Developing New Analysis Tools for Near Surface Radio-based Neutrino Detectors

Steven W. Barwick[a],* for the ARIANNA collaboration

[a]*University of California,*
Department of Physics and Astronomy, Irvine, CA USA

E-mail: sbarwick@uci.edu

The ARIANNA experiment is an Askaryan radio detector designed to measure high-energy neutrino induced cascades within the Antarctic ice. Ultra-high-energy neutrinos above 10^{16} eV have an extremely low flux, so experimental data captured at trigger level need to be classified correctly to retain as much neutrino signal as possible. We first describe two new physics-based neutrino selection methods, or "cuts", (the updown and dipole cut) that extend the previously published analysis to a specialized ARIANNA station with 8 antenna channels, which is double the number used in the prior analysis. The new cuts produce a neutrino efficiency of > 95% per station-year of operation, while rejecting 99.93% of the background (corresponding to 53 remaining experimental background events). When the new cuts are combined with a previously developed cut using neutrino waveform templates, all background is removed at no change of efficiency. In addition, the neutrino efficiency is extrapolated to 1,000 station-years of operation, obtaining 91%. This work then introduces a new selection method (the deep learning cut) to augment the identification of neutrino events by using deep learning methods and compares the efficiency to the physics-based analysis. The deep learning cut gives 99% signal efficiency per station-year of operation while rejecting 99.997% of the background (corresponding to 2 remaining experimental background events), which are subsequently removed by the waveform template cut at no significant change in efficiency. The results of the deep learning cut were verified using measured cosmic rays which shows that the simulations do not introduce artifacts with respect to experimental data. The data driven results in this paper demonstrate that the background rejection and signal efficiency of near surface antennas meets the requirements of a large scale future array, as considered in baseline design of the radio component of IceCube-Gen2.

---

*Speaker
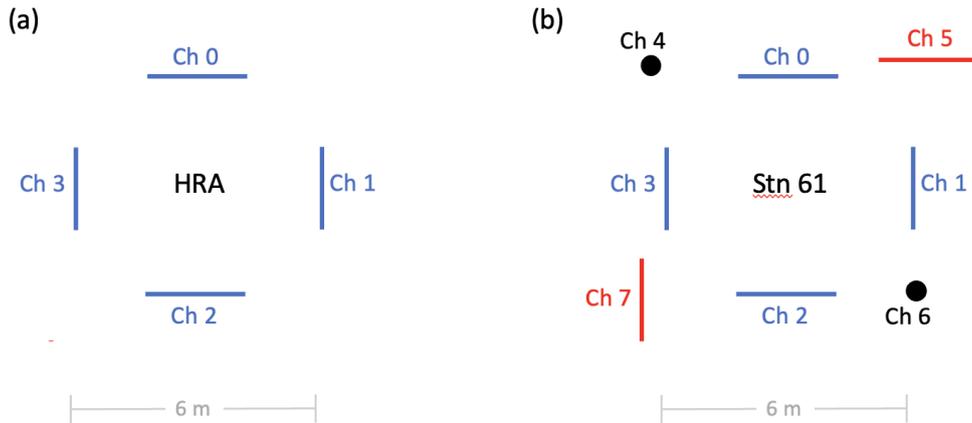
Figure 1: Overhead diagram of the station antenna layouts for (a) one of 7 identical ARIANNA HRA stations and (b) station 61. Black dots indicate fat dipole antenna, blue lines refer to downward facing LPDAs, and red lines are upward facing LPDAs.

1. Introduction

Radio-based neutrino astronomy in the ultra-high-energy regime (UHE, $E_\nu > 10^{17}$ eV) has the potential to provide insight into the sources of cosmic rays with energies in excess of 10^{20} eV[1]. Neutrinos are ideal messengers because they have neutral charge and extremely low mass, so they travel through space without interacting with electromagnetic fields and have a low interaction probability with matter. The latter property requires a large, high density, target volume to detect the expected small neutrino flux. The ARIANNA experiment utilized ice as target material for neutrino interactions and it was designed to observe a short (ns) pulse of radio frequency emission. It is comprised of two components: a hexagonal array of pilot stations with locations in Moore's Bay, Antarctica which is on the Ross Ice Shelf, and two additional stations located at the South Pole, Antarctica. The array at Moore's Bay was augmented by several special purpose stations that investigated a variety of improvements to the baseline technology of the initial hexagonal array. All but one of the deployed hardware systems operated autonomously, with independent power and communication. A recent summary of previous results obtained by the ARIANNA experiment is provided in [2].

In 2020, the ARIANNA collaboration reported[3] a limit on the astrophysical neutrino flux and neutrino detection efficiency, ϵ, defined as the fraction of neutrino events observed by the detector that survive the analysis cuts, for the 7 station array of near-surface detectors (termed the HRA). One station of the HRA is schematically shown in panel (a) of Figure 1). The HRA operated reliably for 4.5 years before it was decommissioned. That study achieved a background-free search with $\epsilon = 0.79$ using a template matching procedure to search for neutrino candidates using data from just 4 downward facing LPDA. The templates were produced by convolving the bipolar electric field emitted by the neutrino-induced particle cascade with the amplifier and LPDA antenna responses for a given arrival direction at the detector. The analysis in that earlier work exploited the fact that most background processes produce waveforms with quite different time dependence in the highly directional LPDA antennas (termed the "LPDA cut" in this paper). The accuracy of the

neutrino waveform prediction was verified by the observation of EHE cosmic rays[4], which serve as a calibration beam due to their similarity to neutrino signals (after applying the antenna and amplifier response to both). However, the expected sensitivity of future radio based high energy neutrino detectors, such as IceCube-Gen2[5] are a factor of $\sim 10^4$ larger than demonstrated by the ARIANNA HRA array, which requires the development of additional analysis tools to meet the challenge of identifying neutrinos at high efficiency.

This paper is focused on developing new analysis tools, or "cuts", that improve the neutrino efficiency and then apply these cuts to data collected by a specialized ARIANNA station with 8 antenna receivers, station 61, deployed at the South Pole in 2018 (see panel (b) of Figure 1). This work describes three new cuts that augment the LPDA cut by incorporating the information from 4 additional antenna channels, which provide significant new information that simplifies the selection of neutrino events and rejection of background. First, we describe two physics-motivated cuts that, when combined with the previously developed LPDA cut, achieve $\epsilon > 90\%$ while residual background is reduced from 1 event per 20 station-years of operation to 1 event per 1000 station-years. Then a third cut is introduced that exploits deep learning tools. Deep learning networks offer the ability to train a neural network model on two distinct data sets to learn their distinguishing features. Additional details on the development, capabilities, and testing of these cuts can be found in Ref.[6].

2. Neutrino station 61

The experimental data were collected by station 61 during the Austral summer months between December 10, 2018 and January 10, 2021, resulting in a total livetime of 1 year. Events were removed during periods of human activity at the detector site. Due to the diffuse neutrino flux limits set by IceCube and the duration of data taking from this one station, no neutrinos are expected in the data set. Thus, the data comprise primarily of thermal noise events, wind generated noise events, and noise generated from the ARIANNA electronics. In total, there are 74,530 (corresponding to a time averaged trigger rate of $\sim 2 \times 10^{-3}$ Hz) experimentally triggered events. At the specified trigger threshold, events due to thermal fluctuations occur at a rate of $\sim 3 \times 10^{-4}$ Hz, so the majority of events collected by station 61 are non-thermal in nature.

Geographically, there are differences between station 61 and the HRA stations. Station 61 is located at an elevation of 2800 m, about 5 km from South Pole Station. Due to its proximity to the South Pole Station, station 61 observes more anthropogenic noise from weather balloon launches, spark plug noise from snowmobiles and other vehicles with motors, and communications to aircraft and other bases in Antarctica. HRA stations are located in a more remote location, so these sources of noise are comparatively small. However, the average winds are stronger at Moore's Bay, so wind generated radio noise from blowing snow particles is more prevalent. In addition, the electrical noise from the battery charging electronics on the ARIANNA station is greatest during the time of year when the position of the sun transitions from above the horizon to below the horizon. At Moore's Bay, this happens daily during sunrise and sunset for a period of about a month whereas at the South Pole, sunrise and sunset occurs only once per year.

NuRadioMC [7] was used to generate a set of expected neutrinos events for the ARIANNA detector. The events are randomly distributed in direction and follow a representative GZK neutrino
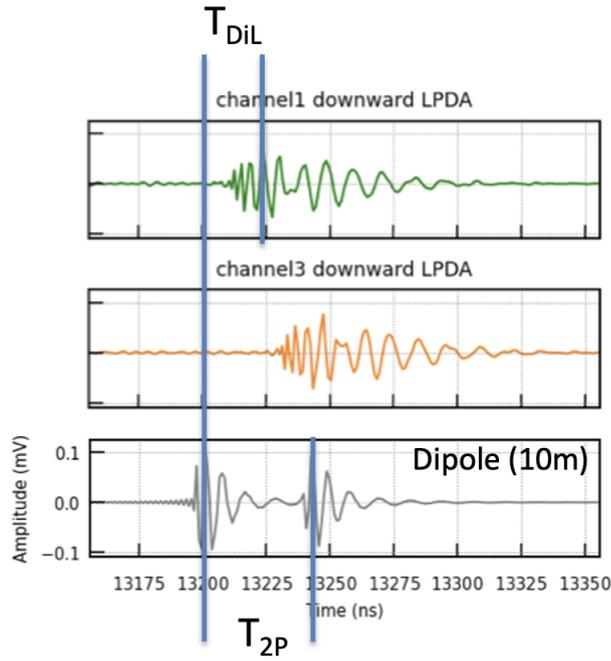
Figure 2: Neutrino waveforms observed by a vertically oriented fat dipole buried 10m beneath the snow surface (bottom waveform), and two parallel downward facing LDPAs (top and middle waveform). The relative timing between the fat dipole and earliest arriving LPDA signal, T_{DiL} is shown, as is the relative timing between the direct and reflected pulse in the dipole channel, T_{2P}.

flux distribution with energies from $10^{17}eV - 10^{19.5}eV$. There is also a weight cut of 10^{-5} performed so events with weights below this value are removed; since simulated neutrinos are generated uniformly in all directions, the weight cut removes lower probability neutrino events (corresponding to arrival angles below the horizon) that would not likely be measurable by ARIANNA. The radio pulses propagate from the interaction vertex through the South Polar ice to the antennas at the detector station. Then the ARIANNA station 61 detector is simulated with NuRadioReco [8] with the exception that the dipole at channel 4 is buried at a depth of 10m instead of the actual depth of 2.6m. With this change, the generated neutrino template for the dipole of channel 4 is inappropriate for a neutrino search (and no results are presented here), but the data from this experimental channel does provide insight on the fraction of background events that would satisfy the dipole cut in a future station where the dipole will be deployed at a greater depth[5], usually suggested between 10-15m. The resulting neutrino radio signals are simulated in all 8 channels by convolving the electric field pulses with the antenna response. Though the signals in the shallow dipole of channel 6 are properly modeled, they are ignored in this analysis.

ARIANNA station 61 was deployed with 8 antenna channels. In addition to the 4 downward LPDAs, the station includes: (1) two upward facing LPDAs to identify cosmic rays, wind, and anthropogenic backgrounds that propagate through the air, and (2) two fat dipoles to observe signals with vertical polarization. It is expected that a future station will include at least one dipole buried to a depth of ~10m to observe the distinct neutrino signature of a double pulse from a neutrino event (the second pulse is due to reflections off the air-firn surface and is only produced by radio emission

from within the ice). The double pulse structure from direct and reflected signal trajectories was cleanly observed in a dipole buried at 8.5m in another specialized ARIANNA station, station 52 [9]. The radio signal from the neutrino interaction that directly propagates to the antenna creates the first pulse of the dipole waveform. The time delayed second pulse is due to the reflection from the ice-firn interface at the surface. For shallow detector stations located at high elevations, this feature can be observed for most neutrino events because radio pulses from neutrino events typically reach at the snow surface with arrival angles (where 180 deg corresponds to vertically up-going propagation) between 90 deg and 130 deg, the largest angle corresponding to total internal reflection (TIR), and the amplitudes of the direct and reflected pulses are expected to be similar due to the symmetry of the response of the dipole antenna relative to the normal direction from the dipole axis. For arrival angles between 130deg - 143deg , reflection is not total, and the relationship between the amplitude of the direct pulse and second pulse is calculated from the Fresnel equations. Approximately 11.4% of the neutrino events fall into this category. The ratio of the amplitudes of the direct and delayed pulse depends on arrival angle, which is well measured by a shallow surface station [10], and to a lesser extent polarization, which is not included in this analysis. Finally, for arrival angles greater than 143 deg, a delayed pulse is ignored because the reflected amplitude is small. Fewer than 1.5% of the neutrino events arrive at these angles. Figure 2 shows an example of the wealth of new information in the dipole channel from relative timing of the waveforms generated by downward facing LPDAs and a fat dipole buried at a depth of 10m. The time difference between the two pulses in a dipole channel is T_{2P}, the signal time difference between the earliest arrival time in the LPDA and the first pulse in the dipole is T_{DiL}. Anthropogenic and cosmic ray radio signals that propagate through the atmosphere prior to entering the ice do not generate double pulse structures with time delays smaller than 100ns.

3. Description of new analysis cuts

3.1 Updown cut

Neutrino signals, by their nature, come from below the ice-firn interface in most cases. On the other hand, many types of backgrounds, including wind-induced and anthropogenic events, come from above the ice surface. Due to the directionality of the LPDA antennas, events from above the surface typically have a larger signal in the upward-facing antennas than downward-facing ones, and vice versa. By studying the ratio between the amplitudes in upward and downward facing antennas, a cut was developed to reject above-the-surface events while efficiently retaining below-the-surface events. It is relatively common for sources of pulsed radio background to have larger signals in the upward facing LPDA, whereas for neutrino signals, the pattern is reversed. However, the difference in signal strength between upward and downward facing LPDA is mitigated for neutrino signals arriving near the horizon because the antenna response is very similar at small angular differences, both positive and negative. Another mitigating effect concerns reflections at the ice-air interface. This step of the analysis procedure reduces the background event sample from 74,530 events to 41,821 events, while ϵ_{updown} >99% for the simulated neutrino signal.

3.2 Dipole cut

The dipole cut procedure requires a neutrino template waveform for the fat dipole antenna to cross-correlate with the observed waveform. The template encodes the expected time delay (T_{2P}) and fractional amplitude of the delayed pulse as a function of arrival angle. Finally, the relative time between the direct pulse in the dipole and earliest arriving signal in the LPDA channels (T_{DiL}) is computed and included in the template analysis by restricting the time window of cross-correlation calculation.

3.3 Deep Learning cut

In contrast to the more traditional techniques used to develop the previous selection criteria, a new cut was developed using a deep learning approach which exploits the unique "chirped" signal from the LPDA. A convolutional neural network (CNN) architecture was chosen, which builds on previous work [11] in which CNNs were found to be more efficient at discriminating between neutrino signal and noise compared to fully connected neural networks.

The training of the Deep Learning cut includes a mixture of experimental data (for backgrounds, which cannot be simulated at the moment) and simulated neutrino events, which presents a special challenge to avoid artificial differences. Since experimental neutrinos are expected to be rare, and have yet to be measured by the ARIANNA detector, building an experimental data set of neutrino signal is not possible. Rather, experimentally observed cosmic rays were used as a proxy for a training set of experimental neutrino events. Since artificial differences are not expected between the experimental cosmic rays and the experimental background training set from the same ARIANNA station, a network that depends on an artifact would improperly identify nearly all the experimental cosmic rays as background events. However, the network properly identified most experimental cosmic ray events as signal, providing evidence that the network is not significantly influenced by artificial differences in the experimental and simulated training sets[6].

4. Results

Figure 3 shows the residual events after applying the updown and dipole cuts (red dots) or just the Deep Learning cut (blue triangles). In both cases, all background is removed after applying the LPDA cut (orange dashed line). We next extrapolate the combination of updown, dipole and LPDA cuts to 1000 station-years of operation, which is representative of the baseline design for the shallow radio component of IceCube-Gen2 [5]. Events that pass the dipole cut are visually consistent with a random distribution within the (predominantly non-thermal) background events. Crucially, they do not cluster on the upper edge of the background distribution closest to the neutrino signal region, which is consistent with the expectation that the dipole, updown, and LPDA cuts are independent. For a background event (assumed to propagate through the atmosphere to the ice surface) to pass the LPDA cut, it should produce a single pulse shape in the backlobe of the horizontally polarized LPDA that has shape consistent with a neutrino signal, whereas background events that survive the dipole cut must produce a double pulse structure in the vertically polarized dipole with the appropriate time delays (T_{DiL}, T_{2P}). Therefore, it is assumed that χ_{Di} and χ_{LPDA} for background events are uncorrelated and that the background events observed by station 61 are typical of the future array
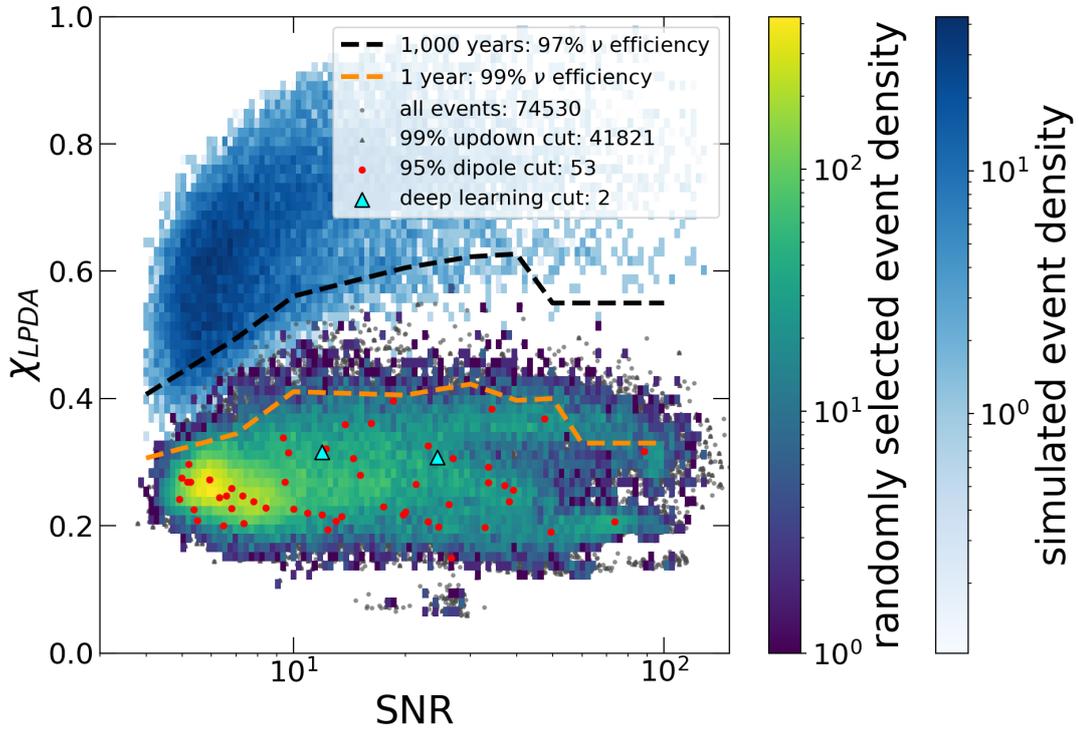
Figure 3: The correlation, χ_{LPDA}, of observed signals in the downward facing LPDA in station 61 as a function of signal amplitude, characterized by the largest signal-to-noise ratio (SNR) in the 4 downward facing LPDAs. The neutrino event density is shown in blue, and all experimental data in gray dots. The color scale shows event densities of experimental data after random selection of all data to scale to 1000 station-years. The region above the orange (black) dashed line contains 99% (97%) of the neutrino events. The neutrino signal efficiency (in percent) of the individual updown and dipole cuts are given in the legend, along with the number of data events that survive the updown cut (gray triangle) and the combined updown and dipole cut(red dot). Blue triangles show the two events that pass the deep learning cut.

of shallow detector stations at the South Pole. The observed background rate of 53/(station-year) of operation prior to the LPDA cut provides the baseline rate. The background rate is extrapolated to 1000 station-years by randomly selecting a total of 53,000 events from the entire background event population, shown in Figure 3 as the blue-yellow colored population. Since the average amplitude of the observed non-thermal background events is approximately a factor 2 larger than the thresholds in station 61 in the ARIANNA station, reducing the thresholds in a future station is unlikely to introduce a new population of non-thermal backgrounds. For the projected background distribution, the black dashed curve was designed to keep 97% of the neutrino signal while rejecting all background.

The combined efficiency of the above three cuts (updown, dipole, and LPDA) for 1,000 station-years is computed from $\epsilon_{tot} = \epsilon_{\text{updown}}\ \epsilon_{\text{dipole}}\ \epsilon_{\text{LPDA}} = (0.99)(0.95)(0.97) = 0.91$. The combination of the Deep Learning cut and LPDA cut appears to be even more efficient, given the factor of 26 reduction in the number of residual events prior to the application of the LPDA cut. Future work

will optimize the efficiency of all the developed cuts, including the Deep Learning cut.

5. Conclusion and Outlook

This work described three new analysis cuts that exploited new features in a special-purpose ARIANNA station with 8 antennas that combined a vertically oriented dipole antenna with upward and downward facing LDPAs. These new analysis cuts, when combined with a waveform shape cut on the downward facing LPDA, remove all background events while retaining 91% of the neutrino signal if projected to 1,000 station-years of operation. This result presents the first evidence that a detector station with near-surface antennas satisfies the requirements for the baseline design of the radio neutrino component of the IceCube-Gen2 project. The new cuts utilize the unique capabilities of all antenna channels. We conclude that deep learning methods will play an important role in future radio-based neutrino detectors that anticipate 1,000 station-years of operation or more. The deep learning cut in this paper shows that all but a few background events per station-year can be rejected by a relatively simple CNN. This could have far-reaching consequences. Given the initial success of deep learning methods to distinguish neutrino signal from background, it may be transformative to incorporate deep learning methods directly into a real-time trigger [12] to reduce the trigger threshold and increase the sensitivity. This idea takes advantage of recent improvements in the digitization speed of low power analog to digital (ADC) devices. Once the digital ADC information is transferred to an FPGA, a real-time trigger based on deep learning networks would efficiently identify and reject most of the unwanted thermal events.

References

[1] S. Barwick and C. Glaser, *Radio Detection of High Energy Neutrinos in Ice*, 2023 [`astro-ph.IM/2208.04971`]

[2] C. Glaser, for ARIANNA Collaboration, 9th ARENA, (2023) [`astro-ph.IM/2304.07179`]

[3] A. Anker et al., *JCAP* **03** (2020) 053.

[4] S. W. Barwick, et al., *Astropart. Phys.* **90** (2017) 50–68.

[5] S. Hallmann et al., *PoS* (ICRC2021) [`astro-ph/2107.08910`]

[6] A. Anker, et al. Submitted to JCAP (2023) [`astro-ph.IM/2307.07188`]

[7] C. Glaser, et al *Eur. Phys. J. C* **80** (2020); DOI=10.1140/epjc/s10052-020-7612-8

[8] C. Glaser, et al *Eur. Phys. J. C* **79** (2019); DOI=10.1140/epjc/s10052-019-6971-5

[9] A. Anker, et al.", *JCAP* **11** (2019) 030. [`astro-ph.IM/1909.02677`]

[10] A. Anker, et al., *JINST* **15** (2020) P09039; [`astro-ph.IM/2006.03027`]

[11] A. Anker, et al. *JINST* **17** (2022) P03007; [`astro-ph.IM/2112.01031v3`]

[12] C. Glaser, A. Coleman, and T. Glüsenkamp, *PoS* (ICRC2023) 1114.

6. Acknowledgement

Full Authors List: ARIANNA Collaboration

Note comment afterwards: Collaborations have the possibility to provide an authors list in xml format which will be used while generating the DOI entries making the full authors list searchable in databases like Inspire HEP.

A. Anker[1,13], P. Baldi[2], S. W. Barwick[1], J. Beise[3], D. Z. Besson[4], P. Chen[6,13], G. Gaswint[1], C. Glaser[3], A. Hallgren[3], J. C. Hanson[6], S. R. Klein[7], S. A. Kleinfelder[8], R. Lahmann[9], J. Liu[6], J. Nam[6], A. Nelles[9,10], M. P. Paul[1], C. Persichilli[1], I. Plaisier[9,10], R. Rice-Smith[1], J. Tatar[11], K. Terveer[9], S.-H Wang[5], C. Welling[9,10], and L. Zhao[1,12]

[1]Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA.

[2]Department of Information and Computer Science, University of California, Irvine, CA 92697, USA

[3]Uppsala University Department of Physics and Astronomy, Uppsala SE-752 37, Sweden.

[4]Department of Physics and Astronomy, University of Kansas, Lawrence, KS 66045, USA.

[5]Department of Physics and Leung Center for Cosmology and Particle Astrophysics, National Taiwan University, Taipei 10617, Taiwan.

[6]Whittier College Department of Physics, Whittier, CA 90602, USA.

[7]Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

[8]Department of Electrical Engineering and Computer Science, University of California, Irvine, CA 92697, USA.

[9]ECAP, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany.

[10]Deutsches Elektronen-Synchrotron DESY, 15738 Zeuthen, Germany.

[11]Research Cyberinfrastructure Center, University of California, Irvine, CA 92697, USA.

[12]The William H. Miller III Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA.

[13]SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA.