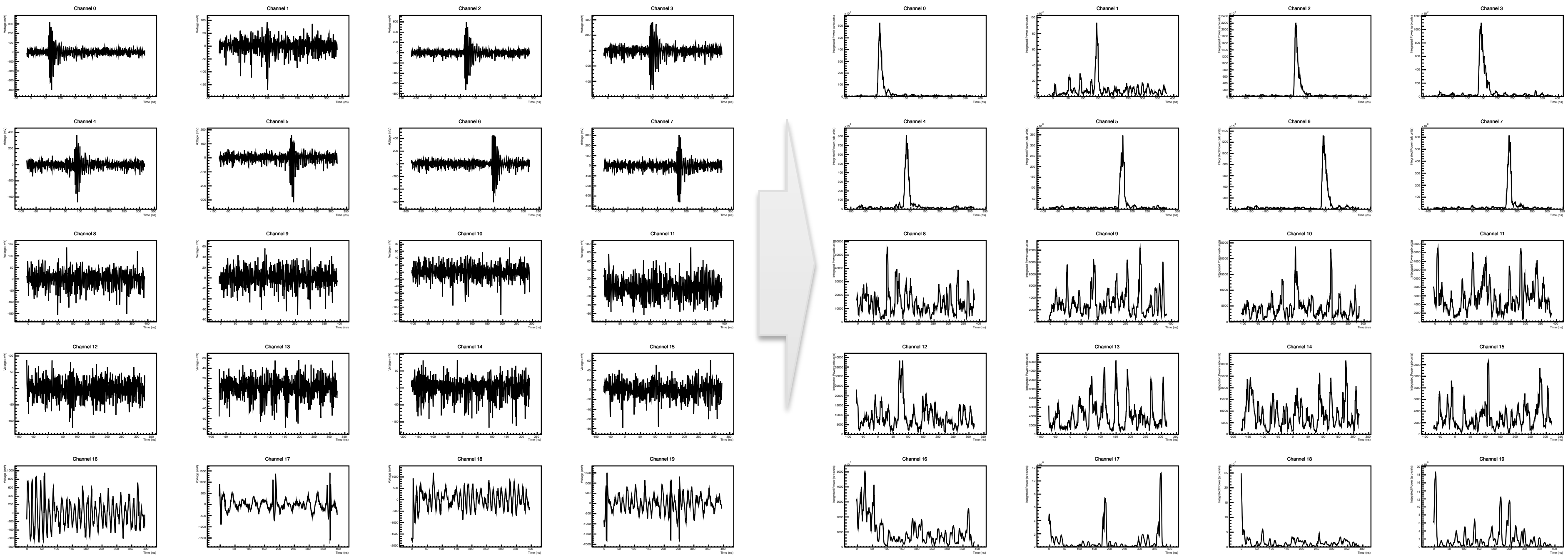
$$\overline{\cos(\theta_A)} = \frac{\cos(\theta_{A,i}) + \cos(\theta_{A,ii})}{2}$$

$$\mathrm{RMS}(\cos(\theta_A)) = \sqrt{\frac{\left(\cos(\theta_{A,i}) - \overline{\cos(\theta_A)}\right)^2 + \left(\cos(\theta_{A,ii}) - \overline{\cos(\theta_A)}\right)^2}{2}}$$

- RMS(cos(θ)) < 0.1 if the arrival directions agree
- Also corrects for differences in baseline lengths

Find “hit times”

Calibration pulser event



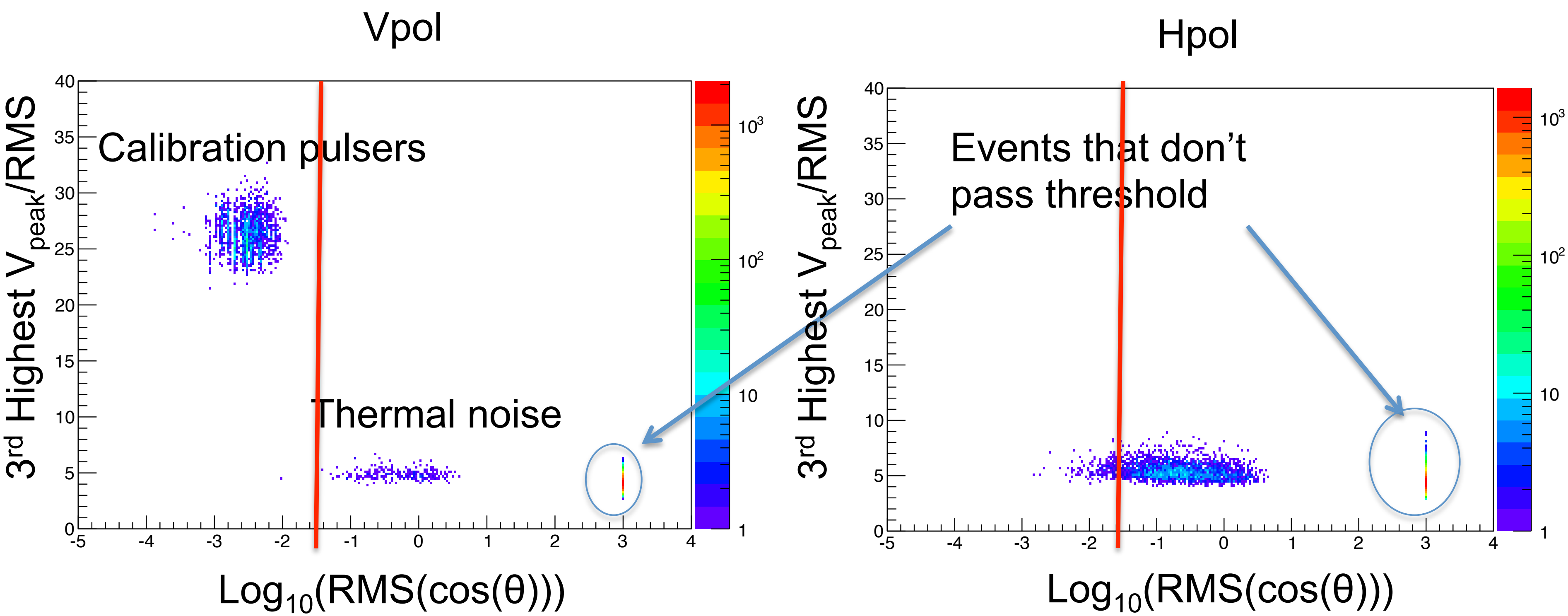
- To decrease noise fluctuations, scan an integrated power window of 5 ns
- Find the two highest peaks, use these as “hit times” for that channel
- Apply a threshold:

$$\frac{\mathrm{RMS}(5\text{ ns around the peak})}{\mathrm{RMS}(\mathrm{waveform})} > \mathrm{Threshold}$$

- Find the face with the timing that agrees best with incoming signal (lowest face RMS)

Preliminary Results - Data

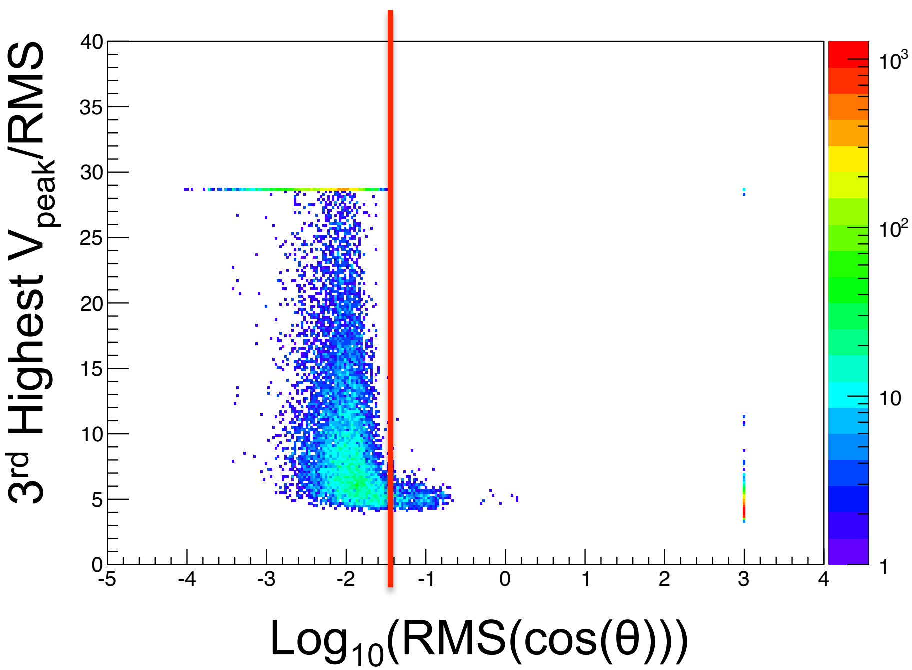
Station A2, Run 1798
Threshold = 2.5



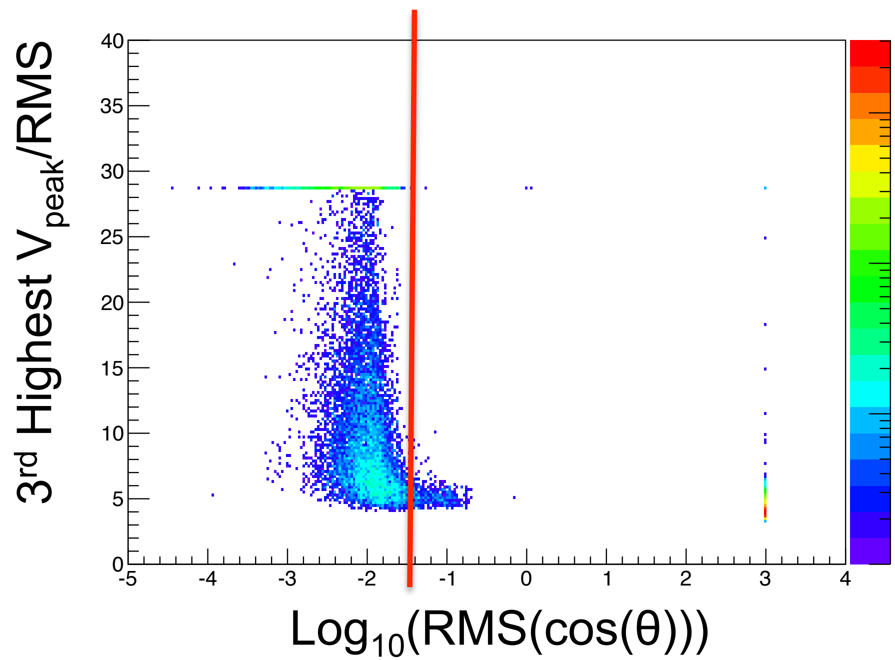
- More event pass threshold in Hpol antennas
  - use separate thresholds for Vpol and Hpol

Preliminary Results - Simulation

Vpol

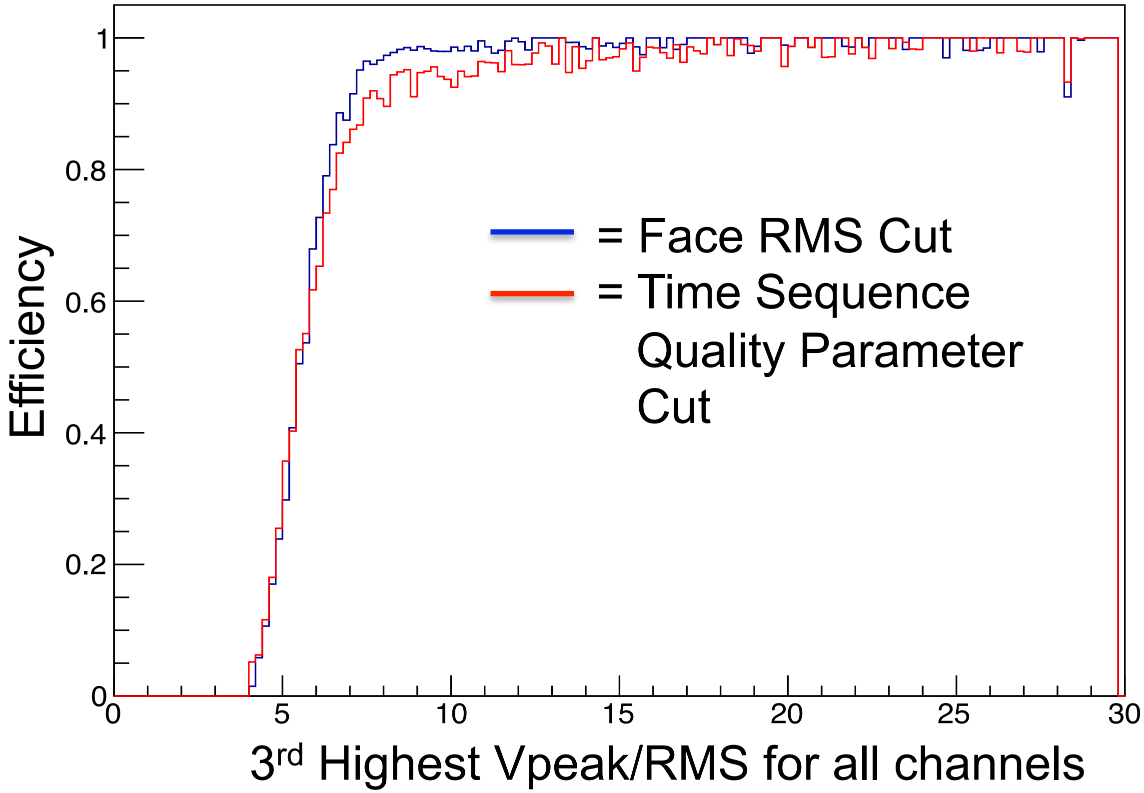


Hpol



- Simulated 10^{19} eV neutrino events generated with AraSim simulation package
- Good separation at high signal strength
- Reasonable separation at lower signal strength
- Noise starts to dominate over low SNR signals – difficult to reconstruct anyway

Efficiency



Vpol thresh = 2.5
Hpol thresh = 2.9
$Log_{10}(RMS(cos(\theta))) < -1.5$

Time Sequence Quality Parameter > 0.6
(value used in 2013 A2/3 analysis, described by Kael in earlier talk)

- Data RF events - Face RMS efficiency = 0.08 %, TSQP = 0.08 %
- Simulation - Face RMS efficiency = 83.1%, TSQP efficiency = 81.6%
- Currently filter algorithms comparable
- Face RMS not optimized, may improve even more

Conclusions

- Testbed GRB neutrino search
  - Optimized search cuts
  - Limiting background search window → cut relaxation
  - New quasi-diffuse flux limit above 10^{16} eV
  - Projected limit for ARA37

- New filter-level cut
  - Efficient in rejecting thermal noise – 0.08% acceptance
  - Efficient in retaining simulated neutrinos
    - > 95% at high SNR
  - Flexible
    - Can characterize individual faces separately
    - Can treat hpol and vpol separately
  - Can improve event selection at the analysis level and maybe even the trigger level
  - Will optimize cut in full analysis (later this year!)

CHEAPR Workshop

Computing in High-Energy Astro-Particle Research

Topics: Genetic programming, analytics, data analysis, feature selection, high-performance computing

Activities: tutorials, lectures, example code packages

Who: Members of ANITA, ARA, LIGO, SKA, others

Experts in genetic programming from industry and academia

When: August 24th – 26th, 2016

Where: Center for Cosmology and AstroParticle Physics (CCAPP), The Ohio State University

Contact: Carl Pfendner pfendner.1@osu.edu or Jordan Hanson hanson.369@osu.edu

Questions?