PROCEEDINGS OF SCIENCE

Boosting the Resolution of the CTAO with Hybrid Machine-learning Maximum-likelihood Event Reconstruction

Georg Schwefer,[a,*] **Robert Parsons**[b] **and Jim Hinton**[a] **for the CTAO Consortium**

[a] *Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany*
[a] *Institut für Physik, Humboldt-Universität zu Berlin, Newtonstraße 15, 12489 Berlin, Germany*

E-mail: georg.schwefer@mpi-hd.mpg.de

Gamma-ray measurements using the imaging atmospheric Cherenkov technique currently achieve the highest angular resolution in astronomy at very high energies, reaching down to arcminute scales at multi-TeV energies. High-resolution measurements provide the key to progress on many of the central questions in high-energy astrophysics, including the sites and mechanisms of particle acceleration up to PeV energies. The huge potential of the next-generation Cherenkov Telescope Array Observatory (CTAO) in this regard can be maximized with the help of improved algorithms for the reconstruction of the air-shower direction and energy.

Here, we present the FreePACT algorithm, a hybrid machine-learning maximum-likelihood reconstruction method for imaging atmospheric Cherenkov telescopes. It employs the neural ratio estimation technique from the field of likelihood-free inference to replace the analytical likelihood used in traditional image-likelihood fitting by a neural network that approximates the charge probability density function for each pixel in the camera.

The significant performance improvements provided by this algorithm are demonstrated using simulations of the planned CTAO southern array. We also discuss implications of the improved angular resolution for the science potential of CTAO using as an example the study of compact X-ray Pulsar Wind Nebulae.

39th International Cosmic Ray Conference (ICRC2025)
15–24 July 2025
Geneva, Switzerland

*Speaker

PoS(ICRC2025)838

1. Introduction

TeV gamma rays are only visible with ground-based observatories through the detection of the extensive air shower produced in their interaction in the atmosphere. Therefore, the performance of ground-based TeV gamma-ray observatories crucially depends not only on the telescope hardware, but also on the algorithms used for the reconstruction of the primary gamma-ray direction and energy from the measured secondary photons or particles.

Currently, the best resolution in the TeV energy range is achieved with Imaging Atmospheric Cherenkov Telescopes (IACTs) capturing and imaging the Cherenkov light emitted by the charged particles in the air showers.

With the current generation of instruments such as the H.E.S.S. observatory, the best resolution in the range of a few arcminutes in the multi-TeV energy range is obtained with likelihood-fitting methods [1]. In these, the shower parameters (energy, source direction, shower core location and height of the shower maximum in the atmosphere $X_{\max}$) are estimated by maximizing the likelihood of measuring the observed charge c_{ij} in each pixel j of each telescope i given the shower parameters $\vec{\eta}$, the telescope location $\vec{r}^{\,i}_{\text{tel}}$ and the pixel coordinate $\vec{r}^{\,ij}_{\text{pix}}$

$$\mathcal{L} = \prod_{i=1}^{N_{\text{tel}}} \prod_{j=1}^{N_{\text{pix}}} p(c_{ij} | \vec{r}^{\,ij}_{\text{pix}}, \vec{r}^{\,i}_{\text{tel}}, \vec{\eta}). \tag{1}$$

The current state-of-the-art implementation of this method, *ImPACT* [1], uses an analytical description of the per-pixel likelihood [2], with the mean value obtained from average-image Monte Carlo templates.

The advent of the next generation of IACTs, especially the Cherenkov Telescope Array Observatory (CTAO), will move the boundary of achievable resolution, both through upgraded hardware and through the development of new reconstruction algorithms.

On attractive option to boost the resolution that was for example recently successfully applied to the reconstruction of astrophysical neutrinos in IceCube [3] are hybrid methods combining likelihood-based frameworks with machine-learning methods. In these, the likelihood does no longer need to be approximated analytically, but is instead learned and approximated by a neural network.

In this work, we present such a hybrid reconstruction method for IACTs called *FreePACT*, short for likelihood-free *ImPACT*.

We first describe the method in section 2, then demonstrate its performance on simulations of the planned CTAO southern array in section 3. These sections are heavily based on the work published in [4]. Finally in section 4 we briefly discuss the science implications of the improved resolution using as an example the morphological study of Pulsar Wind Nebulae (PWNe).

2. The FreePACT method

The *FreePACT* algorithm is based on the method of *neural ratio estimation* (NRE) from the field of likelihood-free/simulation-based inference.

Generally, the idea of NRE [5–7] is to not directly estimate the likelihood of the data x given parameters θ, $p(x|\theta)$, but instead to estimate the likelihood-to-evidence ratio $\frac{p(x|\theta)}{p(x)}$. For the

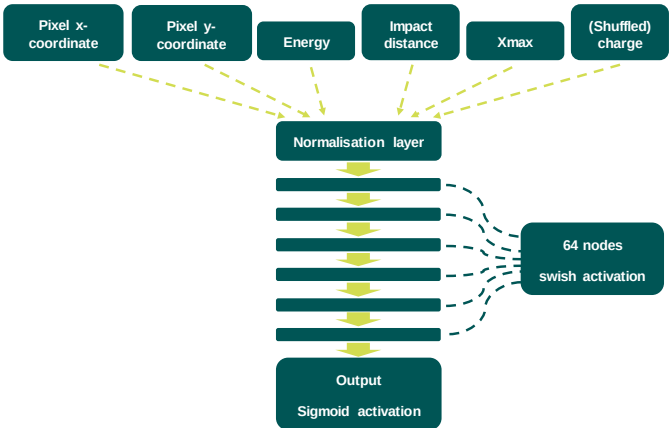

Figure 1: Schematic depiction of the neural network architecture used for the *FreePACT* models.

optimization with respect to the parameters θ, the two are fully equivalent. Using Bayes' theorem, the likelihood-to-evidence ratio can be rewritten as

$$\frac{p(x|\theta)}{p(x)} = \frac{p(x,\theta)}{p(x)p(\theta)} = \frac{d(x,\theta)}{1-d(x,\theta)} \tag{2}$$

with the *decision function*

$$d(x,\theta) = \frac{p(x,\theta)}{p(x,\theta)+p(x)p(\theta)}. \tag{3}$$

This *decision function*, and thus ultimately the likelihood-to-evidence ratio, can nicely be approximated by a binary classifier trained with binary cross-entropy to separate samples from the joint distribution, $p(x,\theta)$, from samples from the product of the two marginal distributions, $p(x)p(\theta)$.

Translating this to the application to IACT event reconstruction, following equation 1, the data x is the measured charge in every pixel c_{ij}, the parameters θ are the shower parameters $\vec{\eta}$ and the telescope locations $\vec{r}^{\,i}_{\text{tel}}$ and pixel coordinates $\vec{r}^{\,ij}_{\text{pix}}$. Samples from the joint distribution $p(x,\theta)$ are easily generated: By construction, these are simulated pixels from shower and telescope simulations, in our case using the CORSIKA/sim_telarray simulation suite [8, 9]. From these, samples from the product of the two marginal distributions $p(x)p(\theta)$ are also generated easily by shuffling the values of the charge randomly between pixels, removing any correlation between them and the parameters θ.

With these samples, we train a binary classifier neural network. The architecture is visualized in figure 1 and is overall quite simple: Following a normalization layer, there are six fully connected layers with 64 nodes each and the `swish` activation function [7] ensuring a smooth output at each node. The last layer features a single node with the `sigmoid` activation function. In this configuration, the model only has around 20000 trainable parameters. We have verified that similar results can be achieved with networks of different sizes. The network is trained with binary cross-entropy and convergence is typically reached within an hour on a CPU.

3. Performance results

3.1 Pixel Charge PDFs

The most basic test for the performance of the *FreePACT* method is to investigate the resulting approximations of the per-pixel charge PDFs. As a reference for this, we simulate gamma rays at a fixed energy and impact distance from a single telescope, select events in a narrow slice of $X_{\max}$, and then make a histogram of the charges measured in specific individual pixels. We then compare these heuristic, "ground truth", distributions to the corresponding *ImPACT* and *FreePACT* PDFs[1].

In figure 2, we show two spot checks for simulations of a CTAO medium-sized telescope (MST) equipped with a FlashCam camera [10]. The left panel shows a pixel in which no Cherenkov photons are expected. The resulting charge distribution is therefore purely a consequence of the night-sky background (NSB) and electronic noise. In contrast to this, the right panel shows a pixel brightly illuminated by Cherenkov light. It is clearly visible that in both cases, the *FreePACT* model gives a better approximation of the true charge PDF compared to the description by *ImPACT*: In the noise-dominated case, it manages to better capture the specific asymmetric shape of the PDF compared to the *ImPACT* model. In the bright signal case, it describes the width of the distribution, arising from shower-to-shower fluctuations, better than the analytical PDF.

It should be stressed that any difference in the reconstruction performance between *FreePACT* and *ImPACT* can directly be traced back to differences in these descriptions of the charge PDFs, as the two methods otherwise work in exactly the same way.

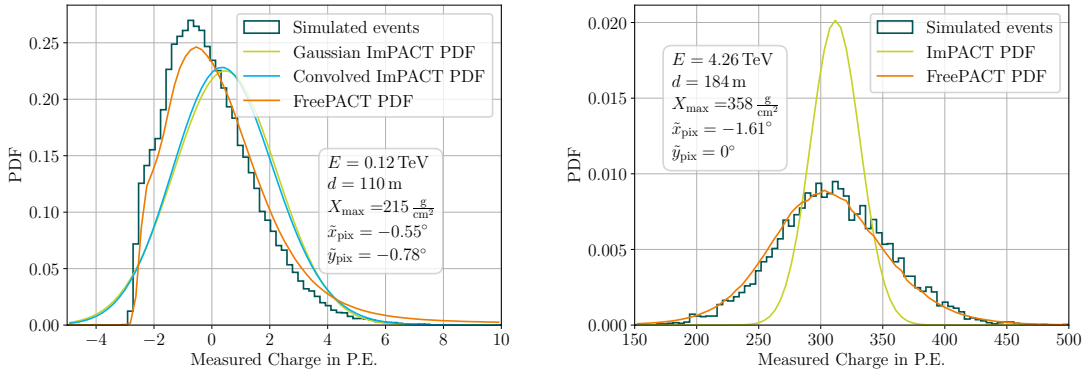

Figure 2: Charge distribution at the given pixel position for the given shower parameters in a FlashCam-MST with the analytical *ImPACT* charge PDFs and the *FreePACT* distribution. On the left, the charge is dominated by noise photons, on the right, the charge is dominated by Cherenkov photons. On the right, as the Gaussian and convolved *ImPACT* PDFs are identical, we only show one curve.

3.2 Resolution curves

The most important measures of performance for any reconstruction algorithm are of course the resulting resolution curves. Here, we show these for simulations of on-axis gamma rays at a

[1] In order to obtain the normalized PDF from the *FreePACT* model, the likelihood-to-evidence ratio needs to be multiplied by the evidence, which is readily available as the total charge PDF of the training data.

zenith angle of 20° with the full planned CTAO southern Alpha configuration array [11] consisting of 14 medium-sized telescopes (MSTs) and 37 small-sized telescopes (SSTs). We select events based on image quality [4]. The resulting effective area is very comparable to the CTAO Prod5 reference [4, 12].

On these simulations, we then compare three different reconstruction algorithms:

- Hillas & Random Forest: This serves as a classical reference algorithm as well as being used as a seed for the *ImPACT* and *FreePACT* fits. For the geometry reconstruction, it uses standard geometrical Hillas reconstruction [13]. The energy reconstruction is handled by a Random Forest trained on a set of image parameters as implemented in `ctapipe` [14, 15].
- *ImPACT*: We use the full convolved form of the likelihood for the reconstruction here.
- *FreePACT* as described above.

Because of the similarity of the effective area, we can additionally compare our resolution curves to the corresponding CTAO Prod5 reference curves [12]. All resolution curves together are shown in figure 3.

First and foremost, figure 3 shows that for both angular and energy resolution, the *FreePACT* method significantly outperforms both the standard Hillas & Random Forest reconstruction as well as the *ImPACT* method over a wide energy range. This clearly demonstrates the potential of the hybrid machine learning-likelihood approach to reconstruction. The power of the likelihood approach in general can be seen by the still sizable difference in reconstruction quality between the Hillas & Random Forest reconstruction and the *ImPACT* reconstruction.

We have also tested the reconstruction performance of the *FreePACT* method for different zenith angles, source field-of-view offsets and levels of NSB. In every scenario, we find it to outperform the *ImPACT* and classical methods [4].

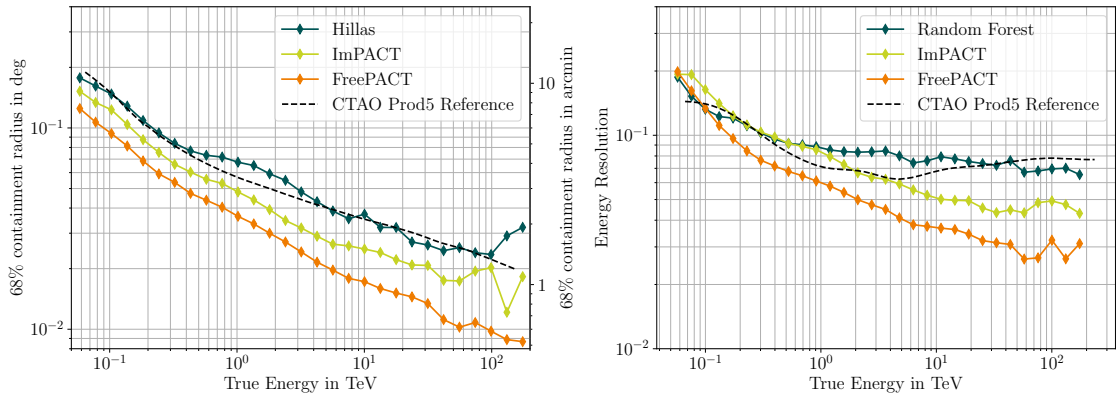

Figure 3: Angular and energy resolution as a function of true gamma-ray energy for the CTA southern array for on-axis photons at 20° zenith angle. Shown in dark green is the standard Hillas/Random Forest analysis that serves as a seed to the image likelihood fits. In lime green, we show the results from a conventional *ImPACT* analysis. The *FreePACT* resolution curves are shown in orange. The official CTAO Prod5 curve (after gamma-hadron separation) is shown in black for reference.

4. Application to morphological study of PWNe

As can be seen in the first panel of figure 3, using the *FreePACT* reconstruction method, it appears possible for the CTAO southern array to reach sub-arcminute angular resolution above energies of around 10 TeV. This of course raises the question which (astro-)physical science questions could be answered with the help of this improved resolution.

We argue that PWNe are natural candidates to profit from these advances because they satisfy two central criteria: First, from source catalogs like those compiled by HAWC [16] and LHAASO [17], they are known to be bright sources in the multi-TeV energy range. Second, they possess plentiful structure on the relevant arcminute scale. This is known from X-ray measurements with observatories such as Chandra [18], XMM-Newton [19] and eROSITA [20].

The connection to the X-ray observations is actually more profound: Assuming the sources to be leptonic, the non-thermal X-ray emission is synchrotron radiation, the intensity of which is proportional to the density of high-energy leptons and the energy density in the magnetic field,

$$I_{\mathrm{X}} \propto n_{\mathrm{e}} \times B^2. \tag{4}$$

The spatial distributions of n_{e} and B are important for understanding the particle acceleration and transport processes in the PWN, but are degenerate in the X-ray observations.

Gamma-ray observations can help lift this degeneracy: The Inverse Compton emission is independent of B and only depends on n_{e} and the energy density of the interstellar radiation field which is typically assumed to be spatially constant on the required scale (e.g [21]).

This, of course, requires sufficient spatial resolution in the gamma-ray observations and that the gamma rays and X-rays are actually produced by the same electrons. Going through the respective calculations [22] and taking into account Klein-Nishina effects and different radiation fields, one finds that in a 10 μG magnetic field typical for these sources, the same electrons that predominantly produce X-rays between 2 keV and 10 keV have Inverse Compton spectra peaking between 10 TeV and 100 TeV. This means that gamma-ray observations in just the energy range where the best angular resolution is possible can indeed be used together with X-ray measurements to infer the morphological properties.

To demonstrate this, we choose HESS J1813−178 as an example source. In X-rays, this middle-aged, powerful PWN [21] has been observed among others by XMM-Newton [23] which has revealed asymmetric, non-gaussian structure. In TeV gamma rays, the source was recently studied using H.E.S.S. [21]. However, it was unable to resolve this structure in the compact nebula.

To test whether this could be possible with CTAO observations, we create an analytical model of the X-ray morphology of the PWN based on and consistent with the XMM-Newton observations from [23]. We then split up this model into different models of the magnetic field and electron distributions according to equation 4 and assume the morphology of the source in multi-TeV gamma rays to be equal to that of the electron distribution. This, in combination with the total flux and energy spectrum of the source which are known from H.E.S.S. observations [21], allows us to create mock observations between 10 TeV and 100 TeV of gamma-ray energy of the source using different sets of instrument response functions (IRFs). In particular, we use public H.E.S.S. IRFs from [24] and the CTAO Prod5 reference and *FreePACT* IRFs described above.

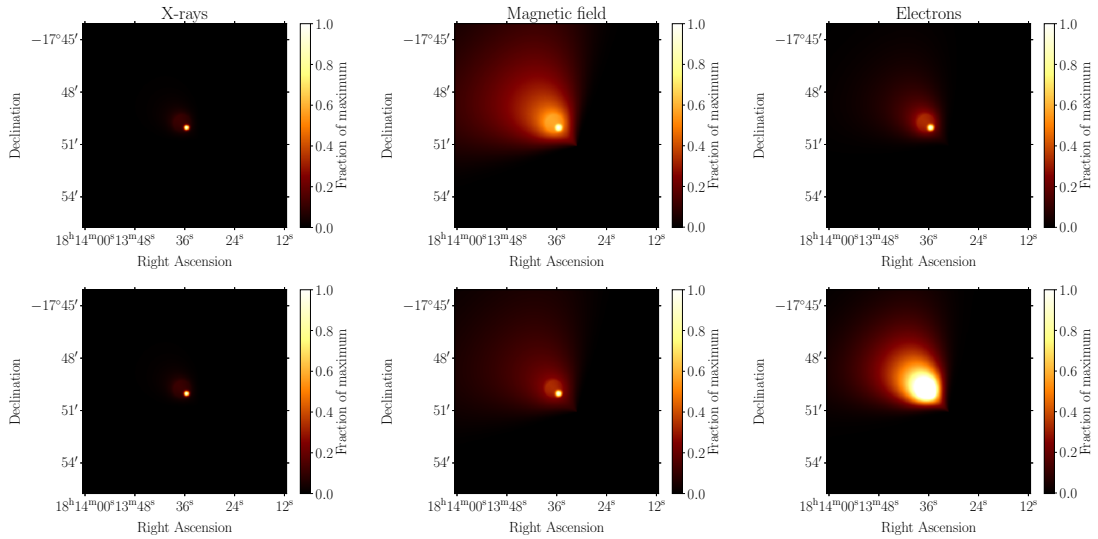

Figure 4: Plot of the two morphological model hypotheses for HESS J1813−178. The first column shows the X-ray morphology, the second column the magnetic field and the third column the distribution of high-energy electrons. The first row shows the models following the assumption $n_e \propto B^2$. In the the second row, the electron energy density is flattened close to the center of the PWN. To emphasize that these are purely morphological models, they are simply normalized to their respective peak values here.

In figure 4, we show two hypotheses for the electron distribution and thus TeV morphology of the source together with the corresponding magnetic field model required to reproduce the X-ray morphology of the source. For the first, we assume a proportionality between electron distribution and magnetic energy density, $n_e \propto B^2$. For the second model, we assume a flat electron distribution near the center of the PWN. In that case, the spike in the X-ray morphology is fully attributed to a spike in the magnetic field.

In figure 5 then, we show corresponding fake (smoothed) count maps for 100 hours of observations for the three sets of IRFs considered. The improved angular resolution going from H.E.S.S. to the CTAO southern array and then using the *FreePACT* method is clearly visible. Note also the large difference in event statistics between H.E.S.S. and CTAO. To quantify these results, we use the method described in [25] to calculate the expected significance with which the second morphological model can be excluded if the first is assumed to be true. While with H.E.S.S., the median significance is less than 1σ, using the CTAO Prod5 reference IRFs we find a median significnace of 1.7σ, corresponding to a 95% exclusion. Then, finally, with the *FreePACT* IRFs the median significance is 2.5σ equivalent to an exclusion at greater than 99%.

References

[1] R.D. Parsons and J.A. Hinton *Astropart. Phys.* **56** (2014) 26 [1403.2993].
[2] M. de Naurois and L. Rolland *Astropart. Phys.* **32** (2009) 231 [0907.2610].
[3] IceCube collaboration *Science* **380** (2023) adc9818 [2307.04427].
[4] G. Schwefer, R. Parsons and J. Hinton *Astropart. Phys.* **163** (2024) 103008 [2406.17502].

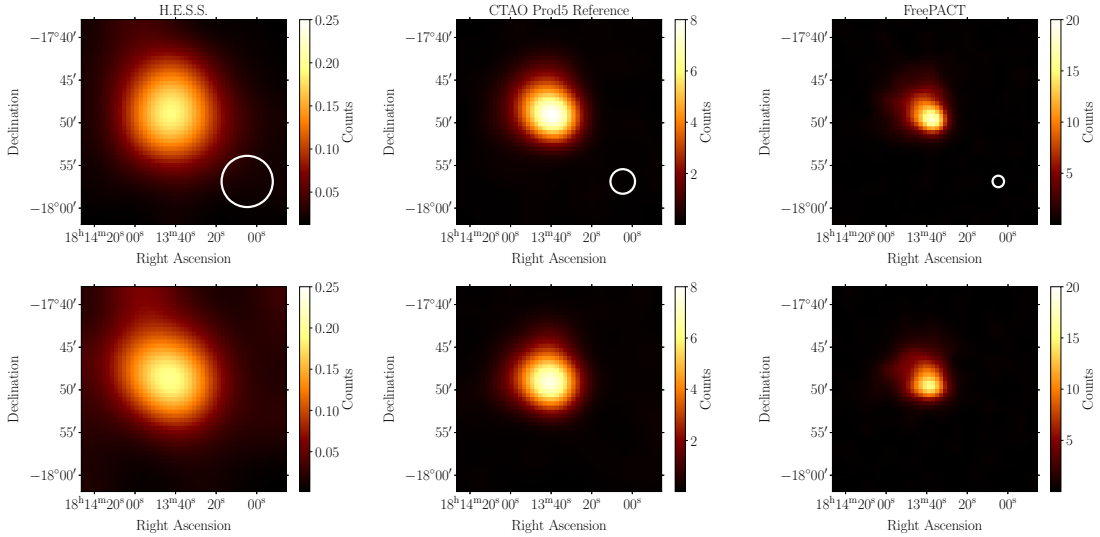

Figure 5: Plot of the smoothed count maps of a mock dataset of the two morphological TeV gamma-ray models of HESS J1813−178 for the three sets of IRFs: In the first column the public H.E.S.S. IRFs, in the second column the CTAO Prod5 reference IRFs and in the third column the FreePACT IRFs. The white circles indicate the size of the respective PSF at 50 TeV, which is also the smoothing radius.

[5] K. Cranmer, J. Pavez and G. Louppe 1506.02169.
[6] J. Hermans, V. Begy and G. Louppe *PMLR* **119** (2020) 4239 [1903.04057].
[7] P. Eller et al. *NIM A* **1048** (2023) 168011.
[8] D. Heck et al., *CORSIKA: a Monte Carlo code to simulate extensive air showers.* (1998).
[9] K. Bernlohr *Astropart. Phys.* **30** (2008) 149 [0808.2253].
[10] G. Pühlhofer et al. *Proc. SPIE Int. Soc. Opt. Eng.* **11119** (2019) 111191V.
[11] "CTAO Webpage." https://www.cta-observatory.org/about/how-ctao-works/.
[12] C.T.A. Observatory and C.T.A. Consortium Sept., 2021. 10.5281/zenodo.5499840.
[13] F. Aharonian et al. *Astroparticle Physics* **6** (1997) 343.
[14] CTA Consortium collaboration *PoS* **ICRC2023** (2023) 703.
[15] M. Linhoff et al. Sept., 2023. 10.5281/zenodo.8348922.
[16] HAWC collaboration *Astrophys. J.* **905** (2020) 76 [2007.08582].
[17] LHAASO collaboration *Astrophys. J. Suppl.* **271** (2024) 25 [2305.17030].
[18] M.C. Weisskopf et al. *Publ. Astron. Soc. Pac.* **114** (2002) 1 [astro-ph/0110308].
[19] XMM collaboration *Astron. Astrophys.* **365** (2001) L1.
[20] eROSITA collaboration *Astron. Astrophys.* **647** (2021) A1 [2010.03477].
[21] H.E.S.S. collaboration *Astron. Astrophys.* **686** (2024) A149 [2403.16802].
[22] J.A. Hinton and W. Hofmann *Ann. Rev. Astron. Astrophys.* **47** (2009) 523 [1006.5210].
[23] S. Funk et al. *Astron. Astrophys.* **470** (2007) 249 [astro-ph/0611646].
[24] H.E.S.S. collaboration 1810.04516.
[25] IceCube collaboration *PoS* **ICRC2023** (2023) 1046 [2308.08233].

PoS(ICRC2025)838