THE EUROPEAN
PHYSICAL JOURNAL C

Background rejection using image residuals from large telescopes in imaging atmospheric Cherenkov telescope arrays

L. Olivera-Nieto[1,a] , **H. X. Ren**[1] , **A. M. W. Mitchell**[2] , **V. Marandon**[1] , **J. A. Hinton**[1]

[1] Max-Planck-Institut für Kernphysik, P.O. Box 103980, 69029 Heidelberg, Germany
[2] Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen Centre for Astroparticle Physics, 91058 Erlangen, Germany

Abstract Identification of Cherenkov light generated by muons has been suggested as a promising way to dramatically improve the background rejection power of Imaging Atmospheric Cherenkov Telescope (IACT) arrays at high energies. However, muon identification remains a challenging task, for which efficient algorithms are still being developed. We present an approach in which, rather than identifying Cherenkov light from muons, we simply consider the presence of Cherenkov light other than the main shower image in IACTs with large mirror area. We show that in the case of the H.E.S.S. array of five telescopes this approach results in background rejection improvements at all energies above 1 TeV. In particular, the rejection power can be improved by a factor $\sim$ 3–4 at energies above 20 TeV while keeping $\sim$ 90% of the original gamma-ray efficiency.

1 Introduction

The Earth's atmosphere is opaque to very-high-energy (VHE) radiation. Incoming gamma-ray photons lose their energy by initiating electromagnetic particle cascades in the atmosphere. The particles in these cascades are highly relativistic and cause the production of Cherenkov light, which is in turn collected by the telescope dishes of Imaging Atmospheric Cherenkov Telescopes (IACTs). Given the huge rates of background cosmic-rays, a factor 10^4 greater than those of gamma-rays [1], differentiating between hadronic and electromagnetic cascades is a task critical to gamma-ray astronomy.

IACTs have superior rejection power to other ground-based arrays in the domain around 1 TeV, exploiting primarily the differences in shower width and substructure between electromagnetic and hadronic showers [2]. At higher energies, of around tens of TeVs the rejection power worsens. High-energy events are often not fully contained in the camera, which makes the determination of the image parameters difficult. Additionally, this parameterisation becomes dominated by the very bright central component of high-energy showers, with additional faint emission that would indicate a hadronic origin being overlooked.

Large numbers of muons, primarily resulting from the decay of charged pions, are usually produced in the hadronic cascades initiated by cosmic-rays. The potential of muons as a tool to separate between these two classes of showers has long been recognized (see, for example [3]). Recently, [4] showed that for very large ($\gtrsim$ 20 m mirror diameter) telescopes efficient identification of muon light can potentially lead to background rejection levels up to 10^{-5} at energies above 10 TeV.

However, muon identification is not a trivial task. Muons that pass directly through the telescope dish produce a distinguishable ring-like image in the telescope camera. But once the ground impact distance increases beyond some tens of meters only a fraction of the ring arc is detected (see top panel of Fig. 1). Therefore muons which arrive far from the telescopes can easily be confused for low-energy shower images, or optical night-sky-background (NSB) noise. Alternative approaches, like making use of the signature of arrival time of muon light are promising, yet even then the task of muon identification remains a challenge [5].

Nevertheless, given the magnitude of the achievable background rejection power shown in [4], less-than efficient muon identification could still result in a significant improvement to the current background rejection power of IACTs. We present an approach in which, rather than being concerned with whether a recorded event contains muon light or not, we simply consider the presence of light other than the main shower image (see bottom panel of Fig. 1). This, of course, has the downside of being less precise in the sense that light
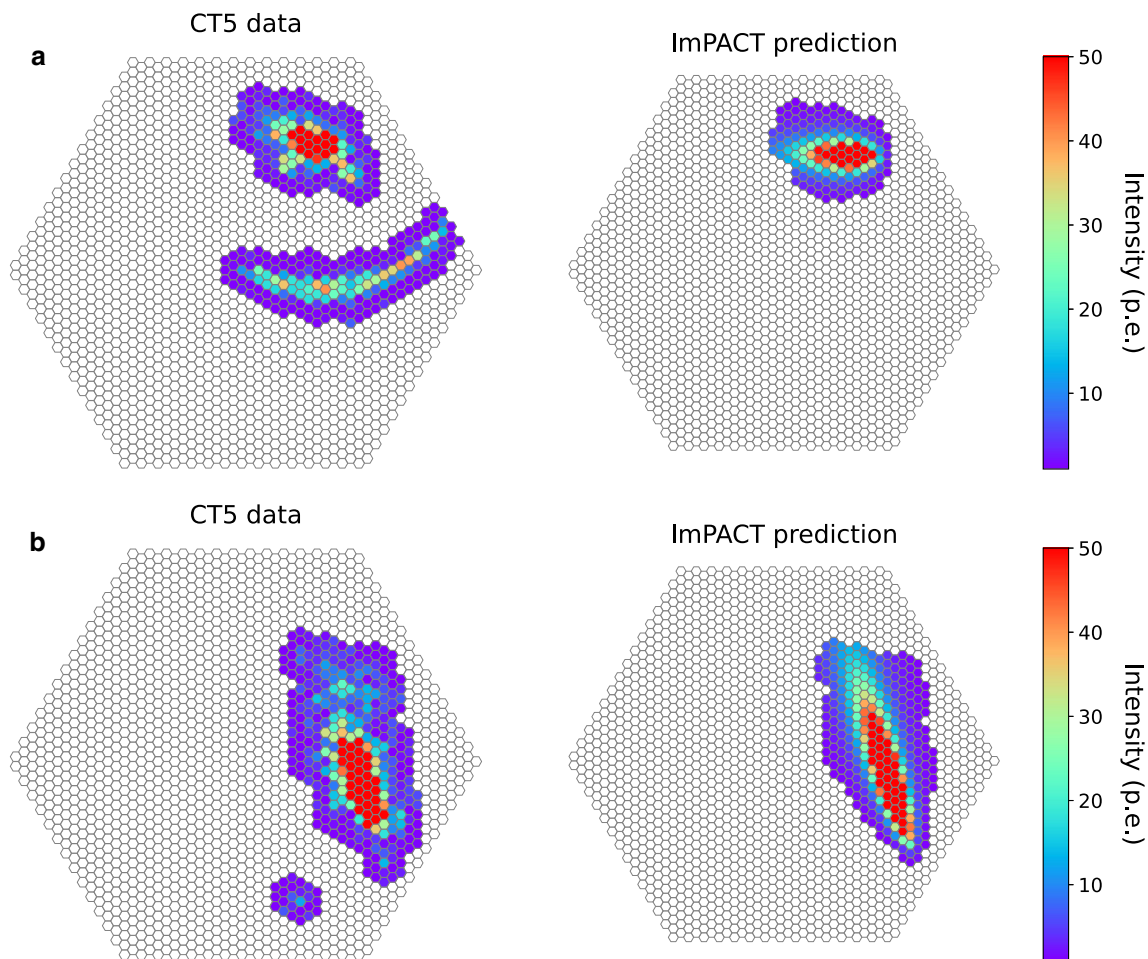
Fig. 1 Large telescope images of rejected events that would be labelled as gamma-ray candidates in the small telescope reconstruction. In both cases, the right panel shows the ImPACT prediction associated with that event, whereas the left panel shows the actual recorded event image. The event image shown in the panel labelled as "a" contains a clear muon arc, whereas the additional feature in the image in panel "b" is much smaller and has a simpler shape. However, the maximum pixel intensity in this feature is more than double what is expected from NSB noise. Note that the colorbar has been restricted so that the fainter features are visible, since the main shower is much brighter in both cases

from particles other than muons – electrons, for example – may be used to reject events. Additionally, unusually high NSB noise could lead to an event being rejected. However, it results in a much simpler implementation which, as shown below, still leads to improved background rejection power.

For this task we used data and simulations from the telescopes in the High Energy Stereoscopic System (H.E.S.S.) [6]. The H.E.S.S. experiment is comprised of a total of five telescopes: four with a dish of 12 m diameter referred to as CT1-4 and a central one, CT5, with a dish of 28 m diameter. The event reconstruction is referred to as *stereo* when only the small CT1-4 telescopes are used, and *hybrid* when data from the entire array is used. In this work we present an alternative approach in which only the data of CT1-4 is used for the event reconstruction, that is, *stereo* mode, but the data of CT5 is used for an extra step of background rejection.

2 Background rejection with image residuals: *ABRIR*

ABRIR (*Algorithm for Background Rejection using Image Residuals*) is a background rejection algorithm which draws additional information from the event image taken by a large telescope, such as CT5 for the case of the H.E.S.S. array. This algorithm is applied after the usual H.E.S.S. *stereo* reconstruction, which includes an initial step of background rejection based on Boosted Decision Trees (BDTs) [2]. In particular, we use two sets of initial cuts: the H.E.S.S. *standard* selection cuts (see Section 4.2 of [2]), optimized for a Crab Nebula-like source and the so-called *hard* cuts, optimized for a faint hard source. The main difference between both sets of cuts is that while the *standard* cuts only require images to have a total intensity larger than 60 photoelectrons (p.e.), the *hard* cuts require a minimum of 200 p.e. This translates into a higher energy threshold but also higher quality events and
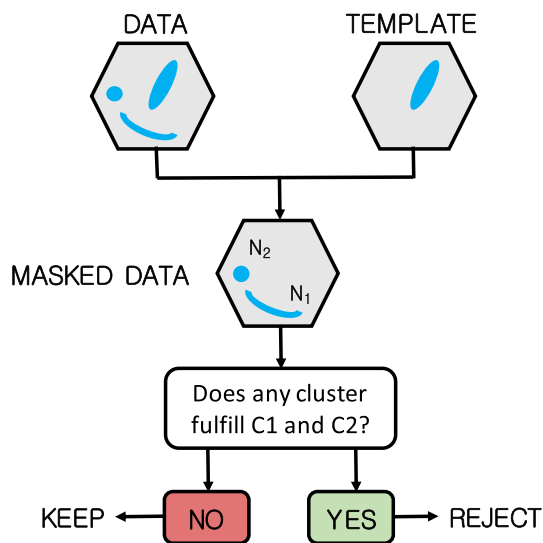
Fig. 2 Schematic diagram representing the algorithm structure

more precise reconstructed parameters, such as core location or energy. Additionally the threshold of the BDT parameter [2] is different, with 0.84 for the *standard* case and 0.8 for the *hard* cut. We will denote the gamma-ray and background efficiency of this first cut as $\eta_{BDT,\gamma}$ and $\eta_{BDT,B}$ respectively. We will apply ABRIR only to the events that survive this initial cut. Only at this step is the data from CT5 used, and it is exclusively CT5 images being considered by the algorithm due to the advantaged muon detection capabilities of large telescopes. Note that the total gamma-ray and background efficiency ($\eta_{tot,\gamma}$ and $\eta_{tot,B}$) needs to be computed as the product of that of the initial cut and that of ABRIR, so:

$$\eta_{tot,\gamma} = \eta_\gamma \cdot \eta_{BDT,\gamma}$$
$$\eta_{tot,B} = \eta_B \cdot \eta_{BDT,B}$$

where η_γ and η_B are the gamma-ray and background efficiency of ABRIR (see Sect. 3).

2.1 The ImPACT templates

In order to identify light as not a part of the main shower, we need to identify the main shower itself. We do this with the help of the *Image Pixel-wise fit for Atmospheric Cherenkov Telescopes* (ImPACT) algorithm [7], which is routinely used by the H.E.S.S experiment. ImPACT is a gamma-ray event reconstruction algorithm that is based on the likelihood fitting of camera pixel amplitudes to an expected image template. These templates are built with a full Monte Carlo gamma-ray air shower simulation, followed by ray-tracing of the telescope optics and simulation of the instrument electronics [8,9]. This fit is computationally expensive, so it is only performed after a first round of background rejection is applied. For each event this process results in an image of what the

most similar gamma-ray event would look like in each of the telescopes. We use this prediction to mask the main component of CT5 image in order to better identify residual features. Note that the CT5 image is not used in this likelihood fit, with the CT5 template being derived only from the best-fit parameters of the CT1-4 fit.

2.2 The rejection criteria

After the main shower is masked, the remaining residual image is searched for clusters of pixels (N). A cluster is only considered if it is comprised of more than three neighboring pixels. For each of the clusters we compute the maximum intensity $I_{\max,N}$, the total intensity $I_{tot,N}$ and the distance to the main shower d_N. In the next step, two conditions are checked against each of the clusters:

(C1) $I_{tot,N} \cdot d_N^2 > I_{tot} \cdot d^2$
(C2) $I_{\max,N} > I_{\max}$

where the threshold values used here for $I_{\max}$ and $I_{tot} \cdot d^2$ are 9 p.e. and 2 p.e $\cdot$ pixel^2 respectively. These values were selected to maintain a gamma-ray efficiency of around 90%, which in turn results in the background rejection performance described in Sect. 3. A different requirement on the gamma-ray efficiency would naturally impact the background rejection power, with better rejection achievable if a higher fraction of gamma-rays is lost. Additionally, these values are specific to the current state of the central telescope in the H.E.S.S. array and should be adjusted as appropriate for application to other telescopes or cameras.

The first condition (C1) selects clusters by their total intensity with a penalization on proximity to the main shower. This is done to reject clusters created solely by a mismatch between the outer row of pixels of the main shower image and the predicted template. Such clusters might have a high total intensity, but they are also very close to the main shower, which means the product of intensity and distance will fall below the required threshold. Additionally, this cut is always survived by relatively bright clusters that are far from the shower, and thus have a high probability of not originating from a mismatch between the data and the template. The second condition (C2) aims to reject clusters resulting from uncleaned noise from the NSB. It does so by requiring the maximum pixel intensity of a given cluster to be above a threshold defined by the 5σ level of the pedestal width of a typical run. Note that this specific value depends on the image cleaning method used, as well as on the camera performance and expected NSB level.

Events that fail either of the two conditions are kept, as well as events with too small or no clusters outside of the main image. A schematic view of the algorithm is shown in Fig. 2. Additionally, events for which no light in the large telescope

was predicted by the template, but something is seen are also marked as rejected, as well as the opposite case, events in which an shower image was predicted, but all pixels in the data are below a minimal threshold of 1 photoelectron.

3 Performance

We asses the performance of the algorithm by testing it on different types of events. Note that both the background rejection power and the gamma-ray efficiency presented in this section are a relative improvement on the value for the initial H.E.S.S. *standard* and *hard* selection cuts. Those curves are shown in Fig. 6.

3.1 Background rejection power

In order to test the performance of the algorithm on the cosmic-ray background, we use what is referred to as *off-runs*, that is, observations in which the telescopes are pointed at fields without a known gamma-ray source. These off-runs are typically the result of dedicated observations of known empty fields, or also observations of extragalactic objects like dwarf spheroidal galaxies that did not yield a gamma-ray detection [10]. These are not completely free of gamma-ray events, due to large-scale diffuse emission like the extragalactic gamma-ray background and also due to possible undetected faint gamma-ray sources. However, the relative fluxes of background cosmic-rays to these gamma-ray sources make it safe to neglect this contribution. It would also be possible to use simulations of proton-initiated showers in order to eliminate this contamination. However, due to uncertainties in the hadronic models used for high-energy particle interactions [11,12], there are significant differences between the properties of measured background events and those that result from simulations. For this reason we decide to use real data only to characterize the background, as it is more realistic and not dependent on the agreement between data and simulations.

Figure 1 shows two example rejected events from one of these off runs which pass the *standard* H.E.S.S. cuts. The right panel shows the ImPACT prediction associated with each event, whereas the left panel shows the actual recorded event image. For both cases, the conditions C1 and C2 are fulfilled.

Figures 3 and 4 show the ABRIR cut efficiency for different samples as a function of the event reconstructed energy for several observation zenith angles. Figure 3 corresponds to the case where events surviving the H.E.S.S. *standard* cuts are given as input, whereas the input events for the curves in Fig. 4 survive the *hard* cuts. The efficiency shown in the figures is calculated simply as the ratio of the number of events before and after the application of ABRIR. As can be seen,

the cut rejects a significant fraction of off-run events in both cases (red line) at all energies. Background rate reductions of up to a factor 2.5 are obtained for energies above a few TeV for the *standard* cuts case, while when using only the higher quality events selected by the *hard* cut as input, this improvement goes up to factors of 3 and even 4 for energies above tens of TeV.

3.2 Gamma-ray efficiency

Besides rejecting a lot of suspected background events, it is important to keep a high fraction of the true gamma-ray events. Some gamma-ray events are expected to be flagged as rejected by the algorithm, and they belong to three different categories:

1. **Gamma-ray events that contain muon light.** As shown by [4], a small fraction of high-energy gamma-ray initiated showers will actually contain muons, which can be detected by the large telescope. These events make up an irreducible set of lost gamma-ray events associated with muon-tagging based rejection.
2. **Gamma-ray events with low-altitude electrons.** Since our approach does not identify muon light as such, the light from any other particle that would create a component in the image besides the gamma-ray shower could lead to an event being rejected. Camera images of gamma-ray showers can contain light from scattered electrons from the shower that emit close to the telescope dish. This can create additional image components separated from the main shower image. The likelihood of this effect drops rapidly as the shower core ground impact point moves away from the telescope position.
3. **Gamma-ray events with unusually high NSB noise.** A fraction of gamma-ray events will contain noise from the NSB that is brighter than the threshold set by condition C2 from Sect. 2.2. This can happen in the simulations due to fluctuations in the NSB level, but it is also important in real data, especially since the NSB is not the same for all regions of the sky. This fraction is thus heavily dependent on the NSB conditions that were assumed when the algorithm's parameters were chosen, and also on the type of cleaning that is used to remove the NSB contribution to the images.

We compute the gamma-ray efficiency of ABRIR by running it on simulated gamma-ray events, using the CORSIKA package [9] for shower and Cherenkov light simulation and the *sim-telarray* package [8] for the telescope response and camera simulation. In order to check the consistency of the result in real data, we also check with a sample of events which reconstructed direction falls within 0.2° of the Crab Nebula and PKS 2155-304, which are known bright gamma-
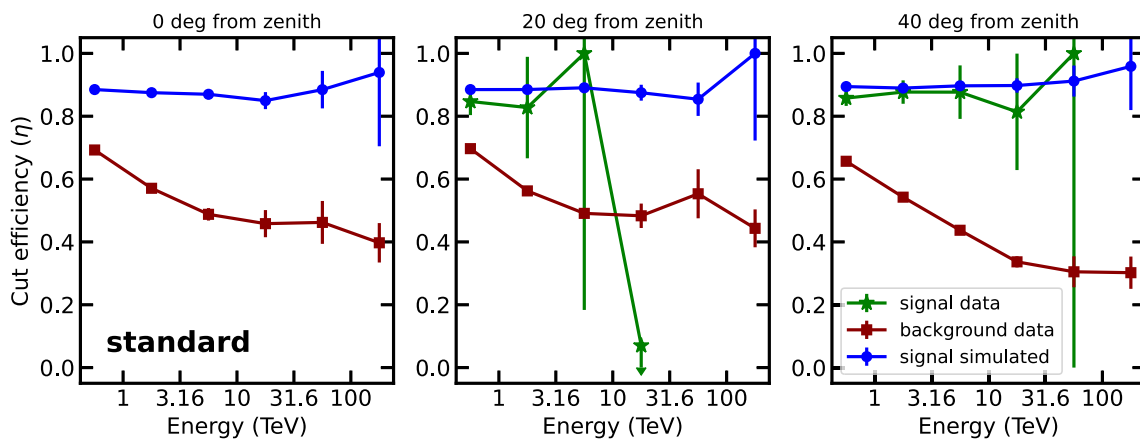
Fig. 3 Fraction of events kept by the ABRIR cut applied after the H.E.S.S. *standard* cuts for simulated gamma-rays (blue dots), background data from off-runs (red squares) and events taken from a radius of 0.2° from bright gamma-ray sources (green stars), in particular the blazar PKS 2155-304 at zenith angles of 20° and the Crab Nebula for the 40° zenith range. Note that since PKS 2155-304 is an extragalactic source, no gamma-rays are expected to arrive from it above a few TeV due to absorption on the extragalactic background light. When zero events survive the cut, the 68% containment limit is drawn assuming Poissonian statistics
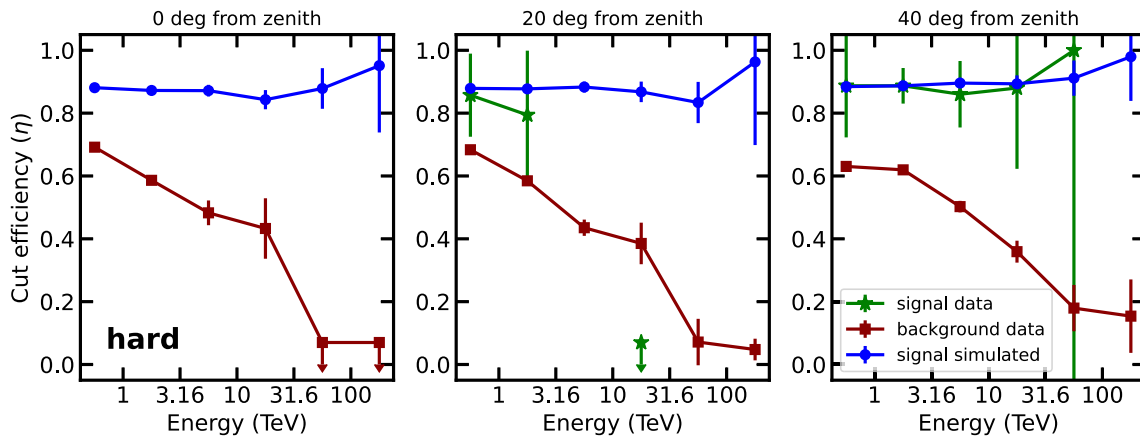
Fig. 4 Fraction of events kept by the ABRIR cut applied after the H.E.S.S. *hard* cuts. Meaning of different panels and colors is the same as in Fig. 3. When zero events survive the cut, the 68% containment limit is drawn assuming Poissonian statistics

ray sources. Note that this sample is expected to contain some small fraction of cosmic-ray background, as cosmic-ray events are distributed roughly uniformly on the sky. However, due to the relatively high gamma-ray flux of these sources, the majority of events in this sample will be true gamma-rays, and thus can be used to check the performance on real data events. Note that this comparison is restricted to the zenith angles in which these sources can be observed by the H.E.S.S. experiment and also by the energies reached by the sources themselves.

Figures 3 and 4 show the ABRIR cut efficiency for the simulated gamma-ray sample (blue lines) as well as the events selected around known gamma-ray sources (green line). Note that the gamma-ray efficiency is mostly flat as a function of energy. This indicates that the group dominating the rejected

gamma-rays are those with residual features caused by the NSB, since, unlike the expected number of muons, this does not depend on the event energy. The gamma-ray efficiency computed with real data is consistent with that found in simulations, which in turn confirms that the simulated NSB level is consistent with that encountered in a typical field.

3.3 Application to a gamma-ray source: the Crab Nebula

In order to verify the expected improved performance, we perform an analysis of the same dataset from a real gamma-ray source with and without the use of ABRIR and compare the results. For this, we use a total of 30 h of observations of the Crab Nebula by the full array of the H.E.S.S. telescopes. Note that since the goal of this paper is to demonstrate the
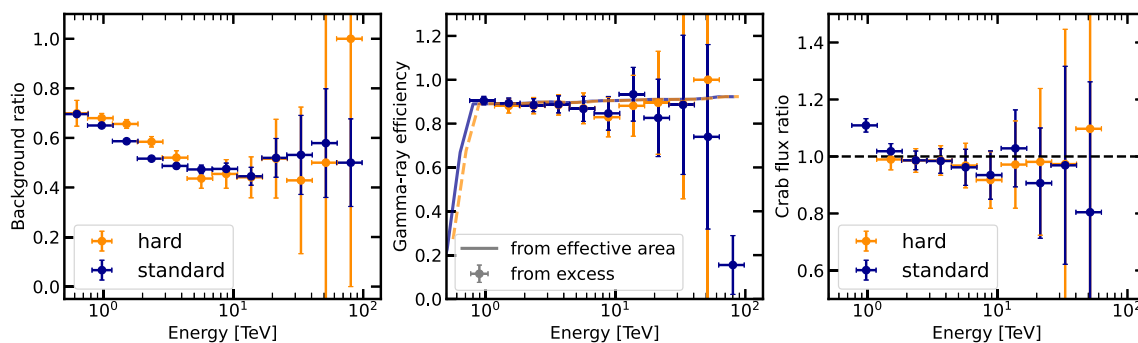
Fig. 5 Verification of the performance using data from the Crab Nebula. All ratios shown in this figure are computed as the quantity after applying the ABRIR cut divided by the same quantity before the cut. *Left*: Ratio of background counts with and without the use of ABRIR for both the *standard* (dark blue) and *hard* (orange) cuts. *Middle*: Gamma-ray efficiency computed as the ratio of the resulting effective area for the datasets with and without ABRIR (solid lines) and as the ratio of the measured excess (data points), again for both sets of cuts using the same color scheme as in the left panel. *Right*: Ratio of the measured flux from the Crab Nebula as a function of energy with and without the use of ABRIR for both sets of initial cuts

relative performance of this analysis technique, only ratios of the relevant quantities will be shown.

In both cases, we first perform a *stereo* reconstruction of the data which excludes the large central telescope. This results in a list of events which, in the case without ABRIR are directly bundled with the relevant instrument response functions (IRFs) via the *Gammapy* [13] package, as described in [14]. When applying ABRIR, an extra step is performed in which the algorithm is applied, the flagged events are removed from the event list and the effective areas are corrected by the gamma-ray efficiency computed from simulations. The new event lists and IRFs are bundled in the same way as before.

Once both datasets are ready, we can extract and compare different quantities from them. First, we estimate the background level in a region of 0.3° radius around the Crab location using the so-called *reflected-region* background method (see [15] for a detailed description). This choice of region size is relatively large for a standard H.E.S.S. point-source analysis, but allows for increased statistics on the background measurement. The ratio of the measured background before and after applying ABRIR as a function of energy can be seen in the left panel of Fig. 5, for both sets of initial cuts. The resulting background efficiency is consistent with the one derived from the off runs shown in Figs. 3 and 4 within statistical errors. Using this measurement of the background, we can compute the excess counts in the source region, and compare it to the ratio of the effective areas. This is an additional verification that the performance on simulated gamma-rays is observed in real data. The result can be seen in the middle panel of Fig. 5. Finally we can compute the flux from the Crab nebula as a function of energy. The rightmost

panel of Fig. 5 shows the ratio of flux measured before and after applying ABRIR as a function of energy for both sets of cuts. As can be seen, they are consistent with a ratio of unity within statistical errors, with the exception of the lowest energy point in the *standard* result. This indicates a possible mischaracterization of the effective area at the lowest energies. Given that effective areas were not produced from scratch for the ABRIR dataset, but rather derived from existing ones using an energy-dependent correction, it is possible that the region near threshold, where the effective area raises rapidly, is less accurately characterized. Dedicated IRFs or an increased threshold would solve this issue.

3.4 Comparison to *hybrid* reconstruction

The background rejection powers achieved with the combination of the baseline *stereo* algorithms and ABRIR are superior than those obtained when data from the large telescope is included in the reconstruction from the beginning – i.e. *hybrid* reconstruction. This can be seen when comparing the combination of the ABRIR efficiencies shown in Figs. 3 and 4 and the baseline *stereo* efficiencies shown in Fig. 6 with those of the *hybrid* analysis shown in Fig. 7. The reason for this lack of performance is that combining information from different telescope sizes at the reconstruction level is a non-trivial task, especially for the algorithms that rely on image shape parameters. Ongoing improvements to the H.E.S.S. *hybrid* chain have shown promising results and may achieve a performance comparable to that of ABRIR in the low energy range. However, at high energies ABRIR will likely continue to provide improved performance beyond that achieved by a hybrid analysis based on the standard image parameters.

4 Discussion

We have presented an algorithm that makes use of large-dish telescopes in an IACT array as a veto step in order to improve the background rejection. Its use improves the background rejection power of the baseline H.E.S.S. *stereo* reconstruction by a factor ranging between 2 and 4, depending on energy and the specifics of the initial cut. The combined efficiency of this extra cut with the baseline rejection power of the *stereo* reconstruction can reach background rejection powers higher than 10^4 for high energies. Improved background rejection is crucial for the detection and characterization of extended and faint sources. Reducing the background rate also reduces the uncertainties associated with it, leading to improved precision at high energies. This improvement can be decisive in order to determine, for example, the presence of energy-dependent morphology.

The different performance at high energies depending on the chosen initial cut for the *stereo* reconstruction is due to their different image amplitude thresholds. The *standard* cuts keep events whose image has a total brightness of more than 60 p.e. in two of the four telescopes at least. The threshold for the *hard* cuts is 200 p.e. The image of an event can have low brightness either because the primary particle had relatively low energy, or because the location of the shower core is far from the array. This means that some events that are below the 200 p.e. threshold will be reconstructed with high energies but large core distances. At large core distances, the number of muons that are detectable by the telescope will be reduced [4], meaning that the power of the veto approach is reduced.

All results shown here are based on images for which the noise has been cleaned using the so-called *tail-cuts* cleaning method, which requires each pixel to have an intensity exceeding a threshold I_1 and a neighbour exceeding a threshold I_2 [16]. In particular we have used $I_1 = 7$ p.e. and $I_2 = 4$ p.e., as required for the ImPACT algorithm [7]. As mentioned in Sect. 3, remaining NSB noise seems to be the leading cause of loss of gamma-ray events when applying ABRIR. A different image cleaning approach, such as those based on the time information of the shower image [17] could improve these results.

The presence of broken or unusable pixels in the camera, that is regions of the camera which do not take data, could in principle impact the ability to identify the presence of additional charge away from the shower [16]. However, in the case of the HESS CT5 camera, the number of such pixels is usually very small. There are just 6 isolated permanently turned off pixels, and for certain sky regions, the presence of bright stars can result in one or two more pixels being turned off [18]. This leads to a very small effect given that the camera has 1764 pixels in total and muons are uniformly distributed across it.

We have shown here the performance of ABRIR when applied to the H.E.S.S. array, which is made up of four middle-sized telescopes and one large, central telescope. The future Cherenkov Telescope Array (CTA) will be made up of telescopes of three sizes, the largest of which is smaller than the central H.E.S.S. telescope yet still large enough to efficiently detect muons. The idea behind the algorithm will thus be applicable to the CTA array, although whether or not will it be a competitive technique will depend on the analysis approaches used for mixed-telescope types and sub-array selection of the observations. Large telescopes are typically seen as an asset only at the low energies due to their reduced threshold. However, as we show here, including large telescopes when observing targets for which the spectrum is expected to extend to high energies is worthwhile given the achievable improvements especially in background rejection.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: The data and simulations used for this work are propriety of the H.E.S.S. Collaboration.]

Appendix A: Efficiencies of the baseline H.E.S.S. cuts

We computed the background and gamma-ray efficiencies by running the H.E.S.S. analysis on the off run events and simulated gamma-rays with and without the baseline Boosted Decision Tree (BDT) cuts. The efficiency is then computed as the ratio of the events passing the BDT cut to the initial number of events. Note that this initial number comprises only events surviving the respective image amplitude threshold for each of the cut sets (60 p.e. for *standard* cuts and 200
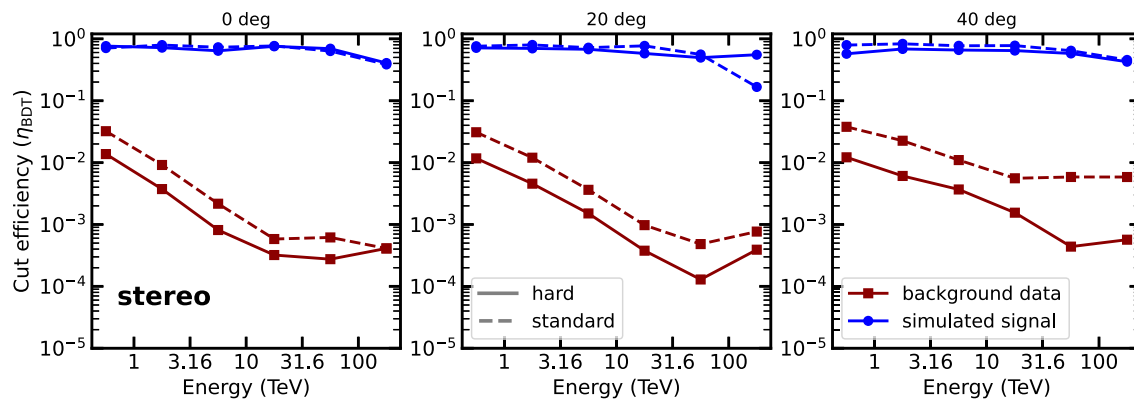
Fig. 6 Cut efficiencies of the *stereo* reconstruction for the *standard* and *hard* cuts
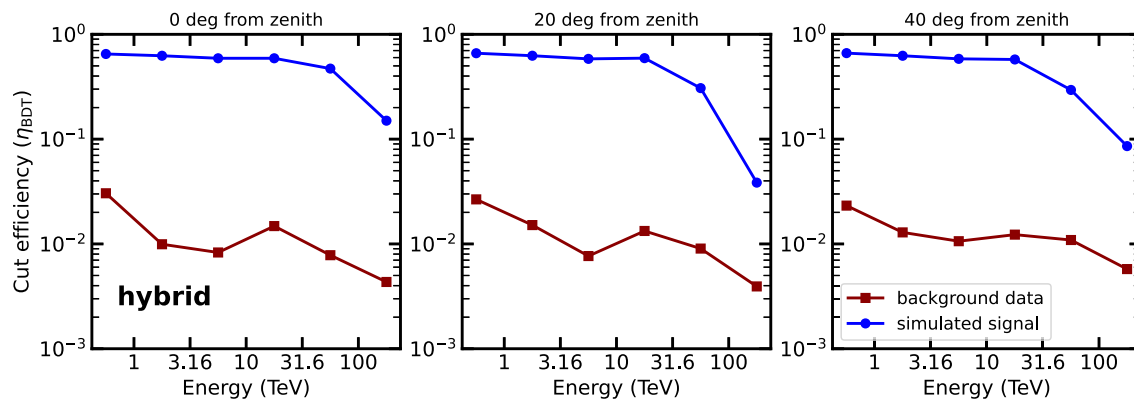
Fig. 7 Cut efficiencies of the *hybrid* reconstruction using the *standard-hybrid* cuts

p.e. for the *hard* cuts). The resulting efficiencies are shown for different energies and zenith ranges in Figs. 6 (*stereo* reconstruction) and 7 (*hybrid* reconstruction). Note that the cuts used in the *hybrid* case, so-called *standard-hybrid* cuts, are different from the *stereo* ones in terms of thresholds, but are optimized for the same science case: a Crab Nebula-like source.

References

1. W. Benbow, H. Collaboration, The H.E.S.S. standard analysis technique. In: Proc. Conf. Towards a Network of Atmospheric Cherenkov Detectors VII, Proc. Conf. Towards a Network of Atmospheric Cherenkov Detectors VII, vols. 163–172 (2005)
2. S. Ohm, C. van Eldik, K. Egberts, γ/hadron separation in very-high-energy γ-ray astronomy using a multivariate analysis method. Astropart. Phys. **31**(5), 383 (2009). https://doi.org/10.1016/j.astropartphys.2009.04.001
3. T.K. Gaisser, T. Stanev, F. Halzen et al., Gamma-ray astronomy above 50 TeV with muon-poor showers. Phys. Rev. D **43**(2), 314 (1991). https://doi.org/10.1103/PhysRevD.43.314
4. L. Olivera-Nieto, A.M.W. Mitchell, K. Bernlöhr et al., Muons as a tool for background rejection in imaging atmospheric Cherenkov telescope arrays. Eur. Phys. J. C **81**(12), 1101 (2021). https://doi.org/10.1140/epjc/s10052-021-09869-0
5. R. Mirzoyan, D. Sobczynska, E. Lorenz et al., Tagging single muons and other long-flying relativistic charged particles by ultra-fast timing in air Cherenkov telescopes. Astropart. Phys. **25**(5), 342 (2006). https://doi.org/10.1016/j.astropartphys.2006.03.006
6. M. de Naurois, H.E.S.S.-II—Gamma ray astronomy from 20 GeV to hundreds of TeV's. In: European Physical Journal Web of Conferences, European Physical Journal Web of Conferences, vol. 136, 03001 (2017). https://doi.org/10.1051/epjconf/201713603001
7. R. Parsons, J. Hinton, A Monte Carlo template based analysis for air-Cherenkov arrays. Astropart. Phys. **56**, 26 (2014). ISSN 0927-6505. https://doi.org/10.1016/j.astropartphys.2014.03.002
8. K. Bernlöhr, Simulation of imaging atmospheric Cherenkov telescopes with CORSIKA and sim_ telarray. Astropart. Phys. **30**(3), 149 (2008). https://doi.org/10.1016/j.astropartphys.2008.07.009
9. D. Heck, J. Knapp, J.N. Capdevielle et al., CORSIKA: a Monte Carlo code to simulate extensive air showers. Forschungszentrum Karlsruhe GmbH (1998)
10. A. Abramowski, F. Aharonian, F. Ait Benkhali et al., Search for dark matter annihilation signatures in H.E.S.S. observations of dwarf spheroidal galaxies. Phys. Rev. D **90**, 112012 (2014). https://doi.org/10.1103/PhysRevD.90.112012
11. R.D. Parsons, H. Schoorlemmer, Systematic differences due to high energy hadronic interaction models in air shower simulations in the 100 GeV–100 TeV range. Phys. Rev. D **100**(2), 023010 (2019). https://doi.org/10.1103/PhysRevD.100.023010
12. Á. Pastor-Gutiérrez, H. Schoorlemmer, R.D. Parsons et al., Sub-TeV hadronic interaction model differences and their impact on air showers. Eur. Phys. J. C **81**(4), 369 (2021). https://doi.org/10.1140/epjc/s10052-021-09160-2

13. C. Deil, A. Donath, R. Terrier et al., gammapy/gammapy: v.0.18.2 (2020). https://doi.org/10.5281/zenodo.4701500

14. L. Mohrmann, A. Specovius, D. Tiziani et al., Validation of open-source science tools and background model construction in γ-ray astronomy. Astron. Astrophys. **632**, A72 (2019). https://doi.org/10.1051/0004-6361/201936452

15. D. Berge, S. Funk, J. Hinton, Background modelling in very-high-energy γ-ray astronomy. Astron. Astrophys. **466**(3), 1219 (2007). https://doi.org/10.1051/0004-6361:20066674

16. F. Aharonian, A.G. Akhperjanian, A.R. Bazer-Bachi et al., Observations of the Crab nebula with HESS. A&A **457**(3), 899 (2006). https://doi.org/10.1051/0004-6361:20065351

17. M. Shayduk, C. Consortium, Optimized next-neighbour image cleaning method for cherenkov telescopes. In: International Cosmic Ray Conference, International Cosmic Ray Conference, vol. 33, 3000 (2013)

18. B. Bi, M. Barcelo, C. Bauer et al., Performance of the new FlashCam-based camera in the 28m telescope of H.E.S.S. In: 37th International Cosmic Ray Conference, 743 (2022). https://doi.org/10.22323/1.395.0743

19. C.R. Harris, K.J. Millman, S.J. van der Walt et al., Array programming with NumPy. Nature **585**, 357–362 (2020). https://doi.org/10.1038/s41586-020-2649-2

20. P. Virtanen, R. Gommers, T.E. Oliphant et al., SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods **17**, 261 (2020). https://doi.org/10.1038/s41592-019-0686-2

21. W. McKinney et al., Data structures for statistical computing in python. In: Proceedings of the 9th Python in Science Conference, vol. 445, pp. 51–56, Austin (2010)

22. J.D. Hunter, Matplotlib: a 2D graphics environment. Comput. Sci. Eng. **9**(3), 90 (2007)

23. J. Pivarski, H. Schreiner, N. Smith et al., scikit-hep/uproot4: 4.2.3 (2022). https://doi.org/10.5281/zenodo.6540484

24. K. Kosack, J. Watson, M. Nöthe et al., cta-observatory/ctapipe: v0.11.0 (2021). https://doi.org/10.5281/zenodo.4746317