Gamma/hadron discrimination using a small-WCD with four PMTs

R. Conceição, a,b,∗ P. Assis, a,b F. Assunção, c A. Bakalová, d U. Barres de Almeida, e C. R. Bom, e J. Correia, c A. De Angelis, f,a L. Dias, c B. S. González, a,b A. Guillén, g G. La Mura, d N. Lourenço, c P. Machado, c S. Marques, a,b L. Mendes, a M. Pimenta, a,b R. Shellard, e B. Tomé a,b and J. Vicha d on behalf of the SWGO Collaboration
(a complete list of authors can be found at the end of the proceedings)

aLIP - Laboratório de Instrumentação e Física Experimental de Partículas, Lisbon, Portugal
bIST - Instituto Superior Técnico, Lisboa, Portugal
cDepartment of Informatics Engineering, CISUC, University of Coimbra, Coimbra, Portugal
dInstitute of Physics of the Czech Academy of Sciences, Prague, Czech Republic
eCBPF - Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil
fINFN and Università di Padova, Italy
gComputer Architecture and Technology Department, University of Granada, Granada, Spain
E-mail: ruben@lip.pt

The Southern Wide-field Gamma-ray Observatory (SWGO) is the next-generation gamma-ray observatory, currently in an R&D phase. The experiment is expected to have a large array of water Cherenkov detectors (WCD) placed at a high elevation (> 4.4 km a.s.l.) in South America. Here we present a WCD concept with reduced surface area and height of stations comprising four PMTs at the bottom. We show that it is possible to reach an excellent gamma/hadron discrimination by analysing the data gathered by this station with machine learning techniques. Such performance can be achieved by analysing the shower patterns at the ground or through the PMTs signal time structure to tag muons. Moreover, it is shown that the station’s performance does not depend on the array configuration (dense or sparse) nor on the shower inclination (θ < 40°). Such a concept reduces the cost associated with the transport of massive amounts of water to high elevation sites while keeping a high physics performance. Therefore, it could be a good candidate station for SWGO, enabling to reach good sensitivities from low energies (∼ 100 GeV) up to the PeV region, covering large ground surface areas (few square km).

37th International Cosmic Ray Conference (ICRC 2021)
July 12th – 23rd, 2021
Online – Berlin, Germany

∗Presenter
1. Introduction

Very-high-energy gamma-rays are essential to understand our surrounding Universe. With the discovery of multimessenger events such as the detection of gravitational waves with electromagnetic counterparts [1] and the simultaneous detection of very-high-energy neutrinos and gamma-rays from a blazar [2], and the observation of ultrahigh-energy photons surpassing the PeV [3], it becomes essential to have wide field-of-view gamma-ray observatories able to survey the sky permanently and covering an energy range going from 100 GeV up to the PeV. Moreover, it is of the utmost importance that this experiment is able to survey the Southern Hemisphere sky, where the Galactic Center is located, and be complementary to the powerful but, with narrow field-of-view, Cherenkov Telescope Array [4]. The Southern Wide-field Gamma-ray Observatory (SWGO) [5] is a collaboration with a well-defined R&D plan to design an experiment that is planned to be placed in South America at an altitude greater than 4400 m above the sea level. To cover from the lowest up to the highest energy, the experiment needs to have a compact array of 80,000 m² and a sparser array covering the biggest possible area allowed by the experiment budget. SWGO station unit will be water Cherenkov detectors (WCD). The WCD is sensitive to the secondary shower particles, including the secondary photons essential to be measured when targeting the lowest energies (∼ 100 GeV). It is a technology that has been shown to be successfully used by many experiments of this kind [3, 6, 7]. However, the exact design of the station is still being studied to cope with the demanding experiment requirements. In this work, we argue that a small-WCD with four photomultipliers (PMTs) at the bottom can be a good candidate station for SWGO, fulfilling all the requirements.

2. The WCD station

The design of a water Cherenkov detector station for a gamma-ray ground observatory needs to cope with stringent physics requirements to be able to accurately reconstruct the gamma induced extensive air shower. Such a station needs to have a good trigger efficiency to the shower secondary particles that reach the ground. The arrival time of these particles has to be measured with a good time resolution, typically of the order of a few nanoseconds. And last but not least, the reconstructed showers have to be identified as having been produced by gamma-rays while not to be mistaken with the huge cosmic ray background. One of the most efficient ways to do this is to identify the presence of muons, which are highly suppressed in electromagnetic showers. The need to tag muons with the water Cherenkov principle led to large stations, like the HAWC tanks, where the electromagnetic particles are rapidly attenuated, producing small signals, while the muons would traverse the whole tank producing a large signal. Another proposed option is to have two independent layers of measurement where the muon is identified by combining the signal recorded in both layers [8, 9]. In this work, we propose to exchange the large volume of water and the complexity of the station structure by the number of photosensors. Muons would be identified by exploring the signal asymmetry between photomultipliers (PMTs), which is larger for muons than for electromagnetic particles.

The station was designed such that the signal asymmetry caused by a vertical muon is maximal while ensuring a complete signal coverage (signal uniformity). The area of the stations is a parameter that should be optimised, taking into account the array’s physics performance and the station cost.
In this work, we have considered cylindrical tanks with a base diameter of 4 m as a reasonable compromise (for reference, the LHAASO WCD has an area of 25 m$^2$).

For stations with a base diameter of 4 m, the water height and the distance of the PMTs to the tank centre, $r_{\text{PMT}}$, will depend on the number of PMTs considered. For 3, 4 and 5 PMTs, we would get an optimal distance of the PMTs to the centre of the tank of 2 m, 1.5 m and 1.25 m, respectively and a corresponding water height of 2.3 m, 1.7 m and 1.25 m. The parameter $r_{\text{PMT}}$ is chosen by placing the PMTs at a position such that together they can always collect the light produced by a vertical entering muon in any position of the tank while minimising the overlap between PMTs, as shown in figure 1 (left).

The water height is obtained, noting that we now aim for a Cherenkov light pool of a radius $r_{\text{opt}}$ and the Cherenkov angle of a relativistic muon in water is roughly $\sim 41^\circ$. The higher the number of PMTs, the lower the needed water height. The exact number of PMTs and thus, of the station dimensions, will depend on the optimisation of several parameters. They will depend on the cost of PMTs and the cost (availability) of water at very high altitude.

The walls of this WCD should be white diffusive to maximise the signal collection, essential to effectively lower the energy threshold. The 2 – 3 ns time resolution needed to accurately reconstruct the shower geometry can be achieved by exploring the direct Cherenkov light while the scattered light can be used to increase the trigger efficiency.

Hence, here we aim to show that the tagging of muons through the signal time trace asymmetry is a viable option. For that, we have considered four PMTs with the corresponding stated above dimensions. Given the arguments before, the paper results should be valid for a slightly lower/higher number of PMTs, provided that the dimensions of the WCD scale accordingly.

The extensive air showers used in this work were simulated with CORSIKA (version 7.5600) [10] and the detector response using the Geant4 toolkit (with version 4.10.05.p01) [11, 12].

To train and assess the machine learning algorithms’ performance, we used proton-induced shower simulations with energies $E_0 \in [4;6]$ TeV. The showers were generated following a $E_0^{-1}$ spectrum, an azimuth angle uniformly distributed, and a zenith angle $\theta_0 \in [5^\circ; 15^\circ]$ (vertical events) and $\theta_0 \in [25^\circ; 35^\circ]$ (inclined events).

The experiment observation level was set at 5 200 m above the sea level. In the energy range used for this study ($\sim 1$ TeV), the number of muons at lower altitudes would be essentially the same, and only the electromagnetic shower component would be attenuated (effectively increasing the energy threshold). Thus, the method to tag/count muons in the WCD proposed in this work is expected to work with similar performance regardless of the experiment’s altitude (provided that it is above about 4 000 m a.s.l.).

Two array configurations of WCDs are compared to investigate the impact of the WCDs distribution at the ground in the results for inclined showers. The array matrices cover an area of roughly 80 000 m$^2$ and have the following characteristics: a dense array with a fill factor of $\sim 80\%$; a sparse array with a fill factor of $\sim 60\%$ (there is no shadowing between stations).

A particular set of simulations was derived in order to enable a sensible comparison between gamma and hadron events. Gamma-induced showers were generated with energies between $E_0 \in [1; 1.6]$ TeV, while proton-induced showers were simulated for energies between $\sim 600$ GeV and 6 TeV. Afterwards, a cut on the total signal recorded at the ground was applied to emulate a typical
gamma/hadron discrimination using a small WCD with four PMTs

R. Conceição

**Figure 1:** (left) Geant4 visualization of the single-layered WCD design with four photomultipliers. (right) Upper scheme visualization of the WCD. A dark circular area with a radius of 1.5 m was drawn around each 8-inch PMT (represented as black circles). Taking into account the height of this WCD, the direct Cherenkov light of a vertical muon that crosses the WCD through one of these areas should be detected by its correspondent PMT (it could be detected by two PMTs in the case that a muon crosses the station through the intersection of two dark areas).

Energy reconstruction. Finally, weights were applied to ensure a realistic power law spectrum of energies (expected to be $\sim E_{0,\gamma}^{-2}$ and $\sim E_{0,\mu}^{-3}$). More details on the simulations can be found in [13].

3. Muon identification in the WCD

The muon identification is done by obtaining the probability of having a muon in the WCD station, $P^{(i)}_{\mu} \in [0; 1]$. This is done through the analysis of each PMT signal time trace using a Convolutional Neural Network (CNN) [14]. As we aim to explore both the temporal (patterns in the signal time traces) and spatial (asymmetry in the PMTs’ integrals) features, the model received as input the following variables:

- Normalised signal time trace of each PMT;
- Integral of each PMTs signal time trace up to 10 ns after the start of the signal $T_0$;
- Cherenkov light measured in the WCD (sum of the four PMTs’ signal trace integrals);
- Normalised integral of each PMTs signal time trace.

The CNN was then trained using solely stations with one single muon (i.e. no electromagnetic contamination). For more details on the Machine Learning method employed and decisions taken, see [13].

In figure 2 it is shown the inverse cumulative function for the probability $P^{(i)}_{\mu}$ extracted from proton-induced showers with $E_0 \sim 5$ TeV. In the left figure, it is shown the impact of the electromagnetic contamination on the identification of the presence of muons in the station. Requiring $P^{(i)}_{\mu} > 0.5$, the number obtained from the CNN upon the analysis of the WCD PMTs signal, one can identify muons more than half of the times (red dashed curve). The chance to misidentify a muon when there is none in the station is less than 10%. As expected, the tagging efficiency of this
method is highly dependent on the electromagnetic contamination in the WCD, becomes ineffective when the fraction of energy carried by the electromagnetic particles is higher than 50%.

This model was also tested for inclined showers and dense vs sparse arrays (see fig. 2). In this figure, it can be seen that the method is nearly independent for any of the chosen configurations. It must be said that for inclined shower samples, the model had to be re-trained, meaning that one would need a CNN model for each zenith angle bin. Nonetheless, the ability to identify muons without being affected by shadowing effects from nearby stations or more inclined particle trajectories is a relevant characteristic for an experiment such as SWGO, where it is foreseen to have both a compact and a sparse array.

![Figure 2: Impact of electromagnetic contamination using the normalised inverse cumulative function for the probability $P_{\mu}^{(i)}$. Stations from proton-induced events with $E_0 \in [4, 6]$ TeV and $\theta_0 \in [5^\circ; 15^\circ]$. Normalised inverse cumulative function of the probability $P_{\mu}^{(i)}$ of stations from proton-induced events with $\theta_0 \in [5^\circ; 15^\circ]$ (dense array) and $\theta_0 \in [25^\circ; 35^\circ]$ (dense and sparse array).](image)

4. Gamma/hadron discrimination

While the accuracy on the identification of muons is important, the ultimate goal is to discriminate gamma induced showers from protons. As such, in this section, we propose a gamma/hadron discrimination strategy that relies on the use of solely the information extracted in the previous sections at the WCD station level, $P_{\mu}^{(i)}$. This information should be combined at the shower event level. A simple and intuitive observable was created for this purpose,

$$P_{\mu} = \sum_{i=1}^{N_S} P_{\mu}^{(i)}$$

where $N_S$ is the number of WCD stations triggered in the event. As mentioned in section 2 the simulations used to perform this test were produced to mimic the realistic experimental conditions: same reconstructed energy at the ground, with realistic flux energy spectra.

In figure 3 (left), it can be seen by eye that there is a large discrimination power and that the computed $P_{\mu}$ is significantly smaller for gamma induced showers than for proton. In fact, by computing the signal (gamma) efficiency $S$ over the square root of the proton background $\sqrt{B}$, one can obtain relatively high values for large gamma selection efficiencies. In particular, if one applies a cut on the distance, accepting only stations above this cut, one can reach $S/\sqrt{B} \sim 4$, values of the order of those reported in [15, 16] for $E_0 \sim 1$ TeV.
Figure 3: Distribution of the probability $P_\mu$ after evaluating all the stations of the array for gamma and proton induced showers with energies $E_0 \in [1; 1.6]$ TeV and $E_0 \in [0.6; 6]$ TeV respectively. Evaluation of the $S/\sqrt{B}$ achieved using selected stations as a function of the selection efficiency $S$ for gamma-rays.

Figure 4: Calibration of the relationship between $N_\mu = \sum_{i=0}^{N_S} N_\mu^{(i)}$ and $P_\mu = \sum_{i=0}^{N_S} P_\mu^{(i)}$ using proton induced air showers. Evaluation of the $S/\sqrt{B}$ achieved using selected stations as a function of the selection efficiency $S$ for gamma-rays. Resolution and bias of the calibration between $N_\mu$ and $P_\mu$.

Finally, it was checked whether the gamma/hadron discrimination model is sensitive to muons, or any other shower feature, by plotting the correlation between the number of muons in the stations, $N_\mu$, and the discrimination variable $P_\mu$. Such plot is shown in figure 4 (left), and it can be seen that there is a strong correlation between the two quantities. In fact, it was possible to derive a linear calibration between $P_\mu$ and $N_\mu$ and compute the bias and resolution obtained using this calibration on the estimation of the number of muons present in the shower event. In figure 4 (right), it can be seen that the bias is essentially negligible while the resolution having 20 muons in the shower event is around 20%. In fact, by fitting to a functional form of $aN_\mu^{-1/2} + b$, one can see that $b$ (the method intrinsic resolution) is $\approx 2\%$.

5. Conclusion

In this work, we have argued that a relatively small single-layered water Cherenkov detector can be a suitable option to the Southern Wide-field Gamma-ray Observatory, which is currently
being planned. In particular, we have shown for $E_0 \sim 1\text{TeV}$ that by placing four PMTs at the bottom of the station in well-chosen positions, it is possible to explore the PMTs signal time trace to tag the presence of muons in the station. This method has a remarkable performance both for vertical and inclined shower events, and for compact and sparse arrays. It is only undermined by the sizeable electromagnetic contamination expected to appear at larger energies and close to the shower core position (for instance, above $E \sim 50\text{TeV}$, only stations about $r \sim 40\text{m}$ should be used). However, the modular nature of the station can be used to explore other discrimination options, such as statistical methods to remove the electromagnetic shower component at the ground and identify the muons. Another interesting idea being currently explored is the analysis of the shower footprint on the ground. This method takes advantage of the different shower development of electromagnetic and hadronic showers, and it has given already promising results [17]. All in all, there is still much work to be done to prove that small single-layered WCD is the right option for SWGO, but the results obtained so far prove that it is an interesting path to be explored.

Acknowledgements

The authors thank also for the financial support by OE - Portugal, FCT, I. P., under project PTDC/FIS-PAR/29158/2017. R. C. is grateful for the financial support by OE - Portugal, FCT, I. P., under DL57/2016/cP1330/cT0002. The SWGO Collaboration acknowledges the support from the agencies and organizations listed here: https://www.swgo.org/SWGOWiki/doku.php?id=acknowledgements.

References


Full Authors List: SWGO Collaboration


1 Laboratório de Instrumentação e Física Experimental de Partículas (LIP), Av. Prof. Gama Pinto 2, 1649-003 Lisboa, Portugal
2 Physics Division, Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM 87545, United States
3 Aix Marseille Univ, CNRS/IN2P3, CPPM, 163 avenue de Luminy - Case 902, 13288 Marseille cedex 09, France
4 University of Padova, Department of Physics and Astronomy & INFN Padova, Via Marzolo 8 - 35131 Padova, Italy
5 Centro Atômico Bariloche, Comisión Nacional de Energía Atómica, S. C. de Bariloche (8400), RN, Argentina
6 Universidad Michoacana de San Nicolás de Hidalgo, Calle de Santiago Tapia 403, Centro, 58000 Morelia, Mich., México
7 FZU, Institute of Physics of the Czech Academy of Sciences, Na Slovaně 1, 180 51 Prague 8, Czech Republic
8 Centro Brasileiro de Pesquisas Físicas, R. Dr. Xavier Sigaud, 150 - Rio de Janeiro - RJ, 22290-180, Brazil
9 Academic Department of Physics – Faculty of Sciences – Universidad Nacional de San Antonio Abad del Cusco (UNSAAC), Av. de la Cultura, 733, Pabellón C-358, Cusco, Peru
10 Instituto de Física, Universidad Nacional Autónoma de México, Sendero Bicicumpa, C.U., Coyoacán, 04510 Ciudad de México, CDMX, México
11 Dipartimento di Fisica, Università degli Studi di Torino, Via Pietro Giuria 1, 10125, Torino, Italy
12 RWTH Aachen University, Physics Institute 3, Otto-Blumenthal-Strabe, 52074 Aachen, Germany
13 University of Maryland, College Park, MD 20742, United States
14 Durham University, Stockton Road, Durham, DH1 3LE, United Kingdom
15 Astronomical Observatory, University of Warsaw, Aleje Ujazdowskie 4, 00478 Warsaw, Poland
16 Facultad de Ciencias en Física y Matemáticas UNACH, Boulevard Belisario Domínguez, Km. 1081, Sin Número, Terán, Tuxtla Gutiérrez, Chiapas, México
17 Facultad de Ciencias Exactas, Universidad Nacional de Salta, Avda. Bolivia N° 5150, (4400) Salta Capital, Argentina
18 Department of Aerospace Science and Technology, Politecnico di Milano, Via Privata Giuseppe La Masa, 34, 20156 Milano MI, Italy
19 Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, C.P. 72592, México
20 Instituto de Astronomía y Física del Espacio (IAFE, CONICET-UBA), Casilla de Correo 67 - Suc. 28 (C1428ZAA), Ciudad Autónoma de Buenos Aires, Argentina
21 Universidad de Guadalajara, Blvd. Giral Marcelino García Barragán 1421, Olímpica, 44430 Guadalajara, Jal., México
22 Instituto de Física de São Carlos, Universidade de São Paulo, Avenida Trabalhador São-carlense, nº 400, Parque Arnold Schmidt -
gamma/hadron discrimination using a small-WCD with four PMTs

R. Conceição

CEP 13566-590, São Carlos - São Paulo - Brasil
23 INFN - Roma Tor Vergata and INAF-IAPS, Via del Fosso del Cavaliere, 100, 00133 Roma RM, Italy
24 Dept. of Physics and CCTVal, Universidad Tecnica Federico Santa Maria, Avenida España 1680, Valparaíso, Chile
25 Universität Würzburg, Institut für Theoretische Physik und Astrophysik, Emil-Fischer-Str.- 31, 97074 Würzburg, Germany
26 Department of Physics, and the Wisconsin IceCube Particle Astrophysics Center (WIPAC), University of Wisconsin, 222 West
Washington Ave., Suite 500, Madison, WI 53703, United States
27 TUBITAK Space Technologies Research Institute, ODTU Campus, 06800, Ankara, Turkey
28 Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen Centre for Astroparticle Physics, Erwin-Rommel-Str. 1, D 91058
Erlangen, Germany
29 Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias, Ave. Eugenio Garza Sada 2501, Monterrey, N.L., 64849, México
30 Max-Planck-Institut für Kernphysik, P.O. Box 103980, D 69029 Heidelberg, Germany
31 Università di Napoli “Federico II”, Dipartimento di Fisica “Ettore Pancini”; and INFN Napoli, Complesso Universitario di Monte
Sant’Angelo - Via Cinthia, 21 - 80126 - Napoli, Italy
32 University of Granada, Campus Universitario de Cartuja, Calle Prof. Vicente Callao, 3, 18011 Granada, Spain
33 IFLP, Universidad Nacional de La Plata and CONICET, Diagonal 113, Casco Urbano, B1900 La Plata, Provincia de Buenos Aires, Argentina
34 University of Utah, 201 Presidents’ Cir, Salt Lake City, UT 84112, United States
35 Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931, United States
36 Dipartimento di Fisica e Astronomia “E. Majorana”, Catania University and INFN, Catania, Italy
37 APC–IN2P3/CNRS, Université de Paris, Bâtiment Condorcet, 10 rue A.Domon et Léonie Duquet, 75205 PARIS CEDEX 13, France
38 University of Leicester, University Road, Leicester LE1 7RH, United Kingdom
39 Department of Physics, University of Trieste and INFN Trieste, via Valerio 2, I-34127, Trieste, Italy
40 Centro de Investigación en Computación, Instituto Politécnico Nacional, Av. Juan de Dios Báez S/N, Nueva Industrial Vallejo,
Gustavo A. Madero, 07738 Ciudad de México, CDMX, México
41 Department of Physics and Mathematics, Universidad de Monterrey, Av. Morones Prieto 4500, San Pedro Garza García 66238, N.L.,
México
42 Department of Physics and Astronomy, University of Alabama, Gallalee Hall, Tuscaloosa, AL 35401, United States
43 Instituto de Tecnologías en DETección y ApSTARtículas (CNEA-CONICET-UNSAM), Av. Gral Paz 1499 - San Martín - Pcia. de
Buenos Aires, Argentina
44 Department of Physics, ETH Zurich, CH-8093 Zurich, Switzerland
45 Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México (ICN-UNAM), Cto. Exterior S/N, C.U., Coyoacán,
04510 Ciudad de México, CDMX, México
46 Departamento de Fisica, Facultad de Ciencias Básicas, Universidad Metropolitana de Ciencias de la Educación, Av. José Pedro
Alessandri 774, Ñuñoa, Santiago, Chile
47 Department of Physics, University of Seoul, 163 Seoulsirripdae-ro, Dongdaemun-gu, Seoul 02504, Republic of Korea
48 Institut de recherche sur les lois fondamentales de l’Univers (IRFU), CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France
49 Laboratoire Univers et Particules de Montpellier, CNRS, Université de Montpellier, F-34090 Montpellier, France
50 Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Luis Enrique Erro 1, Puebla, México