Friedrich-Alexander-Universität
Naturwissenschaftliche Fakultät

Signal Identification & Reconstruction using Correlation

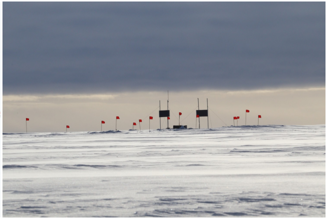
Sjoerd (👞rd) Bouma

April 5, 2024

ECAP

Who am I?

- Sjoerd ≈ 👞rd
- PhD student at Erlangen Centre for Astroparticle Physics (ECAP)
- Part of the Radio Neutrino Observatory in Greenland (**RNO-G**)
- My work: **reconstruction** of (simulated) neutrino signals



This talk

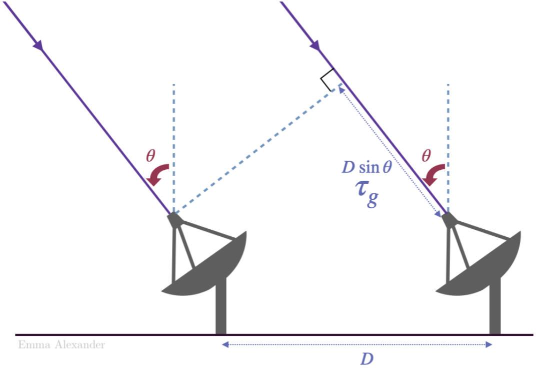
Two common challenges in time-series data:

1. **Identifying** a signal within a noisy time trace of recorded data
2. **Reconstructing** the arrival direction of the signal

- Mostly focus on (2.), though there is often some overlap between both.
- Most examples taken from radio astronomy
- I'm not necessarily an expert! Very interested to hear how **you** tackle similar problems in your experiment/field.

Problem sketch

- N antennas measure a signal coming from somewhere
- How can we tell a signal has been measured?
- How do we know where the signal came from?



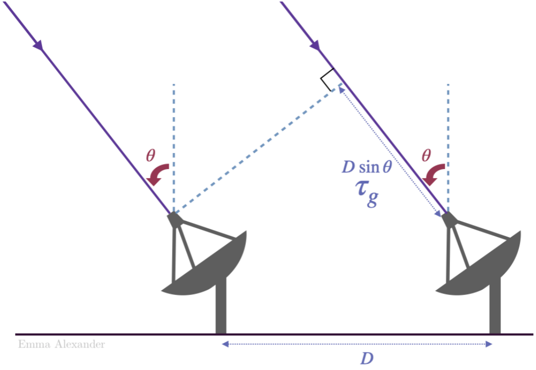
source

Interferometry

- One option is to use **interferometry**
- The signal arrives at antennas separated by $\Delta\vec{x}$ with a relative time delay:

$$\Delta t = \vec{v} \cdot \Delta\vec{x},$$

- Assuming some direction $\vec{v}$, we can sum the signal in different antennas:
  - Signal sums coherently $\rightarrow \times N$
  - Noise is incoherent $\rightarrow \times\sqrt{N}$
- $\Rightarrow$ SNR increases with factor $\sqrt{N}$



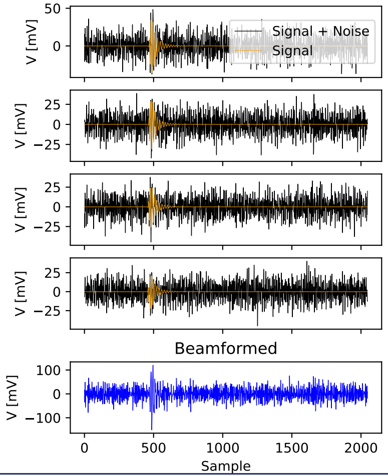
source

Interferometry

- One option is to use **interferometry**
- The signal arrives at antennas separated by $\Delta\vec{x}$ with a relative time delay:

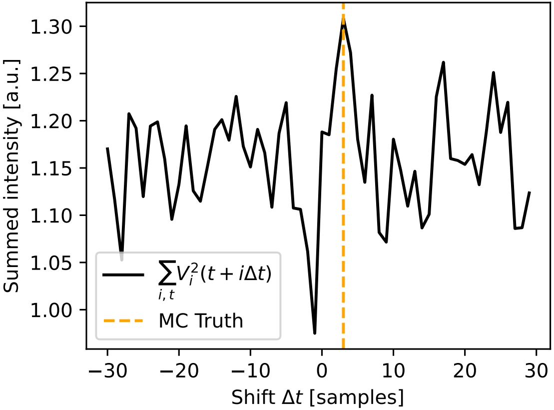
$$\Delta t = \vec{v} \cdot \Delta\vec{x},$$

- Assuming some direction $\vec{v}$, we can sum the signal in different antennas:
  - Signal sums coherently $\rightarrow \times N$
  - Noise is incoherent $\rightarrow \times\sqrt{N}$
- $\Rightarrow$ SNR increases with factor $\sqrt{N}$



Interferometry

- Instead of using a known direction to improve SNR, can also invert this:
- Scan over possible directions and pick Δt that maximizes the amplitude.
- This gives a handle on the arrival direction of the signal



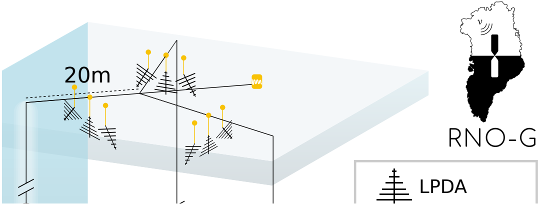
Interferometry

Advantages:

- Don't need an explicit signal model
- Computationally very cheap

Disadvantages:

- To maintain coherence, need a timing resolution better than $1/(2f_{\mathrm{max}})$ (though we can still do intensity interferometry for time-varying signals even if this is not satisfied)
- Signal needs to look similar in all antennas
- Need to correct for non-uniform antenna gain pattern



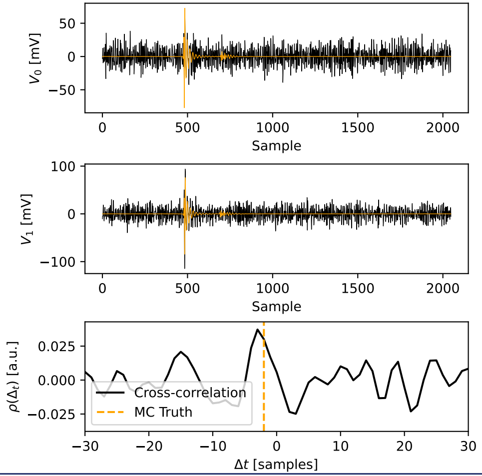
Shallow component of RNO-G detector

Cross-correlation

- An alternative approach is cross-correlation:

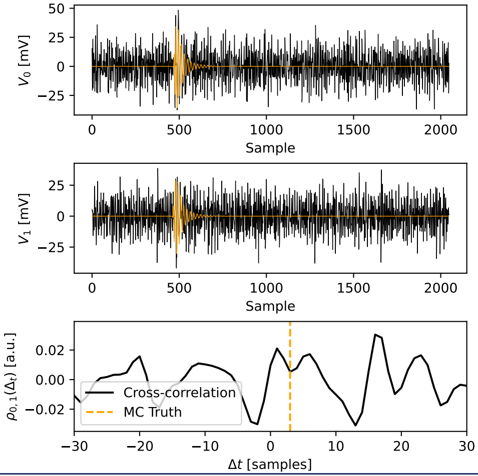
$$\rho_{i,j}(\Delta t) = \sum_t V_i(t) V_j(t + \Delta t)$$

- If both V_i and V_j contain the same signal shifted by $\Delta t'$, expect a peak in $\rho_{i,j}$ when $\Delta t = \Delta t'$
- Often normalize by dividing by $\sigma_i \sigma_j$ such that $-1 \leq \rho_{i,j} \leq 1$.



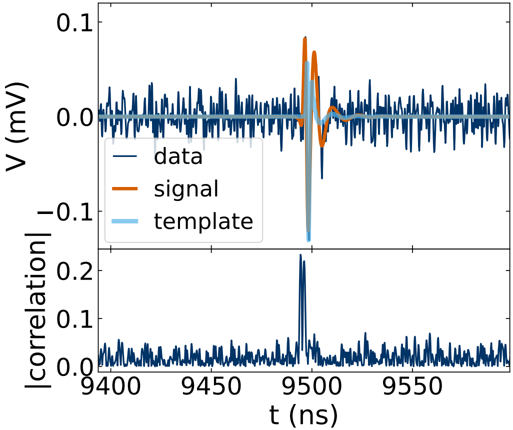
Cross-correlation

- If both voltages contain noise, we are mostly correlating 'noise with noise'
- Can improve on this somewhat by **bandpass-filtering**, i.e. filtering out frequencies with mostly noise contributions
- However, reducing bandwidth also increases maximum correlation of noise with noise.



Template correlation

- Better - correlating noisy data with noiseless **template**
- Need a template that describes the signal (relatively) well
- Can use multiple templates to (try to) account for variation in the signal.



Template correlation

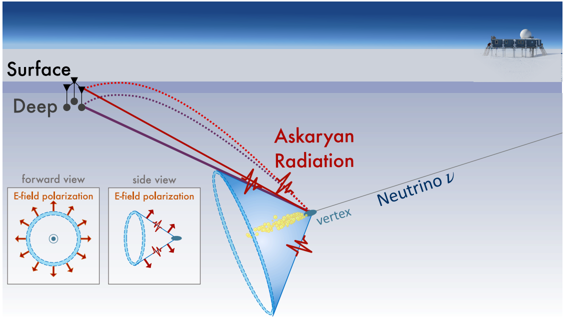
- Template correlation in the time-domain works well for approximately gaussian ($\approx$ white) noise
- If the noise spectrum $S_n(f)$ is not white, can use correlation in frequency domain instead (e.g. gravitational wave template searches):

$$\rho(t) = 4Re \int df \frac{\tilde{V}_i(f)\tilde{h}^*(f)}{S_n(f)} e^{2\pi ift} df$$

Neutrino reconstruction

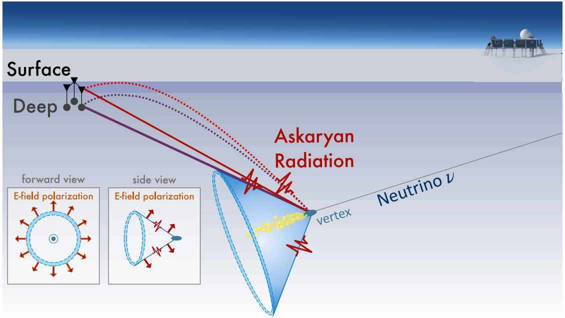
Radio Neutrino Detection

- In-ice shower initiated by UHE neutrino develops a negative charge excess at the shower front, giving rise to **Askaryan radiation**.
- At radio wavelengths ($\mathcal{O}(100 - 1000)$ MHz), **coherent** emission close to **Cherenkov angle** ($\sim 56^\circ$)
- At energies > 10 PeV, strong enough to detect at $\mathcal{O}(1)$ km distances - in-ice radio detector for neutrinos!
- e.g. **RNO-G** in Greenland; ARIANNA, ARA, **IceCube-Gen2 (?)** in Antarctica



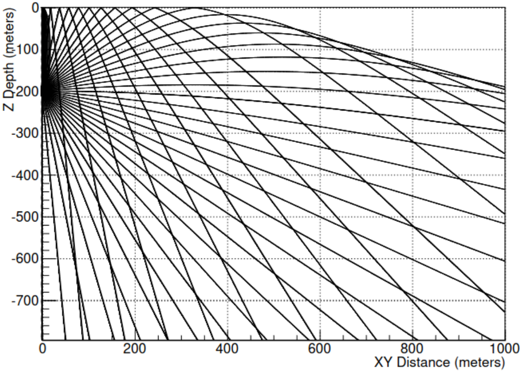
Vertex reconstruction

- The **first** step in reconstructing the neutrino is finding the source of the emission: the **neutrino interaction vertex**
- Use template correlation
- Challenges:
  - **Ice** - refractive index changes ⇒ radio waves 'bend downwards'.



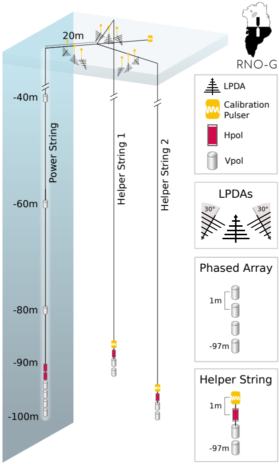
Vertex reconstruction

- The **first** step in reconstructing the neutrino is finding the source of the emission: the **neutrino interaction vertex**
- Use template correlation
- Challenges:
  - **Ice** - refractive index changes ⇒ radio waves 'bend downwards'.
  - This leads to a 'shadow zone'.



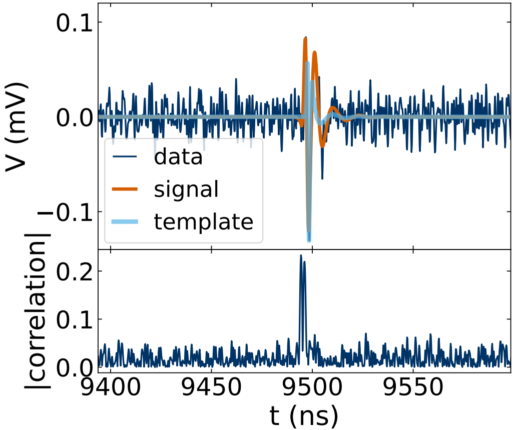
Vertex reconstruction

- The **first** step in reconstructing the neutrino is finding the source of the emission: the **neutrino interaction vertex**
- Currently one of the dominant limitations for neutrino reconstruction (2302.00054)
- Use template correlation
- Challenges:
  - **Ice** - refractive index changes ⇒ radio waves 'bend downwards'.
  - This leads to a 'shadow zone'.
  - Signal not visible in all antennas!



Vertex reconstruction

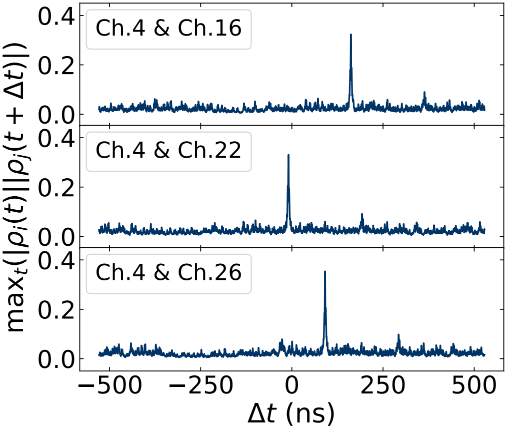
- Use a single template, compute correlation $\rho_i(t)$ for each channel i;
- Take absolute value (to account for polarity)



Vertex reconstruction

- Use a single template, compute correlation $\rho_i(t)$ for each channel i;
- Take absolute value (to account for polarity)
- To determine Δt, multiply them together:

$$\rho_{i,j}(\Delta t) = \max_t(|\rho_i(t)||\rho_j(t+\Delta t)|)$$

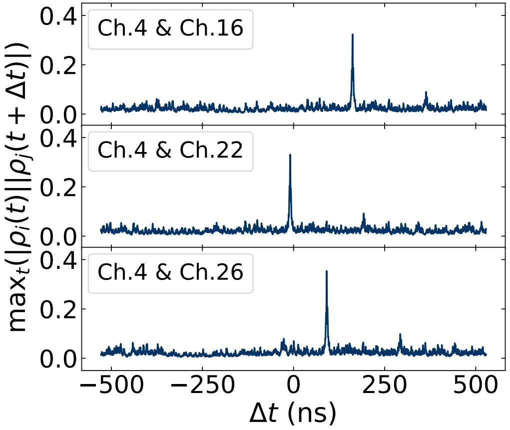


Vertex reconstruction

- Use a single template, compute correlation $\rho_i(t)$ for each channel i;
- Take absolute value (to account for polarity)
- To determine Δt, multiply them together:

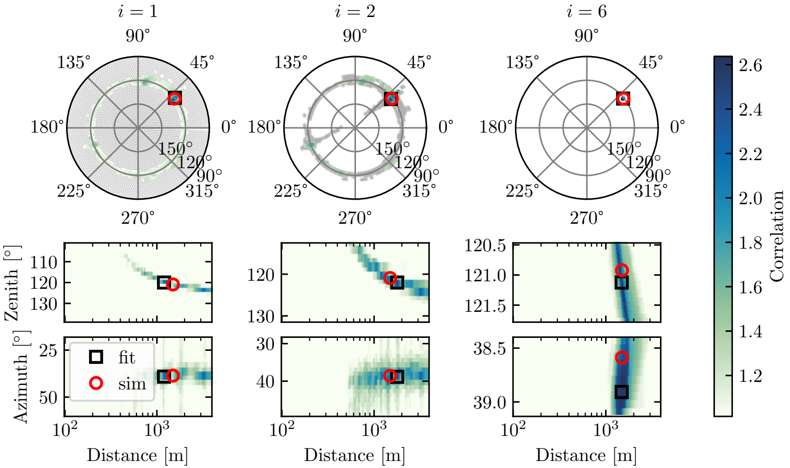
$$\rho_{i,j}(\Delta t) = \max_t(|\rho_i(t)||\rho_j(t+\Delta t)|)$$

- Finally, use a lookup table to convert a vertex position $\vec{x}$ to expected time delays Δt



Vertex reconstruction

- Fit $\vec{x}$ by maximizing total correlation over all antenna pairs i, j
- To avoid local minima, use an iteratively refined brute force search.



Vertex reconstruction

- This works well at high enough SNR, and if the signal is visible in all antennas
- At low SNR, this algorithm will **bias** towards vertex position visible in all antennas (because *some* $|\rho|$ is more than *no* $|\rho|$)

$\Rightarrow$ need to account for the possibility of 'no signal'

Vertex reconstruction

- This works well at high enough SNR, and if the signal is visible in all antennas
- At low SNR, this algorithm will **bias** towards vertex position visible in all antennas (because *some* $|\rho|$ is more than *no* $|\rho|$)

$\Rightarrow$ need to account for the possibility of 'no signal'

- Current strategy: median-subtraction:

$$\rho'_{i,j}(\Delta t) = \max\left\{0, \rho_{i,j}(\Delta t) - \tilde{\rho}_{i,j}\right\}$$

Vertex reconstruction

- This works well at high enough SNR, and if the signal is visible in all antennas
- At low SNR, this algorithm will **bias** towards vertex position visible in all antennas (because *some* $|\rho|$ is more than *no* $|\rho|$)

$\Rightarrow$ need to account for the possibility of 'no signal'

- Current strategy: median-subtraction:

$$\rho'_{i,j}(\Delta t) = \max\left\{0, \rho_{i,j}(\Delta t) - \tilde{\rho}_{i,j}\right\}$$

- **Question**: can we do something better?
- E.g. minimum correlation threshold for inclusion in fit, machine learning magic, ...?

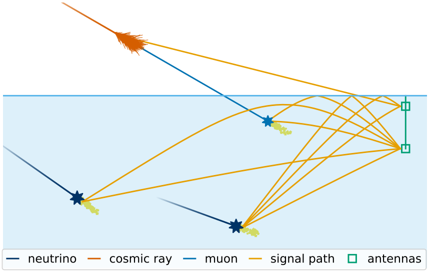
Conclusions

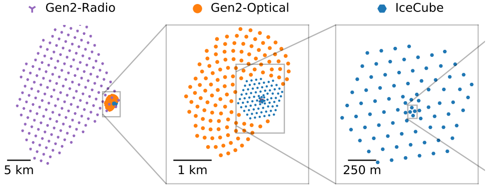
- Discussed two ways to use multiple antennas to find the source of a signal:
  1. **Interferometry** (beamforming): coherently combining signal from N antennas increases SNR by $\sqrt{N}$;
  2. **Cross-correlation**: using a **template** can identify a signal even at low SNR
- For neutrino reconstruction, use template correlation
- This is one of the dominant limits for radio neutrino reconstruction
  - Not all antennas have signal - reconstruction bias ('finding signal where there is none')
- **Question to the audience**: how do you deal with this problem (source reconstruction, signal identification) or these techniques (interferometry, correlation) in your experiments?

Backup

Recap: radio neutrinos

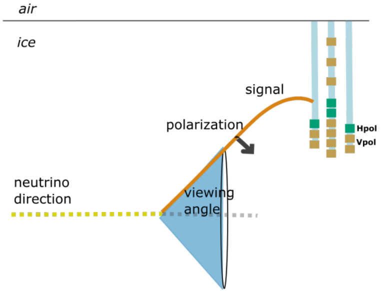
- In-ice shower initiated by UHE neutrino develops a negative charge excess at the shower front, giving rise to **Askaryan radiation**.
- At radio wavelengths ($\mathcal{O}(100-1000)$ MHz), **coherent** emission close to **Cherenkov angle** ($\sim 56^\circ$)
- At energies > 10 PeV, strong enough to detect at $\mathcal{O}(1)$ km distances - in-ice radio detector for neutrinos!
- e.g. **RNO-G** in Greenland; ARIANNA, ARA, **IceCube-Gen2 (?)** in Antarctica





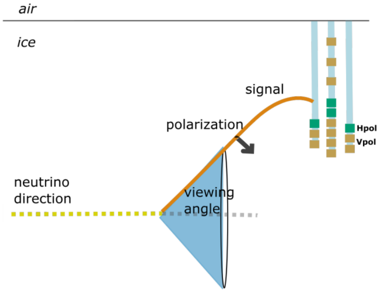
Direction reconstruction: the principle

- Three steps:
  1. **Signal direction** - direction of **emission** at the shower vertex
  2. **Viewing angle** - angle between the neutrino and the emitted signal
  3. **Polarization** - points **towards the shower axis**



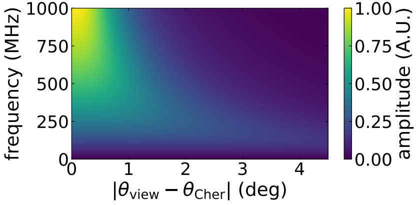
Direction reconstruction: the principle

– Three steps:
    1. **Signal direction**: from 'triangulation'



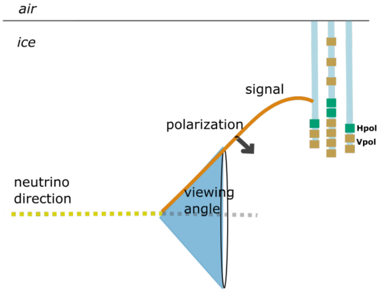
Direction reconstruction: the principle

- Three steps:
  1. **Signal direction**: from 'triangulation'
  2. **Viewing angle**: from shape of spectrum - the emission **loses coherence** further from the Cherenkov angle, with the higher frequencies losing coherence first.



Direction reconstruction: the principle

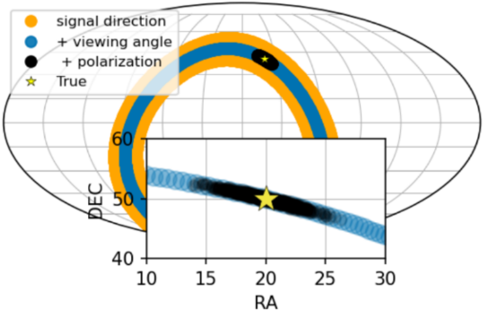
- Three steps:
  1. **Signal direction**: from 'triangulation'
  2. **Viewing angle**: from shape of spectrum - the emission **loses coherence** further from the Cherenkov angle, with the higher frequencies losing coherence first.
  3. **Polarization**: from different antennas ('Vpol' and 'Hpol')



Direction reconstruction: the principle

This is what it looks like...

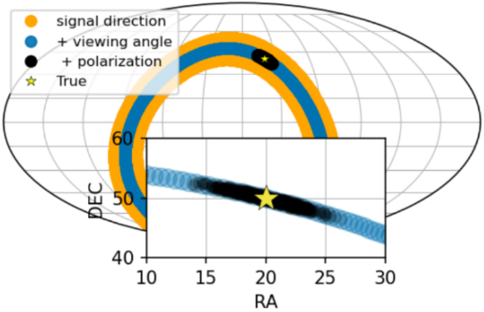
- ...for a **single neutrino**: a small 'ellipse' on-sky.



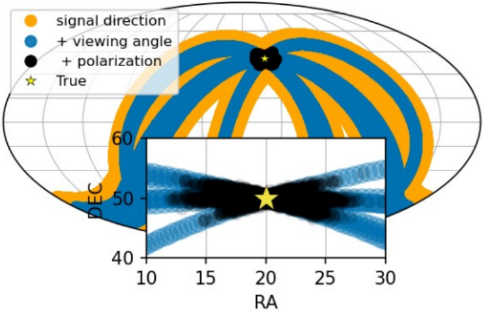
Direction reconstruction: the principle

This is what it looks like...

- ...for a **single neutrino**: a small 'ellipse' on-sky.



- ...for a **source** with multiple neutrinos detected ('point spread function').



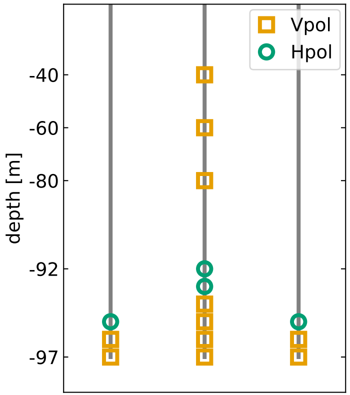
ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

Performance

Simulation setup

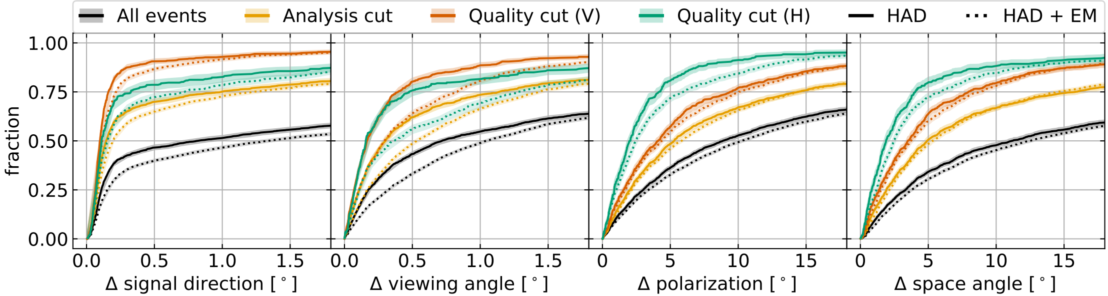
Test case:

- IceCube-like flux + GZK
- RNO-G-like detector:
  - Three strings on a triangular grid
  - Trigger (phased array of 4 Vpols) and Hpol antennas at ∼100 m to maximize sensitivity
  - 3 additional upper Vpols for increased baselines
- Include both hadronic and electromagnetic showers
  - Electromagnetic showers at ultra-high energies more irregular (LPM effect) - harder to fit, & algorithm designed for hadronic showers.



Results

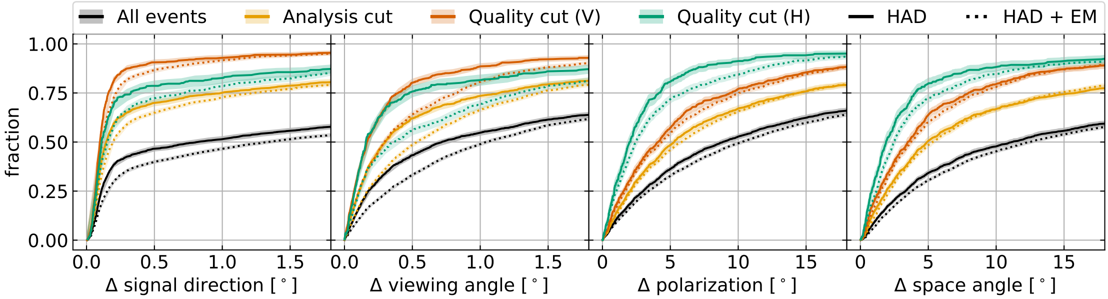
Analysis cut	SNR$>$ 2.5 cut in lower Vpols
Quality cut (V)	SNR$>$ 3.5 in upper Vpols
Quality cut (H)	SNR$>$ 3 in any Hpol



Results

1. Signal direction (vertex reconstruction) limits successful reconstructions

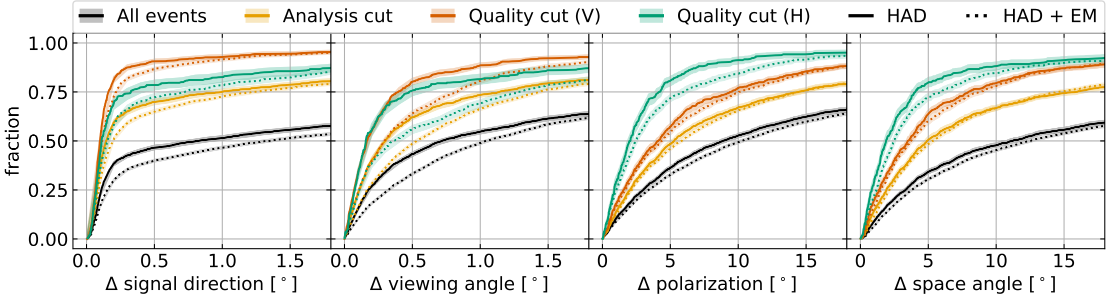
– Mostly (but not exclusively) at low SNR, failure to reconstruct the shower maximum results in 'bad' overall reconstruction.



Results

2. Polarization resolution is the dominant uncertainty

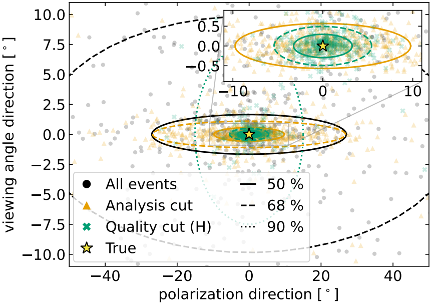
– Larger phase space & relatively less sensitive Hpol antennas lead polarization to dominate the angular uncertainty.



Results

3. Uncertainty contours are strongly asymmetric

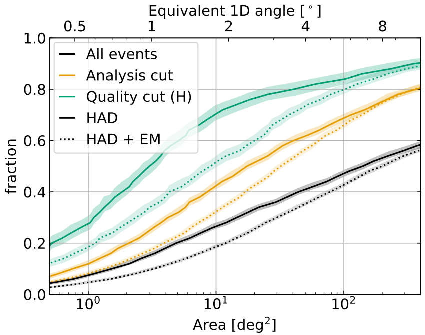
- Dominant polarization uncertainty results in elongated ellipses.
- This means the 1D 'space angle' strongly overestimates the actual uncertainty!



Results

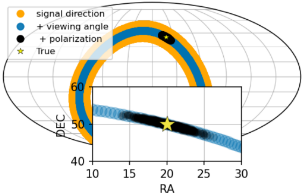
3. Uncertainty contours are strongly asymmetric

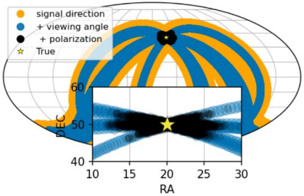
- Dominant polarization uncertainty results in elongated ellipses.
- This means the 1D 'space angle' strongly overestimates the actual uncertainty!
- E.g. median resolution for HAD, analysis cut: 4.9° (space angle) vs. 17 $\mathrm{deg}^2 \approx 2.4°$ 1D-equivalent.



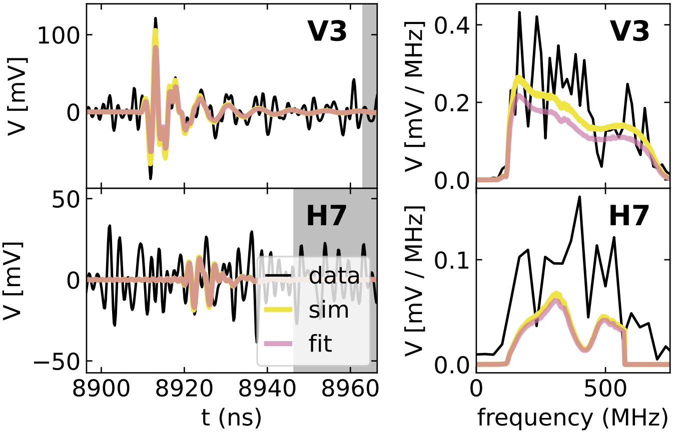
Conclusions

1. We **can reconstruct neutrinos** with a deep in-ice radio detector! (Now we just need to find some...)
2. Resolution limited by **vertex** and **polarization** reconstruction
3. Uncertainty contours are asymmetric - **can not just quote a space angle**!
   - Single event - **ellipse**
   - Point spread function - **bow tie**
4. Improvements expected!
   - Improve vertex reconstruction by better pulse finding at low SNR?
   - Dedicated algorithm for electromagnetic showers?
   - Machine learning?
   - ...

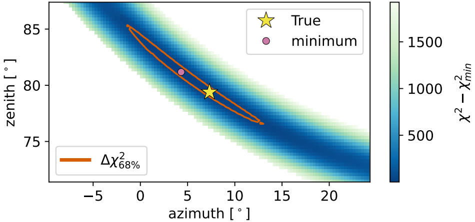


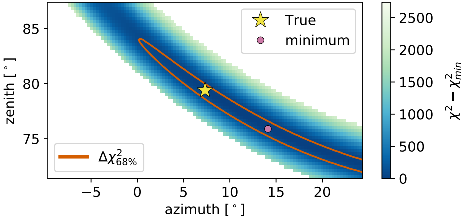


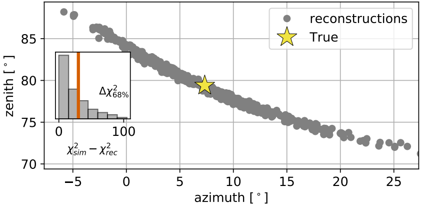
Example reconstruction

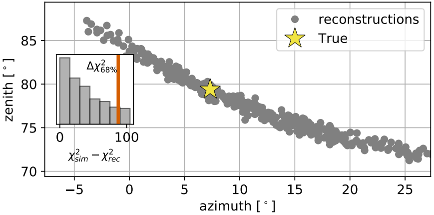


Systematic uncertainties

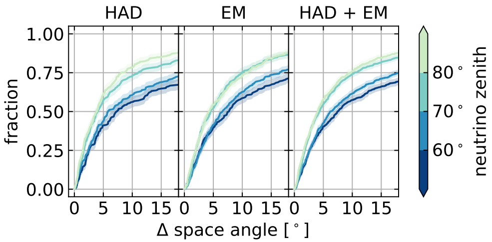


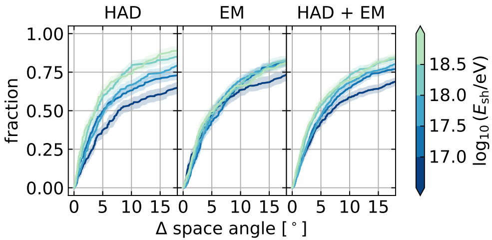






Zenith and energy dependence

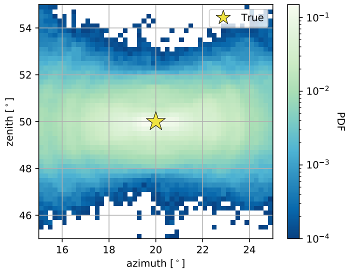




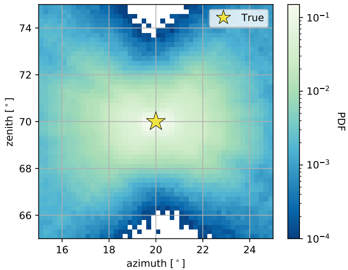
Point spread function

- Shape of the PSF depends on local zenith
- Orientation of the polarization direction geometrically constrained → **bow-tie** shape
- Area larger than single event contour, but smaller than for a symmetric PSF

$\theta = 50^\circ$

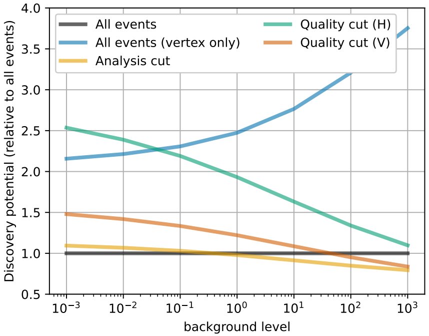


$\theta = 70^\circ$



Discovery potential

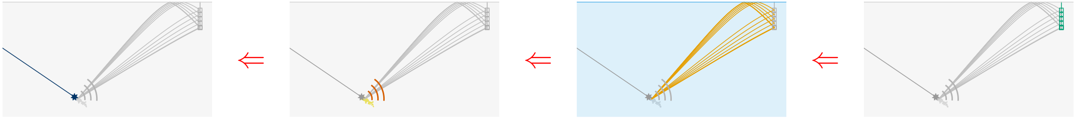
- Can study the source discovery potential for a source at a declination of 20°
- Shown normalized to 'all events' - **lower is better**
- At $\leq$ expected background flux, number of events detected is much more important than resolution.



The algorithm

The forward-folding approach

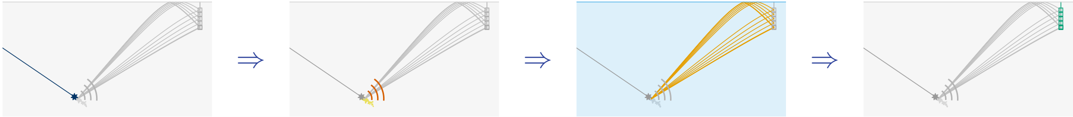
- **Unfolding**: invert the detector response & propagation effects, and fit the **electric field**
- Advantage: (Askaryan) model-independent
- But: inflates noise where detector response is weaker, hard to combine information from multiple antennas



neutrino properties | electric field | in-ice propagation | detector response

The forward-folding approach

- **Forward-folding**: for each direction hypothesis, take the electric field and **forward-fold** it with expected effects from propagation & detector response.
- Fit to measured **voltage traces**.
- Improved accuracy compared to standard unfolding, especially at low SNR [1]

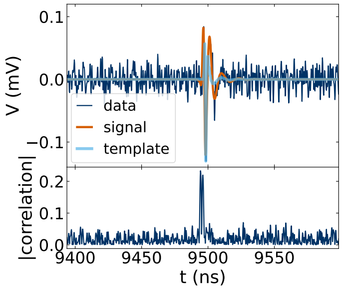


neutrino properties ⇒ electric field ⇒ in-ice propagation ⇒ detector response

[1] arXiv:1903.07023

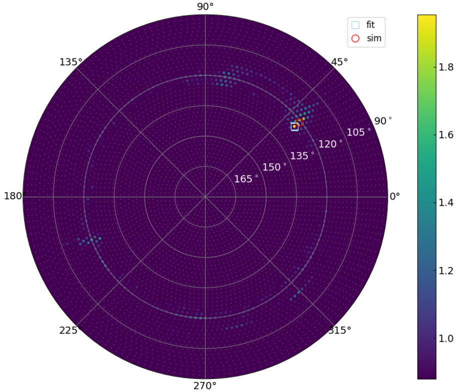
Step 1: Signal direction

- 'Triangulation': use time differences at different antennas to obtain emission vertex ($\approx$ shower maximum)
- Time differences obtained by template correlation



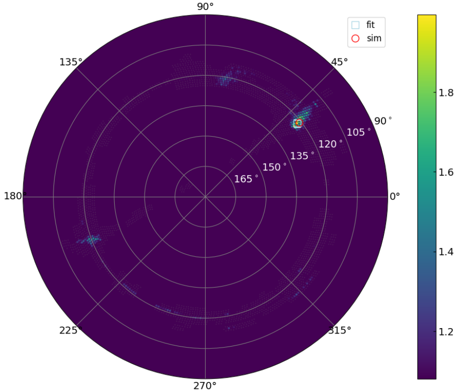
Step 1: Signal direction

- 'Triangulation': use time differences at different antennas to obtain emission vertex (≈ shower maximum)
- Time differences obtained by template correlation
- Maximize total correlation over all channels in iterative grid search



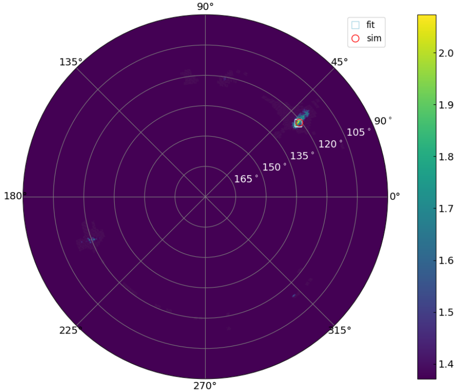
Step 1: Signal direction

- 'Triangulation': use time differences at different antennas to obtain emission vertex ($\approx$ shower maximum)
- Time differences obtained by template correlation
- Maximize total correlation over all channels in iterative grid search



Step 1: Signal direction

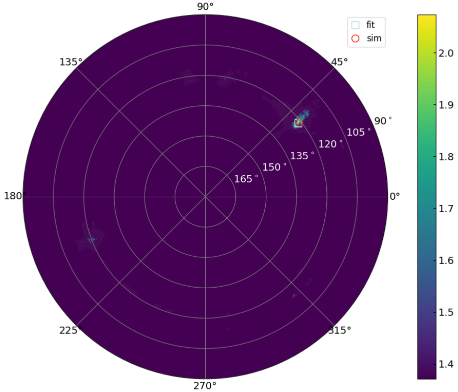
- 'Triangulation': use time differences at different antennas to obtain emission vertex ($\approx$ shower maximum)
- Time differences obtained by template correlation
- Maximize total correlation over all channels in iterative grid search



Step 1: Signal direction

- 'Triangulation': use time differences at different antennas to obtain emission vertex (≈ shower maximum)
- Time differences obtained by template correlation
- Maximize total correlation over all channels in iterative grid search

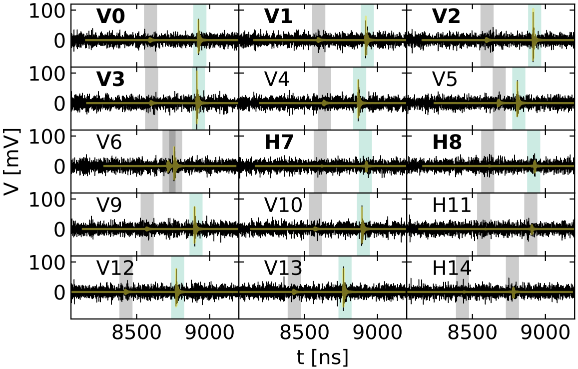
→ Ice model + ray type + vertex position determine **signal direction**



Step 2: Find pulses

- Use emission vertex as input for the direction reconstruction.
- Exact pulse arrival times not known due to uncertainties in vertex, ice model, group delays…
- At low SNR, end up fitting random noise fluctuations.

→ identify approximate pulse windows, and include only those with amplitude $> 3.5\sigma_{\mathrm{noise}}$



Step 3: Fit neutrino properties

For each viewing angle, polarization and shower energy hypothesis:

- Forward-fold expected electric field with propagation & detector effects
- Determine exact pulse arrival time within each pulse window using correlation
- Compute $\chi^2 = \sum_{n=1}^{n_\text{pulses}} \sum_{i=1}^{n_\text{samples}} \frac{(x_i - f_i(\theta_\text{view}, \phi_\text{pol}, E_\text{sh}))^2}{\sigma_\text{noise}^2}$

→ Obtain neutrino properties that minimize χ^2