

**Supplemental Materials for**  
*Ultra-high Energy Air Showers Observed by ANITA-IV*

**CONTENTS**

I. Overview	2
II. ANITA-IV flight	2
A. ANITA-IV instrument.	2
1. Antennas.	2
2. Ground Calibration Stations.	3
3. Preamplifiers.	4
4. Triggering & Digitization.	4
III. Cosmic Ray analysis	4
A. Impulse response calibration.	4
B. Signal quality cuts.	5
C. Cluster identification.	6
D. Cosmic Ray waveform template & background estimate.	7
IV. Results	8
A. CR unblinding.	8
B. Polarity unblinding.	9
1. Special case of three-notch events.	10
2. Assessing the polarity confidence level.	11
C. Assessment of chance observation.	12
1. The near-horizon events.	12
2. Over-the-horizon propagation	14
3. Monte Carlo estimate of overall $p$ -value.	17
V. Discussion	17
A. Possible explanations for observed polarity.	18
B. Particle-physics-based explanations for the observed near-horizon events.	19
VI. Comparison of ANITA-IV to ANITA-III.	19
A. CR and CR-like event waveforms.	21
References	27

## I. OVERVIEW

Ultra-high energy cosmic ray (UHECR) interactions in Earth's atmosphere produce showers of copious secondary relativistic particles known as extensive air showers (EAS). Their secondary particles often impact Earth's surface and can be detected directly, but they also produce secondary optical and coherent radio emission. The latter radio signals appear in the form of nanosecond-scale, highly beamed radio impulses, 100% linearly polarized, with a broadband radio-frequency spectra usually observed in the 1-1000 MHz frequency range, but with components that have been observed out to several GHz [S1].

ANITA was not specifically designed to detect EAS radio emission, but because the impulses bear similarities to ANITA's target signals – neutrino-induced showers within the ice sheets which produce radio impulses via the Askaryan effect – the trigger system is responsive to EAS radio events with reasonable efficiency. The average collecting area for such signals is not large compared to ground-based UHECR observatories, but the signals do provide a non-anthropogenic source of calibration signals. EAS radio signals observed by ANITA show strong horizontal polarization, due to the fact that the geomagnetic field  $\mathbf{B}_{geo}$  in Antarctica is predominantly vertical, and the EAS radio signals arise from Lorentz-force-induced acceleration of the EAS charged particles in the  $\mathbf{v} \times \mathbf{B}_{geo}$  direction, where  $\mathbf{v}$  is the average velocity vector of the shower.

ANITA is insensitive to EAS at low zenith angles  $\theta_z$ , since the antenna array boresights favor a direction closer to the horizon. As a result,  $\theta_z \gtrsim 50^\circ$  for ANITA UHECR events, and these events are seen almost exclusively in reflection off the surface of the ice sheet (or ice shelves in coastal regions). In most of ANITA's frequency range (200 – 1200 MHz) the surface of the ice is relatively smooth on average, and reflection coefficients are high. While there are regions of wind-sculpted surface where the *sastrugi*, or wind-blown formations, can grow to meter scales, these are quite uncommon, and the average root-mean-square surface roughness of the ice is several cm, leading to largely specular reflections without significant distortion.

ANITA also observes a small number of events that beam directly toward the payload from above the horizon, and comparison of these events with the reflected sample shows immediately that the reflected events are phase-inverted with respect to the above-horizon events, due to the effect of the Fresnel reflection coefficient at the ice surface, which induces a  $180^\circ$  flip in the polarity of the waveform. The dominant horizontal polarization makes EAS radio events a completely distinct class compared to neutrino events produced by the Askaryan effect within the ice; these latter impulses emerge from the ice with dominant vertical polarization.

## II. ANITA-IV FLIGHT

NASA long-duration balloon (LDB) payloads in Antarctica launch from the LDB facility adjacent to the Williams field airstrip, about 8 km out on the Ross Ice shelf, east of McMurdo Station on Ross Island, Antarctica. ANITA-IV was launched on Dec. 2, 2016, reaching a float altitude of about 40 km several hours later, and flew in the Antarctic polar vortex for 28 days until the flight was terminated on Dec. 29, 2016, about 160 km from the South Pole.

### A. ANITA-IV instrument.

#### 1. Antennas.

ANITA-IV's receiving antennas are a set of 48 custom quad-ridged horns designed by Antenna Research Associates, covering a passband of 0.18-1.2 GHz. Fig. S1 shows the flight configuration of the payload, with antennas arranged in four rings that share the same central axis. All 48 antennas have their boresights angled down  $10^\circ$  below the horizontal, to maximize their sensitivity to events coming from below the horizon, which appears at a distance of about 650-700 km, and an angle of about  $-6^\circ$  during a typical flight. The two lower aligned rings of 16 antennas are offset in azimuthal angle  $\phi$  by  $22.5^\circ$  each, and each of these  $\phi$ -sectors has an aligned partner in one of the two upper rings; thus for any given direction there are 3 antennas with identical response to that direction.

Each antenna is also dual-polarization with a horizontal (H) and vertical (V) component. The antenna-to-coaxial cable transformer, or *feed* for each polarization is arranged such that the dual-polarization responses of the coincident H and V antennas provide excellent precision for instantaneous measurements of the polarization state of any incident waveform. The angular response function, or *beam* of the antenna is about  $\sim 50^\circ$  in angular diameter on average over the  $\sim 1$  GHz band, providing an average gain of about 10 dB relative to an isotropic antenna (10 dBi).

Thus, a radio wavefront arriving from a plane-wave source will induce detectable signals over 3-5  $\phi$ -sectors, allowing for pulse-phase interferometry between up to 15 antennas to localize the source direction [S27]. The typical angular



FIG. S1: The ANITA-IV payload in front of the LDB hangar, with lower photovoltaic panels deployed as in the flight configuration (C. Miki photo). Each of the quad-ridge horn antennas is about 0.95 m on a side.

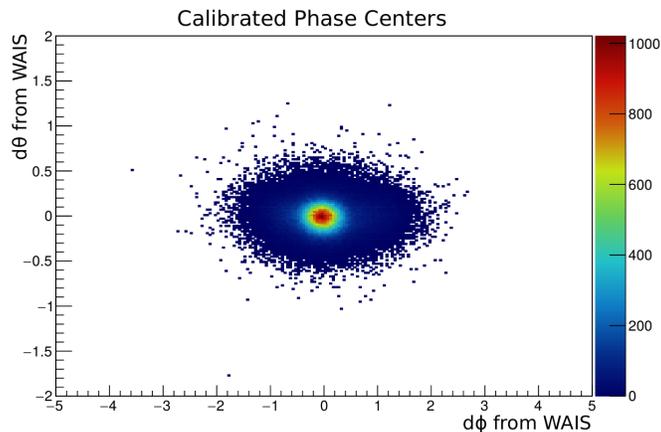


FIG. S2: The pointing resolution determined from fitting using WAIS pulses. There is a mean offset of  $0.01^\circ$  and a resolution of  $0.42^\circ$  in azimuth angle  $\phi$ , and a mean offset of  $0.01^\circ$  and a resolution of  $0.14^\circ$  in elevation angle  $\theta$ .

resolution via this method is  $\sim 0.2^\circ$  in elevation and  $\sim 0.5^\circ$  in azimuth, but it can improve by a factor of two or more for high signal-to-noise ratio (SNR) events.

## 2. Ground Calibration Stations.

To achieve this, ANITA makes use of ground-based radio pulser stations. These are observed by the payload during flight and provide a calibration of residual cable and electronic delays, and payload geometry to the necessary level.

Fig. S2 shows an example of the resolution function determined from one of our ground pulser stations in ANITA-IV, at WAIS divide, for all pulses, independent of the SNR. The fitted resolutions are  $0.14^\circ$  in elevation angle  $\theta$  and  $0.42^\circ$  in azimuthal angle  $\phi$ .

### 3. Preamplifiers.

Antenna signals are amplified and passband-filtered by a low-noise preamplifier module, with an effective noise temperature  $T_{LNA} \approx 65$  K on average across ANITA's band. The steady-state antenna temperature  $T_{ant} \approx 100-120$  K due to the fact that the beam is filled by a composite of sky at  $T_{sky} = 3$  K, and ice with  $T_{ice} \approx 230$  K; the two polarizations see slightly different ice and sky temperatures due to the different Fresnel coefficients for transmission and reflection at the ice surface. The net  $T_{sys} \approx 170$  K varies by 10-20% throughout the flight due to the effects of solar radio noise and varying ice fields-of-view.

### 4. Triggering & Digitization.

Because of the difficulty of continuous sampling and digitization of 96 radio signals at several Gsamples/s, ANITA only captures digital waveforms when a preset *trigger* condition is met, based on combinations of antenna power over several nanoseconds exceeding a threshold. To maintain high signal sensitivity, trigger thresholds are lowered to the point where random radio thermal noise fluctuations trigger the system at a rate of order 50-100 Hz, yielding a data rate that can be stored and later retrieved for offline analysis. For ANITA-IV the effective threshold corresponded to a voltage amplitude in the receiver of about 2.6 times the thermal noise voltage RMS,  $V_{noise} = \sqrt{kT_{sys}\Delta f Z}$  where  $k$  is Boltzmann's constant, the system temperature is  $T_{sys}$ ,  $\Delta f = 1$  GHz is the bandwidth, and  $Z = 50\Omega$  is the system impedance. Further details on the system can be found in the ANITA-IV neutrino analysis paper [S17]

## III. COSMIC RAY ANALYSIS

The initial calibrated ANITA-IV data set contains over 90M triggers, for which 100 ns duration digital waveforms for the entire 96-channel array (48 each of H and V) are stored to 14-bit precision. Of these events, of order 98% are due to thermal noise fluctuations, and these are useful as a real-time measure of the instrument performance, but are reduced to a negligible background prior to the CR or neutrino analyses by requiring waveform and beamforming coherence.

### A. Impulse response calibration.

An important part of the CR and other data analyses involves accurate estimates of the signal receiving system transfer function. The complete system includes the antenna, an initial preamplifier and receiver which includes primary bandpass filters and equalizers, a second-stage receiver which includes tunable notch filters and additional bandpass filters. These are followed by a split in the signal path, with half the received power going to the trigger system and the other half into a 20 ns delay cable followed by the primary ANITA analog-to-digital converters, the Sampling Unit for Radio Frequencies (SURF) modules. For ANITA-IV both the preamplifier and tunable notch filters were completely new units, designed to minimize noise figure, and allow flexible, real-time blocking of offending continuous-wave (CW) signals in several known frequency bands.

For any such system with a composite of many individual complex RF elements joined sequentially, the transfer function, or impulse response, is best measured end-to-end with the complete system intact. In our case, the system can be thought of as a sequence of many individual two-port devices, each of which has its own scattering-parameters matrix, known as the  $S$ -matrix in radio-frequency and microwave usage. The  $S$ -matrix is a  $2 \times 2$ -element matrix of complex, frequency-dependent coefficients for both forward and backward scattering from each of the ports in a 2-port device. Such systems are only reciprocal for passive devices, and this is not the case for ANITA, so the ordering of the sequence of 2-port devices is important.

For most devices, such as amplifiers or filters, they are designed with small back-reflections at each port within the passband, but even when this is the case, the composite response of the system usually deviates substantially from the cascaded sequence of  $S$ -matrices of each 2-port device. In our case, calibrations of the system response were made in Antarctica prior to launch, but could not be done with the complete system, and were instead constructed from a composite of measurements of the antenna and system responses taken separately, and then convolved after the fact

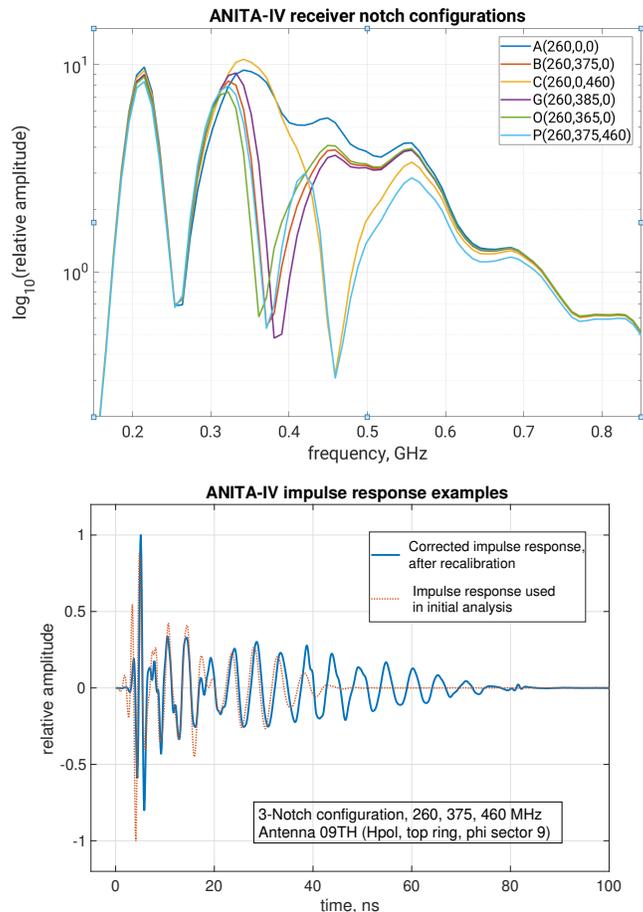


FIG. S3: *Top: the effects of the various notch configurations on the ANITA-IV passband. Bottom: examples of the typical impulse response from the 3-notch case. Orange-dashed line: the initial estimate used for most of the analysis. Blue solid line, measured response made late in the analysis, and used for current results.*

to estimate the impulse response. This approach had been found to be generally accurate in past flights, but was later found to be quite inaccurate for ANITA-IV.

For ANITA-IV the additional complexity due to the tunable notch filter system created unexpected mismatches in the system response. These were discovered well into the analysis stage as we will discuss further below. In Fig. S3(top) we show the effects of the various notches which were used during the flight on the ANITA-IV passband. On the bottom of the same figure, we show an example of a typical time domain response. The dashed line shows the derived impulse response estimated from the composite measurements, convolved together. The blue solid line shows what was actually measured in the lab after recovering ANITA, using a complete end-to-end measurement including the antenna. The actual impulse response showed much more dispersion at low frequencies, and some additional phase distortion at other portions of the spectrum which appear as deviations in the time domain.

## B. Signal quality cuts.

As noted above, the vast majority of recorded triggers are simple thermal fluctuations, and these are easily rejected to the level of  $\sim 10^{-9}$  through routine signal analysis methods and coherence requirements. The degree of coherent impulsive behavior in the signal is one of the strongest discriminators between thermal noise triggers and actual arriving impulsive plane waves. Because the system impulse response has a group delay which varies substantially with frequency over ANITA's band, a deconvolution procedure is used to dedisperse these delays, thus producing a signal that approximates the temporally coherent arrival of the original plane wave at the antenna. Thermal noise triggers, when so deconvolved, tend to rapidly lose whatever instantaneous coherence they had when they triggered

the system, and are thus strongly rejected by straightforward analysis.

Additional non-signal backgrounds come from local RF interference from sporadic sources on the payload itself, which cannot meet the timing requirements for an arriving plane wave, and CW from above-horizon satellites. RF from satellites cannot be impulsive due to the strong dispersion of the ionosphere in ANITA's band, but their CW signals can conspire with thermal noise to produce narrow-band triggers, which are rejected offline for this characteristic.

Once these are removed, about 1M impulsive events that reconstruct to apparent point sources remain in both polarizations. The vast majority of these are of anthropogenic origin, and can be identified due to their tendency to form clusters around centers of human activity. Once two impulsive events are deemed to form a cluster, they are no longer considered as signal candidates; we require complete spatial isolation for both CR and neutrino candidates, regardless of when in the flight they are observed.

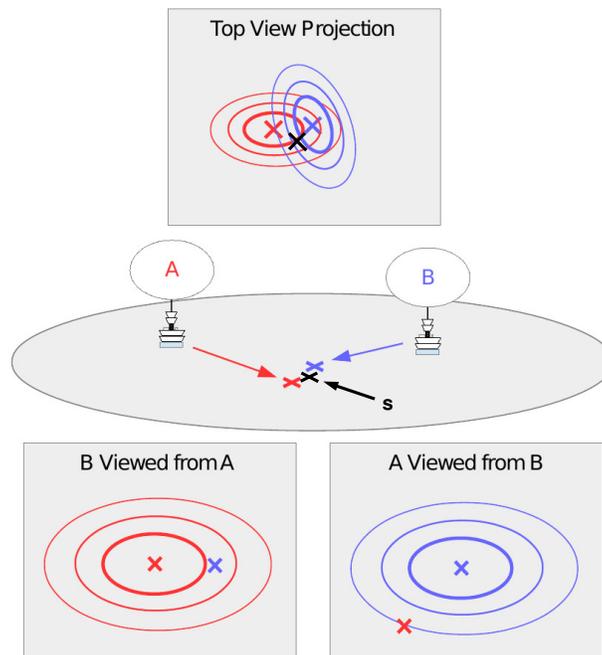


FIG. S4: *Top panel: event A and B error ellipses projected onto the continent, as per the geometry shown in the middle panel. In the bottom panes, the projection of the probability contours for event A with respect to B's position, and the converse are shown, indicating the asymmetry of the clustering problem. Position  $s$  indicates the best fit anthropogenic source position as determined by the angular likelihood method described in the text.*

### C. Cluster identification.

Events are grouped into clusters according to a likelihood methodology in which pairs of events are parameterized by their elevation and azimuth angles ( $\theta, \phi$ ) and respective standard errors ( $\sigma_\theta, \sigma_\phi$ ) from the payload at the time of detection. These directions and their standard error ellipses are then projected onto the Antarctic continent, forming closed quasi-conic curves on the ice surface, depending on the topography of the region around the event, as indicated in Fig. S4. Since no temporal limit is set on the possibility of two events forming a cluster within the flight duration, any pair of events that exceed a pre-set angular likelihood threshold of being associated with one another are considered a cluster. We term this portion of the analysis *clustering*.

ANITA has used several methods for clustering in the past. Initially, a simple overlap threshold was used, for example, if two events had any overlap in their  $3\sigma$  projected contours on the continent, they would be excluded. A more effective method was later developed which used a likelihood approach to quantify the possibility that two events were from a common source. Finally, for this analysis, the likelihood method was refined to correct deficiencies in these prior methods, and was used for the current analysis, as we describe here.

Consider two events, denoted  $A$  and  $B$ , with angular positions and errors  $(\theta_A \pm \sigma_{\theta,A}, \phi_A \pm \sigma_{\phi,A})$ , and  $(\theta_B \pm \sigma_{\theta,B}, \phi_B \pm \sigma_{\phi,B})$ , as shown in Fig. S4 [S6]. We can form the likelihood that event  $A$  is consistent with the position of event  $B$ :

$$-\log(L_{AB}) = \left( \frac{\theta_A - \theta_B}{\sigma_{\theta,A}} \right)^2 + \left( \frac{\phi_A - \phi_B}{\sigma_{\phi,A}} \right)^2 \quad (\text{S1})$$

and similarly the likelihood that event  $B$  is consistent with the position of event  $A$ :

$$-\log(L_{BA}) = \left( \frac{\theta_B - \theta_A}{\sigma_{\theta,B}} \right)^2 + \left( \frac{\phi_B - \phi_A}{\sigma_{\phi,B}} \right)^2. \quad (\text{S2})$$

The  $L_{AB}$  likelihood has been used extensively in prior ANITA analyses. However, because the two events are observed from different payload locations, and the standard errors for events are generally different, depending on the SNR of the events,  $L_{AB} \neq L_{BA}$ . In practice, using this method introduces asymmetry in the clustering, such that whether events cluster can depend on the order in which they are processed, an undesirable feature.

Since the goal of the clustering analysis is to identify whether two events came from a common, presumably anthropogenic, origin, we wish to use a form of likelihood which, when minimized, finds the point on the continent surface which best fits any pair of events taken in any order. Then we use known anthropogenic sources to identify a threshold value for the likelihood which defines whether events should be considered clustered or not. For this we use a sum of the two likelihood terms above.

Following the description in reference [S15], we perform a fit for each pair of events  $A, B$ , minimizing the likelihood  $L(\mathbf{s})_{AB}$  that the two events came from some common source location  $\mathbf{s}$ , given by

$$\begin{aligned} -2\log(L(\mathbf{s})_{AB}) &= \left( \frac{\theta_A - \theta_{sA}}{\sigma_{\theta A}} \right)^2 + \left( \frac{\phi_A - \phi_{sA}}{\sigma_{\phi A}} \right)^2 \\ &+ \left( \frac{\theta_B - \theta_{sB}}{\sigma_{\theta B}} \right)^2 + \left( \frac{\phi_B - \phi_{sB}}{\sigma_{\phi B}} \right)^2. \end{aligned} \quad (\text{S3})$$

The fit varies a position on the surface,  $\mathbf{s}$ , which appears at payload coordinates  $(\phi_{sA}, \theta_{sA})$  for event  $A$ . The event  $A$  angular resolution,  $\sigma_{\phi A}$  and  $\sigma_{\theta A}$ , is estimated from the resolution of calibration pulse events with the same SNR. Events  $A, B$  are considered clustered if

$$-2\log(L(\mathbf{s})_{AB}) < \Lambda$$

where  $\Lambda$  is an angular threshold;  $\Lambda = 6$  for the current analysis.

This angular cluster identification methodology is effective for the large majority of events, but for events observed at locations at steeper upcoming angles, from locations closer to the payload, the projected distances of the angular errors are too small to adequately exclude anthropogenic clusters. Thus in addition to the angular likelihood cut, a conservative fixed radius of 40 km is enforced around each event, based also on measured distributions of events around measured anthropogenic sites, and two or more events within 40 km of each other on the surface are considered a cluster and excluded from the analysis.

The cluster grouping analyses causes an unavoidable loss of efficiency for signal detection, since actual signal events can group with background, and rarely, with each other, and are thus excluded. However, we deem this necessary to ensure that the anthropogenic background for the complete analysis is  $\ll 1$ . In the current analysis the efficiency after cluster removal was 84.5%; this is primarily just due to the geometric removal of area from the observed region of the continent, due to the sum of the individual projected areas under each cluster.

#### D. Cosmic Ray waveform template & background estimate.

In our normal neutrino analysis, the cluster grouping stage affords us the opportunity to make an *a priori* background estimate, and regardless of the details of the analysis the outcome of such estimates is typically an expected background of  $\mathcal{O}(1)$ . However, because ANITA has detected CRs in prior flights, and because there is a large body of work on radio signals from CRs, we develop waveform templates, both from existing prior events, and from computer simulation, to provide waveform template matching to help ensure a relatively pure sample of CRs in our analysis. A waveform template then allows for cross-correlation analysis to identify CR candidates, but also to identify what fraction of the anthropogenic background in clusters also correlate with such templates at a level that they could be mistaken for CRs in analysis, if they appeared as isolated events. This allows us to reduce the expected background by about an order of magnitude.

To avoid a statistical penalty and because of the similarity in waveforms shapes for CRs detected and simulated over a range of parameters, we use a single CR template for the search, which is done in two stages. A trial template is used to identify a set of high-SNR CR events, which are then averaged to form a higher-SNR second-stage template, and this latter template is then used for the complete CR search.

A purely vertical geomagnetic field, acting via the Lorentz force  $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$  on charged particles in a CR air shower, can only produce horizontal acceleration and thus horizontally-polarized radiation. However, there are residual non-vertical geomagnetic field components in Antarctica which yield a small component of vertically-polarized radio emission, and this component is readily identifiable in CR events, and closely correlated to the geomagnetic locale of the event. From prior experience and modeling of the geomagnetic field in Antarctica, the plane-of-polarization of CRs rarely exceeds  $15 - 20^\circ$  and is completely contained within a cut of  $\pm 45^\circ$ , and we thus include only the dominant H-polarization events in this analysis.

Applying both the template and geomagnetic requirements to events in clusters provides a statistical sideband sample from which to determine the anthropogenic background, which is assumed to “leak” out from clusters into isolated single events. The final signal analysis is still blind to polarity at this stage, and a detailed analysis of the anthropogenic sample, including estimates standard errors, gives an *a priori* robust background estimate of  $0.185_{0.08}^{0.14}$  events for each polarity in the final CR sample. Including these cuts in addition to cuts already used to remove thermal-noise events reduces the thermal noise background to a completely negligible level,  $\sim 5 \times 10^{-7}$  events.

## IV. RESULTS

Final CR results were obtained in two steps. Once the cluster grouping and background estimation was complete as described above, we then proceeded to unblind to isolated events, but remain blind to the polarity of these events. At one point after unblinding, we corrected the polarization calculation, but our final selection of the CR events remains the same as the blind selection

Once the CR sample was thus obtained, we refined our polarity estimators while remaining blind to polarity, to avoid bias in the final polarity analysis. The polarity-blinding was accomplished by using random-sign unit multiplication of an event’s waveforms, with a hidden seed based on the event number. Well after the initial unblinding of polarity, when we discovered and corrected errors in the estimation of the system impulse response, we were no longer able to reblind the data or the polarity due to the small signal sample size and the uniqueness of the events. This latter analysis was thus performed in a non-blind fashion, but with a high degree of scrutiny and analysis validation required by the collaboration. We summarize these analyses in the following subsections.

### A. CR unblinding.

The unblinded sample of isolated events showed 30 candidates. Of these events, 29 are consistent with CRs in their template correlation and geomagnetic parameters. For the remaining event, an error in the geomagnetic computation led to its inclusion in the sample, although it is not consistent with the geomagnetic parameters of a CR. The same error affected the geomagnetic parameters for other events, but in those cases the events remained consistent with expectations for CRs and once corrected the results remained unaffected.

A separate, independent analysis [S26] specifically designed to search for neutrino events also found a large subsample of the CRs as a sideband of the main analysis, confirming the isolated, impulsive nature of those that were found in common. However, the efficiency of this analysis was reduced due to the use of an older, less effective clustering method, involving a simple contour overlap rather than a likelihood method. This approach could not be properly normalized for near-horizon events where the contours did not intersect the continent, and thus results on the four near-horizon CRs observed here are not reliable for our purposes.

In this neutrino sideband analysis, two of the CRs accepted in the current analysis, events 19848917 and 50549772, were found to cluster with other anthropogenic events. One additional event 4098827 was found to overlap with a known human encampment at Dome C, but it did not cluster with any other detected events. To ensure that the current analysis had not incorrectly included these events in the sample, each of these two clusters was investigated to see if the candidate CRs were likely to belong to these clusters.

For the case of 19848917, it was found to be well isolated, with a distance of  $> 60$  km to the nearest event, and a likelihood ratio of  $10^3$ , based on a log likelihood (LL) of 13 compared to a threshold of  $LL \geq 6$  for clustering. In addition, it was found to be statistically dissimilar to events in any nearby clusters, and none of the clusters had any events which exceeded 19848917’s figures-of-merit for CR identification. This event is thus deemed to have a very small chance of being anthropogenic.

For the event 50549772, the results are mixed. The event is still well over the isolation threshold (LL=7, a log-likelihood factor of 2.8 above the threshold of 6), and the nearest cluster is 60 km from this event, well over the 40 km minimum. This event correlates only moderately with events in that cluster, with correlation coefficients (CCF) of 0.35 to 0.6, typical of CCF values among most events that have already passed a cut requiring an impulsive character. In this nearby cluster there is however one event with a higher CR figure-of-merit than 50549772, but statistically with 29 observed isolated CR, we already expect several CR to have been “captured” by clusters, so the significance of this other CR-like event is unclear. 50549772 thus appears to have a greater posterior likelihood of being anthropogenic, but it still passes all requirements for a CR, and we do not find any conclusive reason to exclude it in our current analysis.

The current CR analysis did not apply any requirement that events be isolated from known camps as long as there were no other events from that camp that clustered. Thus the final event in question here, 4098827, which is otherwise completely isolated, remains in our current sample.

## B. Polarity unblinding.

In prior analyses, the polarity of ANITA CR events was determined by a phase-referenced, windowed correlation of the candidate deconvolved waveform with the CR template. The normalized, windowed cross-correlation coefficient computed between a measured candidate CR waveform  $w(t)$  and the CR template  $\mathcal{C}(t)$ , evaluated over a window length  $T$  is given by

$$\mathcal{X}(\tau) = \frac{1}{\sqrt{\mathcal{A}_w(0)\mathcal{A}_\mathcal{C}(0)}} \int_{-T/2}^{T/2} w(t)\mathcal{C}(t - \tau)dt \quad (\text{S4})$$

where  $\tau$  is called the *lag*, or time offset of the product of the functions, and  $\mathcal{A}_w(0)$ ,  $\mathcal{A}_\mathcal{C}(0)$  are the autocorrelations of the waveform and template, respectively, at  $\tau = 0$ , which provide the normalization:

$$\mathcal{A}_w(\tau)|_{\tau=0} = \int_{-T/2}^{T/2} w(t)^2 dt \quad (\text{S5})$$

with a similar equation for  $\mathcal{A}_\mathcal{C}$ . In practice these equations are discretized at the sample period of the data after interpolating to a suitably fine reference grid, typically 10 ps/sample. The window is determined by selecting the region around the peak group delay of the signal, typically close to the maximum of the intensity in the coherently beamformed waveform. A window of  $T = 5$  ns centered on this is found empirically to capture  $> 90\%$  of the energy, and the sign of the largest magnitude of the cross-correlation coefficient in this window defines the relative polarity among events.

As noted in a previous section, the impulse response of the ANITA-IV system had increased complexity compared to prior flights, and in fact was incorrectly estimated in the first complete pass of the analysis, including the polarity unblinding. Once the error in the impulse response was found and corrected, it was deemed necessary to reblind the polarity and determine a simpler metric for polarity estimation that was robust in the presence of the effects of the ANITA notch filters, since residual effects from these notches may persist in the waveforms even after deconvolution.

Reblinding in this case was accomplished by first having the data analyst performing a repetitive series of training exercises, observing the 29 CR waveforms and their deconvolutions randomized in order, in event number, and in polarity. These exercises were repeated many times until it was no longer possible for the analyst to identify specific waveforms. Further analysis in setting cuts for the polarity was performed with all of the three different blinding levels in place throughout.

For most simple waveforms with compact temporal support, polarity is straightforward to define. For unipolar waveforms, those dominated by a single large pole or mode, the polarity is defined simply by the sign of the pole. For bipolar waveforms, those dominated by two adjacent sequential poles or modes with roughly equal size, the polarity is defined by the order of those modes, whether positive-to-negative (here denoted positive polarity), or negative-to-positive (denoted negative polarity).

Further complications such as tertiary poles producing quasi-tripolar shapes are possible as well, but in our data the first two cases, unipolar and bipolar, represent the vast majority. In the small number of cases where significant tertiary modes appear, we have considered only the two largest modes in a bipolar shape. These can be identified both by their relative amplitude, and by considering the peak energy flow, or intensity, of the waveform, which localizes the portion of the waveform which contributes the largest power to the signal.

The resulting simple algorithm based on the sign and direction of the primary poles, was found to be the most robust in the reblinded study of the 29 CR. Under the empirical observation of general similarity in CR event waveform

TABLE I: CR or CR-like events seen by ANITA-IV.

event index	event number	UTC of event yyyy mm dd hh mm ss	Source			ANITA			elevation angle, deg.	azimuth angle, deg.	polarity <sup>[b]</sup>
			Lat.	Lon.	Alt. <sup>[a]</sup>	Lat.	Lon.	Alt. <sup>[a]</sup>			
1	4098827	2016 12 03 10 03 27	-75.584	124.074	3197	-80.2	131.2	38864	-6.17	338.314	+1 <sup>[c]</sup>
2	9734523	2016 12 05 12 55 40	-71.862	32.612	19000	-80.9	31.6	39247	-5.681	2.006	+1
3	12131787	2016 12 06 06 00 34	-78.090	3.129	2634	-79.0	6.1	38614	-17.285	324.701	-1
4	15738420	2016 12 07 10 11 00	-76.335	-56.681	40	-77.3	-42.4	39507	-7.633	279.17	-1
5	16821419	2016 12 07 16 37 03	-77.152	-59.623	43	-78.1	-51.7	39447	-11.266	294.333	-1
6 <sup>[d]</sup>	17904564	2016 12 07 23 24 10	-78.809	-54.731	48	-77.8	-58.5	38925	-16.178	144.456	...
7	19848917	2016 12 08 11 44 54	-80.983	-79.974	758	-76.7	-72.9	38969	-6.73	194.344	+1
8	20936205	2016 12 08 20 45 53	-76.022	-88.574	1238	-77.3	-85.2	39443	-13.49	326.905	-1
9	25580797	2016 12 10 08 55 36	-81.443	-121.590	1113	-80.6	-120.8	37702	-22.325	188.272	-1
10	25855454	2016 12 10 11 06 39	-78.991	-119.060	1609	-80.3	-122.5	37505	-13.168	27.039	-1
11	36785931	2016 12 13 06 51 19	-82.508	154.850	2170	-82.1	153.8	38549	-36.559	162.566	-1
12	39236841	2016 12 13 20 49 08	-79.841	123.841	2904	-82.0	130.4	37451	-8.454	331.359	-1
13	40172984	2016 12 14 02 28 22	-82.576	113.675	3006	-82.3	119.5	38613	-21.749	248.545	-1
14	45684620	2016 12 15 10 48 09	-82.975	79.944	3747	-82.0	78.3	37527	-17.351	168.514	-1
15	47396999	2016 12 15 20 31 28	-81.881	62.560	3872	-81.2	68.6	36377	-15.249	230.51	-1
16	50549772	2016 12 16 15 03 19	-83.462	16.004	2572	-81.9	47.3	38519	-6.43	234.078	+1
17	51293223	2016 12 16 19 08 08	-74.800	11.427	18600	-81.7	39.2	37527	-5.470	306.446	+1
18	54063721	2016 12 17 11 25 03	-82.610	11.176	2499	-81.2	22.9	39413	-9.875	224.328	-1
19	64472798	2016 12 20 05 40 44	-85.523	-53.227	1702	-85.2	-60.5	38575	-26.012	124.297	-1
20	64859493	2016 12 20 08 22 26	-84.830	-74.162	1113	-84.5	-64.4	38319	-19.707	246.182	-1
21	64861754	2016 12 20 08 23 18	-85.103	-71.284	1230	-84.5	-64.5	38316	-21.714	223.006	-1
22	66313236	2016 12 20 17 36 12	-84.119	-70.513	865	-85.1	-78.4	39358	-15.956	39.846	-1
23	66509677	2016 12 20 18 50 42	-85.935	-112.563	2377	-85.2	-81.3	39380	-8.789	237.462	-1
24	72164985	2016 12 22 06 28 14	-86.233	3.889	2627	-86.9	-104.3	38575	-6.16	139.848	+1
25	74197411	2016 12 22 19 58 49	-84.825	-123.197	1163	-86.2	-121.6	39228	-14.824	353.971	-1
26	83074427	2016 12 25 06 57 34	-82.122	-162.066	18	-85.6	-171.6	39666	-7.338	20.837	-1
27	88992443	2016 12 26 22 38 52	-85.676	159.113	2569	-85.7	141.7	39277	-14.823	95.834	ind <sup>[e]</sup>
28	91525988	2016 12 27 15 03 36	-86.268	141.217	2864	-86.8	127.8	39118	-18.648	61.283	-1
29	93744271	2016 12 28 05 05 02	-87.897	147.457	3044	-87.4	140.6	39572	-29.964	153.791	-1
30	95576190	2016 12 28 17 16 11	-86.959	150.403	2917	-88.3	138.4	38871	-13.691	27.186	-1

<sup>a</sup> Latitude and Longitude in degrees, altitude in meters.

<sup>b</sup> For event polarity,  $-1$  indicates and inverted waveform,  $+1$  a non-inverted waveform.

<sup>c</sup> This event is one of two 3-notch CR events, and its polarity is demonstrably opposite that of the normal reflected CR event 36785931.

<sup>d</sup> Event 17904564 is included here for completeness, although it should not be considered part of the CR sample proper.

<sup>e</sup> Event 88992443 showed a very distorted waveform with many modes, and was deemed to have indeterminate polarity.

shape, whether reflected or direct, the polarity of the direct events then defines absolutely the non-inverted phase, and reflected events show a  $180^\circ$  phase inversion compared to this. The presence of both direct and reflected CR thus allows us to define unambiguously the order of poles expected in bipolar events as well as the sign of unipolar events.

Table I summarizes the location and angular parameters for all of the CR and CR-like events, and includes the estimate of the polarity, taken from the sign of the average of the four methods described above. Various individual waveforms including the deconvolutions and original coherently-summed waveforms of all 30 events are shown in Appendix A.

### 1. Special case of three-notch events.

Two of the CR candidates identified were detected during periods of excessive anthropogenic radio interference, during which the ANITA instrument had all three notch filters activated, impacting a significant fraction of the events' radio spectra. While our deconvolution methods can reconstitute the missing frequency components with reasonable fidelity, the overall waveform shape for such events retains some of the phase and amplitude distortion caused by the filters. For this reason, comparison with other CR that did not have this level of filtering is no longer a reliable indicator of their CR properties.

Fortunately the two events we observed were found to have conclusively opposite polarity, and the polarity algorithm used and tuned when polarity was blind found that the two events gave consistent and robust polarity measures.

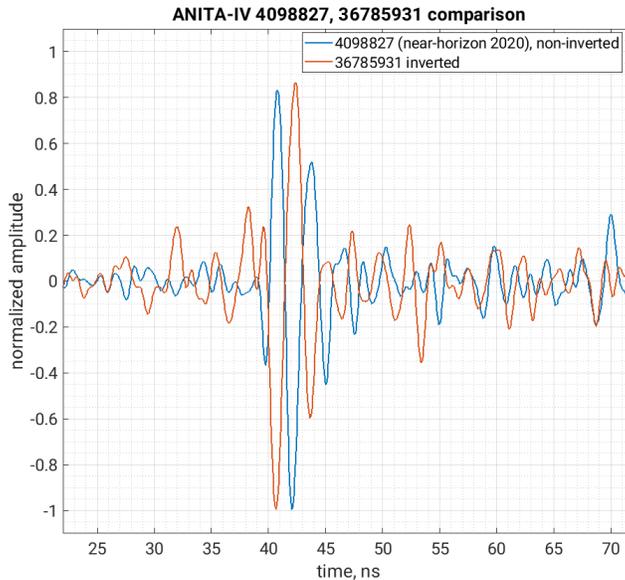


FIG. S5: *Overlay of normalized average deconvolved waveforms for the two three-notch events, showing their anti-correlated waveforms, a clear indication of opposite polarity.*

Figure S5 shows an overlay of these two events as seen with the CLEAN algorithm. The waveforms are clearly anticorrelated, indicating they have opposite polarity, as the algorithms found. In this case, 4098827 is one of the near-horizon anomalous events, and we thus have greater confidence that the polarity is correctly identified, despite the potential residual distortion caused by the triple-notch effects.

## 2. Assessing the polarity confidence level.

We have presented our implementation of the method to determine the polarity for each event, but this does not provide any statistical method of assessing the likelihood that the measured polarity is an accurate estimate of the true polarity of the event. To do this, we have developed a variational Monte Carlo which generates a large ensemble of statistically similar events based on the parameters of each of the CRs, within a range determined by the SNR of the event and the known properties of thermal radio noise.

For each CR event, an ensemble of  $N = 24,000$  similar events were generated based on the CLEAN components of each event, which provide an estimate of the intrinsic field structure prior to convolution with the system response. CLEAN decomposes each event into a set of  $\delta$ -function-like elements, each of which represents the strength of the system impulse response observed at that location in the waveform. By then convolving these with a suitable restoring beam, combined with a statistically independent realization of thermal noise, an event can be generated which can then be independently tested for polarity by the seven methods. For each CR  $i$ , and for each ensemble trial  $k$ , a binary choice of  $p_{ik} = \pm 1$  is made in the polarity, and a cumulative sum

$$S_i = \sum_{k=1}^N p_{ik}$$

is produced. The final normalized value of  $S_i/N$  is the estimate of the statistical polarity for CR event  $i$ .

Fig. S6(Left) shows the resulting statistical polarities including those of the component methods. The two initially indeterminate events, 74197411, and 88992443, are clearly evident.

On the right, we show the fraction of the 24,000 ensemble trials for each CR which produced a polarity opposite to the polarity established for the observed event, on a logarithmic color scale. The method that determines the polarity via the average waveform in this case is insensitive to large differences among the methods and gives a more robust value.

Event 19848917 shows around a 1% chance of polarity flip per trial. We investigated this event in detail, flagging all cases where its polarity was misidentified, and it was found to be due to the effects of a moderate tripolar component

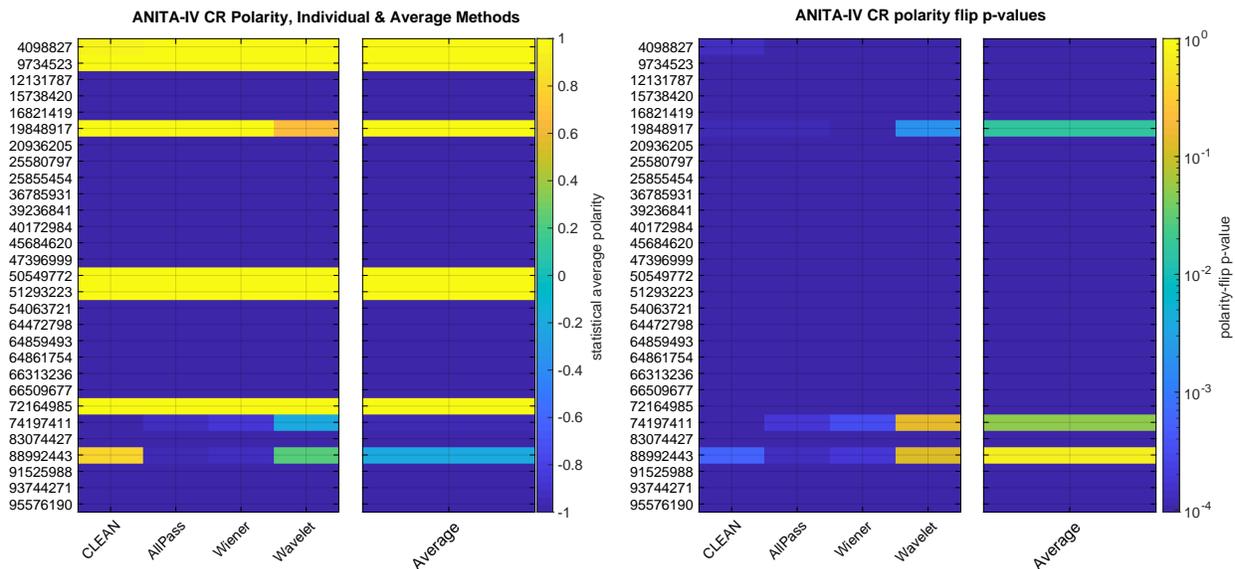


FIG. S6: *Left: the statistical polarity of each of the 29 CR candidates is shown for the four component methods, and the average method which defines the polarity. Right: The statistical p-values for polarity confidence, based on the number of observed polarity flips in 24,000 trials.*

in the event. When noise fluctuations increased this component above the threshold level for the algorithm to consider it a potential bipolar mode, the primary mode order was then spuriously reversed. After extensive investigation, we determined that this was a deficiency of the algorithm, rather than an intrinsic problem with that event's polarity. While further rebinding could have refined the algorithm, the 1% polarity-flip chance for this event had a negligible effect on the overall significance of the near-horizon events, as we show in the next section.

### C. Assessment of chance observation.

It is evident from considering the statistical possibilities that polarity may be misidentified, that the significance of our observation of 4 anomalous non-inverted CRs must be determined by considering all of the possible backgrounds, even before investigating whether there could be physical mechanisms to cause these observations. These backgrounds consist of

1. Anthropogenic events misidentified as CR, with an equal likelihood of either polarity. The background for these has been separately quantified as  $0.37^{+0.27}_{-0.17}$  events in our sample.
2. Events of normal inverted polarity that suffer a statistical fluctuation leading to a misidentified polarity flip; the data displayed in Fig. S6 quantifies this for each event.
3. Above-horizon events that are misreconstructed in their angular location to a position below the horizon.

#### 1. The near-horizon events.

For case 3 above, three of the near-horizon CR are within a few standard deviations of the horizon based on our angular resolution estimates, and could be candidates for this background process. To assess its importance we need

to have accurate estimates of the angular direction of the events, the topography of their locations, their SNR and corresponding angular error estimates, and any systematic angular offsets that may obtain.

ANITA's angular resolution is determined by interferometric imaging of ground-based radio pulsar stations. These are positioned at several locations around the continent with the hope that multiple stations will be observed, although good coverage of one station is adequate to determine the angular resolution. For ANITA-IV, we successfully observed one such station at the West Antarctica Ice Sheet (WAIS) divide station, operated by the US Antarctic Program. This was observed over many days with several hundred thousand events measured in both horizontal and vertical polarizations.

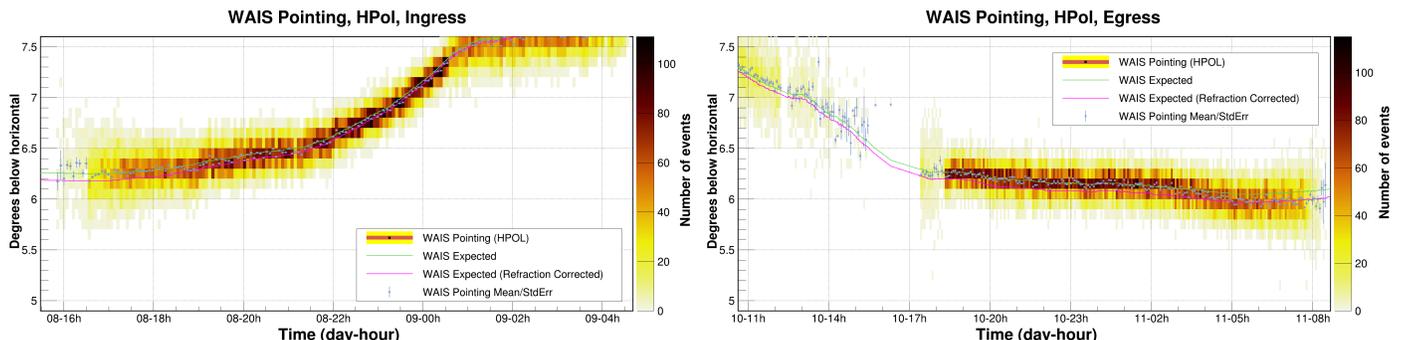


FIG. S7: 2D histogram of reconstructed elevation angles of the WAIS calibration data vs. time for both the ingress of the payload into view of the WAIS pulsar (left) and egress (right). The data were not corrected for refraction, and the uncorrected (expected) curve is shown, along with a curve showing the refraction corrected elevation angles as well.

These events are used to construct a global fit to residual time and phase offsets in the system, producing a set of calibration constants that remove these offsets and yield statistical angular errors in the elevation angle  $\theta$  and azimuth  $\phi$  vs. event SNR. In this case, the elevation angle errors are most relevant; they vary from about  $0.25^\circ$  for events at the lower end of the SNR distribution, to  $\sim 0.1^\circ$  for events of high SNR.

To check for systematics we have observed the near-horizon event distributions of the calibration data. Only about 20% of the WAIS calibration data occupies the angular range of the anomalous near-horizon events, but all of the data is used for calibration, which produces a single set of calibration constants independent of event elevation angle. A summary of these data is shown in Fig. S7. We observe no obvious systematic in these data at  $\leq 0.1^\circ$  level, although the events at the extreme angular range of the egress, just as WAIS moves beyond the horizon, do show a possible upward shift at the  $+0.1^\circ$  level. There is no corresponding effect at ingress, so this may be a time-dependent systematic, or possibly an uncalibrated payload tilt.

We have also used anthropogenic events associated with Amundsen-Scott station at the South Pole at a time that it was very close to ANITA's horizon. These data show a possible systematic lowering of the elevation centroid of events, at about the  $-0.1^\circ$  level. However, from Fig. S8 it is evident in these data that there are many events included in the sample that arise from the movement of at least one and possibly two aircraft with a component of their motion that approaches ANITA, leading to a centroid that moves in from the horizon along the flight path. This could account for the shift in the centroid very near the horizon, making these data inconclusive.

Two other data sets were also investigated for systematics: data from a separate companion experiment to measure the reflectivity of the ice sheets, the HiCal flights, appeared to favor a lowering of the elevation centroid of events, at about the  $-0.05$  to  $-0.1^\circ$  level. However, HiCal was imaged at several degrees above the horizon, and has quite different frequency content and pulse shape from the CR events, so the direct relevance of these data remains uncertain.

A similar offset was seen in centroids of thermal noise events from the Sun, imaged near the edge of our field of view,  $10$ - $15^\circ$  above the horizon. These are events with no impulsive signal present, they are produced by the classical incoherent thermal noise correlation produced by the Solar disk, and thus their relevance as a measure of systematics in the pointing of a coherent impulse is also uncertain.

As a result of these investigations, since we see potential systematic offsets with either sign at the  $0.1^\circ$  level, and we include cases with and without this possible systematic error in assessing the  $p$ -value significance of our 4 anomalous events.

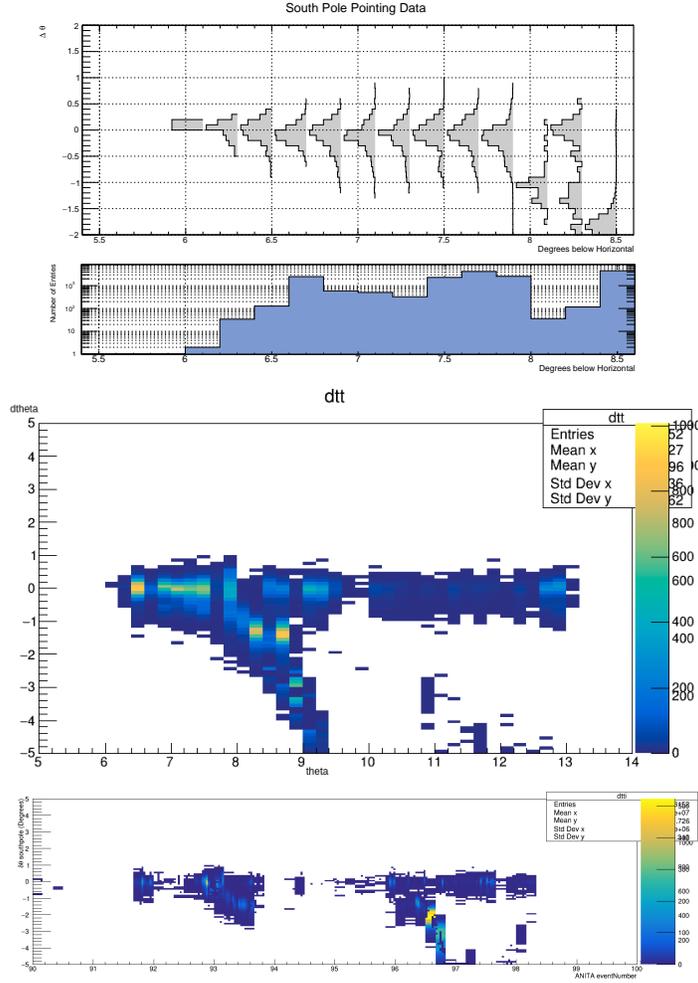


FIG. S8: *Top: histograms of event distributions near South Pole station, showing a possible systematic offset. Middle: a 2D distribution of the events vs. angle, showing evidence for two separate centroids of events in the South Pole data. Bottom: A time series of the event distributions, showing that the centroid offsets are likely the result of aircraft motion in the vicinity of South Pole station. Such aircraft motion is common throughout the Austral summer.*

## 2. Over-the-horizon propagation

In considering whether any of the four near-horizon events could be CRs that interacted beyond the horizon and were reconstructed to positions below the horizon, thus with a much nearer origin than they actually had, the additional propagation distance and path through the troposphere may be important in assessing the likelihood of this possibility. The two evident stratospheric events that ANITA-IV observed were  $0.3 - 0.5^\circ$  above the horizon, and one additional stratospheric event observed by ANITA-III was also in a similar angular range relative to the horizon. Thus we have no examples of events propagating in angular range closer than about  $0.3^\circ$  from the horizon. At a typical horizon distance of about 700 km, this gives a ray path impact parameter of 3.7 km above the ice at nearest approach.

In the physical optics of radio propagation, the wavefront's coherence is largely preserved as long as no obstruction comes within the radius of the first Fresnel zone of the wavefront at some location  $P$ , determined by

$$R_F(P) = \sqrt{\frac{cd_1d_2}{f(d_1 + d_2)}} \quad (\text{S6})$$

where  $c$  is the speed of light,  $f$  the radio frequency, and  $d_1, d_2$  are the distance from the evaluation point  $P$  to the source and receiver, respectively. In ANITA's case,  $d_1 \simeq 500$  km, and  $d_2 \simeq 700$  km for a CR shower wavefront propagation very near the horizon to ANITA at a typical 40 km altitude.

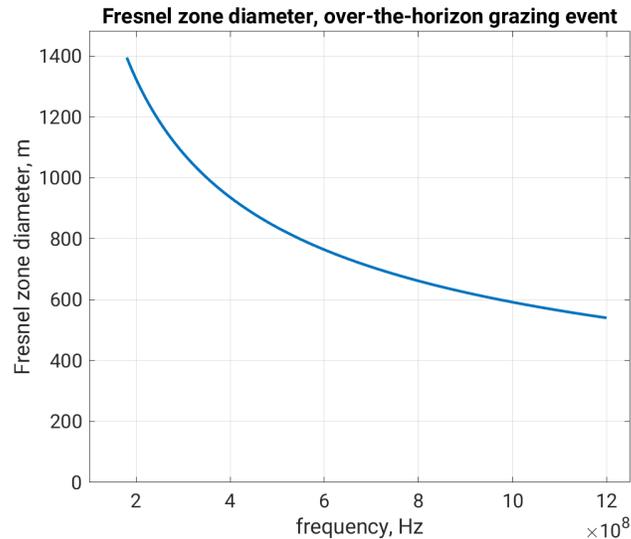


FIG. S9: *Fresnel zone diameter near the point of closest approach to the ice surface for the case of an over-the-horizon CR event with an origin about 500 km distant (based on ray tracing and CR interaction modeling), with ANITA about 700 km distant at an altitude of 40 km.*

The resulting values for the Fresnel zone diameter  $2R_F$  are shown in Fig. S9. If the central wavefront ray is below about 700 m from the ice surface, corresponding to about  $0.06^\circ$ , the surface itself becomes an obstruction at the low end of ANITA's band; for a central ray at 250 m height, the surface is an obstruction at all frequencies within the ANITA band, and at lower frequencies there are likely rather strong effects. Thus we can expect that propagation will be impacted somewhere in this range.

In addition to the issue of Fresnel zone obstruction, the tropospheric radio index of refraction can vary significantly within the first few kilometers above the ice in Antarctica. To illustrate this we show profiles of the radio index of refraction derived from the twice-daily radiosonde data from Amundsen-Scott station at the South Pole in Fig. S10. In these data we have subtracted off the average exponential index of refraction, showing the residual variation around the mean value.

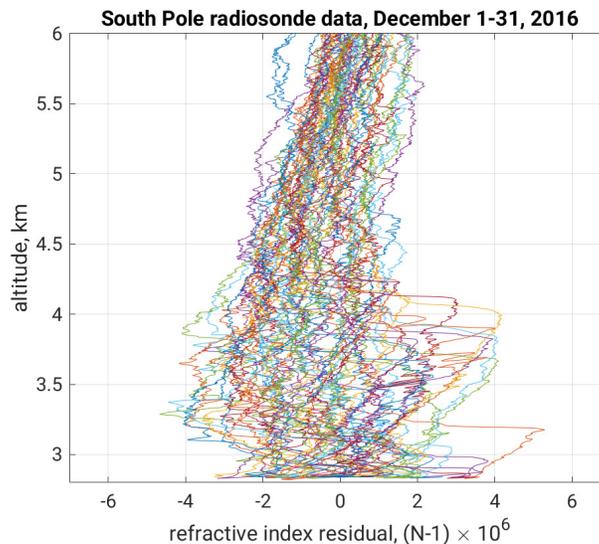


FIG. S10: *Radio index of refraction data derived from South Pole radiosonde data during the time of the ANITA-IV flight. The overall exponential curve of the index of refraction has been subtracted, showing the variation in the residuals relative to the average exponential curve.*

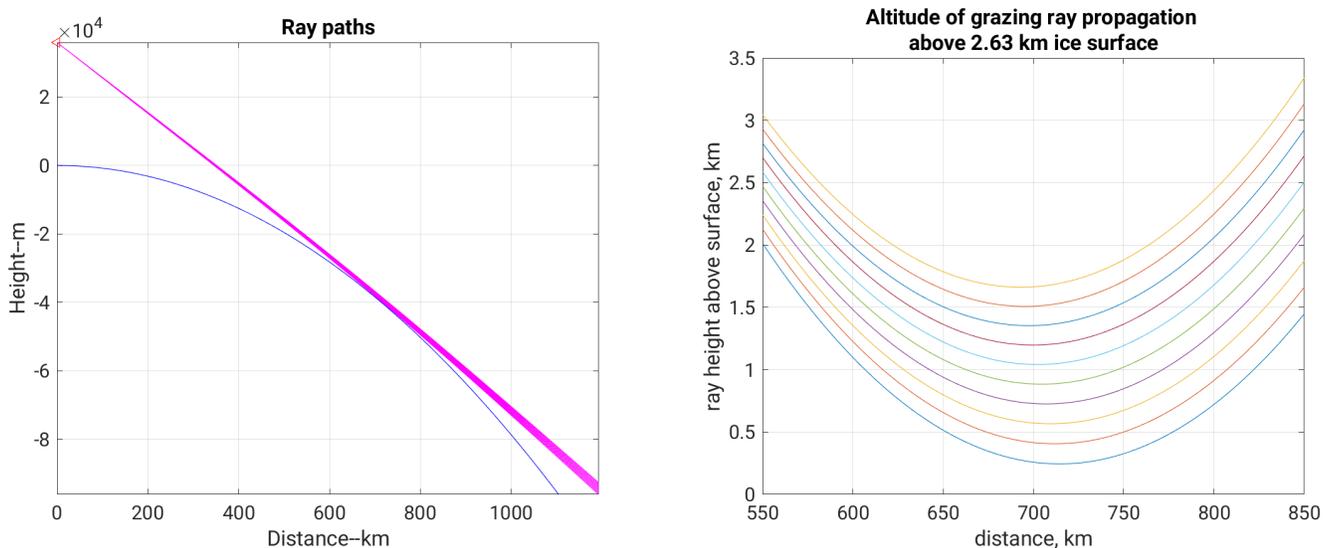


FIG. S11: *Left: Ray paths considered here. Right: altitude of rays with respect to ice surface around the location of closest approach.*

The variations shown in Fig. S10 constitute a phase screen for the propagation of radio wavefronts near the surface. It is evident that the largest variations occur in the last 1.2-1.5 km of the atmosphere above the ice surface at South Pole (altitude 2.8 km).

Occultations of Global Positioning System (GPS) Satellites by the horizon are now routinely observed in many experiments for purposes of atmospheric and ionospheric sounding [S33]. These observations track the carrier signals of the GPS transmitter on a satellite as it disappears below the horizon, and by measuring the phase and phase distortion, can recover detailed profiles of the refractive index of atmosphere, averaged over the region of propagation of the GPS radio signal. Radio rays that approach within several Fresnel-zone radii of the surface are subject to defocusing of the ray bundles due to the near-surface gradient of the index of refraction and multipath effects, both of which can distort the signal. In GPS occultations, the efficiency for detecting signals over Antarctica has been found to drop below 50% for ray impact parameters that fall below 1 km [S34]. The GPS L2 signal at 1227 MHz, which is tracked in both phase and amplitude during GPS occultations, is a factor of 4 in frequency above the  $\sim 300$  MHz mid-band for ANITA CR detection; thus the corresponding Fresnel zone radii are a factor of two smaller, indicating that such effects may be important at even larger surface impact parameters than for GPS signals.

To get an idea of the possible effects of the Tropospheric phase screens, we assume a frozen-in pattern over several hundred km, and estimate the dispersion of the delay across the wavefront for propagation over this distance. To determine the appropriate distance for this propagation, we use refractive radio ray tracing with geometric parameters appropriate to event 72164985, which is closest to the horizon in our data. A set of refracted ray paths, and their corresponding altitudes around the location of closest approach is shown in Fig. S11.

From these results, using the rays in the 250-500 m minimum altitude, we take a near-surface propagation distance of 200 km to be typical for the region with the largest variations. Fig. S12(Left) shows a time series of the twice-daily delay dispersion values (the root-mean-square deviations of the wavefront delay in ns) over December 2016, the period of the ANITA-IV flight. Although these are only measured at a single location, South Pole station is on a representative portion of the Antarctic ice sheet, and it is reasonable to assume that such variations are representative of other regions at similar altitudes, as are the locales of our near-horizon events.

From Kirchoff physical optics theory, knowledge of the wavefront distortion, in this case in the form of delay dispersion, can provide an estimate of the loss of intensity due to such distortion, given the wavenumber  $k = 2\pi f/c$  of the radio signal. For our conditions, the intensity loss factor  $I/I_0$  is given by

$$I/I_0 = \exp(-k^2\sigma_\phi^2) \quad (\text{S7})$$

where in our case the wavefront phase dispersion parameter  $\sigma_\phi = f\sigma_\tau$  and  $\sigma_\tau$  is the observed delay dispersion in seconds.

In Fig. S12(Right) we show the average loss vs. frequency for ANITA's band for rays in this grazing propagation regime, along with a  $\pm 1\sigma$  band indicating the variations from day to day. The average intensity loss of 6 dB in the

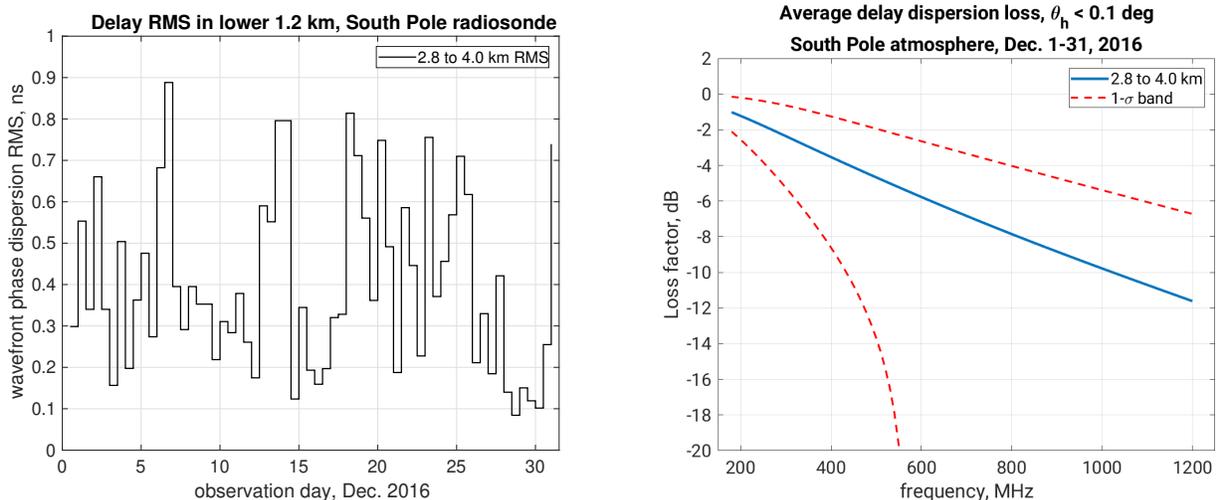


FIG. S12: *Left: daily average group-delay RMS across the lowest 1.2 km above South Pole station for December 2016, the period of the ANITA-IV flight. Right: Average phase-screen propagation losses for events with ray paths close to the horizon, assuming the observed South Pole atmosphere data and event parameters similar to 72164985.*

mid-band is substantial, and does not appear consistent with what is observed in the near-horizon events, which do not show a significant loss of power at high frequencies.

For this reason it appears appropriate to add an effective buffer zone of order  $0.1^\circ$  for near-horizon propagation, such that, if these events are to arise from misreconstructed over-the-horizon CR interactions, their true direction is likely to be at least  $0.1^\circ$  above the horizon.

### 3. Monte Carlo estimate of overall $p$ -value.

We have now considered all three background sources, anthropogenic, polarity flips, and mis-reconstructed above-horizon normal CRs. For the latter, we have assessed the potential effects of near-horizon propagation. While these effects do not preclude this background, they indicate caution in assessing the angular range of the misreconstruction.

To estimate the chance that all of these backgrounds taken in any combination with the 29 observed CR and CR-like events could produce the four anomalous events by accident, we use a Monte Carlo method again, now using statistical random variables to simulate the three types of backgrounds, according to their individual distributions.

For the anthropogenic background, we draw from the actual measured distribution (not shown here), which had the mean values quoted above, to generate realizations of a Poisson mean, which we then draw on via a Poisson random number generator. For the polarity flip background, we test a uniform random variable with a threshold defined by the chances determined in the polarity analysis for the 29 CRs. Finally, to assess the chance of near-horizon misreconstruction, we use Gaussian random variables with the appropriate parameters for each event, allowing its position to vary. We do this for two cases, one in which the event can propagate on rays that skim the horizon, and another case where the central ray must be at least  $0.1^\circ$  above the horizon.

The results of this Monte Carlo are shown in Table II. We consider several variations of the results, two for the size of the buffer used for the near-horizon propagation analysis, and two for which mean of the polarity flip, geometric or arithmetic, is used to determine the chance. For each result here we also report the mean of the Poisson fitted to the resulting distribution of typically  $10^5$  trials per CR per configuration.

## V. DISCUSSION

Our two prior reports of anomalous non-inverted events involved apparent cosmic-ray-like signatures observed at relatively steep arrival angles at the payload, about  $-27^\circ$  and  $-35^\circ$ , far below the apparent horizon. ANITA-IV data do not show any further such steeply arriving non-inverted events, although the polarity for 88992443, arriving from  $-14.8^\circ$  is indeterminate due to its low SNR and incommensurate waveform with the other CR. ANITA-IV does observe four non-inverted events close to, but below the horizon: 4098827, 19848917, 50549772, and 72164985.

TABLE II: Table of estimated  $p$ -values for background to produce the four near-horizon anomalous events.

Backgrounds assumed	Horizon buffer <sup>1</sup>	Pointing bias <sup>2</sup>	P-value	Normal statistics $\sigma$
(1) Anthropogenic non-inverted; (2) pointing errors lead to below-horizon misidentification; (3) polarity misidentification. → Combined chance of producing 4 or more non-inverted, below-horizon events	0.0	0.0	3.9e-4	3.36
	0.1	0.0	1.1e-4	3.69
	0.0	-0.1	3.7e-3	2.68
	0.1	-0.1	3.9e-4	3.36
	0.0	+0.1	1.2e-4	3.67
	0.1	+0.1	7.5e-5	3.8

Table notes:

1. Horizon buffer: minimum above-horizon distance required for a source
2. Pointing bias: uncorrected systematic offset in direction reconstruction (either sign is considered possible).

### A. Possible explanations for observed polarity.

The high degree of correlation of the anomalous events reported here with CR waveforms and parameters leads us to also consider the possibility that the polarity is correctly identified but has been inverted by some other natural effect. We consider here a number of non-typical physical effects which might lead to the observed polarity for a normal, reflected CR. We have discussed other effects in prior reports [S16], so this list only includes several possibilities that have not been investigated previously.

*a. Ice surface or subsurface structure or defects.* While it is possible to construct ice surface shapes in the form of multiple reflecting surfaces (for example, acting in a manner analogous to a corner reflector) such structure would require large scale fracturing and perhaps also uplift of ice, or a topographically mountainous region. Both effects can only appear in a very small fraction of the ice sheets and shelves and none of the events here appear in locales that present such effects. To explain the total of 6 non-inverted CR-like events from this flight and our other reports, out of about 70 total CRs, requires the fractional topographic area of such structures would have to approach 10% of the area of the continent observed, an improbable fraction for which there is no evidence in precise radar altimetry data.

It is also possible to get coherent reflections from subsurface layers below the ice surface on either the ice sheets or ice shelves; these effects are seen in radar reflections, and in some cases such layers may appear coherent over many km in radar tracks. There are several difficulties with using a subsurface layer reflection to invert the phase of a pulse. First, the dielectric contrast of the Fresnel coefficient from air to surface ice or snow will always be comparable to, or larger than, any subsurface dielectric contrast between layers, and thus causality favors the arrival of the initial reflection before any subsequent reflection. ANITA does not average pulses temporally, and the polarity of the leading portion of the electric field of each waveform is primarily what determines the polarity. We do often see trailing electric field components in CR events, and these may well be due to subsurface reflections, or just due to effective optical aberration of the reflecting surface over the first Fresnel zone, but the polarity is not affected by these components.

*b. Concave ice surface ellipsoids.* Coherent ice surface structure over multi-kilometric scales can produce a phase inversion under one scenario: if a region of the ice surface of order a Fresnel zone diameter in projection forms a coherent optical surface in the shape of an ellipsoidal reflector, with one focus of the ellipse beyond the distance to ANITA, then ANITA would observe a non-inverted reflected pulse, since the optics itself inverts the signal phase. We have inspected the ice surface features in the locale of each event, and we do not find any coherent structure that is remotely congruent to this requirement. Since the Fresnel zones for reflection in the ANITA-IV anomalous events are many km in scale, the required coherence region is also quite large, and would require concordance of many thousands of modes of the typical surface features on the ice sheets.

*c. Coherent backscattering from CRs.* We have considered a scenario where ANITA is actually observing a coherent backlobe from a CR arriving from the opposite direction from what is implied by the measured direction, which would lead to a relative inversion of the signal phase. We have simulated possible configurations where this might apply using finite-difference time domain methods via the Remcom XF7 software suite, and we can find no plausible surface state or structure that could lead to a significant chance (of order 7%) of this phenomenology.

*d. Stopping radiation.* Two previous anomalous events from ANITA-I and ANITA-III arrived at steep angles. At these angles, it has been suggested [S22] that a cosmic-ray shower which terminates at the ice surface may develop a form of transition radiation with geomagnetic correlation, leading to a pulse which mimics a non-inverted CR radio pulse under some conditions. However, this hypothesis has not yet been confirmed by observations, and in any case would evidently not apply to the four near-horizon events observed here, for which the proposed geometry, which

requires steeper angles, will not work.

### B. Particle-physics-based explanations for the observed near-horizon events.

The four anomalous events observed in ANITA-IV do not stand in their own right at a discovery-level significance. However, in each of the three of four ANITA flights, ANITA-I, -III, and -IV, which included a dedicated trigger for CR detection, CR-like events with anomalous non-inverted polarity have been detected, in each case with estimated background probabilities of  $10^{-2}$  or less. We are thus motivated to consider whether all of these events arise from possible physics signals.

In prior work on ANITA's observed steeply upcoming non-inverted cosmic-ray like events [S16], we considered the possibility that neutrino interactions in the ice sheets could lead to secondary  $\tau$  lepton decays in the lower atmosphere, then producing upcoming air showers. This scenario naturally explains the CR-like characteristics of the radio signals, but suffers from the extreme attenuation that neutrinos would undergo during propagation through the Earth at the angles observed.

In fact, this scenario is virtually ruled out for quasi-isotropic high energy neutrino fluxes with Standard Model cross sections, as the implied neutrino flux is in tension with other constraints by IceCube and Auger [S23].

In the case of the current results, however, an SM neutrino-based origin is harder to rule out by the same argument. At the angles of the observed near-horizon anomalous events,  $\tau$  leptons have a  $\sim 1\%$  chance of exiting the ice for original neutrino energies of  $\sim 10^{18-19}$  eV, entering the Earth along the implied chord of these directions. While this exit efficiency appears low, it is not so low as to exclude a neutrino origin for these events, particularly if the source was transient, as is often the case in high energy astrophysics.

It is also possible that such events could be produced by direct air showers from penetrating particles that enter the atmosphere from beyond the horizon, but do not interact until they are within the payload horizon. The  $\sim 1^\circ$  opening angle of radio emission from a CR air shower can lead to a slightly upcoming radio beam relative to the parent particle track that is slightly downgoing. Neutrinos interacting in the atmosphere are candidates for such a process, but the column depth of the atmosphere is far less than that of the ice and Earth, and should disfavor this process statistically. In-ice interactions, producing Askaryan emission, as an explanation for these events is also disfavored due to the high degree of horizontal polarization of these events, which severely restricts the phase space for escape of suitably polarized emission from below the ice.

## VI. COMPARISON OF ANITA-IV TO ANITA-III.

ANITA-IV used low-noise amplifiers with a noise temperature of 65 K compared to 100 K for ANITA-III. For both experiments, the combination of ice and cold sky in the  $\sim 50^\circ$ -diameter primary beam of the quad-ridged-horn antennas (along with Fresnel effects at the ice surface) leads to an estimated antenna temperature of about 115 K, and the resulting system temperature for ANITA-IV is of order 20% lower than for ANITA-III, a substantial improvement.

ANITA-IV's trigger system also required a coincidence of both left- and right-circular polarization (LCP/RCP), thus limiting triggers to highly linearly polarized signals with uniform sensitivity to any plane-of-polarization. Under ideal conditions, this trigger feature was therefore equally sensitive to both neutrino signals from the Askaryan effect, and of the almost completely horizontally polarized CR signals in Antarctica. This also led to a modest improvement of 5% in the threshold for event detection.

Fig. S13 shows distribution of arrival angles for the ANITA-III (A3) and ANITA-IV (A4) CRs. In the top pane, the event distributions are shown, and in the bottom, the normalized cumulative distribution function for the two flights is shown. In the final analysis, ANITA-III detected 28 CR events compared to ANITA-IV's 29 events, suggesting that the two flights had similar sensitivity to CRs.

ANITA-III suffered a significant loss of exposure due to radio frequency interference, in two  $\sim 50$  MHz bands centered at 260 MHz and 380 MHz. This led to an effective livetime that was less than half of the ANITA-IV livetime. The use of tunable notch filters in ANITA-IV to overcome the high levels of RF interference seen by ANITA-III improved the livetime, but led to a loss of effective bandwidth, since these filters remained in place most of the flight due to limitations of the control system. The decision to use these filters was made based on simulations of neutrino sensitivity, which was not adversely impacted by the filters since the spectral content of Askaryan events is very broadband. However, the loss of low-frequencies in ANITA-IV did decrease the sensitivity to CR events, which are low-frequency dominant.

The effect of these differences in payload response is illustrated in Fig. S14, where we show histograms of the distributions of CRs for each flight. For ANITA-III we use a subsample here of 20 events with a more complete

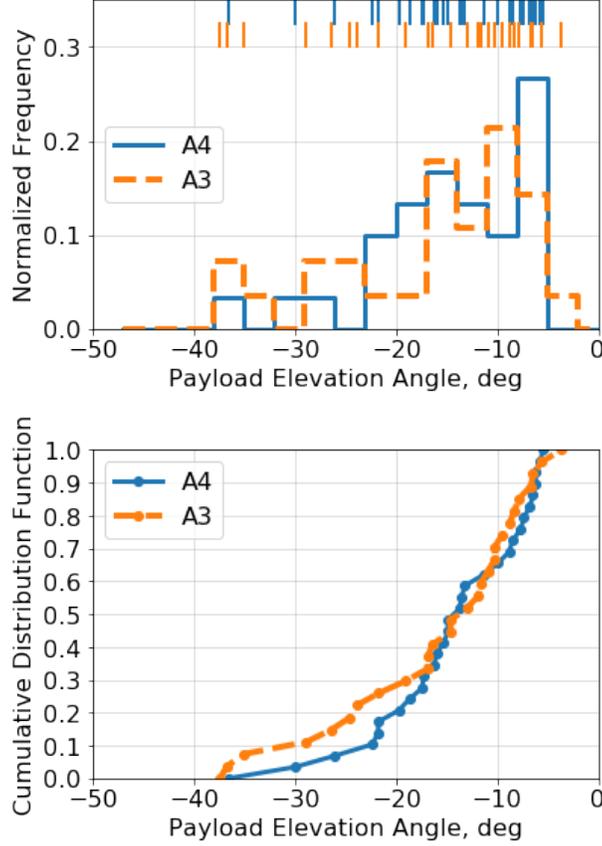


FIG. S13: *Top: Event elevation angle distributions as seen from the payload at typical 39 km altitude. Bottom: normalized cumulative distribution function for the two event samples.*

calibration, since the determination of spectral parameters is sensitive to this. In each event we fit the measured radio amplitude spectral density for an exponential

$$A(f) = A_{0.3} \exp \gamma(f - 0.3)$$

where  $A(f)$  is the frequency-dependent amplitude spectral density (the square root of power spectral density) as a function of frequency  $f$  in GHz,  $A_{0.3}$  is the amplitude at 0.3 GHz, and  $\gamma$  is the exponential slope. The upper plot shows the distribution of  $A_{0.3}$ , along with a log-normal distribution fitted curve for each set of amplitudes, with the fitted mean in the legend. ANITA-IV shows a 20% lower turn-on amplitude for the CRs, consistent with the lower noise figure of the receivers, and the improved trigger.

The lower pane of Fig. S14 shows the distribution of the exponential slopes, fit in this case to an extreme-value distribution (also known as Gumbel distribution), which is appropriate for a parameter which in this case is constrained to be  $\gamma \leq -1.0$  in theory[S10]. The distribution of fitted slopes exceeds this in some cases due to statistical variation in the spectral fitting. The means of the two fitted distributions differ by  $> 3.2\sigma$ , with ANITA-IV tending toward much shallower slopes, giving more high-frequency power in the spectral density.

The slope distribution for ANITA-IV is consistent with CR events being observed closer to the peak of the angular intensity (typically a  $\sim 1^\circ$  cone about the incoming CR direction) than ANITA-III. This was expected due to the loss of low-frequency power due to the always-on notch filters. We can translate the mean exponential slope into a typical angular width about the cone, resulting in an implied annulus of angular width  $\Delta\theta \simeq 0.1^\circ$  for ANITA-IV and  $\Delta\theta \simeq 0.25^\circ$  for ANITA-III. Since the acceptance solid angle scales with this quantity, we conclude that ANITA-III had an acceptance which was on the order of  $2.5\times$  that of ANITA-IV.

The ANITA trigger is best modeled as a cut on integrated power over a wide bandwidth, which implies that the lowest  $A_{0.3}$  events seen by ANITA-IV would also have the lowest exponential slope, which is observed. This correlation means that we can similarly compare the typical energy of ANITA-IV CRs to ANITA-III by translating the mean  $A_{0.3}$  and  $\gamma$  for each flight into a typical CR energy, using the prescription in reference [S10]. This, in fact, indicates that the ratio of the typical CR energy between ANITA-III and ANITA-IV is  $E_{A4}/E_{A3} \simeq 1$ , and thus we would not

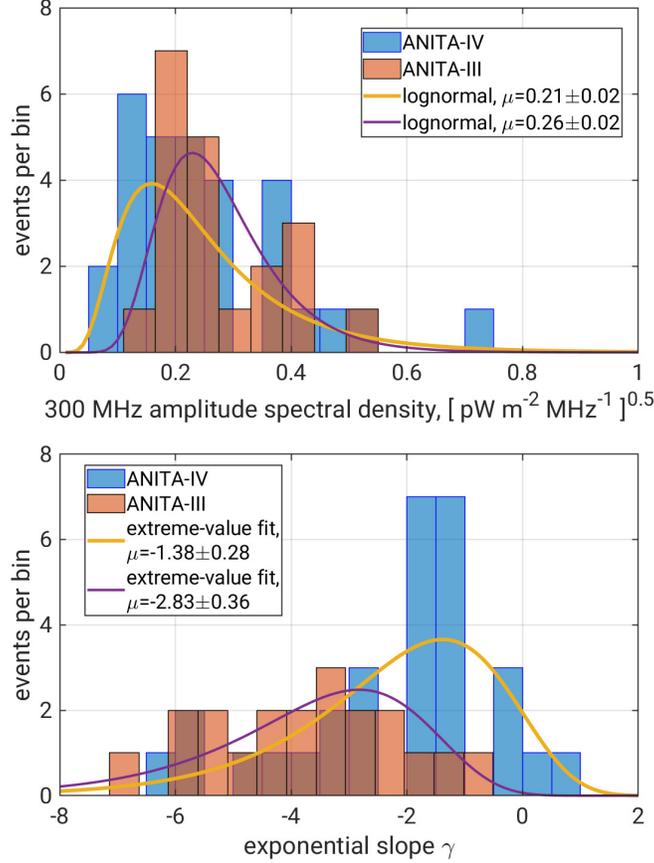


FIG. S14: *Top: Distribution of amplitude spectral density at 300 MHz of ANITA-III and -IV CRs, with fitted curves. Bottom: Distribution of radio amplitude spectral slopes for the same two event samples, with fitted curves.*

expect any increase in the number of CRs due to the cosmic ray flux. Thus, the larger acceptance of ANITA-III is essentially canceled by its much lower livetime relative to ANITA-IV.

The acceptance for detection of CRs depends crucially on how much of this radio emission cone is visible, and it appears that ANITA-IV, while having more sensitivity overall, had substantial restrictions on its CR acceptance. We conclude that the lack of detections of steeply upcoming events in ANITA-IV do not pose any significant constraints on the type of anomalous event observed in ANITA-III. Conversely, it appears that three of the four near-horizon anomalous events observed in ANITA-IV would have likely been below ANITA-III's detection threshold. So given the low statistics of the anomalous events, we do not have clear evidence of any tension yet between the results of the two flights.

#### Appendix A: CR and CR-like event waveforms.

The following five figures show the four different realizations of waveforms for each of the 30 CR and CR-like events, including the event 17904564 for completeness, although it is not part of the CR sample proper. In each case the CSW is shown with the corresponding impulse response overlain, followed (in row order) by the CLEAN waveform, the corresponding CLEAN components, the Allpass deconvolved waveform, a cross-correlation of the CSW and impulse response, and the Wiener-deconvolved waveform.

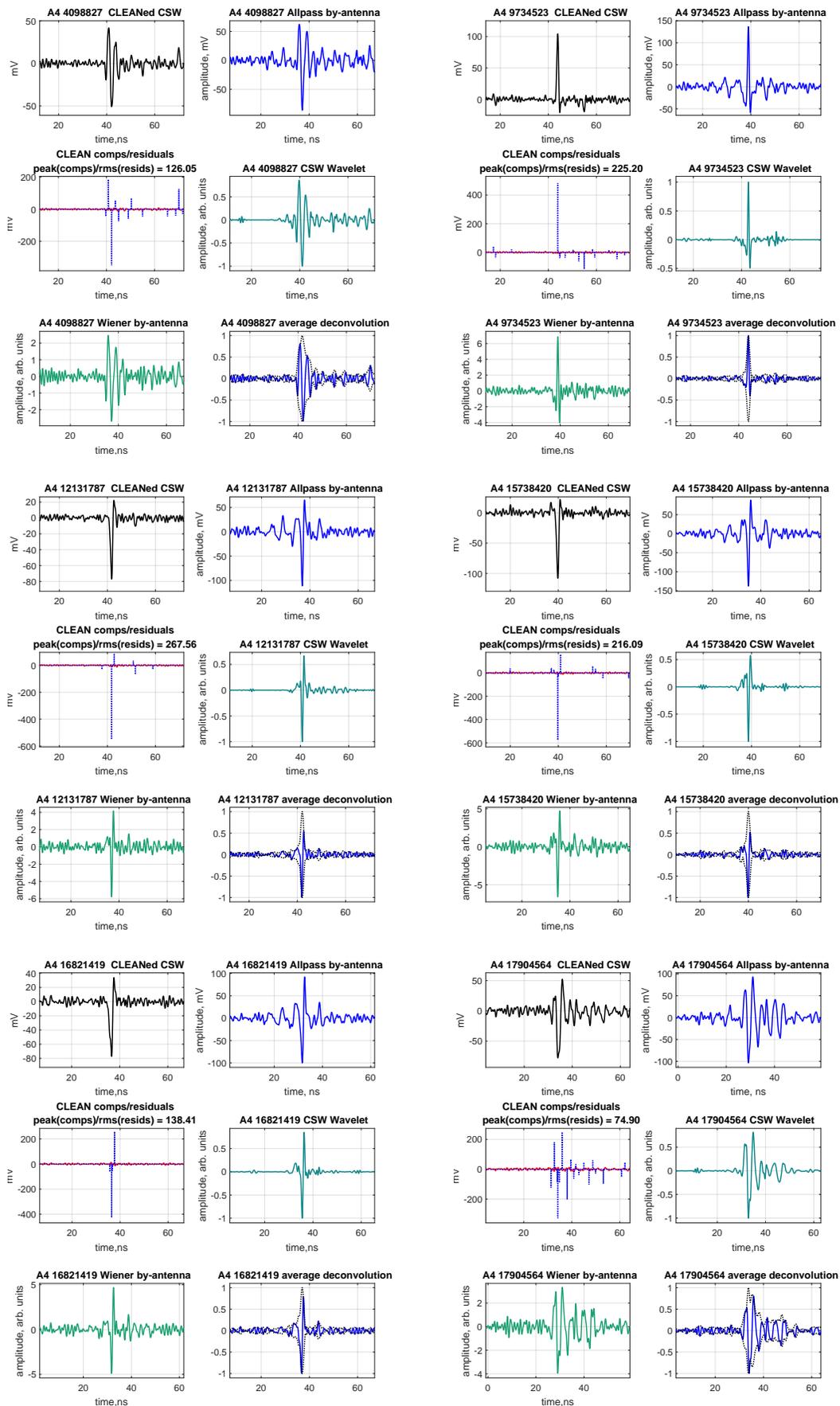


FIG. S15:

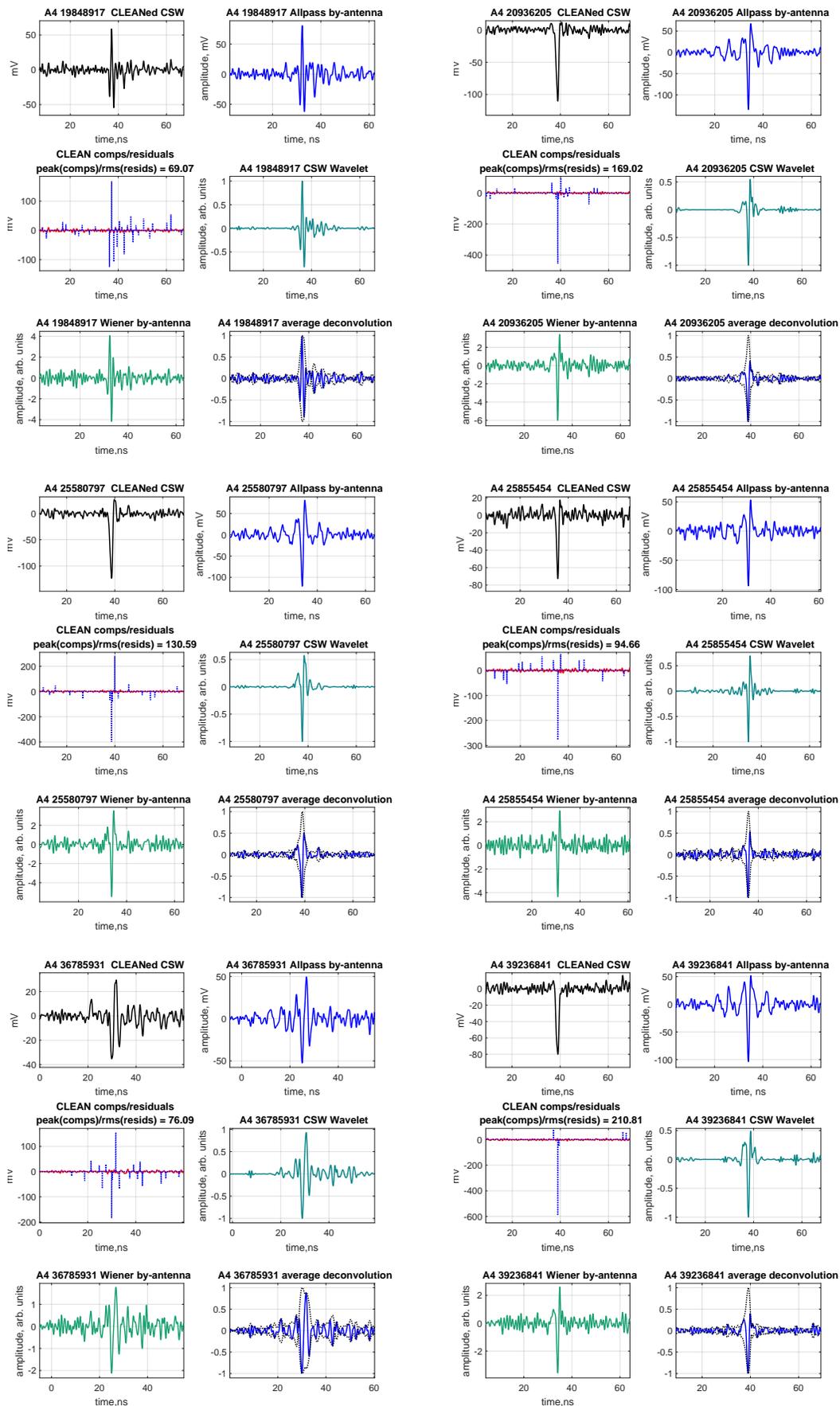


FIG. S16:

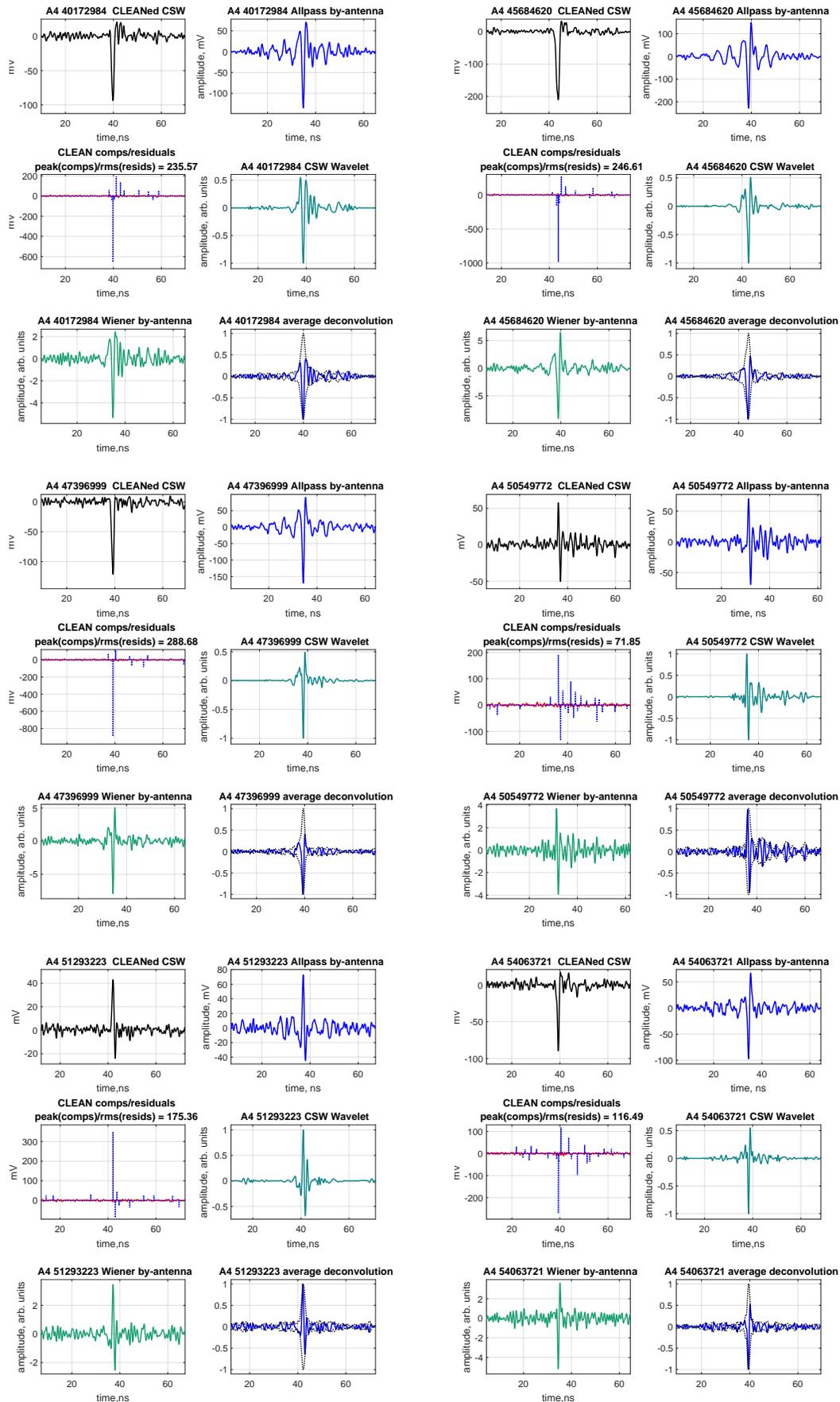


FIG. S17:

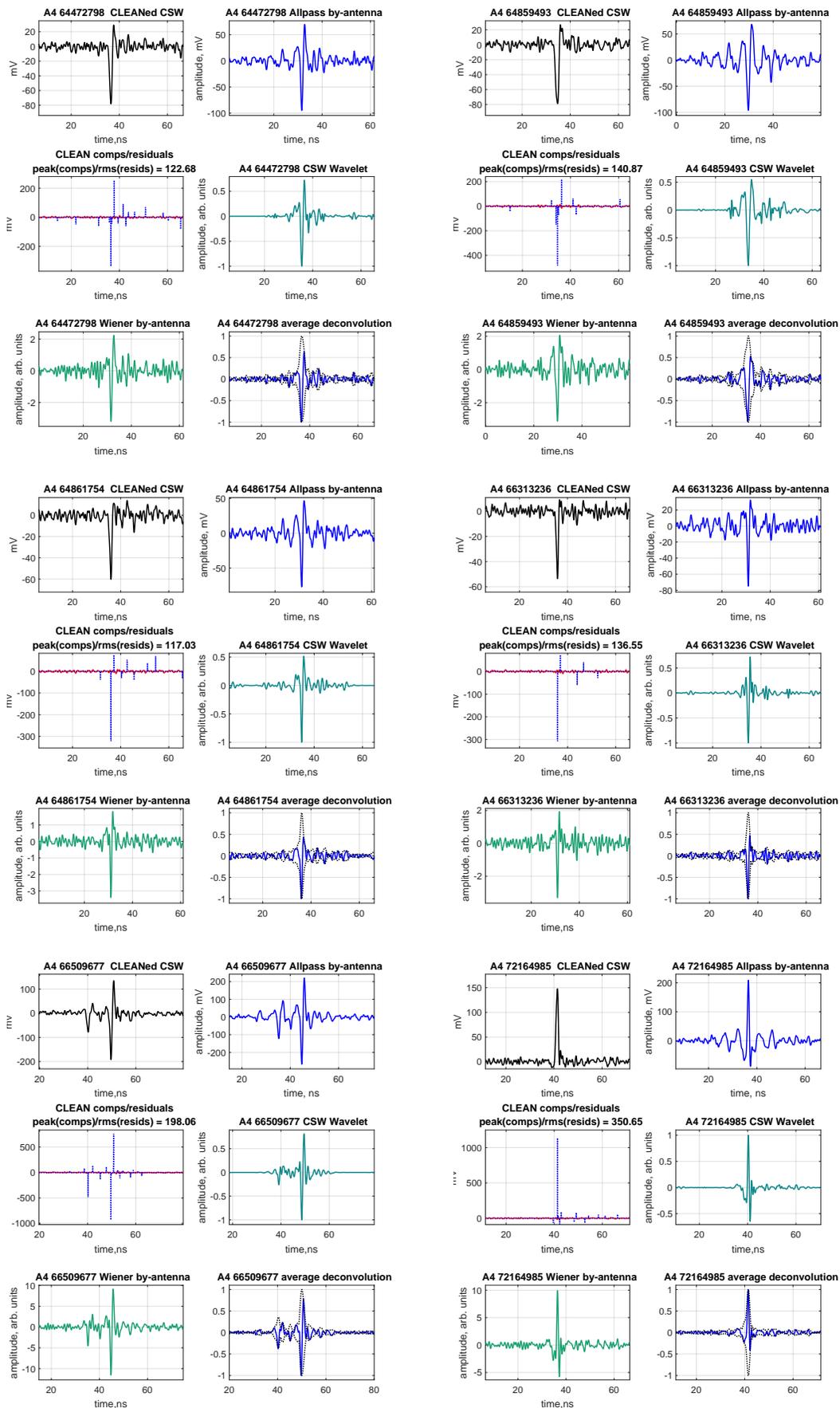


FIG. S18:

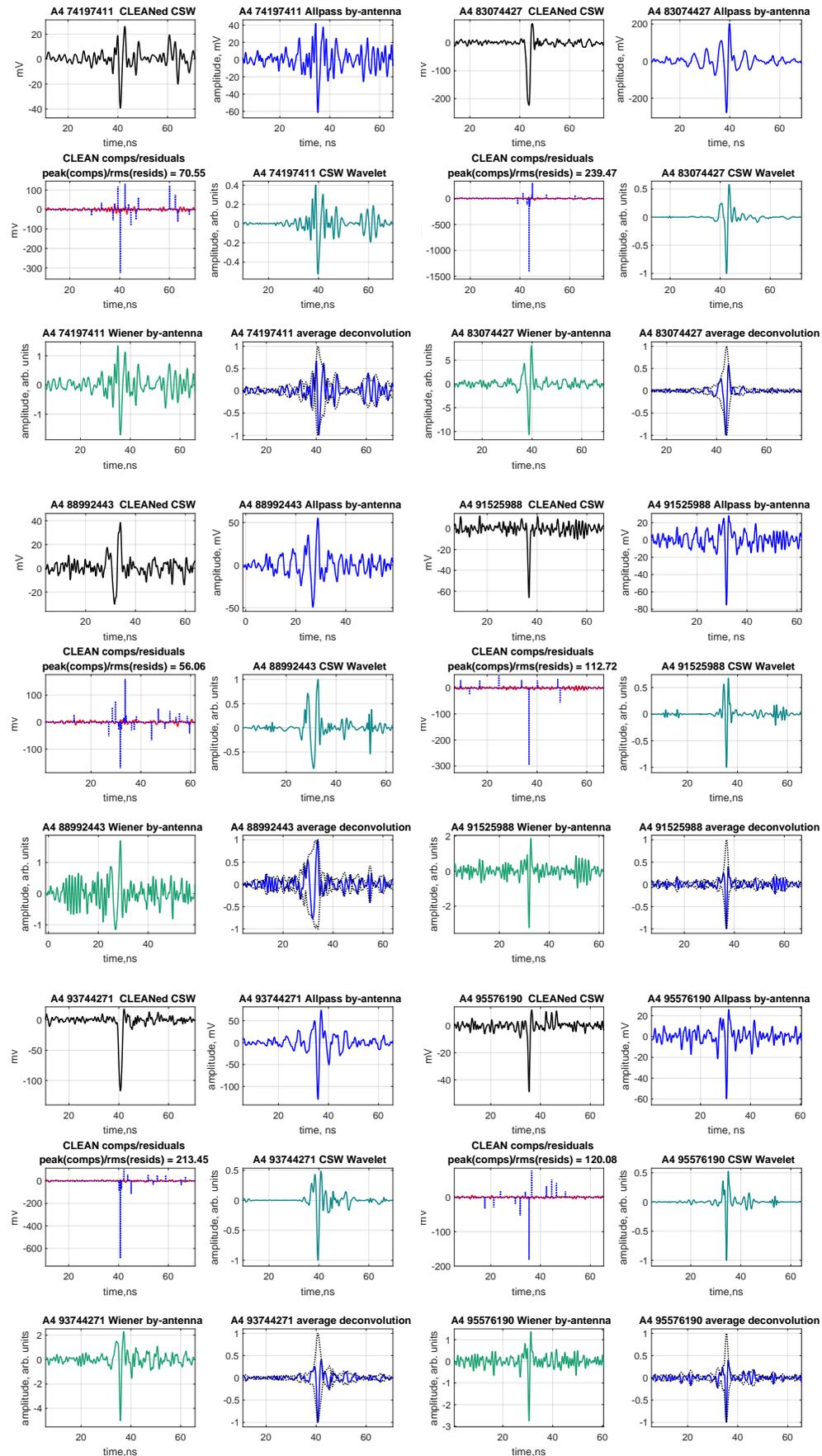


FIG. S19:

- 
- [S1] “First Experimental Characterization of Microwave Emission from Cosmic Ray Air Showers,” R. Smída *et. all*, Phys. Rev. Lett. 113, 221101, 2014.
- [S2] G. A. Askaryan, Excess Negative Charge of an Electron-Photon Shower And Its Coherent Radio Emission, *JETP* **14**, 441 (1962); also *JETP* **21**, 658 (1965).
- [S3] P. Gorham, *et al.* [ANITA Collaboration], “Radio-detected Anomalous Cosmic Rays Observed by ANITA-IV,” Phys. Rev. Lett., in preparation, 2019. [companion paper to current report]
- [S4] D. Saltzberg, *et al.* Observation of the Askaryan Effect: Coherent Microwave Cherenkov Emission from Charge Asymmetry in High-Energy Particle Cascades, Phys. Rev. Lett. 86 (2001) 2802.
- [S5] The ANITA Collaboration, P. W. Gorham, *et al.* [ANITA Collaboration], “Observations of the Askaryan Effect in Ice,” Phys. Rev. Letters 99, 171101 (2007); <http://arxiv.org/abs/hep-ex/0611008>.
- [S6] B. Rotter, unpublished dissertation fulfilling PhD requirements, Univ. of Hawaii at Manoa, 2017.
- [S7] Gorham, P. W. *et al.*, The Antarctic Impulsive Transient Antenna ultra-high energy neutrino detector: Design, performance, and sensitivity for the 2006-2007 balloon flight. *Astropart. Phys.* **32**, 10-41 (2009).
- [S8] S. Hoover *et al.* [ANITA Collaboration], Phys. Rev. Lett. 105, 151101 (2010).
- [S9] Constraints on the ultra-high energy cosmic neutrino flux from the fourth flight of ANITA, P. W. Gorham, *et al.* [ANITA collaboration], Phys. Rev. D 99, 122001 (2019).
- [S10] H. Schoorlemmer, *et al.*, [ANITA collaboration], “Energy and Flux Measurements of Ultra-High Energy CRs Observed During the First ANITA Flight” *Astropart. Phys.* 77, 32,(2016); also arVix:1506.05396.
- [S11] Porter, N. A., Long, C. D., McBreen, B., Murnaghan, D. J. B. & Weekes, T. C., Phys. Lett. 19, 415-417 (1965); Vernov, S. N., Abrosimov, A. T., Volovik, V. D., Zalyubovskii, I. I. & Khristiansen, G. B., *Pisma v ZhETF* 5, 157-162 (1967) [*Sov. Phys. JETP Letters* 5, 126-130 (1967)]; Barker, P. R., Hazen, W. E., & Hendel, A. Z., Phys. Rev. Lett. 18, 51-54 (1967); Fegan, D. J. & Slevin, P. J., Nature 217, 440-441 (1968); Hazen, W. E., Hendel, A. Z., Smith, H., & Shah N. J., Phys. Rev. Lett. 22, 35-37 (1969); Hazen, W. E., Hendel, A. Z., Smith, H., & Shah N. J., 24, 476-479 (1970); Spencer, R. E., Nature 222, 460-461 (1969); Fegan, D. J. & Jennings, D. M., Nature 223, 722-723 (1969); Allan, H. R., *Progress in Elementary Particles and Cosmic Ray Physics*, 10, edited by Wilson, J. G. & Wouthuysen S. G. (North-Holland, Amsterdam, 1971), 171-304.
- [S12] Ardouin, D. *et al.*, *Astropart. Phys.* 31, 192-200 (2009); Nehls, S. *et al.*, Nucl. Instrum. Meth. A589, 350-361 (2008); W.D. Apel, *et al.* (LOPES Collaboration) *Astropart. Phys.* 32, 294, (2010), Anna Nelles *et al.*, Proceedings of the 33rd International Cosmic Ray Conference, (2013), The Pierre Auger Collaboration, Proceedings of the 33rd International Cosmic Ray Conference, (2013)
- [S13] K. Belov *et al.*, “Accelerator measurements of magnetically-induced radio emission from particle cascades with applications to cosmic-ray air showers,” Phys. Rev. Lett. 116, 141103 (2016); arXiv:1507.07296.
- [S14] P. W. Gorham *et al.* [ANITA Collaboration], Characteristics of Four Upward-pointing Cosmic-ray-like Events Observed with ANITA, Phys. Rev. Lett. (2016), 117, 071101, <http://arxiv.org/abs/1603.05218>
- [S15] P. Allison *et al.* [ANITA Collaboration], “Constraints on the diffuse high-energy neutrino flux from the third flight of ANITA,” Phys. Rev. D 98, 022001 (2018), <https://arxiv.org/abs/1803.02719>
- [S16] P. Gorham, *et al.* [ANITA Collaboration], “Observation of an Unusual Upward-going Cosmic-ray-like Event in the Third Flight of ANITA,” Phys. Rev. Lett. 121, 161102 (2018), <https://arxiv.org/abs/1803.05088>.
- [S17] P. W. Gorham *et al.*, [ANITA collaboration], Constraints on the ultra-high energy cosmic neutrino flux from the fourth flight of ANITA, Phys. Rev. D 99,063011 (2019).
- [S18] Alvarez-Muñiz, J.; Carvalho, W. R.; Zas, E., *Astroparticle Physics*, 35, 325-341. (2012); Ludwig, M., Huege, T., *Astroparticle Physics* 34, 438-446 (2011); Vincent Marin, Benoit Revenu, NIM A. 662, Supplement 1, 11 Pages S171 – S174, (2012) Werner, Klaus; Scholten, Olaf, *Astroparticle Physics*, 29, 393-411. (2008)
- [S19] J. Alvarez-Muñiz, W. Carvalho, A. Romero-Wolf, M. Tueros, and E. Zas, “Coherent radiation from extensive air showers in the ultrahigh frequency band,” Phys. Rev. D, 86, 123007 (2012).
- [S20] A technical description of atmospheric sounding by GPS occultation, G.A. Hajj, E.R. Kursinski, L.J. Romans, W.I. Bertiger, S.S. Leroy, *Journal of Atmospheric and Solar-Terrestrial Physics* 64 (2002) 451–469.
- [S21] Josep M. Aparicio<sup>1</sup>, Estel Cardellach, and Hilda Rodriguez, Information content in reflected signals during GPS Radio Occultation observation, *Atmos. Meas. Tech.*, 11, 1883-1900, 2018.
- [S22] Coherent transition radiation from the geomagnetically-induced current in cosmic-ray air showers: Implications for the anomalous events observed by ANITA, Krijn D. de Vries, Steven Prohira, submitted to PRL; [arxiv.org/abs/1903.08750](http://arxiv.org/abs/1903.08750).
- [S23] A. Romero-Wolf, *et al.*, [ANITA collaboration], A comprehensive analysis of anomalous ANITA events disfavors a diffuse tau-neutrino flux origin, arXiv:1811.07261v3, 2019.
- [S24] P. Gorham *et al.* [ANITA Collaboration], Phys. Rev. Lett. 117, 071101 (2016).
- [S25] P. Gorham *et al.* [ANITA Collaboration], Phys. Rev. Lett. 121, 161102 (2018).
- [S26] Peng Cao, PhD dissertation, Univ. of Delaware, 2018.
- [S27] A. Romero-Wolf, S. Hoover, A.G. Vieregg, *et al.*, *An interferometric analysis method for radio impulses from ultra-high energy particle showers*, *Astropart. Phys.* **60**, 72, (2015).
- [S28] F. Halzen & D. Saltzberg Phys. Rev. Lett. 81, 4305, (1998); V. Bugaev *et al.* *Astropart. Phys.* 21, 491, (2004); O. Blanch Bigas, *et al.* UHE tau neutrino flux regeneration while skimming the Earth, Phys. Rev. D 78:063002, (2008).

- [S29] “Calculation of High Energy Neutrino-Nucleon Cross Sections and Uncertainties Using the MSTW Parton Distribution Functions and Implications for Future Experiments,” A. Connolly, R. S. Thorne, D. Waters, Phys. Rev. D 83:113009, 2011.
- [S30] L. A. Anchordoqui *et al.*, Phys. Rev. D 74, 043008, (2006); N. Armesto, C. Merino, G. Parente, E. Zas, Phys. Rev. D 77, 013
- [S31] P. Allison *et al.*, [ANITA Collaboration], “Dynamic tunable notch filters for the Antarctic Impulsive Transient Antenna (ANITA),” Nucl. Instr. Meth. A 894, 47 (2018). <https://arxiv.org/abs/1709.04536>.
- [S32] P. W. Gorham, *et al.* [ANITA Collaboration], “Antarctic Surface Reflectivity Measurements from the ANITA-3 and HiCal-1 Experiments,” J. Astron. Instrum. 06, 1740002 (2017), <https://arxiv.org/abs/1703.00415>.
- [S33] A technical description of atmospheric sounding by GPS occultation, G.A. Hajj, E.R. Kursinski, L.J. Romans, W.I. Bertiger, S.S. Leroy, Journal of Atmospheric and Solar-Terrestrial Physics 64 (2002) 451-469.
- [S34] Josep M. Aparicio<sup>1</sup>, Estel Cardellach, and Hilda Rodriguez, Atmos. Meas. Tech., 11, 1883-1900, 2018.
- [S35] S. Prohira, *et al.* “Antarctic Surface Reflectivity Calculations and Measurements from the ANITA-4 and HiCal-2 Experiments,” Phys. Rev. D 98,042004 (2018), <https://arxiv.org/abs/1801.08909>.
- [S36] M. G. Aartsen *et al.*, Phys. Rev. Lett 117, 241101 (2016).
- [S37] A. Aab *et al.* (Pierre Auger Collaboration), Phys. Rev. D 91, 092008 (2015)
- [S38] Greene, Chad A., *et al.* “Antarctic Mapping Tools for Matlab.” Computers & Geosciences, vol. 104, Elsevier BV, July 2017, pp. 151157, doi:10.1016/j.cageo.2016.08.003.
- [S39] Fretwell, P., *et al.* “Bedmap2: improved ice bed, surface and thickness datasets for Antarctica.” The Cryosphere 7.1 (2013). <http://dx.doi.org/10.5194/tc-7-375-2013>
- [S40] G.A. Hajj, E.R. Kursinski, L.J. Romans, W.I. Bertiger, S.S. Leroy, “A technical description of atmospheric sounding by GPS occultation,” Journal of Atmospheric and Solar-Terrestrial Physics 64 (2002) 451–469.
- [S41] J. Hanson, *et al.*, Radar absorption, basal reflection, thickness and polarization measurements from the Ross Ice Shelf, Antarctica, Journal of Glaciology, Vol. 61, No. 227, 2015.