Published by IOP Publishing for Sissa Medialab

Received: *November 8, 2024*
Revised: *February 15, 2025*
Accepted: *March 19, 2025*
Published: *April 23, 2025*

TECHNICAL REPORT

Gamma/hadron separation using the arrival time distribution of particle cascades at TeV energies for SWGO

D. Luzquiños, A. Colán and J. Bazo*

Sección Física, Departamento de Ciencias, Pontificia Universidad Católica del Perú, Av. Universitaria 1801, Lima, Peru

E-mail: jbazo@pucp.edu.pe

Abstract: Given the success of high-altitude wide-field gamma-ray detectors, such as HAWC and LHAASO, we explore a new gamma-hadron separation variable for the future Southern Wide-field Gamma-ray Observatory (SWGO), currently in the R&D phase. SWGO will be a ground-based, high duty cycle, extensive air shower water Cherenkov detector array with a high fill factor core, expected to be located in the Atacama Astronomical Park, Chile, at an altitude of 4770 m. To identify gamma ray astrophysical sources, primary particles need to be reconstructed from the air showers reaching the detector array using their characteristics to distinguish between gamma rays, considered as signal, and hadrons (i.e. cosmic rays) that are considered background. We use CORSIKA to simulate the development of air showers in the atmosphere up to the arrival of secondary particles at the array of water Cherenkov tanks. We propose the arrival time distribution of secondary particles reaching the detector array as an alternative gamma/hadron separator variable. To evaluate its performance we simulated photons and protons, as primary particles, in the energy range from 1 to 100 TeV for vertical events (i.e. zenith angle = 0º) reaching the center of the array. The optimal separation parameter found, given the above constraints, is the time of the 15% percentile of arriving particles inside a ring of 100 to 150 m. The recognized signal is $\gtrsim$ 88% on average and the background rejection is $\gtrsim$ 79%. Nevertheless, the overall time resolution of the tanks, estimated at 3.2 ns, is comparable to the average time separation between photons and protons, which is above 3.7 ns. Consequently, the actual efficiency of this variable is expected to be lower.

Keywords: Particle identification methods; Detector modelling and simulations I (interaction of radiation with matter, interaction of photons with matter, interaction of hadrons with matter, etc); Gamma detectors; Simulation methods and programs

*Corresponding author.

https://doi.org/10.1088/1748-0221/20/04/T04008

2025 JINST 20 T04008

Contents

1 Introduction

Observatories based on arrays of Cherenkov water tanks at ground level have revealed significant scientific potential for studying and discovering astrophysical sources of the highest energy gamma-rays. The Southern Wide-field Gamma-ray Observatory (SWGO) [1–3] will be such a detector located in the Southern hemisphere. It aims to complement the work of other experiments located in the Northern Hemisphere, such as HAWC (High-Altitude Water Cherenkov) [4] in Mexico and LHAASO (Large High Altitude Air Shower Observatory) [5, 6] in China, and will also be a counterpart to CTAO [7].

High energy particles reaching the Earth's atmosphere from astrophysical sources generate Extensive Air Showers (EASs). EASs generated by cosmic-ray protons and nuclei are the main source of background for discovering astrophysical gamma-ray sources. With the goal to distinguish gamma primary particles various gamma/hadron (G/H) separation techniques are employed [8–16], such as analyzing the shape (profile) of the cascade, the penetration depth of particles in the atmosphere, the compactness of the shower C, the muon count, etc. as well as machine learning methods that use multiple variables simultaneously to achieve a more accurate categorization.

The main method for G/H separation involves counting muons reaching the Earth's surface, achieving a signal-to-noise ratio of $\frac{\epsilon\gamma}{\sqrt{\epsilon_{\text{proton}}}} = 12$, where ϵ is the efficiency [17]. Gamma-rays generate particle showers with relatively few muons, while hadronic showers, produced mainly by protons, produce more muons due to nuclear interactions.

In this work, we aim to find a new variable to separate EASs generated by gamma rays (signal) from those induced by cosmic-rays (background) at TeV energies using the timing and location information from secondary particles that reach the SWGO detector. Using CORSIKA [18] simulations we examine the distribution of particle arrival times within specific radii of the detector reference configuration to find the best definition of a new G/H separation variable and its associated cut value.

This work is divided as follows: section 2 gives an overview of the SWGO experiment and reference detector configuration, section 3 focuses on simulations using CORSIKA software to analyze extensive air showers (EAS), section 4 searches for a new G/H time separation variable finding the best cut for this variable and its performance. Finally, in section 5 we present the conclusions and limitations of this work.

2025 JINST 20 T04008

2 The Southern Wide-field Gamma-ray Observatory (SWGO)

SWGO [1–3] will be an astrophysical wide field-of-view observatory that relies on the Cherenkov effect in water for particle detection. It will be located in Pampa La Bola at the Atacama Astronomical Park in Chile, at 4770 m above sea level. Its duty cycle is close to 100%, with an instantaneous field-of-view of the order of steradians. SWGO will cover an energy range from 100 s of GeV to PeV and is currently in the R&D phase.

The scientific goals of SWGO include, among others, observing the galactic center, that can only be done from Southern latitudes, galactic and extragalactic sources, specially PeVatrons, the Fermi bubbles emanating from the center of our galaxy, regions of stellar formation and diffuse galactic emission, transient events as Gamma-Ray Bursts, etc. as well as cosmic-ray observations and physics Beyond the Standard Model [1, 2].

For this work, we will use one possible detector array configuration [3] consisting of 6589 unit detectors arranged in a circular area covering just under 0.30 km^2. The array has an inner core of 160 m radius with 80% tank occupancy and an outer ring surrounding the core with a maximum radius of 300 m, with an occupancy of 5% [19]. The water tank used for this work [3] is double-layered with a diameter of 3.82 m and total height of 3 m. It has 2 PMTs pointing upwards and downwards, respectively, each located in the center of the tank at the height of the separation level (0.50 m) between the two chambers. The lower chamber is used for muon identification.

For this study, the backup site of Imata, in Arequipa-Peru, was chosen to perform the simulations. The Imata site is located at 4450 meters above sea level at a latitude of 15.9ºS. However, the primary site in Pampa La Bola at the Atacama Astronomical Park in Chile would not differ much from these results given its relative similar characteristics.

3 CORSIKA simulations

We use CORSIKA [18] (version 7.7420 with QGSJET-II-04 for high energy hadronic simulations and URQMD 1.3cr for low energy hadronic simulations) to simulate the propagation of the EASs produced by primary particles to the detector surface. The geomagnetic conditions and observation level of the Imata site were used to simulate the atmospheric showers.

We simulated as primary particles photons and protons with fixed energies in the range from 1 to 100 TeV, vertically incident to the center of the detector array (i.e. zenith=0º), consisting of 1000 showers for each particle type and energy.

A geometrical filter based on the tanks reference configuration, seen in figure 1, was used. A minimum energy threshold of 50 MeV [1] is applied to secondary particles falling into the tanks to have been detected. However, a detailed detector simulation, including its time resolution and detection efficiency, is not included in this study.

4 Search for a G/H separation variable based on the arrival time distribution

There are several implemented algorithms to differentiate between gamma- and hadron-initiated air showers in various previous gamma-ray observatories. For instance, in the case of HAWC 99.9% of the detected events are generated by cosmic-rays [20]. Some of the most common variables include the muon count, the 68% containment radius, the lateral distribution, the compactness parameter [21] as well as PINCness [21], LIC, disMax (Maximum Distance between PMTs) [9], etc.

2025 JINST 20 T04008

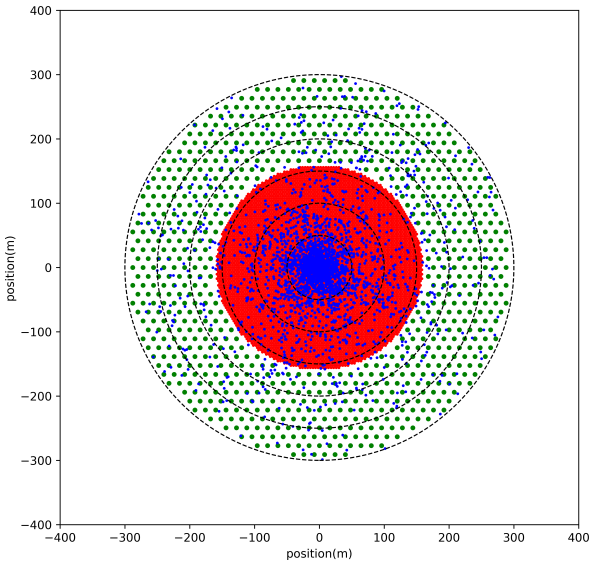
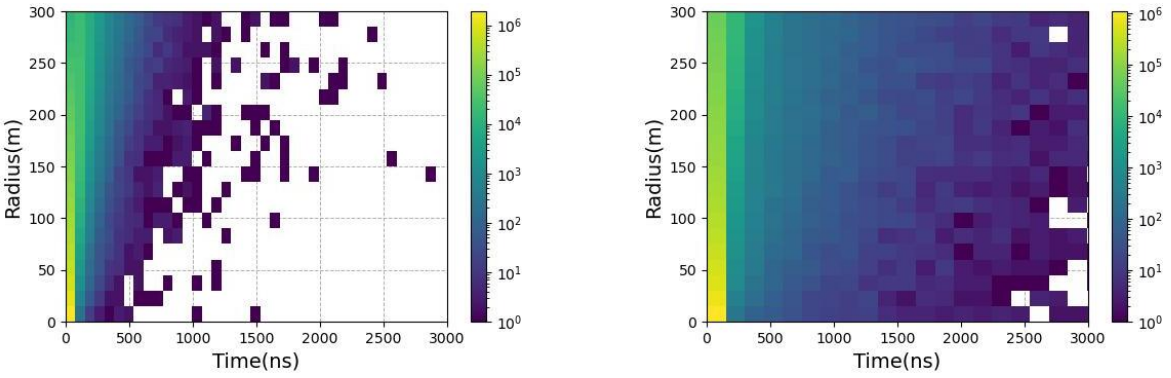


Figure 1. Geometry of the SWGO detector array tank distribution used for this analysis. Inner core tanks in red have 80% area coverage and in green are the outer array tanks with 5% filling factor. The blue points represent the arriving particles reaching the ground for a 100 TeV photon using CORSIKA. The concentric black circles, increasing by 50 m, divide the array into six sectors to be studied.

We analyze the time of the secondary particles reaching the sensitive area of the tanks as a possible new G/H variable. We simulate 1k shower for each energy and primary particle using CORSIKA. We set to zero the time when the first secondary particle reaches the observation level. As an example, in figure 2 the arrival time versus the incidence radius of secondary particles for photons and protons for a 100 TeV event is shown for a maximum radius of 300 m, corresponding to one possible SWGO layout. The photon event, as expected, is less dispersed than the hadronic event.



(a) Photons (b) Protons

Figure 2. Time distribution of secondary shower particles reaching the observation level for 100 TeV primary energy, inside the maximum radius of one possible SWGO layout.

To characterize the time distribution, we define $t_{N\%}$ as the elapsed time between the first particle impacting the tanks until when N% of the total particles reach the detector. After simulating several showers, we calculate the mean $\overline{t_{N\%}}$ and its corresponding standard deviation (σ), as shown in figure 3.

2025 JINST 20 T04008

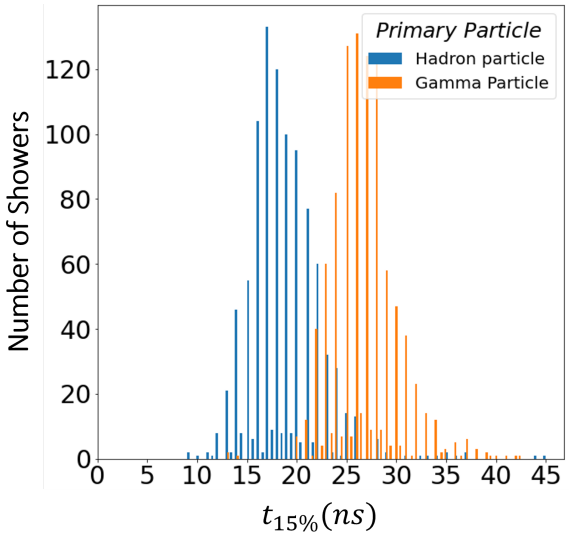


Figure 3. Distribution of $t_{N=15\%}$ for 1000 simulated 100 TeV photon and proton showers.

Furthermore, to add hit position information, since SWGO is a detector with radial symmetry, we divide the tank array into 6 concentric rings (i.e. sectors) of 50 m radius, to get significant particle statistics, from 0 to 300 m, as shown in figure 1. Since we are only simulating vertical events centered in the origin, there is no need for a variable geometrical division, which would depend on the primary particle zenith angle and core position inside the tank array and could also be adjusted for the inner and outer cores.

We explore $\overline{t_{N\%}}$ for each energy, particle type and sector the N% percentile of the time distribution in the range from 5% to 95% in steps of 5%, seeking to identify the best variable that enables differentiation between the two types of atmospheric showers. The last percentile is not considered due to the stochastic nature of the last few secondary particles.

To quantify the gamma-hadron separation for each energy (E), sector (r) and percentile (N%) we use:

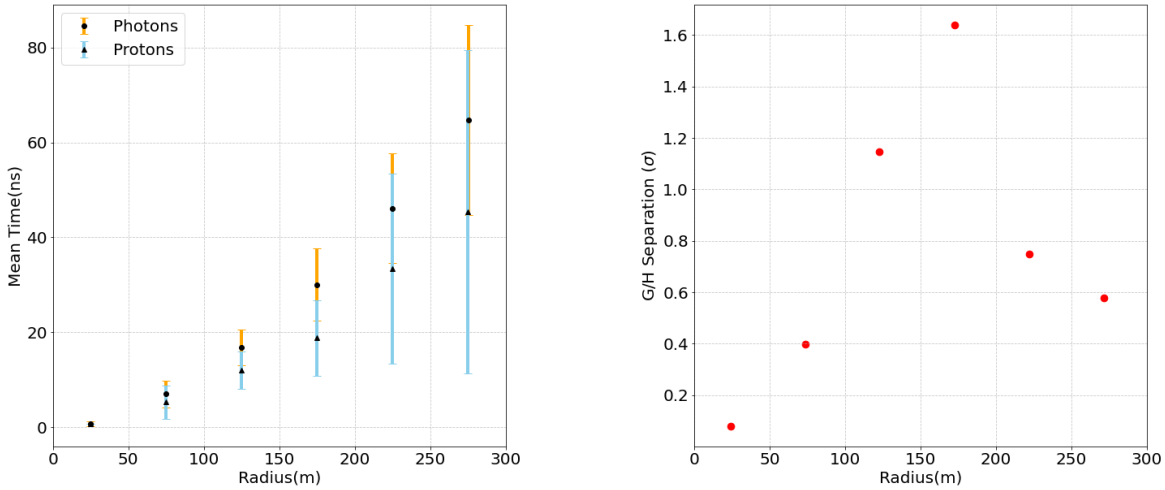
$$\sigma_{G/H}(E, r, N\%) = \frac{|\overline{t_{\text{hadron}}} - \overline{t_\gamma}|}{\sqrt{\sigma^2_{\text{hadron}} + \sigma^2_\gamma}} \tag{4.1}$$

where $\overline{t_{\text{hadron}}}$ and $\overline{t_\gamma}$ are the mean times arrival times for a given energy, sector and percentile for hadrons and photons, respectively, and σ_{hadron} and σ_γ are their associated standard deviations.

As an example, the results for 100 TeV particles and $N_{15\%}$ as a function of the radius are shown in figure 4(a) for the mean time and in figure 4(b) for the $\sigma_{G/H}$. We see that the error increases as the radii of the 50 m radial bins (i.e. rings) increase, since less particles are found in the outer array, given the sparse density of tanks. Furthermore, the dispersion of the hadronic events is higher and increases more rapidly than that of photon events. The separation has a maximum at 150 m radius, which suggests that, in this range, the optimal distance for separation is found, since for larger radius the outer array with less tank density is found. In this simulation, the shower core is assumed to coincide with the center of the array. Thus, for events in which the shower core is located far from the array center, the conclusions may differ.

However, it is necessary to search for different percentiles and energies in order to find the best variable. For example, the percentile scan for the first four rings (the last rings have too low

2025 JINST 20 T04008



(a) Average time for gamma and hadron showers.

(b) Gamma/Hadron separation based on eq. (4.1).

Figure 4. Gamma/Hadron separation using the 15% percentile at 100 TeV as a function of the radius in 50 m bins for SWGO, using 1000 simulated showers.

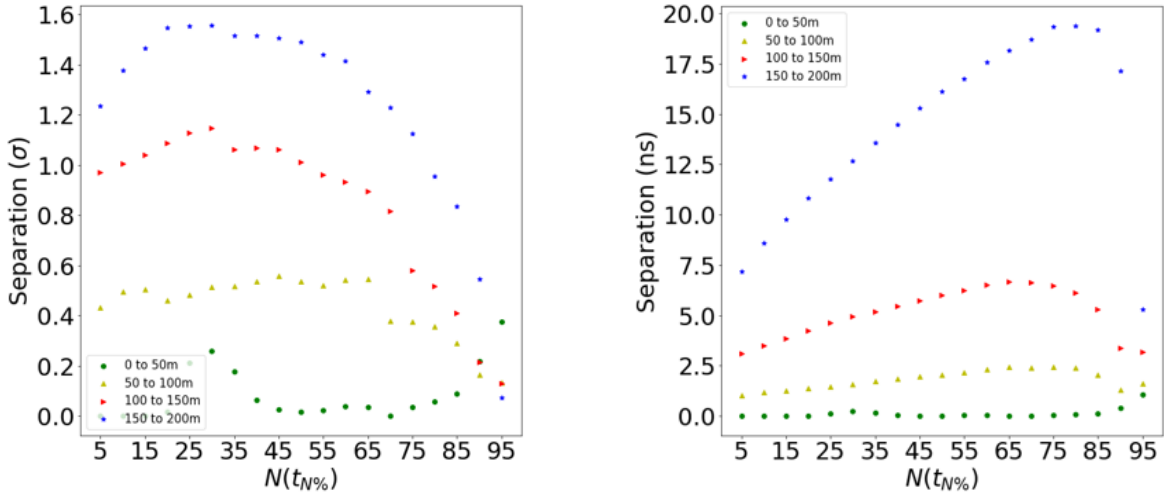
statistics) at 100 TeV is shown in figure 5(a). The best separations are found at larger radius around the lower percentiles. However, the rings farthest from the center of the shower have very low statistics compared to the inner rings and, therefore, it is better to use the inner array (in the range from 0 to 150 m). Thus, the best candidates for a G/H separator are found in the ring from 100 to 150 m in the percentiles below 50%.

On the other hand, the difference between mean times $|\bar{t}_{\text{hadron}} - \bar{t}_{\text{gamma}}|$ for the same percentile scan and the first 4 rings is shown in figure 5(b). As expected, the time differences increase as the radius increases. The same behaviour is observed as the percentile increases for larger radius, since for smaller radius most particles arrive almost at the same time. It is worth mentioning that the time differences for radius larger than 100 m are greater than 2 ns, the minimum time resolution of standard PMTs. Specifically, the time difference for ($t_{15\%}$ in the 100 to 150 m ring at 100 TeV is 3.7 ns.

However, it is necessary to estimate the time resolution of the tank given the photon propagation inside the water until reaching the PMT. For this we used secondary particles of the gamma-initiated air-showers reaching ground from CORSIKA and passed them through a water Cherenkov tank using the AERIE software [22] from the SWGO Collab. The arrival time distribution of photo-electrons (PE) reaching the upper PMT is shown in figure 6. The energy of the primary particle does not change the distribution, while there is a small difference between photons and protons. Given the maximum standard deviation of the distributions of 0.61 ns, we take 1.2 ns as the tank time uncertainty for vertical showers. Thus, the overall (PMT + tank) time resolution is 3.2 ns, which is slightly less than the ($t_{15\%}$ time difference of 3.7 ns. Therefore, an improvement of the overall time resolution would be desirable to use this time variable without constraints as a G/H separator. Nevertheless, we carry on with the analysis to estimate its full potential.

Next, we analyze the percentiles in the energy range from 1 to 100 TeV in only two rings (i.e. from 50 to 150 m), discarding the sector from 0 to 50 m because in it almost all particles arrive at the same time and also the rings in the outer array given the low filling factor. In each ring, the percentile with the greatest separation for each energy is found. Since we do not observe a percentile

2025 JINST 20 T04008



(a) Gamma/Hadron separation based on eq. (4.1).

(b) Mean time difference: ($|\bar{t}_{\text{hadron}} - \bar{t}_{\text{gamma}}|$).

Figure 5. Gamma-Hadron separation at 100 TeV in SWGO for different percentiles and rings, using 1000 simulated showers.

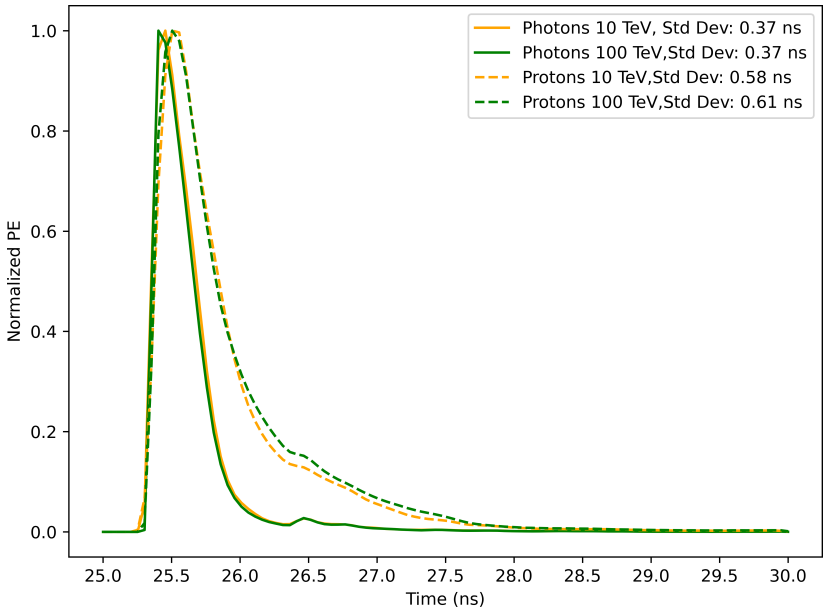


Figure 6. Normalized arrival time distribution of PE reaching the upper PMT inside a cylindrical double-layered tank of 3.82 m diameter and 3 m height for different energies of primary vertical photon- and proton-initiated showers, with 50 events per energy. The secondary particles from the CORSIKA output are injected at the center of the tank and transported using the AERIE code [22]. The following number of (photons: secondary particles, PE; protons: secondary particles, PE) are generated: (146, 327 k; 1.7 k, 690 k) at 10 TeV and (2.2 k, 4697k; 25 k, 8662 k) at 100 TeV. The standard deviation is given for each energy representing the uncertainty.

dependence with energy, a constant fit is applied, considering a 5% error for each percentile, since it is its bin size, which is presented in figure 7. The ring with the best fit (i.e. lowest reduced χ^2 = 2.44) corresponds to the 100 to 150 m ring for N%=16. Thus the percentile $t_{15\%}$ in the 100 to 150 m ring is chosen as the best G/H separator.

2025 JINST 20 T04008

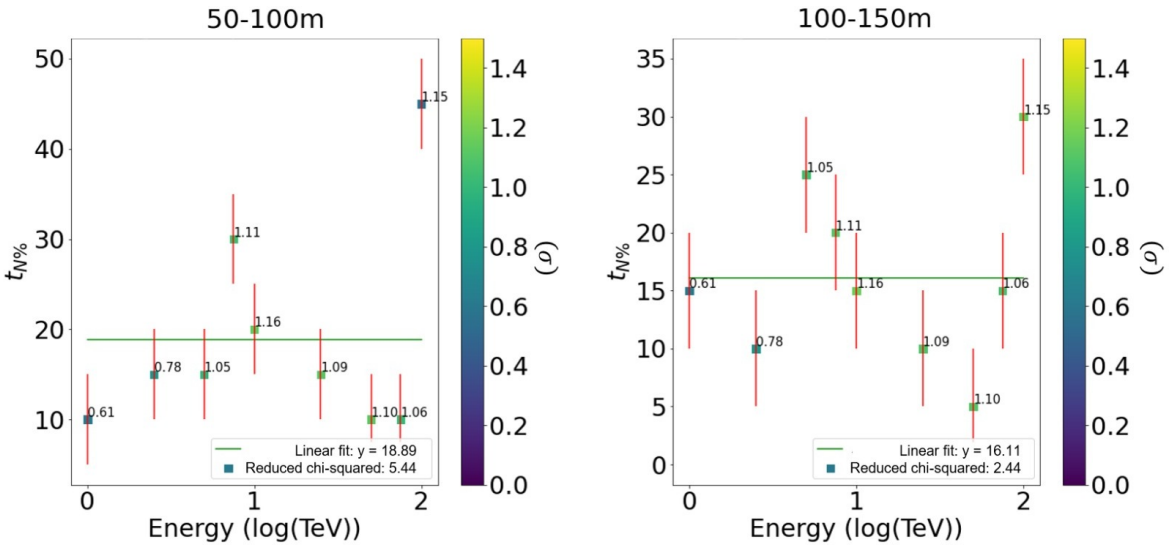


Figure 7. Time percentile ($t_{N\%}$ which gives the greatest G/H separation (color coded as a number of sigma) as a function of energy, for two SWGO rings. 1000 simulated showers are used for each energy. A constant fit is applied for each ring.

For the $t_{15\%}$ in the 100 to 150 m ring cut value we analyze the mean time difference between photons and protons as shown in figure 8 and compare it to the overall water Cherenkov tank resolution. The average difference is above the time resolution across all energies. However, when the spread of the distribution is considered, this is no longer the case, limiting the performance of this variable.

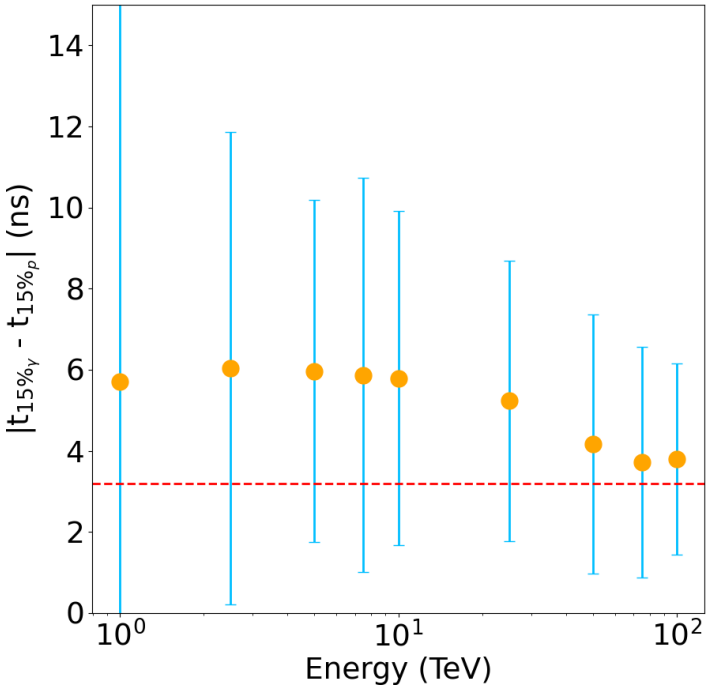


Figure 8. Mean time separation for $t_{15\%}$ in the 100 to 150 m ring between photons and proton initiated showers for different energies, using 10k events for each energy. The red dashed line shows the overall water Cherenkov tank resolution.

2025 JINST 20 T04008

Then, we search for the best cut value to obtain the highest gamma-ray efficiency and cosmic-ray rejection. For this study, we increase the statistics to 10k simulated showers for each primary and energy. From these showers 70% will be used to define the cut value and the remaining 30% events to evaluate the chosen cut efficiency.

Since we know that the mean arrival time for the whole array for a given percentile is lower for hadronic showers than gamma-initiated showers (see figure 3 and figure 9), any shower that has mean arrival time greater than the cut value will be considered a gamma event, and consequently, if its mean time is less than the cut value, it will be a hadronic event.

To simplify the search, we only evaluated two cuts around the values where the signal and background superpose: a lower value $t_{\min} = \overline{t_{15\%p}} - \sigma_p$ and an upper value $t_{\max} = \overline{t_{15\%\gamma}} + \sigma_\gamma$. Figure 9 shows the $t_{15\%}$ distribution for photons and protons at different energies, showing with a black line a possible cut.

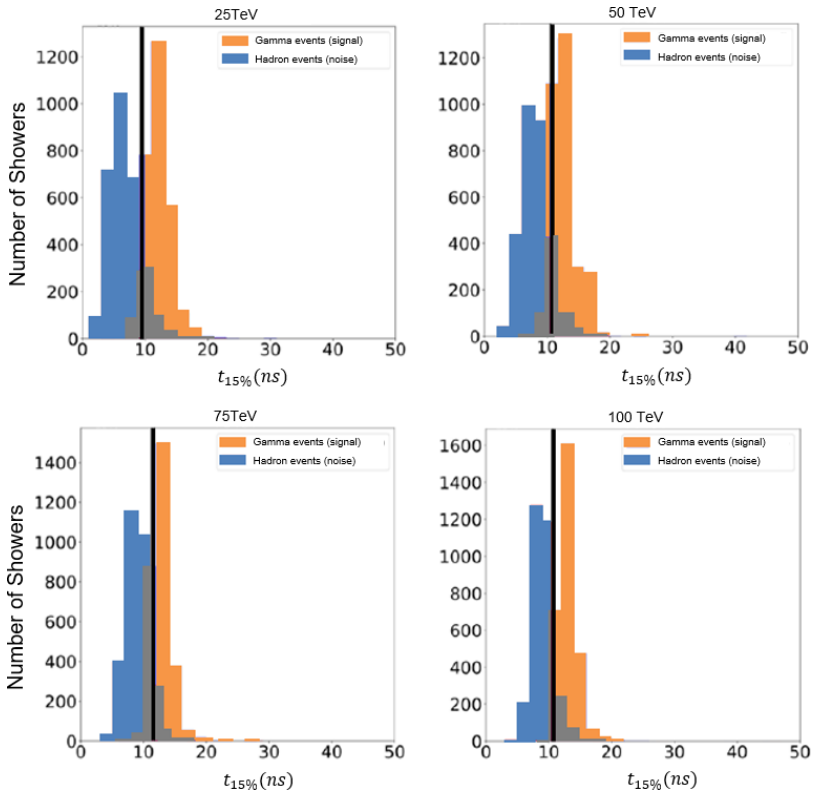
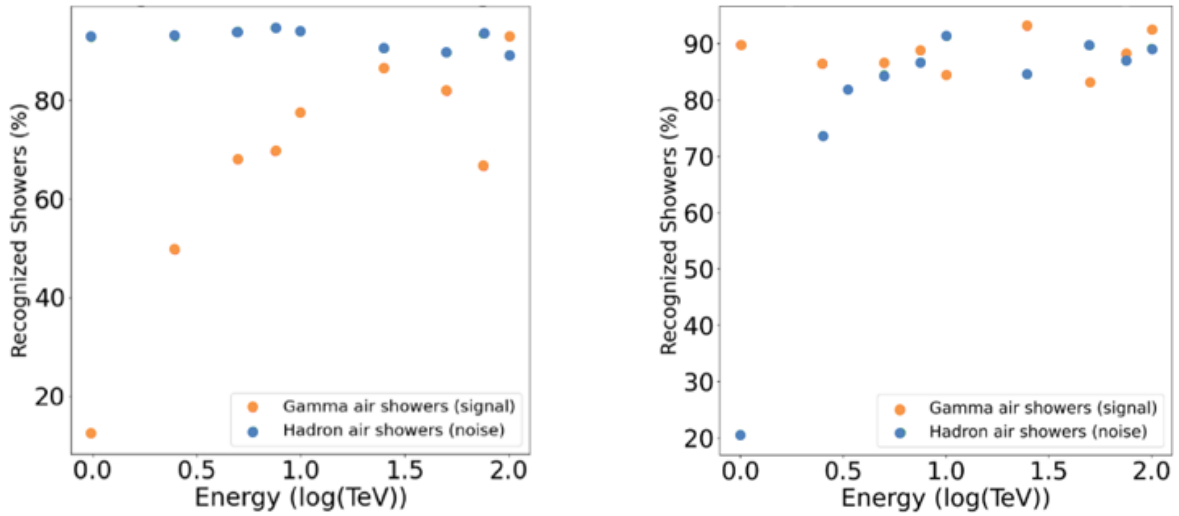


Figure 9. Distributions of $t_{15\%}$ for photons and proton initiated showers in the 100 to 150 m ring for different energies, using 3k events for each energy. The black line represents the $t_{\min}$ cut value and was calculated using 7k events for each energy.

Using both cuts, the signal and background efficiencies are calculated with the rest (i.e. 30%) of the simulated sample for each energy, as shown in figure 10). For the $t_{\min}$ cut the background efficiency exceeds 89% across all energies, whereas the signal efficiency is considerably lower at lower energies (1–5 TeV) but increases above 60% at higher energies. For the $t_{\max}$ cut, an increase in signal efficiency can be observed. However, background rejection decreases. Since we are interested in good signal efficiency, we chose $t_{\max}$ as the best cut value, obtaining an average of 88% recognized signal, which is comparable to the muon count variable [12] that obtained 92%. On the other hand, 78.6% of the background is rejected, which is lower than the muon count variable [12] that obtained 99.97%.

2025 JINST 20 T04008



(a) Using t_{min} as cut value.

(b) Using t_{max} as cut value.

Figure 10. Signal and noise efficiency applying a cut on $t_{15\%}$, using 3k events for each energy.

5 Conclusions

A new candidate for a gamma-hadron separation variable based on the arrival time and spatial distribution of secondary particles in the context of the SWGO experiment has been presented. The optimal variable corresponds to the 15th percentile of the time distribution ($t_{15\%}$) in the 100 to 150 m ring for energies between 1 and 100 TeV, without an important energy dependence.

Using as cut value $\overline{t_{15\%\gamma}} + \sigma_\gamma$ the signal and background efficiencies obtained were on average $\gtrsim$ 88% and $\gtrsim$ 79%, respectively, sufficiently high to be comparable with the results from muon counting. Thus, this variable is a good candidate for G/H separation.

However, the overall water Cherenkov tank resolution, including the photon scattering inside the tank and the PMT resolution, is estimated to be around 3.2 ns. When this time resolution is compared to the time separation between photons and protons across all energies, we see that even if the average is above the resolution, the spread of the distribution is not. Thus, an improvement in the overall time resolution is needed to ensure that this time variable can be used as a good G/H separator.

The study simulated perfect conditions without any detector associated efficiency or random noise due to atmospheric muon background, using air-showers in the center of the detector without any inclination. A different geometry rather than the simple rings with fixed radius would be necessary to take into account an inclined event shifted from the center. Another simplification of this study is that we only use protons as the cosmic-ray background instead of the real heavy nuclei composition. In addition, showers with a core far from the central high-density array might not have enough statistics for this method to work. Therefore, the efficiency of this G/H separation variable on real events, will most probably be worse.

Finally, this new proposed G/H separation variable could be integrated into more complex machine learning algorithms that include other known variables to increase the sensitivity of the SWGO experiment.

2025 JINST 20 T04008

Acknowledgments

D.L. and J.B. appreciates the support from the Peruvian National Council for Science, Technology and Technological Innovation (CONCYTEC) under Project Grant 078 – 2021. J.B. thanks the Dirección de Fomento de la Investigación (DFI-PUCP) for funding under Grant No. DFI-2021-C-0020. A.C. thanks the Faculty of Science and Engineering (FCI-PUCP) for funding under Grant No. 001 4121 AG0346 AB23 07. The authors thank our colleagues within the SWGO Collaboration for useful suggestions and the AERIE software framework used in this work.

References

[1] A. Albert et al., *Science Case for a Wide Field-of-View Very-High-Energy Gamma-Ray Observatory in the Southern Hemisphere*, arXiv:1902.08429.

[2] P. Abreu et al., *The Southern Wide-Field Gamma-Ray Observatory (SWGO): A Next-Generation Ground-Based Survey Instrument for VHE Gamma-Ray Astronomy*, arXiv:1907.07737.

[3] R. Conceição, *The Southern Wide-field Gamma-ray Observatory*, *PoS* **ICRC2023** (2023) 963 [arXiv:2309.04577].

[4] A.U. Abeysekara et al., *Observation of the Crab Nebula with the HAWC Gamma-Ray Observatory*, *Astrophys. J.* **843** (2017) 39 [arXiv:1701.01778].

[5] LHAASO collaboration, *The Large High Altitude Air Shower Observatory (LHAASO) Science Book (2021 Edition)*, *Chin. Phys. C* **46** (2022) 035001 [arXiv:1905.02773].

[6] F. Aharonian et al., *The observation of the Crab Nebula with LHAASO-KM2A for the performance study*, *Chin. Phys. C* **45** (2021) 025002 [arXiv:2010.06205].

[7] CTA Consortium collaboration, *Science with the Cherenkov Telescope Array*, World Scientific (2018) [DOI:10.1142/10986] [arXiv:1709.07997].

[8] Z. Hampel-Arias and S. Westerhoff, *Gamma Hadron Separation using Pairwise Compactness Method with HAWC*, *PoS* **ICRC2015** (2016) 1001 [arXiv:1508.04047].

[9] HAWC collaboration, *Gamma/hadron separation with the HAWC observatory*, *Nucl. Instrum. Meth. A* **1039** (2022) 166984 [arXiv:2205.12188].

[10] X. Wang, *Gamma Hadron separation using traditional single parameter method and multivariate algorithms with LHAASO-WCDA experiment*, *PoS* **ICRC2019** (2020) 820.

[11] S. Westerhoff et al., *Separating gamma and hadron induced cosmic ray air showers with feed forward neural networks using the charged particle information*, *Astropart. Phys.* **4** (1995) 119.

[12] N.M. Gonzalez, F.A. Snchez, M. Roth and A. Etchegoyen, *A Muon-based Observable to Detect Photons at Ultra-high Energies*, *PoS* **ICRC2019** (2020) 271.

[13] R. Conceição et al., *Muon identification in a compact single-layered water Cherenkov detector and gamma/hadron discrimination using machine learning techniques*, *Eur. Phys. J. C* **81** (2021) 542 [arXiv:2101.10109].

[14] R. Conceição, L. Gibilisco, M. Pimenta and B. Tomé, *Gamma/hadron discrimination at high energies through the azimuthal fluctuations of air shower particle distributions at the ground*, *JCAP* **10** (2022) 086 [arXiv:2204.12337].

[15] B.S. González et al., *Tackling the muon identification in water Cherenkov detectors problem for the future Southern Wide-field Gamma-ray Observatory by means of machine learning*, *Neural. Comput. Appl.* **34** (2022) 5715.

2025 JINST 20 T04008

[16] J. Glombitza et al., *Application of graph networks to a wide-field water-Cherenkov-based Gamma-ray Observatory*, *JCAP* **02** (2025) 066 [arXiv:2411.16565].

[17] S. Kunwar et al., *A double-layered Water Cherenkov Detector array for Gamma-ray astronomy*, *Nucl. Instrum. Meth. A* **1050** (2023) 168138 [arXiv:2209.09305].

[18] D. Heck et al., *CORSIKA: A Monte Carlo code to simulate extensive air showers*, FZKA-6019 (1998).

[19] U.B. Barres de Almeida, *Status of the southern wide-field gamma-ray observatory (SWGO)*, *Frascati Phys. Ser.* **74** (2022) 123.

[20] T. Capistrán et al., *Use of Machine Learning for gamma/hadron separation with HAWC*, *PoS* **ICRC2021** (2021) 745 [arXiv:2108.00112].

[21] T. Capistrán, I. Torres and E. Moreno, *Gamma/hadron separation in HAWC using neural networks*, *Proc. SPIE Int. Soc. Opt. Eng.* **9908** (2016) 990845.

[22] HAWC collaboration, *Data Acquisition Architecture and Online Processing System for the HAWC gamma-ray observatory*, *Nucl. Instrum. Meth. A* **888** (2018) 138 [arXiv:1709.03751].

2025 JINST 20 T04008