

## $S_b$ for gamma / hadron separation in SWGO

---

**I. D. Vergara Quispe,<sup>a</sup> P. M. Hansen,<sup>a</sup> L. Nellen,<sup>b,\*</sup> A. G. Mariazzi<sup>a</sup> and D. G. Melo<sup>c</sup> for the SWGO collaboration**

<sup>a</sup>*Instituto de Física La Plata,  
Diagonal 113 e/ 63 y 64, La Plata, Argentina*

<sup>b</sup>*Instituto de Ciencias Nucleares, UNAM  
Cto. Exterior S/N, 04510 Coyoacán, México*

<sup>c</sup>*ITeDA  
Av. Gral. Paz 1499 Edificio 42, Sector C, Buenos Aires, Argentina  
E-mail: [hansen@fisica.unlp.edu.ar](mailto:hansen@fisica.unlp.edu.ar), [ivergara@fisica.unlp.edu.ar](mailto:ivergara@fisica.unlp.edu.ar),  
[lukas@nucleares.unam.mx](mailto:lukas@nucleares.unam.mx), [mariazzi@fisica.unlp.edu.ar](mailto:mariazzi@fisica.unlp.edu.ar),  
[diego.melo@iteda.cnea.gov.ar](mailto:diego.melo@iteda.cnea.gov.ar)*

The Southern Wide-field Gamma-ray Observatory (SWGO) is an international collaboration working on a ground-based gamma-ray observatory that will be located in the southern hemisphere. A crucial step in the analysis is to identify the showers produced by gamma rays and separate them from the abundant background of hadronic showers. In this work, we propose to adapt the observable,  $S_b$ , used successfully to composition studies in the Pierre Auger Observatory to the SWGO detector. This observable takes into account the signal and the position of each triggered detector. It characterizes the shape of the lateral distribution of the signal, which depends on the nature of the primary particle. The value of  $S_b$  is therefore suited to identify gamma induced showers and to reject the more frequent hadronic showers. This analysis has been performed using a scaling factor to link the observable,  $S_b$ , between the central and outtrigger array. Also we explore how this new observable improves the separation of primary proton and gamma-induced air showers in terms of the merit factor.

38th International Cosmic Ray Conference (ICRC2023)  
26 July - 3 August, 2023  
Nagoya, Japan



---

\*Speaker

## 1. Introduction

In this study, we propose the adaptation of the observable  $S_b$  [1, 2], which is currently utilized in the Pierre Auger Observatory (PAO) [3, 4] for the purpose of gamma hadron separation, to be employed in the upcoming SWGO (Southern Wide-field Gamma-ray Observatory). The modification of  $S_b$  is required to take into account the higher, varying density of detector stations in SWGO.  $S_b$  captures the combined influences of the muon and electromagnetic components on the lateral distribution function of air showers, originating from different primary species at any given primary energy. It is defined as follows [1]:

$$S_b = \sum_{i=1}^N \left[ S_i \times \left( \frac{r_i}{r_0} \right)^b \right] \quad (1)$$

where the sum runs over all the triggered stations,  $S_i$  is the signal recorded in the  $i$ th station,  $r_i$  is the distance of this station to the shower axis,  $r_0$  is a distance of reference, and the parameter  $b$  is a free parameter, allowing for the maximization of the discriminating power of this observable.

In the original work, the  $S_b$  observable has been investigated for different array configurations, including square and triangular grids. In the case of the triangular grid, various separation distances between stations were considered, ranging from 500 m to 2000 m.

This observable was initially proposed for proton-iron discrimination [1] and has also been utilized within the Pierre Auger collaboration for gamma-hadron separation [2].

The Surface Detector of the Pierre Auger Observatory consists of a triangular configuration with a station separation distance of 1500 m to detect cosmic rays with energies above  $10^{16.5}$  eV [3]. In this study, we propose the utilization of the  $S_b$  observable for gamma discrimination in a detector designed to operate within a lower energy range (100s GeV to 100s TeV) and employing a denser array configuration. This implies taking into account smaller station separations and core distances than previously studied.

To assess the discrimination power of this observable, we computed the merit factor (equation 2) and the Q factor (equation 3, [5]). These factors are defined as follows:

$$MF = \frac{\mu_\gamma - \mu_p}{\sqrt{\sigma_\gamma^2 + \sigma_p^2}} \quad (2)$$

where  $\mu$  and  $\sigma$  are the mean and the standard deviation of values of  $S_b$ ; the subscripts  $\gamma$  and  $p$  correspond to distributions for gammas and protons.

$$Q = \frac{\epsilon_\gamma}{\sqrt{\epsilon_p}} \quad (3)$$

where  $\epsilon_\gamma$  is the fraction of gamma ray events retained after a cut, and  $\epsilon_p$  is the fraction of proton induced events retained after the same cut applied.

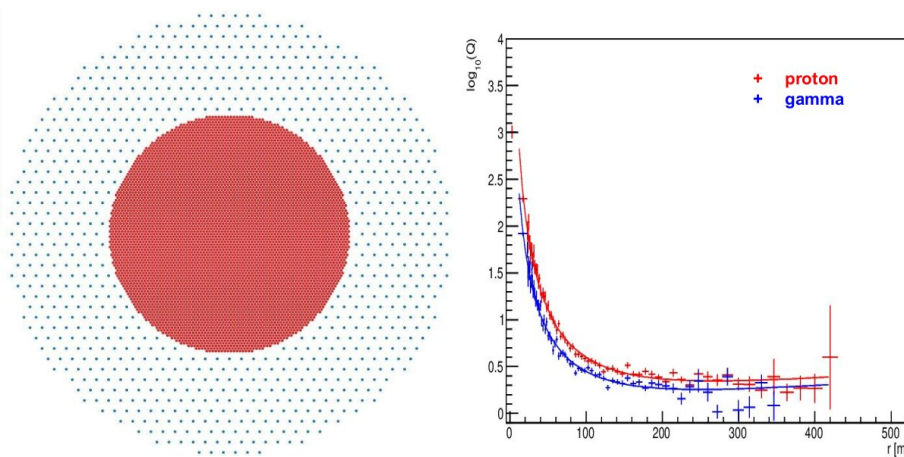
## 2. SWGO detector

The SWGO detector is a ground-based gamma-ray detector, which boasts a wide field of view and a high duty cycle. It is planned to be situated in South America at an approximate altitude of

4.4 km. The detector's scope spans primary particles within an energy range from the hundreds of GeV up to the PeV scale [6]. The design primarily relies on water Cherenkov detector stations, characterized by a dense core area extending 160 m radius, followed by a lower density outer array or outrigger area reaching up to 300 m (radius) (see Figure 1 left). Each station will be equipped with two Photo Multiplier Tubes (PMTs), with one positioned to observe the top of the tank, thereby capturing the electromagnetic component of the shower, and another situated to observe the bottom, enabling the detection of the muonic component.

In our analysis, we simulated proton and gamma-induced showers using the CORSIKA Monte Carlo simulation program [7]. We then applied the detector simulation and reconstruction framework developed by the SWGO collaboration, which is based on the code developed by the HAWC Collaboration. The simulation library follows a  $E^{-2}$  spectrum to evenly distribute the computational time across each energy decade.

The parameter  $S_b$  is derived from the total signal measured in each triggered detector for a given event. Consequently, it relies on the normalization and shape of the lateral distribution function of the total signal. Figure 1 (right) displays the logarithm of the total charge of the PMT at the top of the tank as a function of the distance to the shower axis for protons and gamma primaries. The plotted lines represent fits of the lateral distribution function to an NKG-like function. This figure clearly illustrates the distinction between the proton and gamma lateral distribution functions observed in the SWGO standard configuration (Figure 1 (left)).



**Figure 1:** Left: Schema of the SWGO detector design: central array in red, outrigger array in blue. Right: Lateral distribution signal for gamma and proton-induced showers with zenith angle ( $5^\circ$ - $7^\circ$ ) and similar number of triggered stations (4000-5000).

### 3. Scaling Factor

Previous studies of  $S_b$  focused on uniform array configurations. However, in the case of SWGO, the array exhibits varying densities, which changes the relative contribution of station to  $S_b$  in different regions. To address this, a scaling factor is employed to compensate.

A preliminary method for calculating the scaling factor involves utilizing the ratio of detector densities between the central and outrigger arrays. The ratio of the area occupied by stations to the total number of stations in each configuration is found to be 16, serving as the scaling factor.

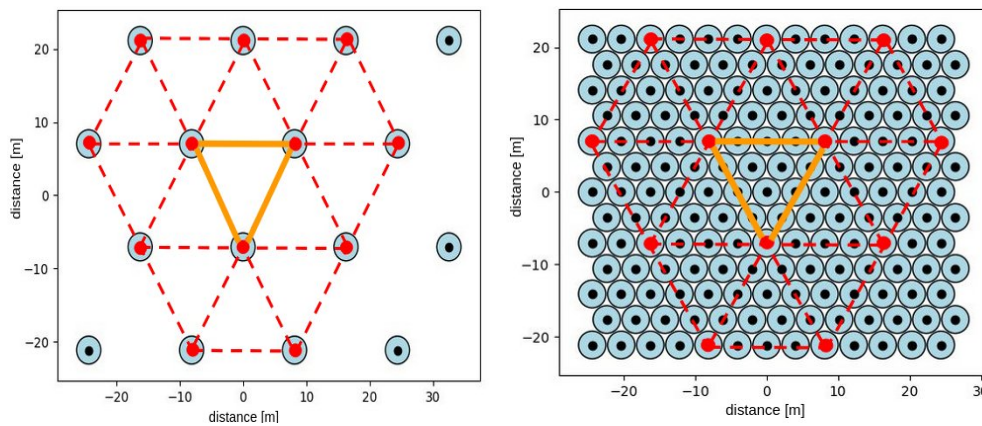
Another approach for calculating the scaling factor involved constructing a unit cell using the three closest stations from the outrigger array. Figure 2 (left) illustrates the unit cell under investigation, highlighted by the yellow-colored sides, along with neighboring unit cells formed by stations surrounding it in the outrigger array.

Each station in the unit cell shares a common vertex with six neighboring cells. Thus, the analysis shows that the three stations in the unit cell collectively contribute only half a station per cell.

By applying the same procedure to the central array with a unit cell size similar to the outrigger array (see Figure 2), one must consider common side stations, inside unit cell stations, and stations with common vertices:

- The contribution from the common vertices remains the same as in the outrigger array, accounting for half a station per cell.
- Three common side stations are shared by two unit cells, resulting in a contribution of  $3/2$  stations per side per unit cell. Each cell has 3 sides contributing  $9/2$  stations per cell.
- Additionally, each cell has 3 inside stations, not shared with other cells, contributing 3 stations per cell.

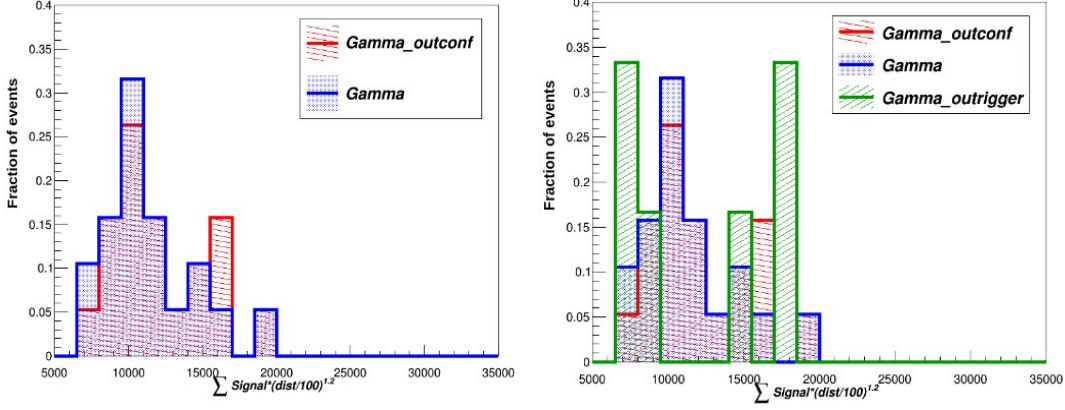
Therefore, the total contribution of stations per unit cell in the central array is equal to  $1/2 + 9/2 + 3 = 8$ . By comparing the central array's contribution of 8 stations to the outrigger array's 0.5 stations, it becomes evident that there are 16 stations in the central array for every station present in the outrigger array. This comparison confirms the scaling factor to be 16.



**Figure 2:** The red lines outline the boundaries of the unit cells, while the yellow lines indicate the unit cell under observation for different configurations of the SWGO array. On the left side is the outrigger configuration, while on the right side is the central array configuration.

We verify the scaling method by extending the outrigger configuration to the central array. This is done by removing stations in the central array to generate a sub-array of the density of the

outrigger array. In this part of the analysis, the observable  $S_b$  is calculated for both configurations



**Figure 3:** The histograms display the distributions of  $S_b$  values obtained through standard and weighted calculations, considering the denser (blue histogram) and outrigger configurations (red and green histograms), respectively. On the left side, the histograms represent events with the core located in the central array, shown by the red and blue histograms. On the right side, the analysis expands to include events with the core in the outrigger array, represented by the green histogram.

of the central array: first for the original configuration, then for the modified configuration, where the central array is adjusted to be similar to the outrigger array, but including the scaling factor.

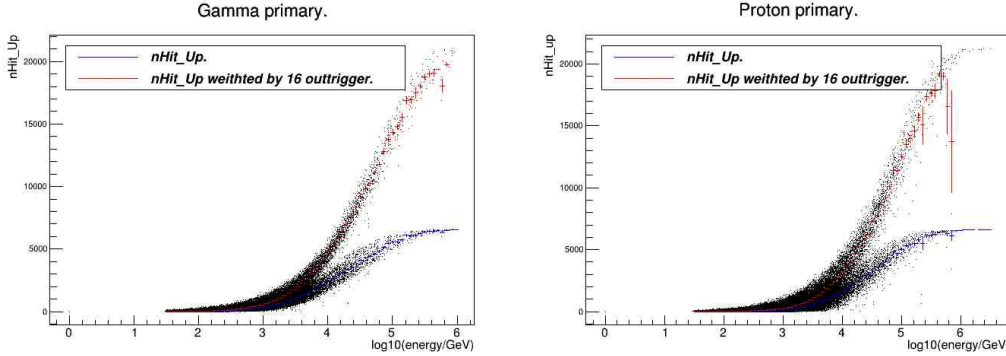
In order to ensure that a significant portion of the shower is well contained within the array and to avoid potential irregularities at the array boundaries, the analysis focused on events with a core distance between 70 to 90 meters away from the center of the array. Additionally, only triggered stations within a distance of less than 50 meters from the shower core were considered for calculating the observable  $S_b$  by two methods:

- **Standard calculation:** This method considers all the PMTs triggered (with a distance of less than 50 meters from the core) corresponding to the normal configurations of the central array. The calculation of  $S_b$  is defined as follows:  $S_b = \sum_{i=1}^N \left[ S_i \times \left( \frac{r_i}{100} \right)^b \right]$ .
- **Weighted calculation (outconf):** In this method, the PMTs triggered in the central array are considered, but with the outrigger configurations. To account for the dismissed stations and change the array configuration, a scaling factor of 16 is applied. The definition of  $S_b$  for this method is:  $S_b = \sum_{i=1}^N \left[ S_i \times \left( \frac{r_i}{100} \right)^b \times 16 \right]$ . The same events used in the Standard calculation are utilized for comparing the results.

In the left plot of Figure 3, the results for  $S_b$  obtained using the two methods described above are presented. The plot demonstrates that both sets of results closely match each other, indicating that the scaling factor of 16 successfully reproduces the results obtained in the central array when using the outrigger configuration.

After this, the weighted calculation was applied to events with a core located in the outrigger region (220 - 240 meters from the center of the array), considering only the stations within a distance of less than 50 meters from the core. The results of this calculation are presented in the

green histogram in the right plot of Figure 3. These new results align with the previously discussed results, demonstrating consistency (within statistical fluctuations).



**Figure 4:** The station multiplicity ( $nHit_{up}$ ) as a function of primary energy is displayed as a blue profile curve. The red curve represents the scaled  $nHit_{up}$ . The left panel corresponds to gamma primaries, while the right panel corresponds to proton primaries.

#### 4. $S_b$ observable

The total number of triggered stations in an event is a crucial parameter that reflects the size of the shower.

Due to the distinct densities present in the SWGO array, it is recommended to adjust the total number of triggered stations in an event using a scaling factor that accounts for these variations in density. Consequently, sample events generated by proton and photon primaries are corrected with a scaling factor of 16 for stations located in the outrigger region.

In this analysis, the signal from the PMT positioned at the top of the tank is utilized to compute the  $S_b$  values. Figure 4 displays the relationship between the total number of hits in the top PMT ( $nHit_{up}$ ) and the energy for gamma (right) and proton (left) primaries. The blue curve represents the overall number of up PMTs triggered in an event, while the red curve accounts for the scaled  $nHit_{up}$  (When a station is in the outrigger array, its  $nHit_{up}$  count is multiplied by 16 before being included in the calculation).

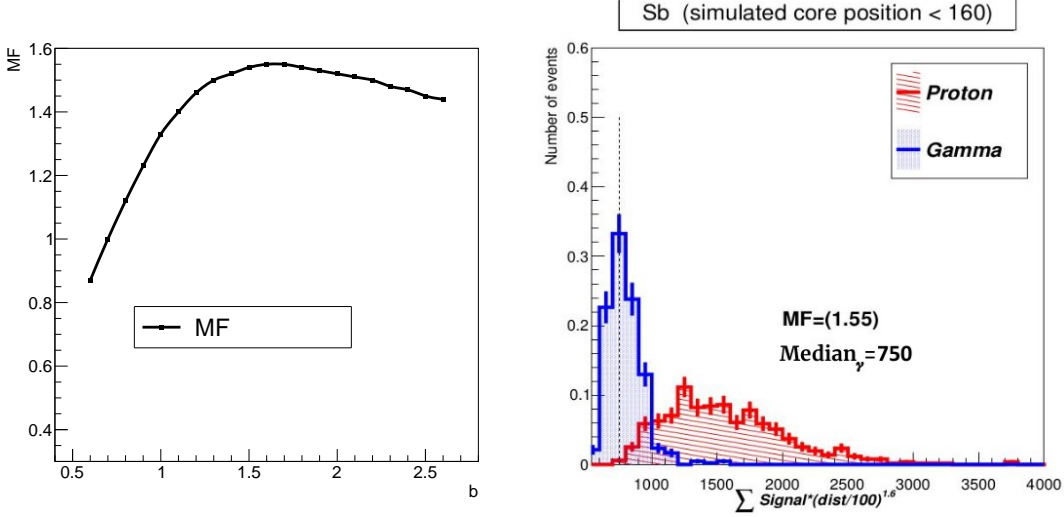
Another step in the analysis of the power discrimination of the  $S_b$  observable involves grouping our data based on the scaled  $nHit_{up}$  in an event. The entire sample has been divided into 8 bins of the scaled  $nHit_{up}$ : (1000-2000), (2000-3000), (3000-4000), (4000-5000), (5000-6000), (6000-8000), (8000-11000), and (11000-15000). These intervals were selected to ensure sufficient statistics in all the bins. As the flux of particle primaries decreases with increasing energy, the higher energy bins, as shown in Figure 4, correspond to wider intervals.

For each bin of scaled  $nHit_{up}$ , the merit factor of  $S_b$  is calculated as a function of the parameter  $b$ . The results of this analysis are illustrated in the left panel of Figure 5, where the optimal value of  $b$  that maximizes the merit factor can be observed.

Figure 5 (right) shows the distribution of  $S_b$  values, with the blue curve representing gamma primaries and the red curve representing proton primaries. These curves are obtained using the

optimal value of  $b = 1.6$ , as determined from the left plot of the same figure. This figure corresponds to the range of  $nHit_{up}$  from 4000 to 5000. Moreover, the median value of the  $S_b$  distribution is indicated by a vertical line to evaluate the contamination at 50% efficiency, resulting in a value of 0.2%. The median value for  $S_b$  is 750, and the associated Q factor is determined to be 11.3.

This analysis has been applied to each  $nHit_{up}$  bin, and the results are summarized in Table 1. We have included the results of both the merit factor and the Q-factor as a function of  $b$  for the



**Figure 5:** Left: Discrimination power of  $S_b$  measured by the merit factor for different  $b$  values (optimum value at  $b = 1.6$ ). Right: Distribution of  $S_{1.6}$  for gamma (blue) and proton (red) primaries. The median value of the gamma distribution is represented by a vertical line.

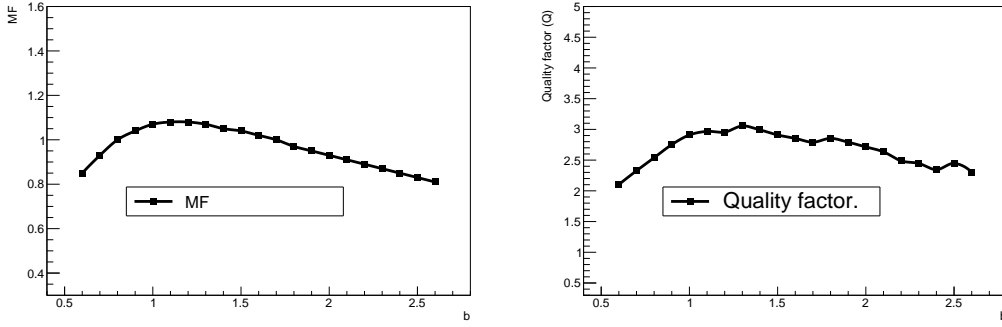
$nHit_{up}$ range	optimum $b$	MF	contamination at $\epsilon 50\%$	Q
(1000 – 2000)	1.1	1.08	2.8%	2.97
(2000 – 3000)	1.3	1.32	1.3%	4.42
(3000 – 4000)	1.3	1.33	1.3%	4.37
(4000 – 5000)	1.6	1.55	0.2%	11.3
(5000 – 6000)	1.8	1.46	0.3%	8.82
(6000 – 8000)	2.3	1.48	0.8%	5.53
(8000 – 11000)	2	1.22	1.9%	3.63
(11000 – 15000)	3.1	1.02	6.2%	2.0

**Table 1:** Discrimination power  $S_b$  observable:  $nHit_{up}$  range, optimum  $b$ , merit factor, contamination at 50% efficiency and quality factor Q are shown in this table.

first bin used in our analysis (see figure 6). Notably, these two quantities exhibit a similar behavior. Based on this observation, we have opted to utilize the merit factor for the optimization of  $b$ .

The  $S_b$  observable proves to be effective in discriminating photons, as indicated by the Q factor values listed in Table 1. The majority of Q factor values are greater than 3.

The highest Q factor, 11.3, was achieved in the  $nHit_{up}$  range of 4000 to 5000, corresponding to an energy of approximately 10 TeV.



**Figure 6:** The discrimination power of the observable  $S_b$  as a function of  $b$  is shown on the left, measured by the merit factor, and on the right, measured by the quality factor  $Q$  for  $nHit_{up}$  in the range (1000-2000).

On the other hand, the lowest  $Q$  factor value, 2.0, was obtained in the last bin of the  $nHit_{up}$  range. This can be attributed to the wider bin mixing showers from a larger range of energies and therefore different characteristics.

## 5. Conclusions

The  $S_b$  observable in the SWGO detector was adjusted using a scaling factor that accounted for the different configurations of the outrigger array. This scaling factor served as a weight to compensate for the lower density of the outrigger array in the calculation of station multiplicity ( $nHit_{up}$ ).

We investigated the optimal value of the parameter  $b$  for  $S_b$  in different  $nHit_{up}$  intervals. The maximum separation was achieved with a merit factor of 1.55 and a contamination rate of 0.2% (where one proton contaminates the sample) at energies around 10 TeV ( $nHit_{up} = [4000, 5000]$ ), corresponding to a  $Q$  value of 11.3.

The  $S_b$  observable demonstrates good performance in separating gamma-ray and hadron showers in a dense array, particularly within the energy range detected by SWGO. However, the discrimination power decreases for larger  $nHit_{up}$  values, as wider bins are used, which mix showers from different energies to obtain more statistical data.

## References

- [1] G. Ros, A. D. Supanitsky, G. A. Medina-Tanco, L. del Peral, J. C. D’Olivo, and M. D. Rodríguez Frías, *A new composition-sensitive parameter for Ultra-High Energy Cosmic Rays*, *Astroparticle Physics*, 35, 140-151, 2011.
- [2] G. Ros, A. D. Supanitsky, G. A. Medina-Tanco, L. del Peral, and M. D. Rodríguez Frías, *Improving photon-hadron discrimination based on cosmic ray surface detector data*, *Astroparticle Physics*, 47, 10-17, 2013.
- [3] <https://www.auger.org/>
- [4] Pierre Auger Collaboration, *A Search for Photons with Energies Above  $2 \times 10^{17}$  eV Using Hybrid Data from the Low-Energy Extensions of the Pierre Auger Observatory*, *The Astrophysical Journal* (2022) 933:125.
- [5] R. Atkins *et al* Milagro Collaboration, *Observation of TeV Gamma Rays from the Crab Nebula with Milagro Using a New Background Rejection Technique*, *Astrophys.J.*595:803-811,2003.
- [6] <https://www.swgo.org/SWGOwiki/doku.php>
- [7] D. Heck, J. Knapp, J.N. Capdevielle, G. Schatz, T. Thouw, *CORSIKA: a Monte Carlo code to simulate extensive air showers*, Forschungszentrum und Universität Karlsruhe, Karlsruhe